52 5225

nchs

Proceedings of the 1993 Public Health Conference on Records and Statistics

Toward the year 2000

Refining the measures

Mayflower Hotel • Washington D.C.

July 19–21, 1993

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
Centers for Disease Control and Prevention
National Center for Health Statistics

CDC
CENTERS FOR DISEASE CONTROL AND PREVENTION

5225

nchs

Proceedings of the 1993 Public Health Conference on Records and Statistics

Toward the year 2000

Refining the measures

Mayflower Hotel • Washington D.C.

July 19–21, 1993

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
Centers for Disease Control and Prevention
National Center for Health Statistics

Each article has been prepared by the author.

Toward the year

2000

Refining the measures

Manning Feinleib, M.D., Dr.P.H.
Director
National Center for Health Statistics

1993 PUBLIC HEALTH CONFERENCE ON RECORDS AND STATISTICS PROGRAM COMMITTEE

CHAIRPERSON

Jack R. Anderson
Acting Deputy Director
National Center for Health Statistics

MEMBERS

Lester R. Curtin
Peter L. Hurley
Sandra S. Smith
James A. Weed
P. Douglas Williams
Ronald W. Wilson

PUBLIC HEALTH CONFERENCE ON RECORDS AND STATISTICS STAFF

Jacquelyn Adler
Robin T. Barnes
Barbara P. Butler
Elizabeth T. Darling
Don R. DeRocha
Nancy G. Hamilton
Barbara W. Hetzler
Thomas S. Vissman
Brenda L. Wolfrey

EXHIBITS AND PUBLICATIONS SUPPORT STAFF

Joan D. Burton
Annette F. Gaidurgis
Sarah M. Hinkle
Patricia E. Keaton-Williams
Sabra F. Lloyd
Stephen L. Sloan

114

TABLE OF CONTENTS

ACKNOWLEDGMENTS v

OPENING CEREMONY
- Introductory Remarks by Manning Feinleib, Walter R. Dowdle, and Clifton R. Gaus 3

FIRST PLENARY SESSION: SPAN OF HEALTHY LIFE
- Valuing Lives, Valuing Health - Donald L. Patrick 12
- Tough Decisions: Tradeoffs Involving Years of Healthy Life - Jane E. Sisk 17

Session A
- Penniffer Erickson *
- The Calculation of Health-Adjusted Life Expectancy for Ontario Using a Health Status Index - Roger Roberge*
- Accounting for Deaths When Measuring Health Status Over Time - Paula Diehr* 34

Session B
- Underlying Causes of Fetal Death: Effect of an Additional Panel of Check-Boxes on the Wisconsin 1989 Revised Fetal Death Report Form - Russell S. Kirby*..... 41
- Status and Effectiveness of Perinatal Regionalization in Virginia - Sherry Allison-Cooke 46
- A Methodological Approach to Improving Pregnancy Outcomes in African American Women - Cecile H. Edwards*..... 52
- Intermediate Variables as Determinants of Adverse Pregnancy Outcome in High Risk Inner City Populations - Ravi K. Sharma 55
- Discussant: Michelle A. Williams., ... 60

Session C - CHILDREN AND YOUTH
- The "Bright Smiles, Bright Futures" Curriculum Evaluation: Designing Health Education Evaluations Which Address the Methodological Challenges of Collecting Data From Young Children - Robert S. Gold*..... 63
- Effects of Maternal Education and Access to Medical Care on Analysis of Injury Risk According to Patterns of Child Care - Mary D. Overpeck* 67
- Drug and Alcohol Findings: New York State Students in Fifth and Sixth Grades - Blanche Frank* 73
- Homeless Street Youth: Strategies and Obstacles in Gathering Data on Antecedents to Homelessness - Therese van Houten 78

Session D - MENTAL HEALTH INDICATORS-FROM EPIDEMIOLOGY TO SERVICE DELIVERY
- The Social Costs of Anxiety Disorders - Andrew C. Leon* 87
- Reflections on Appropriate Mental Health Questions for the Health Interview Survey - James S. Larson..... 93
- Specialty Mental Health Service Utilization in the United States - Marilyn J. Henderson..... 99

Session E - FAMILY AND HEALTH
- Parental Factors Influencing Patterns of Prenatal Care Utilization - Peter T. D'Ascoli* 105
- Family and Cultural Factors in Mental and Physical Comorbidity in Children - Jacqueline L. Angel* 109
- Assessing Family Strengths and Youth Behavior: Analyses of Three National Databases on Children and Families - Barbara W. Sugland* 115
- Discussants: Martha Farnsworth Riche and Karen Smith Thiel..... 120

Session F - NUTRITION
- Assessing Fruit and Vegetable Intakes: Toward the Year 2000 - Susan Krebs-Smith*..... 127
- Methodological Issues in Analyzing School Menus - Patricia McKinney..... 128
- Monitoring Progress Toward Meeting Year 2000 Objectives: NHANES III, 1988-91 - Clifford L. Johnson* 130
- Monitoring the Nation's Health: The 1991 Health Survey for England-Results and Methodology - Amanda White 133

*Speaker of coauthored paper.

Session G - INJURIES AND VIOLENCE
-Reporting External Cause of Injury and Poisoning Codes in Medicare Hospital Discharge Data - Harry L. Savitt* 141
-Emergency Department Weapon-Related Injury Surveillance - Victoria Vespe Ozonoff* 145
-Surveillance for Nonfatal Firearm Injuries - Roberta K. Lee* 149
-Developments in the Surveillance of Central Nervous System Injuries - David J. Thurman*... 155

Session H - MEASURING THE HEALTH AND DISABILITY OF OLDER AMERICANS
-Health Expectancy: The Canadian Experience - Russell Wilkins 161
-Compression of Disability: Evidence from the National Long Term Care Surveys - Larry Corder 167
-Compression of Disability? Results from the Established Populations for Epidemiologic Studies of the Elderly - Marcel Edward Salive* 173
-Discussants: William Scanlon and Lois M. Verbrugge..... 177

SECOND PLENARY SESSION - MEASURING HEALTH DISPARITIES
-Using Data on Health Disparities in the Formulation of Public Health Policy - Mohammad N. Akhter..... 181
-Closing Gaps in Health and Health Care: Improved Data for a New Health System - David Mechanic., ... 182

Session I - WHAT'S NEWS IN HEALTH STATISTICS?
-Telling the Truth From the Trash - Victor Cohn..... 189
-Other Panelists: Alan L. Otten and Rhonda Roland..... 190

Session J - RACE-ETHNICITY CLASSIFICATION
-American Indian Infant Mortality Review in Wisconsin: Background, Problems and Process - Raymond D. Nashold*..... 193
-Tribe-Specific Coding and Reporting in New Mexico - Corazon M. Halasan 199
-Racial Disparities in Birth Outcomes: Race, Ethnicity, or Nativity - T. Joseph Sheehan*..... 205
-Underestimation of Racial Minorities in U.S. Army Records - Claire C. Gordon* 210

Session K - PEOPLE WITH DISABILITIES
-Contribution of Developmental Disabilities to Childhood Mortality in the United States: A Multiple Cause-of-Death Analysis - Coleen A. Boyle* 213
-Definitions of Disability and Their Measurement and Operationalization in Survey Data - Barbara M. Altman 219
-Applications of the CHART: Craig Handicap Assessment and Reporting Technique - Gale G. Whiteneck 225
-Applications of the International Classification of Impairments, Disabilities, and Handicaps: From Concepts to Taxonomy to Data - Scott Campbell Brown 228
-Discussant: Mary Chamie..... 234

Session L - HEALTH STATISTICS FROM RECORD SYSTEMS
-Access to Prenatal Care Following Major Medicaid Eligibility Expansions - Trude Bennett* . . 239
-County-Wide Small Area Analyses: Estimation of Denominators for Users of a County Health Care Agency and Their Application in Health Status Measures - Arthur J. Davidson* 244
-Measuring Work-Related Health Disparities for Minority Populations - Carol A. Burnett*. ... 249
-Cognitive Issues in Improving the Quality of Mortality Statistics - Jared B. Jobe* 255

Session M - NUMERATOR-DENOMINATOR LINKAGE
-The Relationship Between Maternal HIV Seroprevalence and the Hospitalization of HIV Infected Children in Their First Year of Life - Vito M. Logrillo* 263
-Surgical Procedure Rates with Numerators from Hospital Data and Denominators from Several Different "At Risk" Populations - Jane F. Gentleman* 269
-Generalized Linear Models and Simultaneous Surveys - Wilfred L. Rosenbaum*..... 275
-Benefits of Probabilistic Linkage of Electronic Patient Records for Highway Safety - William H. Walsh, Jr..... 281

* Speaker of coauthored paper.

Session N - INTERNATIONAL COMPARISONS
-Trends in Mortality in East and West Germany During the 3 Decades Before Unification - Karl E. Bergmann* 287
-Mortality in the United States and France: 1988 - Alvan 0. Zarate* 288
-The International Rank of the United States in Infant Mortality: Are We Really 24th? - Robert B. Hartford 293
-Health Expectancy of the Elderly in Xichang, China and Canada: Comparisons Based on Measures of Autonomy in Various Activities of Daily Living - Edward Ng*. 294

Session 0 - ISSUES IN CONFIDENTIALITY
-Privacy Policy in an Electronic World - Joan Turek-Brezina 303
-Towards Compiling Better Health Indicators through Data Linkage - Brian K. Gallagher*, ... 304
-The Unique Record Number: An Alternative to Traditional Person Identifiers - Andrew Asher* 308
-California's Experience: SSN and Confidentiality - Eleanor Meux 312

Session P - METHODOLOGICAL ISSUES FOR NATIONAL MINORITY DATA
-Design of the National Health Interview Survey, 1995-2004 - James T. Massey.. 317
-Estimation and Analytic Methods for Minority Statistics Based on Multiple-Source Survey Data - Lester R. Curtin* 318
-Use of Behavioral Risk Factor Surveillance System to Monitor Risk Factors in Minority Populations - Emma L. Frazier* 319
-Obtaining Data for Asian Pacific Americans from National Surveys - Larry H. Shinagawa.. ... 323

Session Q - SOCIOECONOMIC INDICATORS
-Measuring Socioeconomic Inequality in Health: An Update on Methodological Issues - Elsie R. Pamuk* 329
-Unemployment and Fertility: A Transfer Function Analysis With Additional Interventions - Andreas Muller 335
-A Comparison of Socioeconomic Indicators Across Three Decades - Alan B. Humphrey*. 341

Session R - MINORITY HEALTH RESEARCH
-Health Data on Hispanic Women: Methodological Limitations - Hortensia Amaro 347
-Cooperative Agreements to Advance the Understanding of the Health of Asian and Pacific Islander Americans - Dong Suh 352
-Social Cognition and Responses to Health Survey Questions Among Minority Populations: Preliminary Evidence - Timothy P. Johnson*. 357
-Health Status and Access to Care Among Rural Minorities: What Do We Know? - Keith J. Mueller* 363
-Improving Health Statistics Among American Indians by Data Linkages with Tribal Enrollment Registries - Jonathan R. Sugarman*. 368
-Discussants: Ruth Zambrana and P. Douglas Williams 372

Session S - INTERCENSAL POPULATION ESTIMATES
-Estimating the Population of Rural Communities by Age and Gender: A Case Study of the Local Expert Procedure - David A. Swanson* 375
-Intercensal Estimates of the Population of States and Counties by Age, Sex, Race, and Hispanic Origin: 1980 to 1989 - Sam T. Davis 381
-Preparing Population Estimates for the Post-1990 Period: The Combination of Data from Various Sources - Signe Wetrogan 384

Session T - SPECIAL FOCUS SURVEYS
-Comparing Two Methods of Surveying Sparse Minority Populations -Jay S. Buechner* 387
-Reducing Sampling Costs: A Multi-Stage Design Using Replicated Quota Sampling to Yield a Sample of Hispanic and African American Households in Los Angeles - Laurie M. Anderson* 393
-Issues of Methodology for Agricultural Health and Hazard Surveys - Lorraine L. Cameron*. . 396
-Evaluation of Methods Employed in the Assessment of Health Care Needs and Access to Care in Dade County Florida Following Hurricane Andrew - Nancy D. Barker* 397
-Dual Frame Sample Surveys Based on NHIS and State RDD Surveys - Monroe G. Sirken*... 400

THIRD PLENARY SESSION - ASSESSING PREVENTIVE AND PRIMARY HEALTH SERVICES
-Kenneth E. Thorpe 409
-Access to What and Why? Toward a New Generation of Access Indicators - Lu Ann Aday... 410

* Speaker of coauthored paper.

Session U - COST-BENEFIT ANALYSES
-A Model of Cost-Effectiveness and Injury Prevention - David P. Hamburger* 419
-Expansion of Medicaid Eligibility for Family Planning Services from 25 to 133 Percent of Poverty: A Cost-Benefit Analysis - Douglas R. Murray 425
-The Impact of Dissemination of Smoking Cessation Methods on Health Care Costs Among Pregnant Medicaid Recipients - Richard A. Windsor* 431
-The Oregon Medicaid Plan: Prioritizing Health Services Using Data and Values - Barry F. Anderson 435
-Discussion: Ethical Aspects of Rationing Based on Cost-Benefit Analyses - Paul T. Menzel.. . 439

Session V - COMPARABLE METHODS AND DATA
-Monitoring and Prioritization of CDC's Consensus Set of Health Status Indicators by State Health Departments in the United States - Sharon L. Zucconi* 445
-A Nonlinear Regression Model for the Projection of Expected Values for Selected Health Status Indicators in the Year 2000 - Chun-Lo K. Meng* 450
-Computerized Database for Strengthening Capacity for Local Area Needs Assessment, Surveillance, and Monitoring of Year 2000 Objectives: The Illinois Experience - Merwyn R. Nelson* 454
-A Telephone Survey Methodology for Local Health Departments' Community Health Status and Risk Factor Assessments Related to Healthy People 2000 - Bruce M. Brock 459

Session W - MEDICAL EFFECTIVENESS OUTCOMES RESEARCH
-Integrating Values into Clinical Care - David B. Pryor 467
-Outcomes Research for Effectiveness and Quality: The Major Issues - Sheldon Greenfield.. .. 468
-Effectiveness of the New Jersey Healthstart Initiative on Prenatal Care and Infant Outcome, 1988 to 1991 - Maryanne J. Florio* 469

Session X - HEALTH DATA FOR PROGRAM MANAGEMENT
-Public Law 99-457 Child Find and Tracking System - Joyce M. Eatmon 477
-The Competitive Initiatives Program-A New Jersey Response to the Escalating Costs of Health Care - Marcia M. Sass* 479
-Development of a Data System for the Management of a Comprehensive Cancer Screening Program - Patricia M. Jamison 483
-Using Medicare's End Stage Renal Disease Program Management and Medical Information Database for Program Monitoring and Policy Development - Joel W. Greer 484

Session Y - DATA SUPPORT FOR MEDICAL CARE POLICY
-Data Sources Used in Prioritizing Health Services Under the Oregon Health Plan - Darren D. Coffman* 493
-Using HCFA's National Claims History File for Health Policy Development - Stephen T. Parente* 499
-Assessments of Outcomes for Coronary Artery Bypass Graft Surgery in New York State - Harold Kilburn* 505

Session Z - PREVENTION EFFECTIVENESS
-Cost Effectiveness of Worksite Cholesterol Screening - Robin D. Gorsky* 511
-An Economic Evaluation of Preventing or Delaying the Onset of the Complications of Diabetes - Erik Dasbach.. 513
-Cost of Foodborne Illness and Prevention Interventions - Tanya Roberts 514

Session AA - BARRIERS TO CARE
-Using Hospital Records to Evaluate the Performance of the Outpatient Delivery System - John Billings.. 521
-Sociocultural Barriers to Appropriate Care - Sarita L. Karon* 522
-Evaluation of Barriers to Prenatal Care: A Model Project - Mary E. Losch* 528
-Discussant: P. Ellen Parsons.. 530

Session BB - CDC'S RECOMMENDATIONS ON PUBLIC HEALTH SURVEILLANCE SYSTEMS
-National Public Health Surveillance: Where Do We Go From Here? - Elaine D. Eaker.. 533
The Future of Microcomputers in Epidemiology - Andrew G. Dean.. 534
-The Electronic Information Highway: A Pathway for Public Health Information - David A. Ross 535

*Speaker of coauthored paper.

APPENDIX

Program Session Organizers 539
Program Session Chairpersons 541
Program Session Speakers 543
Exhibitors 549

First Plenary Session

Opening Ceremony

Increasing Span of Healthy Life

Toward the year

2000

Refining the measures

TOWARD THE YEAR 2000--REFINING THE MEASURES

Dr. Manning Feinleib, National Center for Health Statistics

Welcome and Greetings

It is a pleasure to officially open the 1993 Public Health Conference on Records and Statistics and welcome all of you to the 24th national meeting of the premier forum for vital and health statistics. There is a record attendance for the conference, with preregistration of more than 1300 individuals from virtually every state in the nation. There are representatives from 22 other countries and I would particularly like to extend a warm welcome to the large delegation from Canada. We have a busy program covering important topics with informative presentations, but another value of this conference is the time for greeting friends and exchanging ideas with colleagues. I hope to see you all at our welcoming mixer this evening.

Conference Theme

"Toward the Year 2000--Refining the Measures" is the theme of this year's conference. It recognizes the importance of data in the nation's disease prevention and health promotion policy.

The importance of data in setting health policy and directing health programs has grown -- as illustrated by expansion of this conference from a few hundred who represented the vital statistics programs to the many who deal with all aspects of health information. Health reform offers a new arena for the use of data in formulating policies, developing and implementing programs, and measuring their success.

Conference Program

The conference is organized around the data support and measurement issues for the major goals of Healthy People 2000:

- -- Increase the span of healthy life
- -- Reduce health disparities
- -- Achieve access to preventive services

But our conference program is a broad umbrella covering many issues:

- -- the health of special groups --. minority, children, aged
- -- major concerns -- nutrition, injury, disability
- -- methodological issues -- race and ethnicity, linkage, confidentiality

The conference serves as a bridge linking prevention and health reform and we deal with issues related to the measurement of the quality, cost, access, and effectiveness and outcomes of health care.

I extend our appreciation to the 122 speakers presenting at our opening and plenary sessions and in the 28 concurrent workshops who are bringing the results of their latest work and research to us. They represent Federal, State and local government; universities; associations and research centers.

Exhibits offer another way to learn of the current program activities and data resources of'many of the major health programs. There are many interesting exhibits and demonstrations from other parts of the Centers for Disease Control, Public Health Services, the Department of Health and Human Services and others.

Update - Health Statistics Developments and Plans

As the conference begins I would like to take this opportunity to provide a brief overview of the some of the latest plans and developments in health statistics.

Healthy People 2000 -- Since the last conference there has been much activity at NCHS and CDC, other Federal agencies, and throughout the states to meet the data needs for Healthy People 2000. We are now tracking all objectives at the national level. Soon we will publish the first progress report, Healthy People 2000 Review. This report provides the first assessment of the program. It shows that baseline data are still needed for about 10 percent of objectives.

In another important development, CDC has funded community assessment grants in states to build state data capacity. Knowledge gained in those eight assessment states will be shared with others. The Robert Wood Johnson Foundation and the Maternal and Child Health Block Grants are also aimed at increasing state data capacity.

Virtually all states are using health status indicators -- first presented at the last PHCRS -- to provide a uniform assessment of progress in key areas. A recent survey of states showed a high level of activity in building data capacity and carrying out assessment initiatives.

Data Coordination and Infrastructure -- Data coordination through the PHS Task Force on State and Community Data ensures that PHS data activities involving the states and communities take place within the context of uniform policies and a coordinated and comprehensive framework. The newly-formed Joint Council Of official Public Health Agencies Task Force (ASTHO, NACHO, USCLHO) will look at the relationship between Federal, state and local needs from the state and local perspective.

We expect to be involved in a major effort to revitalize the state centers for health statistics. Working with the ASTHO, AVRHS and others, we will review and update guidelines for state centers; have states review and designate a center under the new guidelines; and provide a program of technical assistance

Health Reform -- The data needs for health reform have taken on new importance recently. NCHS has moved to complete the implementation of the National Health Care Survey. All originally planned components covering health care in all settings will be in place by 1995. We are also implementing many of the recommendations from the Institute of Medicine (IOM) study calling for expansion of the survey through patient followup studies, links to population-based surveys, and enhanced health expenditures components. IOM has also developed a set of access indicators to review access to personal health care in America.

Like the health status indicators, these standard measures will provide a base for uniform data collection and comparison.

At NCHS we are also expanding the National Health Interview Survey, adding components on health insurance, access to care, and expenditures tracking. The 1993-94 National Mortality Followback Survey 1993-94 will look at health care and its costs in the last year of life.

A new initiative is the National Employer Health Insurance Survey to be fielded early next year. This survey, a joint project of NCHS, the Health Care Financing Administration and the Agency for Health Care Policy and Research, is intended to produce national and state estimates of health care spending. The survey will collect data on employer and employee insurance premiums, benefits, and administrative costs and provide information not previously available but essential to measure the effects of some of the changes being proposed.

In addition to new types of data, new methodologies will be needed to provide the array and complexity of information needed at the speed it will be required. Some of the approaches under discussion are: Quick response tracking and monitoring surveys, expanded telephone interview surveys, and surveys to be fielded in a modular fashion. New conceptual approaches will be necessary, for example a set of indicators for health reform, to measure "real" change in a changing system.

Introduction of Opening Session Speakers

I am pleased to introduce Dr. Walter Dowdle, Acting Director of the Centers for Disease Control and Prevention and Acting Administrator, Agency for Toxic substances and Disease Registry. Dr. Dowdle was appointed Deputy Director in 1987 and served as acting director from 1989-90.

Dr. Dowdle assists in leading the agency of the U.S. Public Health Service responsible for promoting health and preventing disease, injury and premature deaths. CDC's 10 Centers, Institute and Program Offices work closely with local, state and other federal agencies to protect the public health.

Dr. Dowdle has a long and distinguished career in public health. He directed the Center for Infectious Disease from 1981 to 1986, became deputy director for AIDS in 1986, and was assistant director for science from 1979 to 1983. Earlier he was director of the Virology Division in the Bureau of Laboratories. He was director of the World Health Organization Collaborating Center for Influenza for almost a decade.

Dr. Dowdle is well-known to public health leaders across the country and in many nations. He brings us greetings from CDC and ATSDR and focuses on "Data Needs in an Era of Change."

(presentation by Dr. Dowdle, see page 5)

Thank you Dr. Dowdle for setting the stage for this conference. Prevention and health reform certainly mean new requirements and new opportunities for health statistics.

For a preview of health reform for consumers, administrators, and researchers we turn to our next speaker, Dr. Clifton Gaus. Dr. Gaus has a diversified background in health care policy, health care research, academia and private business. His academic associations were with Johns Hopkins University and Georgetown University. He worked on health policy and research for the federal government during the Nixon, Ford and Carter administrations and later directed the Center for Health Policy Studies at Georgetown University. He has spent the last ten years as a health consultant and business leader, and was responsible for establishing the Association for Health Services Research.

Today he serves as senior adviser in the Office of the Assistant Secretary for Health and has been closely involved in the deliberations of the white House Task Force on Health Care Reform, particularly considering the new demands for data.

(presentation by Dr. Gaus, see page 7)

Thank you Dr. Gaus. Never before have the challenges or opportunities for health statistics to contribute to the nation's health been so great. It is a pleasure to have so many attending what may be the biggest and most important PHCRS. I wish everyone a very productive meeting.

DATA NEEDS IN AN ERA OF CHANGE

Walter R. Dowdle
Centers for Disease Control and Prevention

The Crucial Role of Health Information

Good morning and thank you, Dr. Feinleib. I have heard great things about this conference for a number of years and I am pleased to be here. This conference continues to be an excellent opportunity for representatives of government, academia, and professional associations to come together for the purpose of sharing the latest advances in public health statistics.

This year more than ever, information will play a crucial role in dictating the public health agenda. At the very core of such issues as prevention, health reform, and childhood immunization is the need for better health information. Whether forming a strategy for overhauling the country's health system or implementing a widespread vaccination program it is crucial, that we have good data on which to base sound decisions.

Those of us attending this conference today have a shared responsibility for producing the health information this nation needs. Providing such data will take the efforts of federal and State governments, as well as the public and private sectors.

We are grateful to the leadership that NCHS has provided. In fact, I don't know what we would do without an organization like NCHS to serve as a focal point for developing data capacity. Nowhere is this leadership more essential than in developing the data support -- that is, Objective 22 -- for "Healthy People 2000."

This year's conference focuses on the Year 2000 objectives, and specifically, the data needed to ensure that we reach those goals. "Healthy People 2000" has been extremely useful as a "compass", to guide us in setting the public health agenda for the remainder of this century. The mission of this conference is to refine the measures of these objectives so that progress in reaching our goals can be determined. This information is needed to adopt, change, or refocus programs. This information is needed to target problem areas. and this information is needed to gauge our successes.

I'd like to briefly say a word about the importance of "Healthy People 2000" in terms of what it has meant to the CDC mission. Last year the word "Prevention" was officially added to CDC's title, mandated by Congress. However, prevention has been the very soul of this agency since it's inception and it was very fitting that Congress took steps to formally recognize this.

For almost a half century CDC has been dedicating to preventing disease, injury, and disability. CDC has lead responsibility for many of the year 2000 objectives. But we don't do it alone. Disease prevention and health promotion is a collaborative, cooperative endeavor, and we are all united in this common goal.

Within CDC, we have taken several steps recently to expand our prevention efforts. We have established a new National Center for Injury Prevention and Control in response to the growing epidemic of unintentional injuries and violence, particularly among our youth. Unintentional injury is the leading cause of potential years of life lost and intentional injury that is, (homicide and suicide) ranks third and is still rising.

We have established a National Immunization Program, elevating this activity within our agency to focus our attention and efforts on protecting children from unnecessary disease, in keeping with the President's Comprehensive Child Immunization Initiative.

We have identified several areas which are critical to the success of this Initiative, all involving very close coordination with the States.

- The development of a national public awareness campaign to educate parents about immunization.
- Vaccine Supply/Purchase Program
- The collection of information on existing state-wide systems for the purchase and distribution of childhood vaccines.
- The Monitoring and Development of Immunization Action Plans
- The Measurement of Immunization Coverage--how well are we accomplishing our goals?
- The Building of Community Coalitions-- to create grassroots support for immunization activities.

- The Surveillance of Vaccine Preventable Diseases--using this information to guide vaccination policy development.

Of course, the component which has been discussed by many of you at this conference design and development of a system for monitoring immunization records. Ideally this would consist of immunization registries in each of the 50 States. From there, conceivably a network could be established to connect with these registries to exchange immunization records on the local, State, and national level.

Such a system is still speculative at this point. Much yet remains to be done. This is a major public health challenge.

Of course, the overriding issue facing the country right now is health reform, and while this is centered around health care and the problems associated with cost containment, access, and quality, there are also many exciting possibilities for incorporating public health and prevention concepts into the foundation of health reform. In a society of diminishing resources and increasing needs, prevention makes a valuable contribution to the health care of Americans, not only in a social context in terms of lives saved, but in an economic context in terms of dollars saved.

The future challenges us and our public health partners to develop alternative strategies for melding public health with health care reform so that prevention services can be within the reach of all Americans.

The role that data will play in testing these strategies and shaping this process is particularly compelling. Many of the things we have traditionally measured will be changing more rapidly, and in ways we haven't been attuned to monitoring. There will be new requirements for health information - and new opportunities for obtaining it. Current data systems will be modified, and new systems will be designed along the way. As always, the needs for State data have been increasing and will continue to do so. CDC recognized this need in establishing the Community Assessment Initiative, which focused on data for prevention. We're now a year into this program to build State data capacity with active programs in 8 States and with plans to share the results with others.

There will also be a need for the Federal government to meet some of the State-level data requirements during the phase-in period of health reform. I know that NCHS is giving a great deal of thought now as to how these data needs may be met.

Ultimately, the approach may be patterned after the model successfully employed for Healthy People 2000. For the Healthy People 2000 system, NCHS draws on 700 data points on over 300 objectives from 117 different organizations.

Again, health reform is still in its formative stages. A lot of this is speculative and subject to change as the overall reform plan changes. But it is clear that we can gain a lot from some of the successes we've had with "Healthy People 2000," and I am anxious to see what transpires here at this conference as the week progresses.

Best wishes to all of you for a most successful conference.

OPENING CEREMONY
Keynote Address

Clifton R. Gaus, Office of the Assistant Secretary for Health

Thank you, Dr. Feinleib. I bring greetings from Dr. Philip Lee, the Assistant Secretary for Health. He sends his regrets. I'm sure you will understand his health care reform meetings at the White House have played havoc with his schedule. However, that leaves me with the distinct honor and pleasure to be here to talk about health care reform and preview what it means to you as citizens, consumers, and the health statistics professional community.

As citizens we seem to know something needs to happen, but we don't know why. Recent polls tell us that 90% of Americans believe, "The Nation has a health care crisis." However, 81% said they were satisfied with their own health care quality; 70% were satisfied with health insurance coverage; and 49% were satisfied with costs.

These polls also show that a high percentage of Americans are willing to pay higher taxes for health care, but a far less amount than is needed to provide universal coverage of comprehensive benefits. What all Americans would like -- is better coverage for less money, and polls show most believe they will in fact, get just the opposite.

It seems like the country's mood goes like the old saying, "most people are in favor of progress; it's the changes they don't like."

In President Clinton's report, A Vision of Chancre for America, he wrote to Congress:

"Throughout our history, at every critical moment, Americans have summoned the courage to change, to adapt our nation's policies and institutions to address new problems in a changing world. Today we must once again find the courage to change."

AS the President focused on the economy and the Federal deficit, it became abundantly clear, there is no area where we need to change more than our health care system.

Skyrocketing health care costs are driving the Federal deficit. Unless the present pattern is changed, 50 percent of the growth in the deficit between now and the year 2000 will be in health care costs.

Not only are rising health care costs depleting our government's resources, and sapping our ability to pay for other societal obligations -- they are pricing American companies out of the health insurance market. Our companies can't continue to hide the inflated costs of health insurance in the price tags of goods and services. We are losing jobs because the high cost of health benefits for employers is raising the cost of American goods. And our products are unable to compete price-wise in international markets.

While the economic costs of our health care system are profound, the human costs are immeasurable. Americans experience far too much preventable human suffering. Our system is sickness oriented. It rewards treatment of illness and injury, rather than keeping people well.

For example, 30 percent of all cancers in the U.S. are caused by tobacco. Over 1000 people die each day from diseases attributed to smoking. An estimated 27 billion dollars in health care costs were attributed to tobacco-related illness in 1990. Society seems quite willing to underwrite the extraordinary costs of surgery. Yet, we are not willing to do all we could to see that young people don't become addicted or that those who are hooked get unhooked.

The system is broken and deteriorating. Fewer Americans are receiving the full benefits of our health care system at an increasing cost. And, no one is absolutely guaranteed continued access to health care in this country. Not even the fully insured. Because of reasons like pre-existing conditions, change in employment, high risk ratings, over the next two years, one of four Americans will lose the health insurance they have today. All Americans are vulnerable with a system that is bankrupting our nation's future.

Reform -- has been part of the National debate for nearly thirty years now. But in the past, we couldn't come to consensus, we couldn't find the courage to change and we let opportunities for reform pass unchallenged. We stayed with an imperfect, sometimes dysfunctional system that promotes spiraling health care costs and excludes millions of Americans who need care the most.

History cannot repeat itself. The health of the people, as well as the nation's economy are riding on what we decide and how soon we do it. After the most thorough and inclusive policy-making process in American history, President Clinton will soon finalize a proposal for comprehensive health reform. The President has promised to announce his Health Care Reform plan as soon as the Congress finishes work on the Budget. Policy-makers in the White House and the Department of Health and Human Services are nearing completion of the plan and it will be ready as the President has promised.

The plan is founded on six principles -- security, comprehensiveness, choice, quality, controlling cost, and simplicity,

The President's plan will provide you and your family the security of knowing you will have health coverage even if you switch jobs, lose your job or have a preexisting condition. All Americans will have a comprehensive benefits package that can never be taken away regardless of circumstances. Everyone, regardless of age, health status or occupation will be able to purchase a health care policy from one of several competing local health plans. All policies would include a comprehensive set of government-established standard benefits including physician services, inpatient and outpatient hospital care, medically necessary emergency care, short-term and crisis-intervention mental health care, some alcohol and drug abuse treatment, diagnostic laboratory and radiological services and some home health care.

Consumers would pay one annual fee that would cover most medical needs instead of worrying about the cost of each procedure. And, no plan can charge more to a consumer due to illness, pre-existing conditions or age.

Providing universal financial access won't solve all the problems that the Nation has in public health or health care financing. Providing financial access for 37 million uninsured Americans will highlight problems on the capacity side of our system. As part of the President's health care plan, investments in infrastructure in poor urban and rural areas and in public health will be increased. Incentives will be provided for health professionals to settle in and stay in these underserved areas. And, Regional Health Alliances will be responsible for building health networks where none exist.

These Regional Health Alliances are a key component of the President's Health Care Plan. They will serve as the collective insurance purchasing agents for large groups of individuals. The Health Alliance will then provide you a variety of health plans to choose from. Under the President's proposal, your employer or insurance company won't pick your health plan. You will. And, once a year you'll have the option to choose or change health plans.

To help you make your choice, the Health Alliance will collect data on each plan's price, on how its patients fared and on consumer satisfaction. The Alliance will then make this analysis available to you in a report card format. This "managed competition" will shift the market for health insurance from competition based on risk to competition based on prices and quality.

Also to improve quality, the President's plan will create standards and guidelines for all practitioners to follow. These will reorient quality measurement toward outcomes and away from process and retrospective case-by-case review. It will increase our commitment to medical research and promote the availability and use of primary and preventive services. We believe that if we help control costs up front with an investment in prevention, we will save costs that would be incurred to repair the damage. An investment in prevention is a principle that runs through out the President's proposal and it is a principle that is efficient in both human and financial terms.

In addition to prevention, controlling the growth of health care costs will be accomplished by increasing competition, reducing administrative costs and imposing budget discipline. Competition, as I mentioned before, will be increased by focusing on cost effective care and encouraging both providers and consumers to make cost effective choices. Malpractice reform will reduce spending prompted by "defensive medicine." And, stiff new penalties for those who abuse and cheat the system will root out billions of wasteful health care spending. Bureaucracy will be slashed, drastically reducing paperwork, hassles, and administrative costs. States will have a strong role in reform, with greater flexibility than they now have. Providers will see a streamlining and scaling back in their regulatory, billing, and reporting requirements. And, ultimately consumers will participate in a simpler system with

understandable plan descriptions, standardized benefits, and-no fine print about exclusions from coverage.

Should the plan's competitive and administrative savings not achieve desired results, a budget will ensure that cost growth will be nevertheless controlled. Under the President's plan, the health care system will operate under a budget which will eventually limit health care spending growth to growth close to gross domestic product.

Health care reform will provide a unique opportunity for the field of health statistics, We will be developing organized system of care focussing on populations. For once, we have a built in denominator. New information system technologies have made possible what was unthinkable just a few short years ago. And there will be tremendous new data needs because consumers for once will be making decisions about their plans.

Data systems are the linchpins of the President's plan. We have envisioned as part of health care reform an electronic highway of data with uniform health data standards and shared responsibilities. This new framework for health data will present you the nation's health data professionals many exciting challenges and opportunities.

To meet data and information needs for the President's plan there are three specific but highly interrelated types of data needs:

- data to support the administration and operations of the health care system

- survey data to monitor the implementation of health care reform and assess its impact on the population, and

- data to support core public health functions under health care reform.

Under health care reform, it will be possible to greatly simplify and streamline administrative operations. This will be done through the adoption of national uniform health data sets for enrollment, claims and encounters. Indicators of access, utilization, quality and outcomes will be developed. And, electronic data networks and regional clearinghouses will link health plans, health alliances, States and the federal government. These links are needed to share the data necessary to provide the report card type information for consumers and purchasers of health care services. Sharing will obviously occur with appropriate safeguards for confidentiality, privacy and security.

Health surveys at the national, state and plan level will play an important role in health care reform. Health care reforms are likely to be phased in over time. Therefore it will be critical to develop surveys to monitor the implementation of health care reform and to provide feedback to policy makers for corrections and improvements.

Similarly, it will be important to assess the impact of health care reform on the general population and on providers, as well as on potentially vulnerable populations, States and the health system in general. In the long term, administrative data will provide some of this information, but national and state level surveys will be conducted on a frequent basis to monitor access, utilization, costs, quality and outcomes. In addition, to prepare performance report cards for each health plan, consumer access and satisfaction surveys will be needed on a plan by plan basis.

Examples of some of the performance measures we are thinking about for States and health plans include --

- Percent of enrollees able to enroll in their "first choice" plan.

- Percent of enrolled population satisfied with their plan care.

- Percent of enrolled population satisfied with their primary care provider

- Percent of enrolled population satisfied with access to specialist care.

- Time delays in obtaining primary care appointments

- Percent of eligible population receiving covered preventive services -- immunizations, mammography, prenatal care

- And, rates of adverse events indicating poor quality of covered services -- for example inpatient readmission rates; unplanned return to surgery rates; morbidity and mortality rates.

The list is illustrative not exhaustive.

While one goal of health care reform is to provide universal financial access to health care, a variety of core public health programs and activities will continue to be needed. Each will have associated data requirements. And, in the new data framework, public health

data systems will be strengthened and more closely integrated with administrative and survey data systems.

Health care reform will mean significant changes. Your participation in this conference signifies a tremendous commitment to improving and assuring sound and timely health data systems. I invite you to assist in developing the systems that will translate the framework of health care reform into a working reality.

A sample of some of the issues we will have to address are:

- What are the best measures and indicators of access, appropriateness, costs, quality and outcomes of care that could be used to inform consumers in choosing among plans?

- What should be the content of uniform minimum health data sets for enrollment, for claims processing and for encounters?

And, how will the content of the uniform data sets --

-support administrative operations?
-support performance monitoring?
-And, support broader public health objectives and purposes?

- What kind of data on practice and utilization patterns would be useful for feedback to providers to support improvements in the health care they administer?

- What polices need to be developed to promote appropriate access to data while protecting confidentiality of individuals?

- What data will States need to carry out specific responsibilities under health care reform? For example, what data are needed to develop and monitor state health expenditure targets, or to carry out quality assurance responsibilities?

- What kinds of unique identification numbering system for plans, providers and persons would be most appropriate for use in health care reform?

- What are the most appropriate measures of utilization in managed care plans and capitation plans, where individual claims data and events are less meaningful?

- What aspects of health care reform are most important to monitor through surveys?

-Improvements in insurance coverage in the population?
-Increased access to regular sources of care?
-Appropriateness of care?
-Increased utilization of services?
-Public attitudes and perceptions?
-Provider behavior?

- What kinds of key measures and indicators should be included in a tracking system to monitor the implementation and impact of health care reform?

-At what geographic levels?
-What existing data systems can provide those indicators at national and state levels?
-What new systems would be needed?

- In addition to tracking surveys, what kinds of in-depth surveys will be needed to help understand the forces behind trend changes in key indicators?

- How do we best monitor health expenditures on a continuing basis?

- Are there benefits of more longitudinal data?
-How do we link surveys with administrative data?

Over the years we have developed a number of data systems to support public health programs and activities such as

-block grant application and reporting requirements
-categorical program reporting requirements
-disease and risk factor surveillance systems
-vital statistics systems and disease and treatment registries, and
-general purpose health surveys.

These data systems have served us well for many decades, and they provide much of the data and indicators to measure progress in the Healthy People 2000 framework.

- When everyone has insurance coverage providing a basic set of benefits, including clinical preventive services like immunizations, what will the core functions of public health be?

- what data will be needed to support those core public health functions?

- what will be the role of our current public health data systems in the framework of health care reform?

These are just some examples of the challenges we face in health statistics. As we meet the data needs of health care reform, Healthy People 2000 provides us valuable lessons and concepts. Concepts and lessons such as -- setting national health objectives, defining indicators to measure baselines and progress in meeting those objectives, periodic progress reviews, the consensus process employed in developing Healthy People 2000 priority areas and indicators -- these concept and lessons all will be valuable in planning health statistics to support health care reform.

In addition, Healthy People 2000 provides us a solid foundation for identifying public health objectives in the framework of health care reform. For example, many of the objectives and indicators outlined in Healthy People 2000 might be suitable as performance measures for health plans and alliances, especially in the area of preventive services.

As you meet over the next three days, I hope you address some of the issues I've raised for health statistics in health care reform.

I see the next seven years -- "Toward the Year 2000" as the most exciting, thought provoking and rebuilding time of our generation. I look forward to working with as many of you as possible to chart the course of our Nation's health data systems.

VALUING LIVES, VALUING HEALTH

Donald L. Patrick, University of Washington

Goals and objectives for improving the nation's health include (1) increasing the span of healthy life for Americans to at least 65 years from an estimated baseline in 1980 of 62 years; (2) reduce health disparities among Americans by increasing years of health life among African Americans to at least 60, among Hispanics to at least 65, and among people aged 65 and older to at least 14 years of health life remaining; and (3) achieve access to preventive services for all Americans to increase years of healthy life to at least 65 (Department of Health and Human Services, 1990). All three objectives address years of healthy life; in fact, increasing years of healthy life cuts across the 300 goals and 700 targets of the Year 2000 objectives.

This paper discusses the rationale behind measuring increases in the span of healthy life using measures of life expectancy and estimates of the health during remaining years of life. The paper addresses the following questions: what are years of healthy life as contained in the Year 2000 Objectives? How do we calculate years of healthy life for the U.S. population? What are the problems and promises in assessing the year 2000 objectives using the indicator of years of healthy life?

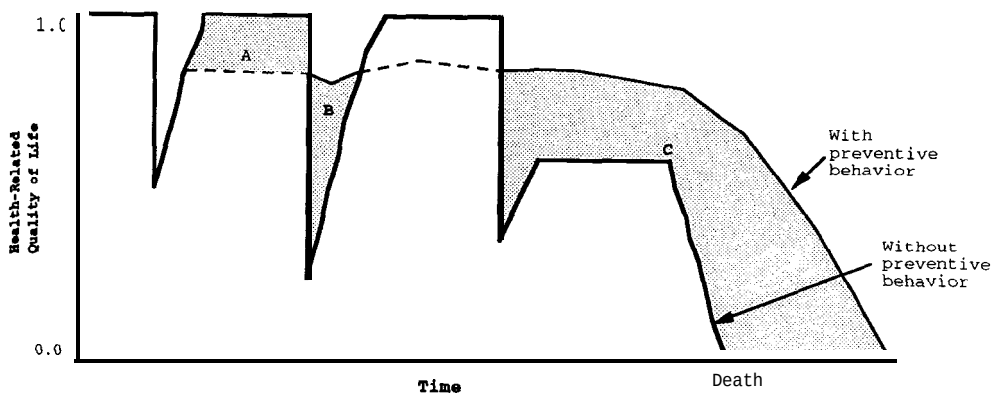
Rationale for Years of Healthy Life. By adopting years of healthy life as a measure of healthy life span, the drafters of ***Healthy People 2000*** recognize the importance of the tradeoffs between quality of life and quantity of life. One reason for measuring the span of healthy life relates to observed changes in the human survival curve. Between 1900 and 1980 a larger number of persons in our population are remaining alive to the end of life, leading to a "rectangularization" of the survival curve. Experts expect a larger percentage of persons to remain alive to age 75 and 85, after which there will incur a rapid rate of mortality. Increasing the span of life, however, does not necessarily mean increasing the span of **healthy** life. As Lois Verbrugge (1984) has asked: "Longer Life? Worsening Health?". Using life expectancy adjusted for disability, Manton and colleagues (1993) suggest, using data from a recent cohort of older adults, that as persons live longer they are experiencing less disability. This important question must be addressed in monitoring the ***Year 2000*** Objectives.

Why is our life expectancy and perhaps our healthy life expectancy changing? First there has been a dramatic decrease since 1960 in infant mortality rates per thousand live births from 26 to 9.1 in 1990 (National Center for Health Statistics 1991). A larger number of infants are surviving their birth experience. At the same time then, we see an increase in life expectancy from 69.7 in 1960 to 75.4 in 1990. The question is: with decreasing infant mortality rates and increasing life expectancy; what is the experience or the health experience of that cohort? Trends in the prevalence of any activity limitation by age group would indicate small but important increases in activity limitation between 1970 and 1980 in the population age 65 or less (Pope and Tarlov 1991).

Years of healthy life will be used to evaluate the effects of primary preventive behavior. As we change our habits to exercise regularly or to reduce the amount of fat in our diet, the nation hopes that these behavioral changes as well as preventive services will keep us healthy to a very long, old, ripe age. Years of healthy life are particularly important in health promotion and disease prevention because they are sensitive to both short-term and long-term tradeoffs in well-being that are likely to occur as a result of prevention initiatives. Figure 1 illustrates these tradeoffs for a prototypical individual. The dashed line illustrates the impact of health promotion and disease prevention activities on years of healthy life for this individual. The sizes and locations of the areas shown are hypothetical. Within a given population individuals will have alternative patterns. The areas as drawn, however, suggest the relationship between health promotion and disease prevention activities and years of healthy life.

In the year 2000 objectives, the years of healthy life equalled 62 years of a 73.7 life expectancy, with 11.7 years of dysfunctional life. How do we monitor our progress to the year 2000? Of crucial importance is to recognize that the gains in life expectancy are not shared equally by all Americans. Average life expectancy and infant mortality rates, in fact, vary considerably for different population groups. Recent data indicate a widening gap in the death rate between blacks and whites, in that deaths per thousand in 1986 among persons 25-64 years old was higher in almost every instance for blacks than for whites. Regardless of race or regardless of income, persons that have higher educations and persons that have higher incomes, experience a lower death rate (Pappas et al. 1993). Therefore we must pay close attention to class, race, and death in looking at our years of healthy life indicators.

Figure 1. Years of Healthy Life With and Without Preventive/ Promotion Behavior for a Prototype Individual



Legend:

A - Primary Prevention may avoid the onset of illness but may result in lowered health-related quality of life because of the inconvenience of changing behavior or engaging in a preventive action.

B - Secondary prevention may reduce the incidence of secondary conditions

C - Tertiary prevention may reduce disability and disadvantage.

source: Patrick and Erickson. 1993

Multiple terms exist for characterizing span of healthy life, and all are important to understanding the dynamics of health during the life course (Robine et al. 1993). First is the concept of *life* expectancy, which is an estimate of the number of years a population can expect to live. Secondly, if one takes life expectancy and adjusts for health, one can obtain an estimate of the number of years a population can expect to live in a specific health state usually defined as disability free-life years or health expectancy. Concepts of active life expectancy, total life expectancy minus disability years, will produce an estimate of a disability-free life expectancy. Finally, if one wants to go beyond measuring the amount of time spent in a disabled state, there is *quality adjusted life expectancy,* which is the estimate of the number of years a population can expect to live in a set of discrete health states that are weighted for "quality". Examples of quality-adjusted life expectancy include quality adjusted life years; well years; weighted life expectancy; and finally years of healthy life.

A year of healthy life is a calculation that discounts how long one lives, that is life expectancy, by how well one lives as a measure of quality of life. For example, persons living ten years in perfect health are calculated as having ten years of healthy life while persons in less than perfect health, let's say .4 who live twenty years, would have eight years of healthy life. Years of healthy life therefore combine estimates of life expectancy and health related quality of life.

Years of healthy life indicate the number of years a population can expect to live with and without quality adjusted function or dysfunction. Estimates of years of healthy life can be calculated using a standard life table adjusted by the level of the health related quality of life of the population. These estimates are used in two ways in the Year 2000 Objectives. Each objective has some intervention in mind. We can calculate the amount of time an individual or cohort spend in a disability-free state or disabled state. If we compare that with a program that may be a preventive service as compared to without the program, we will get years of healthy life gained.

To interpret years of healthy life, we observe the proportion of time, on average, a population spends in an optimal state of health over a given time interval. For example, two years spent in a health state valued at .50 would be equal to one year spent in optimal or perfect health valued at 1.0. For a cohort over its lifetime the expected duration of the life remaining is adjusted by the average number of years estimated to be spent in an optimal health state and in states less than optimal. Assume that there are 100 persons with a life expectancy of 100. They might be expected to live for 75 years in a state of optimal health. But in the last 25 years, they are in states with an average weight of 0.80 yielding an average of 95 years of healthy life (100 X 75(1.0) + 25 (0.80)/100).

Calculation of Years of Healthy Life for the Nation. To calculate years of healthy life, the analysis needs a relevant life table and estimates of the quality of life for the same age groups as contained in life table and for the same socio-economic target population. Although life expectancy is not a problem to estimate, estimating the optimal level of health and less than optimal level to adjust life expectancy is less straightforward. The values or weights to assign to less than optimal health are needed to calculate years of healthy life.

Some persons think of quality of life as the serenity of a pastoral setting. Others think of it as the value that is derived from a close family life. Yet when we fall ill, quality of life begins to be focused around that illness and our attempts to be able to modify what happens to our life and the quality of our life as a result of that illness. We try various forms of treatment. We try a number of different types of changes in our setting, in our health care and in our behavior. Quality of life is also influenced by the reaction of society to that illness.

Some agreement exists on a definition of health-related quality of life, that is, the value that is assigned to how long we live as modified by the symptoms, functional status, perceptions and opportunity that are influenced by disease, injury treatment or policy (Patrick and Erickson, 1993). This definition covers five broad concepts that fall along a continuum of health-related quality of life anchored at the top by an optimal level assigned arbitrarily the value of 1.0 and at the bottom by a minimal level assigned arbitrarily the value of 0.0. Often the two ends of the continuum are labeled with the optimal level defined as "perfect health" and the minimal level defined as "death". Perfect health is an abstract notion meaning "health as good as it can be imagined". Although death may be considered the minimal level of health for many people, states of health such as coma, constant pain, or severe cognitive dysfunction may be considered worse than death, or some negative value below 0.0

The specific domains of survival impairment, functional state, perceptions, and opportunities fall along this broad continuum. A complete representation of health-related quality of life involves (1) specification of the relevant states and/or domains; (2) values or preferences assigned to these states: and (3) duration or probability of survival in different states.

Once we have defined quality of life, an important question arises of how to compare the years of healthy life to be anticipated by newborns and young people as compared to older adults or persons with disabilities. How do we calculate years of healthy life for a person that is experiencing disability throughout his or her total lifetime. Does disability necessarily imply lower quality of life? These two important ethical issues must be considered in constructing measures of years of healthy life.

The challenge for the National Center for Health Statistics is to develop a method of measuring years of healthy life that (1) takes into account younger and older persons and permits comparisons among different age groups; (2) takes into account persons with and without disability, and does not necessarily "downgrade" the quality of life of persons with dysfunction; (3) is applicable to the entire United States population also defined by socioeconomic status, race, and ethnic origin; and (4) can be traced over time to the Year 2000 in order to monitor progress in achieving in objectives.

Looking at available national data sets, there were several possible candidates for inclusion in this measure. First and foremost are role and activity limitations that have been measured as part of the National Health Interview Survey for many years. Many health definitions focus on the capacity for or performance of usual social roles such as holding a job, going to school, parenting, managing a house, taking care of oneself, and maintaining relationships with friends. These concepts are contained in many health status and health related quality of life measures. The National Health Interview Survey routinely includes questions concerning limitations in activities caused by chronic conditions. Two types of activity are covered: major and other. The term *major activity* refers to a person's participation in daily activities in the workplace, in the household, or at school. Social role identification and major activity are classified according to age groups that correspond to the natural status transitions from infancy to old age. The term *other activities* also relates to age and refers to all activities not classified as major that a person in a particular age group might do on any given day. These might be

Table 1. Distribution of Perceived Health Status by Role

Limitation Category		Percentage				
% of Pop.	Limitation	Excellent	Very Good	Good	Fair	Poor
86.2	Not Limited	44.2	30.5	21.1	3.8	0.3
5.2	Limited-Other	11.5	21.2	34.6	25.0	7.7
3.4	Limited-Major	14.7	20.6	38.2	20.6	5.9
1.9	Unable-Major	5.3	10.5	26.3	31.6	26.3
2.0	Limited in IADL	5.0	10.0	25.0	30.0	30.0
1.2	Limited in ADL	8.3	8.3	16.7	25.0	41.7

Data derived from Erickson et. al. (forthcoming)

participating in clubs, athletics, extracurricular activities, and so forth.

In addition to these limitations in major and other activities, the National Health Interview Survey also contains questions concerning ability to carry out activities of daily living or the capacity to perform daily routines and tasks. These are activity limitations and activity restrictions. Activity limitations are long-term reductions in a person's capacity to perform the average activities associated with his or her age group and activity restrictions refer to selected behaviors associated with reduced activity because of long- or short-term conditions.

In developing the NCHS measure of years of healthy life, six different levels of role\activity limitations are used (Erickson et al. forthcoming): (1) not limited; (2) limited in other activities; (3) limited in major activity; (4) unable to perform major activity; (5) unable to perform instrumental activities of daily living, i.e., everyday household chores, and (6) unable to perform activities of daily living, i.e., bathing, dressing, eating, etc.

If we take these levels and automatically assign a score of 1.0 to "not limited" and 0 to "death" or if we arbitrarily assign role/activity restrictions a low, we would be assuming that the health-related quality of life of persons with restrictions and activity limitations was lower than that for persons without restrictions, a common assumption in many measures of health status and quality of life. Disability and dysfunction do not necessarily mean a lower "quality of life" to the persons with those restrictions. Persons with disabilities may be limited in activity, but enjoy a high quality of life. Disability and disadvantage are not only experienced by the individual but also are societal problems that are influenced by environment and available services. Even if a person is unable to bathe, eat, dress and take care of themselves, with the right personal assistance the person with disability can work, can enjoy a high quality of life. Equating disability with lower quality of life can therefore be discriminatory. Hence some means is needed to permit quality of life not to be defined by disability alone.

Last summer at a meeting at the National Center for Health Statistics, George Torrance and I hit upon the idea that perceived health status, that is responses to the question "HOW would you in general rate your health: excellent, very good, good, fair, poor", could be used as a means by which to quality-adjust disability or the role\activity limitations. We know that almost 90% of the population consider themselves to be in good to excellent health. We also know that distribution of perceived health varies by populations, i.e. for black persons, a higher percentage, around 14%, consider themselves in fair or poor health as compared to around 9% of white persons.

Crossing perceived health status (excellent, very good, good, fair, poor) with the six categories of activity/role restriction yields a 5 X 6 matrix. In this matrix, we find that there are persons who are not limited who consider their health to be fair or poor and persons that are totally limited in activities of daily living that consider themselves to be in excellent, or very good health. As shown in Table 1, about 86.2% are not limited and 1.2% are limited in ADL (activities of daily living).

Of the persons who report not being limited at all in role\activity limitations, 44.2% consider their health to be "excellent". For those limited in ADL, about 8.3% consider their health as "excellent", and another 8.3% consider their health as "very good". A small percentage of persons who report that they are in "fair" or "poor" health also report not being limited in any role or activity limitations.

If we conceive of values that range from 1.0 for an optimal state to 0.0 for being dead, the challenge for calculating years of healthy life is coming up with values or weights that are reflected in the six levels of limitation and the five levels of self-perceived health status. The National Center of Health Statistics has used multi-attribute scaling and correspondence analysis to assign values that range between 1.0 and 0 to the different combinations or health states defined by perceived health status and activity limitations (Erickson et al. forthcoming). As shown in Table 2, we observe that a value of .5 or .47 is assigned to a person who is limited in activities of daily living but consider their health as excellent. The same value is assigned to a person who is not limited but considers their health to be "poor".

Table 2. Sample Values for the Interim Measure (Years of Healthy Life)

Role\Activity Limitation	Perceived Health Status: Excellent	Very Good	Good	Fair	Poor	Dead
Not Limited	1.00	0.92	0.84	0.63	0.47	
Limited-Other		0.79				
Limited-Major			0.67			
Unable-Major				0.38		
Limited in IADL					0.17	
Limited in ADL	0.47	0.41	0.36	0.31	0.10	
Dead						0.0

Adapted from Erickson et. al. (forthcoming)

Table 3.. Calculation of Years of Healthy Life for the Total U.S. Population, 1990

Age	I x	n K x	Q	Q* n L x	T x'	YHL Remaining	Life Years Remaining
0-5	100,000	495,073	.94	465,369	6,403,747	64.0	75.4
5-10	98,890	494,150	.93	459,560	5,938,379	60.1	75.1
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
80-85	47,168	197,857	.63	114,650	223,347	b.7	10.9
85+	31,992	193.523	.51	99,697	99,697	3.1	9.3

Adapted from Erickson et al. (forthcoming)

Using these quality weights against the distribution by age bands in the population, one can postulate the years of healthy life remaining using the standard life table and the quality adjustments defined by the value matrix as shown in Table 3. It is important to note in Table 3 that at 85 plus we still have 6.3 years of life remaining but 3.1 years of healthy life remaining.

In 1990 life expectancy at birth was 75.4 years and the corresponding years of healthy life were 64.0, with 11.4 years spent in less than optimal states of health-related quality of life. To meet the target of the Year 2000 of 65 years means a relatively small increase for the total population, assuming the same mortality rates and health experience in 1990. However, for black persons, there is a difference in life expectancy of 69 years as compared to 76 for white persons. If we look at years of healthy life, there is a difference between 65 years of healthy life for white persons and 56 years of healthy life remaining for black persons. How will we increase the years of healthy life for black persons in the United States? To achieve parity, we will have to increase both life expectancy and the health-related quality of life of our disadvantaged populations.

Adjusting life expectancy for disability and quality shows a pattern as illustrated in Table 4 using data from Wilkins and Adams (1983) on health expectancy in Canada in the late 1970s. As shown in Table 4, Canadians living in urban areas had a higher life expectancy than those living in rural areas, 75.3 for communities 1,000,000 or more versus 72.5 for communities less than 1,000 persons.

Table 4. Distribution of Health Expectancy by Community Size, Canada, Late 1970s (in years)

Community Size	Life Expectancy	Disability-Free Life Expectancy	Quality-Adjusted Life Expectancy
Less than 1,000 & rural	72.5	60.1	67.4
1,000-9,999	74.0	59.4	68.0
10,000-99,999	73.8	61.2	68.6
100,000-999,999	73.9	60.6	68.4
1,000,000 & over	75.3	64.0	70.6
All size groups	73.8	61.3	68.6

Source: Wilkins and Adams, 1993

If you adjust for disability, or active life expectancy, you can see in Table 4 that these estimates of life expectancy are lowered considerably. In rural areas of 1,000 persons or less, there are only 60 years of life remaining that are disability free and 64.0 for the areas of 1,000,000 and over. Observe in Table 4 what happens when you quality-adjust life expectancy. Quality-adjusting brings the estimates of life expectancy to the mid-range between the 60 years of disability free life expectancy and the total life expectancy of 72.5 years. Quality-adjustedlife expectancy for persons living in areas less than 1,000 persons are 67.4 and for persons in areas 1,000,000 and over there are 70.6. Assuming that disability is equated with lower quality of life lowers the estimate and quality-adjusting life expectancy raises that estimate. This is also important in terms of income level in that the highest income level has a longer life expectancy, a longer disability free life expectancy and a longer quality adjustment. At the same time, the same pattern holds when you start adjusting for socio-economic status.

Conclusion. Our future use of years of healthy life will involve developing and expanding measures that include role/activity limitations and perceived health status. These questions are contained in the CDC-sponsored Behavioral Risk Factor Survey supplement that is being used by a number of states permitting these states to estimate years of healthy life for their population. These state estimates can be compared with national estimates. Both state and national comparisons will be possible cross-sectionally and over time. Years of healthy life will also permit comparisons among different socio-economic and demographic groups. They are also important in the evaluation of prevention effectiveness.

Problems remain in the use of years of healthy life for monitoring the health of the nation and the health of subpopulations. We need definitions of quality that are more sensitive to well populations. At present, 63% of the population are in "excellent, very good, or good" perceived health status without any role or activity limitations. Definitions of health and quality of life need to be included that address the health status of these well persons over time. These refinements are likely to be in the area of mental health and in physical resilience or capacity. Agreeing on definitions of what constitutes "good" mental health, however, will not be straightforward. Personal mental health also does not have the same social implications as role or activity limitations.

Quality adjustment, that is the value that is placed on life or health, requires refinement of both the methodologies used to estimate these values and in the populations to which these methodologies are applied. In Canada, health-adjusted life expectancy has been calculated for the population of Ontario, Canada's largest province using a Comprehensive Health Status Measurement System developed by McMaster University (Berthelot et al., 1993). Using a multi-

attribute utility function to produce an overall health status index score for each individual, the Health-Adjusted Life Expectancy at age 15 in Ontario is estimated to be 53.6 years for men, 5.7 years less than their life expectancy, and 57.9 years for women, 7.6 years less than their life expectancy. This study illustrates how values can be incorporated into estimates of years of healthy life on large populations.

Information on the values assigned to different health states are needed from the U.S. population in general. We must also examine the question of the conflict between social values for disability-free life and individual values of persons with disabilities. There will always be conflict between how an individual values their quality of life as compared to the way in which society values that quality of life. This conflict may not be solved by no matter what kind of analysis or refinement we make upon measure.

We also need life tables that can be applied to special populations, particularly minority groups in the population, children with developmental disabilities, and persons with mental disabilities. The life expectancy of these populations must be estimated before we can quality-adjust and estimate years of healthy life.

Years of healthy life are not easily observed in empirical studies. One of the major problems is the idea that these measures are obtained by statistical modeling and thus are easily influenced by methods of calculation. The resulting measures must also be interpreted. No matter how good a measure we have, we are also challenged with interpreting trends in years Of healthy life.

Regardless of these problems, the advances made in adjusting life expectancy for disability and quality of life play an important role in evaluating population health levels. The Department of Health and Human Services continues to endorse the notion of raising population health levels through the promotion of health and the prevention of disease. In keeping with the Healthy *People* 2000 objectives, years of healthy life are the appropriate measure for tracking the overall impact of the nation's efforts to implement more intensive and extensive community-based health promotion and disease prevention. Advances in treatment and increases in the use Of clinical preventive services will influence this measure. As we achieve risk reduction and services and protection objectives, we should observe an increase in the years of healthy life. As we focus on the reduction of health disparities, gaps in years of healthy life between population groups should narrow. The metric for evaluation must be matched by national, state, and local efforts to implement *Year* 2000 programs and activities to achieve the goals of longer life with improved health.

Reference8

1. Department of Health and Human Services (1990). Heal *thypeople* 2000. Washington DC: U.S. Government Printing Office.

2. Verbrugge LM (1984). "Longer life but worsening health? Trends in health and mortality of middle-aged and older persons." *Milbank Mem* Fund Q:62(3), 475-519.

3. Manton KG, Corder LS, Stallard E (1993). Estimates of change in chronic disability and institutional incidence and prevalence rates in the U.S. elderly population from the 1982, 1984, and 1989 National Long Term Care Survey. *Journal of* Gerontology, 48(4):5153-5166.

4. National Center for Health Statistics 1992. Health United States, 1991. liyattsville, Md: National Center for Health Statistics.

5. Pope AM, Tarlov AR (eds.) 1991. *Disability in America: toward a national agenda for prevention.* Washington, DC: National Academy Press.

6. Pappas G, Queen S, Hadden W, and Fisher G 1993. The increasing disparity in mortality between socioeconomic groups in the United States, 1960 and 1986. N *Engl J Med,* 329:103-9.

7. Robine JM, Mathers CD, Bucquet D (1993). Distinguishing health expectancies and health-adjusted life expectancies from quality-adjusted life years. Am *J Public Health* 83(6):797-798.

8. Patrick DL, Erickson P 1993. *Health Status and Health Policy: Quality of Life in Health Care Evaluation and Resource Allocation. New* York: Oxford University Press.

9. Erickson P, Wilson R, Shannon I (forthcoming). Years of healthy life: charting improvements in the nation's health. In J Dimsdale, and A. Baum (eds.), *Quality* of *Life in Behavioral Medicine.* Lawrence Erlbaum Publishers.

10. Wilkins R, Adams OB 1983 Health expectancy in Canada, late 1970s: demographic, regional, and social dimensions. Am *J Pub Health* 73:1073-1080.

11. Berthelot RJ, Roberg R, Wolfson MC 1993. The calculation of health-adjusted life expectancy for a Canadian province using a multi-attribute utility function: a first attempt. In Calculation *of health expectancies: harmonization, consensus achieved and furture perspectives.* Eds. JM Robine, CD Mathers,MR Bone, I Romieu. Collogue INSERM/ John Libbey Eurotext Ltd. Vol. 226, 161-172.

TOUGH DECISIONS: TRADEOFFS INVOLVING YEARS OF HEALTHY LIFE

Jane E Sisk
Columbia University School of Public Health

(Not available for publication)

Session A

Years of Healthy Life

Toward the year

2000

Refining the measures

YEARS OF HEALTHY LIFE: A MEASURE OF HEALTHY LIFE SPAN FOR HEALTHY PEOPLE 2000

Pennifer Erickson, National Center for Health Statistics

Ronald W. Wilson, J. Fred Seitz and Ildy Shannon

In formulating health promotion and disease prevention strategies for the next decade, policy makers have captured Americans' desires for long and healthy lives in the overall goals set forth in Healthy People 2000. The first goal calls for increasing the healthy life span as part of a national strategy for improving the health of the Nation. Specific objectives for educational and community-based programs, diabetes and chronic disabling conditions, and clinical preventive services are designed to increase years of healthy life by three years by 2000 (Department of Health and Human Services (DHHS) 1991). The goal of increasing the healthy life span and the specific targets for reaching 65 years of healthy life by 2000 focus attention on the importance of quantifying both length of survival and health-related quality of life for understanding the health of the Nation and its sub-populations.

As this suggests, improvements in the healthy life span for the Nation come from two sources. One source is through the implementation of health promotion and disease prevention programs that save lives that might otherwise be lost prematurely or needlessly. One way of measuring the benefit of these programs in extending life is through the use of the concept of life expectancy which is measured in a life table. The other source is through the reduction of unnecessary suffering, illness, and disability. Tracking progress toward increasing the healthy life span requires a measure that combines quantity of life, or survival, with quality of life, or the population's health experience.

The National Center for Health Statistics (NCHS) is responsible for monitoring the progress toward reaching this goal starting in 1990. To meet this responsibility, it has been necessary to develop a measure using information that was available in 1990 and would be collected annually until 2000. The specific measure for tracking progress in increasing the healthy life span in response to Healthy People 2000 is a measure called "years of healthy life." This paper explains data sources and methods for calculating years of healthy life and presents estimates for selected years beginning in 1984 through 1990.

Years of Healthy Life

Years of healthy life, also referred to as "quality-adjusted life years," "well life years," or "health-adjusted life years," represents the duration of life as modified by decrements in health and well-being. Years of healthy life uses a life expectancy model in which standard life table data are adjusted by the level of health-related quality of life of a population. Measures of health-related quality of life represent many different concepts and domains of life quality, including symptoms and subjective complaints, mental, physical, and social functioning, general health perceptions, and social opportunity (Patrick and Erickson 1993).

Combining measures of different domains of health-related quality of life into a single score that can be combined with mortality requires the adoption of a conceptual model that considers health as a continuum that ranges from perfect health to death. In between these two endpoints lie a number of discrete and mutually exclusive health states that are defined in terms of one or more concepts and domains of health-related quality of life. To convert this conceptual model into an operational definition, death is assigned a value of 0.0 and optimal health is assigned a value of 1.0. Health states falling between these two endpoints are assigned numbers that represent the values that either society as a whole or single individuals place on being in each health state. When years of healthy life are to be used for national policy, as in Healthy People 2000, the values placed on each health state should represent society's values for health and well-being. Various methods from economic and psychometric theory have been used to determine values for different health states (Patrick and Erickson 1993; Torrance 1986).

These values can be interpreted as representing the proportion of time, on average, that an individual spends in optimal functioning over a given time interval. For example, if the person experiences a health state for one year that has a value of .75, we say that the person spent 75 percent of the year in full function; the remaining 25 percent of the year was spent in various states of dysfunction due to either chronic or acute conditions or both. For groups of

individuals, averaging the values assigned to health states experienced by persons in the group gives the mean health-related quality of life for that population for the year.

Years of healthy life for a population over its lifetime can be estimated by adjusting the expected duration of remaining life, i.e., life expectancy, by the average health-related quality of life observed in the population. This gives the average number of years that the population is expected to spend in a state of optimal function and the average number of years spent in states of less than optimal health.

Measuring Years of Healthy Life for Healthy People 2000

Of the many measures of health-related quality of life that are available, three measures, the EuroQOL, the Health Utility Index (HUI), and the Quality of Well-Being (QWB) scale, are based on the conceptual model described above. In addition, each of these measures can be used to calculate years of healthy life and is suitable for monitoring the health of the general population. The EuroQOL defines health in terms of mobility, self-care, main activity, pain, mood and social relationships (EuroQOL Group 1990). The HUI defines health in terms of physical, role, and social and emotional function as well as health problems (Torrance 1982). The QWB defines health in terms of mobility, physical activity, social activity and symptoms and problems (Bush 1984; Kaplan 1988).

None of the existing methods for calculating years of healthy life, however, has been adopted for use in national surveys, such as the National Health Interview Survey, on a regular basis. Although analogues of these measures can be created using national data (Erickson et al. 1988, 1989, 1993), such retrospective analyses usually require information on a wide range of concepts and domains, data that are not available annually. As a result, it has been necessary to develop a interim measure of health-related quality of life using data that were collected in 1990, would be available for each year until 2000, and could be combined with information available in national life tables.

The need for annual measurement points to the use of the National Health Interview Survey (NHIS) as the major source of health-related quality of life data. Two types of information from the National Health Interview Survey, namely role limitation and perceived health, are used to form an operational definition of health-related quality of life. Information about perceived health status is asked for all respondents using the following question: "Would you say your health in general is excellent, very good, good, fair, or poor?"

Role limitation captures a person's limitation in usual social role that is usually associated with his or her particular age group, e.g., working, keeping house, or going to school (Adams and Benson 1992). For the interim measure of health-related quality of life, each person is classified into one of six categories based on his or her age and ability to perform his or her major activity: not limited; not limited in major activity but limited in other activities; limited in major activity; unable to perform major activity; unable to perform instrumental activities of daily living without the help of other persons; and unable to perform activities of daily living without the help of other persons.

The operational definition of health-related quality of life based on the five categories of perceived health and the six categories of role limitation has 30 possible health states ranging from the optimal level of not limited in role performance and in excellent health to the lowest health state which is defined as needing help to perform self-care activities of daily living and being in poor health. These health states are shown in Table 1 along with the percent of persons in each health state based on the National Health Interview Survey. In 1990, over 202 million persons, almost 83 percent of the non-institutionalized population, had no role limitation and were perceived to be in excellent, very good or good health. Approximately 1 percent of the population was in the lowest three health states, that is, unable to perform their major activity and in good, fair or poor health status.

According to the 1990 census, approximately 3.3 million Americans were living in institutions, including correctional facilities, nursing homes, long-term stay hospitals and homes for children, or were serving in the military. Since health status data were unavailable for persons living in institutions, these persons were assigned to health states based on existing information and assumptions about their role limitations and perceived health. Using the work of Colsher and colleagues (1992), persons in correctional facilities were considered to be limited in major activity and in very good health. Nursing home populations were assumed to be limited in activities of daily living and in fair health. Persons in long-term stay hospitals, a category that includes treatment for drug and alcohol

abuse, care for chronically ill, mentally ill, mentally retarded, and handicapped persons, were considered to be limited in instrumental activities of daily living and in good health. Children in long-term care facilities were considered to be limited in other activities, but not in their major role, and in good health. The military population of 1.7 million persons in 1990 comprises a basically healthy, that is, in very good or excellent health and with no limitations in usual social role, group of young individuals; their health status is assumed to be represented by the civilian noninstitutionalized population.

Values were assigned to each of the 30 cells in the matrix defined by perceived health and role limitation using multiattribute utility scaling (Keeney and Raiffa 1976); a detailed discussion of the procedures for defining the values in the matrix is given in a Statistical Note (Erickson et al. 1993). Values range from 1.00 for persons who have no role limitation and are in excellent health to 0.10 for persons who are limited in activities of daily living and are in poor health, Table 2. According to these values, if a person lives one year in excellent health and has no limitation in role, then he or she has one year of healthy life. Other health states result in less than a full year of healthy life. For example, a health state that is defined as being limited in major activity and in good health represents a person having 67 percent of one year of healthy life.

Although no other set of health states exists that assigns values to both role limitation and perceived health, selected research studies allow some comparison for validation. Role limitation is comparable to Role Function of the Health Utility Index (HUI) developed by Torrance (1982). The HUI defines performance of usual social role in terms of being able to eat, dress, bathe, and go to the toilet with or without health and the presence or absence of limitations in playing, going to school or working. These functions are similar to those used in the National Health Interview Survey limitation of activity questions. Scores for the three lowest HUI Role Function levels when rescaled from 0 to 1 and those derived for the role limitation dimension are similar.

With regard to perceived health status, Stewart and colleagues (1988) have shown that the five categories result in a non-linear scale with excellent, very good, and good health being closer in scores than are fair and poor health. This result supports findings from the National Health Interview Survey (NHIS) that also indicated that the levels within the perceived health status scale are non-linear.

Validity of the NHIS health-related quality-of-life scores has also been assessed by examining their relationship with other health variables. For example, health-related quality of life was hypothesized to decrease with age and with increasing number of days spent in bed. For persons with 7 or fewer 12-month bed days, scores ranged from 93.0 for persons less than 18 years of age to 77.9 for persons 65 years and older, Figure 1. For persons who spent more than one month in bed during 1990, the scores ranged from 71.3 for the youngest age group to 35.9 for the oldest age group. Mean scores were also higher for men than for women, a finding that is also consistent with other studies of health status.

Years of Healthy Life

Following the method for adjusting life expectancy using national health status data that was first implemented by Sullivan (1971), years of healthy life for 1990 have been calculated using an abridged life table and age-specific estimates of health-related quality of life of the U.S. population (Erickson et al. 1993). For the total population, life expectancy at birth in 1990 was 75.4 years. The corresponding years of healthy life was 63.9, Figure 2, indicating that approximately 85 percent of the life span of the U.S. population will be spent in a state of optimal function, assuming the same mortality and health situations as experienced in 1990. This dysfunction represents the sum of impacts of chronic and acute conditions that occur throughout the population's lifetime as measured by role limitation and perceived health. For persons 65 years and over, the average number of life years remaining was 20.8; the corresponding years of healthy life was 11.9. That is, of the 20.8 expected life years remaining, 8.9 years, or approximately sixty percent of the life years remaining, are expected to be lived in the optimal health state.

Years of healthy life have also been calculated for Black, White and Hispanic persons. In 1990 Black persons had a life expectancy at birth of 69.1 years with 56.0 of those being years of healthy life. The disparity between life expectancy and years of healthy life for Black persons, 13.1 years, indicates that approximately 80 percent of the average life span for Black persons will be spent in the optimal health state, namely no role limitation and in excellent perceived health. Hispanic persons had a life expectancy

at birth of 79.1 years with 64.7 of those being years of healthy life, indicating that Hispanic persons spend approximately the same'proportion of their life span in optimal function as do Black persons. In contrast, White persons have a life expectancy of 76.1 and 65.0 years of healthy life which indicates that slightly more than 85 percent of the life span of White persons is spent in the optimal state of excellent health and no role limitation.

The higher than expected life expectancy at birth for the Hispanic population is attributed to problems with reporting deaths for this population. There are two major problems with calculating years of healthy life for Hispanics. One problem is that not all state data are included; the notable omission is New York state which is the residence of a high proportion of all Hispanics in the United States. Another major problem is the that the lower mortality experience, reflected by the higher life expectancy, may be the result of misclassification of ethnicty on the death certificate (Sorlie et al. 1992). One source of misclassification may arise if the responsible official does not ask about race and ethnicity but rather completes the death certificate based on subjective judgment.

Trends in Years of Healthy Life

Trends in years of healthy life can be estimated for the U.S. population using data from the National Health Interview Survey from 1984 to the present since the questions used for obtaining information on role limitation and perceived health status for these years are identical. Comparable data for the institutionalized population are not available for intercensal years, however. Thus, analysis of trends is based on the non-institutionalized population only. Also, trends in years of healthy life can not be assessed for Hispanic persons since official life tables are unavailable for this population prior to 1990. For the total, non-institutionalized population, years of healthy life rose from 63.4 years in 1984 to 64.5 years in 1990, Table 3. Comparison of the data for the total population shown in Figure 2 and in Table 3 indicates that the inclusion of the institutionalized population results in a decrease of years of healthy life in 1990 of approximately one-half year.

For the White population, approximately 85 percent of the life span over the 1984 to 1990 interval was spent in a fully functional health state. Between 1984 and 1987, years of healthy life rose from 64.3 in 1984 to 65.0 in 1987. The increase of 0.7 in years of healthy life from 1984 to 1987 was due to a 0.3 years increase in life expectancy and a 0.4 years increase in quality of life resulting in a rise in healthy life years from 64.3 to 65.0. From 1987 to 1990, the 0.5 increase in years of healthy life is due to an increase in life expectancy with no apparent improvement in health-related quality of life.

For the Black population, approximately 80 percent of the life span was spent in a fully functional health state. Between 1984 and 1987, life expectancy for this population decreased by 0.4 years. This was offset by an increase in health-related quality of life so that the overall years of healthy life for the two years remained the same, 55.7 years. Life expectancy in 1990 was the same as in 1987. The increase in health-related quality of life, however, resulted in an increase of years of healthy life of 0.7.

Trends by gender were similar to those observed for the above two racial groups. The level of both life years and years of healthy life was lower for males than females between 1984 and 1990. Males, however, spent a slightly higher proportion of their lives in a fully functional health state than did females, 86 percent compared to 84 percent.

Although the change in years of healthy life from 1984 to 1990 seems small, even a change of 1.1 years of healthy life for the total population is meaningful. Very few of the disease prevention and health promotion activities or even the medical care interventions have major impact on the total population even though specific treatments may have very large impacts on small numbers of patients. For example, a new treatment, such as coronary artery bypass graft surgery, that prolongs life for persons with coronary artery disease, may have a large effect on the number of life years gained for those persons with this disease. When this group is included as part of the total population, the magnitude of the change is lost.

Summary

This analysis indicates that the measure of years of healthy life that has been developed for monitoring progress for Healthy People 2000 is sensitive to change in both quantity and quality of life over time. The increase of approximately one year for the total population over the interval from 1984 to 1990 suggests that the target of 65 years will be met by the year 2000 if no major changes occur in the structure of the health care delivery system or in

the methods for obtaining data on perceived health and role limitation.

The redesign of the National Health Interview Survey, which is scheduled to be implemented in 1996, may have significant impact on the levels of health-related quality of life as measured by the same items during the second half of the decade. The survey design factors that may impact the level are: all self report, rather than self and proxy report combined; item placement in the questionnaire; and changes in item wording or skip patterns.

Changes in data collection procedures are one reason that the current measure of years of healthy life is considered to be an interim measure. Another reason is that additional concepts may be included in the measure in order to make it more sensitive to changes in population health, especially among persons with no role limitation who are in excellent, very good or good health. The National Center for Health Statistics has an active research agenda that is designed to examine the impact of these changes as well as test the impact of the various assumptions used in developing this measure of years of healthy life. Better understanding of this interim measure of years of healthy life will indicate how responsive the health care system is to giving Americans the longer and healthier lives they desire.

References

Adams P, Benson V. Current estimates from the Health Interview Survey, 1991. Vital and Health Stat 10(184).

Bush JW. General Health Policy Model/Quality of Well-Being (QWB) Scale. in Assessment of ouality of life in clinical trials of cardiovascular theraoies. Wenger NK, Mattson ME, Furberq CD, Elinson J (editors) New York:LeJacq Publishing, 1984 pp.189-199.

Colsher PL, Wallace RB, Loeffelholz PL, Sales M. Health Status of Older Male Prisoners: A Comprehensive Survey. Am J Pub Health 82(6):881-884, 1992

Drummond MF, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. New York: Oxford University Press, 1987.

DHHS (1991) Healthy Peoole 2000. U.S. Public Health Service:Office of Disease Prevention and Health Promotion.

Erickson P, Anderson JP, Kendall EA, Kaplan RM, Ganiats T. Using retrospective data for measuring quality of life: National Health Interview Survey data and the Quality of Well-being scale. Quality of Life and Cardiovascular Care 4(4):179-84, 1988.

Erickson P, Kendall EA, Anderson JP, Kaplan RM. Using composite health status measures to assess the nation's health. Med Care 27(3):S66-77, 1989.

Erickson P, Kendall EA, Odle MP, Torrance GW. Assessing health-related quality of life in the NHANES I Epidemiologic Followup Survey. Working Paper Series, 1993a.

Erickson P, Wilson RW, Shannon I. Years of Healthy Life. Statistical Notes. Hyattsville, Maryland: National Center for Health Statistics, 1993b.

EuroQOL Group. EuroQOL--A new facility for the measurement of health-related quality of life. Health Policy 16:199-208, 1990.

Kaolan RM. New Health Promotion Indicators: The General Health Policy Model. Health Promotion 3(1):35-49, 1988.

Keeney RL, Raiffa H. Decisions with multiple objectives: Preferences and value tradeoffs. New York: John Wiley and Sons, 1976.

Patrick DL, Erickson P. Health status and health policy: ouality of life in health care evaluation and resource allocation. New York: Oxford University Press, 1993.

Sorlie PD, Rogot E, Johnson NJ. Validity of demographic characteristics on the death certificate. Epidemiology 3(2):181-184, 1992.

Stewart AL, Hays RD, Ware JE Jr. The MOS short-form general health survey: Reliability and validity in a patient population. Med Care 26(7):724-735, 1988.

Sullivan DF. A single index of morbidity and mortality. HSMHA Health Reoorts 86(4):347-355, 1971.

Torrance GW. Multiattribute utility theory as a method of measuring social preferences for health states in long-term care. in Values and Loncr-Term Care. Kane RL; Kane RA (editors) Lexington, Massachusetts:LexingtonBooks, 1982 pp 127-156.

Torrance GW. Measurement of health state utilities for economic appraisal. Journal of Health Economics 5(1):1-30, 1986.

Table 1
Percent of Persons in Health States Based on
Role Limitation and Perceived Health Status,
National Health Interview Survey, 1990

Role Limitation	Perceived Health Status				
	Excellent	Very Good	Good	Fair	Poor
Not Limited	38.1	26.3	18.2	3.3	0.3
Limited-Other	0.6	1.1	1.8	1.3	0.4
Limited-Maior	0.5	0.7	1.3	0.7	0.2
Unable-Maior	0.1	0.2	0.5	0.6	0.5
Limited in IADL	0.1	0.2	0.5	0.6	0.6
Limited in ADL	<0.1	0.1	0.2	0.3	0.5

Table 2
Values for the Measure of Years of Healthy Life Used
for Monitoring Progress for Healthv People 2000

Role Limitation	Perceived Health Status					
	Excellent	Very Good	Good	Fair	Poor	Dead
Not Limited	1 .oo	0.92	0.84	0.63	0.47	
Limited-Other	0.87	0.79	0.72	0.52	0.38	
Limited-Major	0.81	0.74	0.67	0.48	0.34	
Unable-Major	0.68	0.62	0.55	0.38	0.25	
Limited in IADL	0.57	0.51	0.45	0.29	0.17	
Limited in ADL	0.47	0.41	0.36	0.21	0.10	
Dead						0.0

Table 3
Years of Healthy Life for the Civilian, Non-Institutionalized Population
for 1984, 1987, and 1990, National Health Interview Survey

Population Group	1984		1987		1990	
	YHL	LY	YHL	LY	YHL	LY
Total	63.4	74.7	63.8	74.9	64.5	75.4
Whites	64.3	75.3	65.0	75.6	65.5	76.1
Blacks	55.7	69.5	55.7	69.1	56.4	69.1
Males	61.1	71.1	61.7	71.4	62.3	71.8
	78.2	66.0	78.3	66.6		78.8

Figure 1
Mean Health-Related Quality of Life Score:
By Age and 12-Month Bed Days

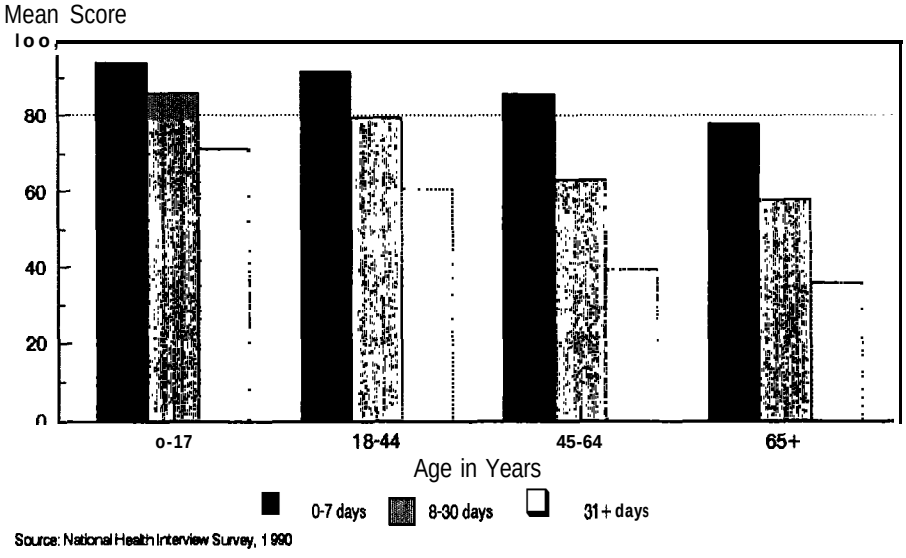
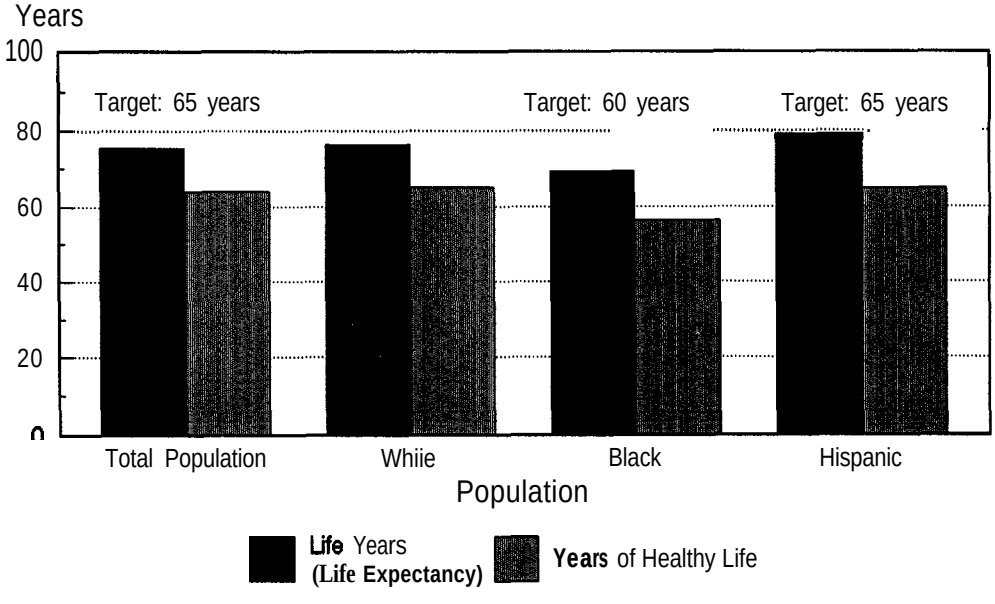


Figure 2
Life Expectancy and Years of Healthy Life
By Race and Ethnicity, 1990



THE CALCULATION OF HEALTH-ADJUSTED LIFE EXPECTANCY FOR ONTARIO USING A HEALTH STATUS INDEX

Jean-Marie Berthelot, Statistics Canada
Roger Roberge Jr. and Michael C. Wolfson

ABSTRACT

Background: Life expectancy of Canadians has increased by approximately 5 years in the last quarter century. It is not clear if the addition of quantity to life is also adding quality.
Method: Health-Adjusted Life Expectancy is calculated by combining a life table stationary population with cross-sectional population estimates of self-reported health status used in the calculation of a health status index. Health status is obtained from the Comprehensive Health Status Measurement System developed by McMaster University and measured in the 1990 Ontario Health Survey for the household population. Health status for the institutionalized population was imputed using a combination of the 1990 Ontario Health Survey and the 1986-1987 Canadian Health and Activity Limitation Survey.
Results: Application of the Sullivan method to the 1985-1987 Ontario life table stationary population in conjunction with the provisional index, resulted in a Health-Adjusted Life Expectancy at age 15 of 53.6 years for men, 5.7 years less than their life expectancy, and of 57.9 years for women, 7.6 years less than their life expectancy.
Conclusions: Even though, data limitations were encountered, the health utility index provides a valuable framework for estimating Health-Adjusted Life Expectancy.

Introduction

Increased life expectancy is a trend observed in much of the developed and developing world. In Canada, life expectancy increased by 4.6 and 5.5 years[1], respectively, for males and females between 1961 and 1986[2]. In the United States, the increases were 4.4 and 4.5 years between 1964 and 1985[3], while in Australia, the increases were 5.4 and 5.3 years between 1965 and 1988[4]. The simple increase in the quantity of life, however, does not give any indication of whether there is an accompanying increase, decrease or maintenance of the health-related quality of life.

In 1971, Sullivan[5] was one of the first to develop and apply a health-adjusted indicator. His Disability-Free Life Expectancy (DFLE) indicator combines cross-sectional age/sex specific disability prevalence data with an appropriate life table. Two health states are used with a disability state having a value of 0 and a disability-free state a value of 1. The prevalence of the disability-free health state is applied to a stationary population from a life table resulting in an estimate of disability-free life expectancy.

Summary indices of population health have evolved in several directions from the original Sullivan work. These directions include indicators which identify various other health states such as impairment or handicap[6] while another provides explicit and more realistic descriptions of the health status dynamics by modeling the underlying health and disability processes (incidence patterns and progression)'. These, and other efforts are reviewed in Wolfson and Manton[8].

The extension of Sullivan's binary representation of health status (e.g., disabled or not) to a continuum of health states is another direction being taken. Torrance[9] suggested the use of a health expectancy index based on multi health states whereby each year of life would be weighted by a utility or value function. The reference values for the index are 0.0 for death and 1.0 for "fully healthy". States considered worse than death, such as very painful terminal cancer or total paralysis, may be valued at less than 0.0. While the conceptual framework for calculating Health-Adjusted Life Expectancy is well understood, the practical difficulties of measuring health state utilities are still being addressed. In 1983, Wilkins and Adams[10] derived the first estimates of the Health-Adjusted Life Expectancy for a national population, by combining the Sullivan method with the cross-sectional prevalence for six disability states and by assigning arbitrary values, corresponding to perceived health-related quality of life, to each of the six disability states.

This paper presents an extension of the Sullivan method for the calculation of the Health-Adjusted Life Expectancy (HALE) at age 15 for the province of Ontario.

Methods

Even though health status can be viewed in thousands of different facets, in practice it is based on a limited number of distinct attributes, with a limited number of functional levels within each attribute. The Comprehensive Health Status Measurement System (CHSMS) is used for this paper[11] and is comprised of eight independent health attributes with five to eight levels of functional ability within each attribute. These attributes consist of: Vision, Hearing, Speech, Mobility (ability to get around), Dexterity (use of fingers), Cognition (memory and thinking), Emotion (feelings) and Pain and Discomfort. The vector of an individual's attribute specific levels of functional ability is defined as a persons health state. The prevalence of each health state is obtained from the 1990 Ontario Health Survey (OHS) and the 1986-1987 Health and Activity Limitation Survey (HALS). Utility weights for the different functional ability levels have to be derived before calculating the health status index.

The analysis presented in this paper is the responsibility of the authors and does not necessarily represent the views of Statistics Canada or the Canadian Institute for Advanced Research.

The utility weights are evaluated on a scale that arbitrarily sets the reference state of death to 0.0 and the reference state of "fully healthy" to 1.0. The evaluation of the preference weights of individuals, specific to the OHS version of the CHSMS, is in development and will not be available before the end of 1993. It is thus necessary to use the results of previous work to obtain provisional utility weights for the OHS. McMaster University had already developed a set of utility weights and a multi-attribute utility function (MAUF) to derive a health status index for a Childhood Cancer Study[12], using a precursor of the CHSMS. Statistics Canada commissioned McMaster University to adapt the results of the Childhood Cancer Study (CCS) utility weights and multi-attribute utility function to the OHS suite of questions, namely the CHSMS.

The OHS CHSMS questionnaire response combinations are transformed into the format used in the CCS to provide health state levels for each CCS attribute. Utilities derived for the CCS are then combined through the CCS-MAUF to produce an overall individual health status index score. Previous work conducted by Torrance[13] determined that the multiplicative form was the most appropriate. Consequently, the CCS-MAUF, used to compute an individual's health status index score, is in the form of a multiplicative utility function and is defined as:

$$HSI_i = \left(\alpha * \prod_{k=1}^{7} w_k(h_{i,k})\right) - \beta$$

where HSI_i is the health status index utility score for individual i, and $w_k(h_{i,k})$ is the preference weight associated with the observed health status level $h_{i,k}$, of the ith individual for attribute k. The values α and β are parameters determined by the data and provide the range for the index scores. For the calculation of the provisional index, α was 1.06 and β was 0.06, such that the resulting health status index scores ranged from -0.02 to 1.0, where -0.02 is the worst state of health and 1.00 is the fully healthy state. Torrance et al.[11] state that "the negative score for the worst health state reflects the fact that on average respondents judged it to be worse than death."

The health status index derived from the OHS represents only the household population for Ontario. Since people residing in institutions are more likely to be disabled[14] and the proportion of people residing in institutions increases with age (less than 1% at age 15-24 and more than 30% at age 85 and over), the impact of the institutionalized population on the calculation of HALE for Ontario can be important, especially at older ages. With no survey providing data on health status compatible with the CHSMS for the institutionalized population, an imputation methodology was derived to reflect the health status of this segment of the population in the overall results.

The methodology consists of combining results from the OHS and the 1986-1987 HALS, to provide health status index scores for people residing in institutions. The HALS was a post-censal survey of approximately 200,000 persons designed to provide a comprehensive picture of persons with disabilities in Canada. The survey covered persons with disabilities in both households and health-related institutions[15]. The adult institutionalized portion of HALS consisted of a sample of approximately 18,000 persons 15 years of age and over. Health status index estimates for the institutionalized portion of HALS were obtained by relating similar individuals in the OHS to those in the HALS.

The mean health status index score for the entire Ontario population by 10 year age groups and by sex was then derived by taking the weighted average of the mean health status index scores for households and institutions, given the proportions of people residing in institutions and households by age group and sex for Ontario. The 1985-1987 Ontario life table stationary population person years were multiplied with the population average health status index score by age group and sex resulting in Health-Adjusted Person Years. The HALEs at age 15 for each sex are calculated using the Health-Adjusted Person Years.

Application of the method

Responses from the 1990 OHS were used to estimate the population prevalence of the different health states. The OHS measured, using an interview and self completed questionnaire, the health status for a sample of over 62,000 individuals residing in households[16]. The eight attributes of the CHSMS were measured in the OHS for all sampled persons aged 15 and over. All of the questions comprising the CHSMS were examined for a basic consistency before being combined into functional levels within each of the health status attributes. Taking the survey design into account, the percentage of respondents assigned to the "fully healthy" category or state was calculated by attribute for the Ontario household population and is presented in Fig. 1.

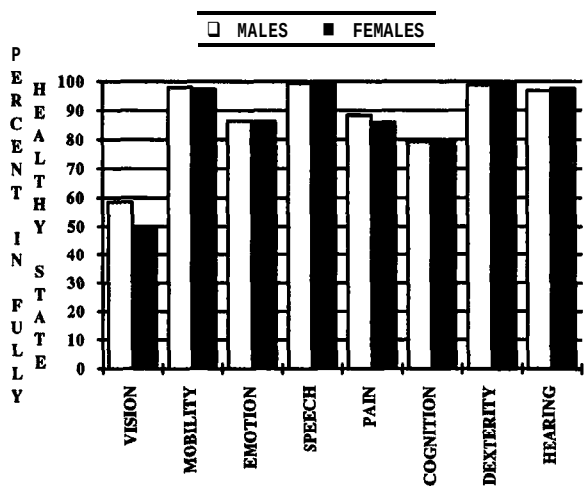


Fig.1 Percentage of household population assigned to "fully healthy" in each health attribute, Ontario 1990, by sex, estimated from a sample of 21,315 males and 24,411 females aged 15 and over.

This graph shows that most people in the Ontario household population are assigned to a "fully healthy" state for each of the attributes, except for vision. The vision attribute, however, is explained by the fact that people wearing corrective glasses are not considered to be 'fully healthy" on that dimension (over 44% of the population required glasses to correct their vision).

The health status index score was calculated for each individual in the OHS. Figure 2 shows the health status index distribution observed in the OHS for males and females, all ages combined. It can be observed from this figure that the majority of the individuals in the household population enjoy a relatively high level of health. Approximately 73% of males and 70% of females are in a health state valued at more than 0.9., with a large proportion (35.2% for males and 29.6% for females) fully healthy. It should also be noted that nobody was assigned to a state worse than death.

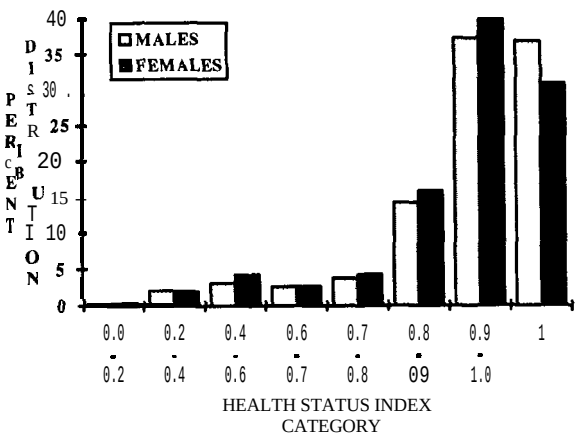


Fig.2. Health status index distribution for males and females in the household population, Ontario 1990, estimated from a sample of 21,315 males and 24,411 females aged 15 and over.

The health status index distribution was compared to the well-being index derived from the self completed portion of the OHS. The well-being index provides a subjective measure of well-being known to be positively correlated with physical health. It is derived from a well-being score developed by Dupuy[17] and has been used in many health surveys[16]. The well-being index values range from a low of 1 (poor) to a high of 4 (excellent). The sex specific distribution of the health status index by the well-being index is presented in Fig. 3.

The different grey shading represents ranges of health status index scores; the darker the shading, the smaller the score of the health status index. The individual health status index scores were averaged within each well-being level and are represented in Fig. 3 by a line connecting asterisks in each vertical bar. This figure shows that the distribution of the health status index scores is broadly consistent with the well-being index: the mean health status index score increases when the well-being index increases, and the proportion of people in the darker (lower) category of health status index decreases when the well-being index increases. At the same time, it is evident that within each level of the well-being scale there is considerable dispersion among health status scores.

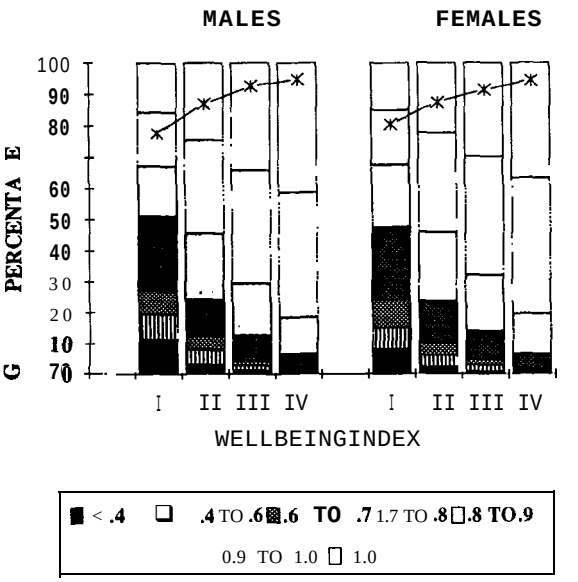


Fig.3. Health status index distribution by Well-Being Index for the household population of Ontario 1990, by sex, estimated from a sample of 21,315 males and 24,411 females aged 15 and over.

The OHS health status index distribution by 10 year age groups for males and females is presented in Fig. 4. This figure illustrates the well known declines in health with respect to aging. A relatively high proportion of the household population is fully healthy (around 60%) in the age group 15-24, while a very small proportion (around 3%) is fully healthy in the 85 + age group. Also, the proportion of the household population in the darker (lower health status) areas increases with age. The individual health status index scores were averaged to provide a mean health status indkx score by sex for 10 year age groups. These mean scores are represented in Fig. 4 with a line connecting asterisks in each vertical bar.

These results represent the health status index for the Ontario household population only and do not incorporate those persons residing in institutions. For the institutionalized portion of the population, the methodology consists of combining results from the OHS and the 1986-1987 HALS. Health status index scores were estimated by relating "similar" individuals in the OHS to those in the HALS. "Similar" was defined by the number of disabilities.

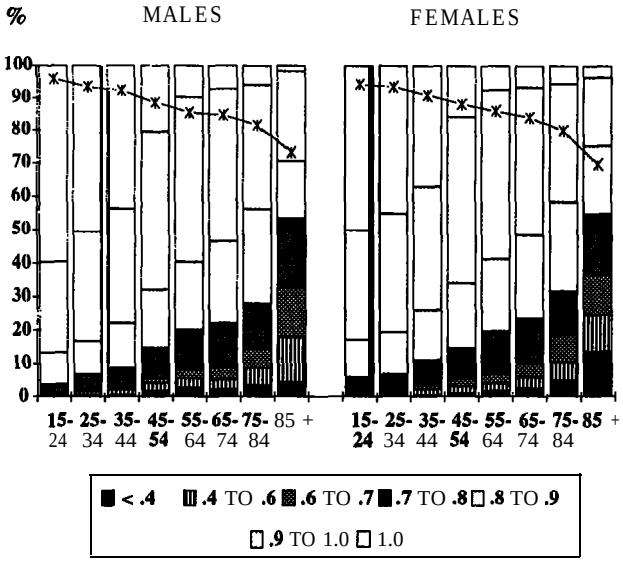


Fig.4. Health status index distribution within age groups for the household population of Ontario 1990, by sex, estimated from a sample of 21,315 males and 24,41 1 females aged 15 and over.

In order to provide for a common basis between the OHS and HALS, the HALS definition of disability was used: individuals were not considered disabled if they used a technical aid that completely eliminated the limitation[14]. In HALS, a temporal condition is also factored in with the requirement that a disability or limitation be a minimum duration of 6 months. OHS disabilities for the 8 attributes of the CHSMS were standardized to the HALS definition of a disability by considering each attribute's health status level. The comparison between the OHS and HALS could only accommodate 6 of the 8 OHS attributes: Vision, Hearing, Speech, Mobility (ability to get around), Dexterity (use of fingers), and Cognition (memory and thinking). HALS had no counterpart for the OHS pain and discomfort, and emotion attributes. Also, the concept of the duration of a disability is not applicable to the OHS and, as such, has been ignored.

The use of a general household population to impute health status index scores for people residing in institutions has many limitations. The household and institutional populations are very different. Most of the severely disabled people reside in institutions, while most of the mild and moderately disabled people reside in households. Even though the methodology partially takes the number of disabilities into account, it does not fully consider severity or disabilities that are not captured by the CHSMS. The mean health status index scores imputed for people in institutions are therefore likely to be inflated.

The mean health status index scores for the combined household and institutionalized populations by age group and sex are presented in Fig. 5. Figure 5 demonstrates that although there might be concern that health status index scores for institutions have been over-estimated, they are still relatively low as compared to the household ones. Health status index scores for the global population are nearly equal to the household population, except for older age categories, where institutional-ization begins to have a greater impact.

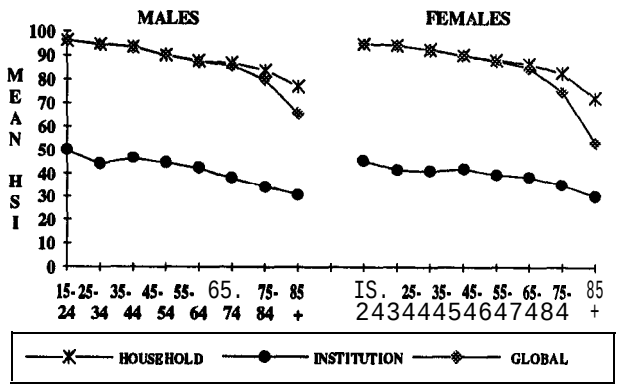


Fig. 5. Mean health status index scores by 10 Year Age Group. The household mean health status index scores are. estimated from a sample of 21,315 males and 24.411 females aged 15 and over.

The HALEs at age 15 for the Ontario population by sex were obtained by applying the average population health status index scores by ten year age group to a 1985-1987 life table stationary population for Ontario. Figure 6. presents the survival and the health-adjusted survival curves, in addition to the difference between the two curves. The ten year age grouping for the health status index scores combined with the single year population data from the life table account for the "saw tooth" patterns in the HALE and LE - HALE curves.

The differences between the survival and health-adjusted survival curves increase with age for males until age 65 and then decrease. For females, the phenomenon is quite distinct, since the differences are increasing until the last age group. This reflects the fact that, beginning at the age group 65-74, the proportion of females residing in institutions is nearly twice that of males[18]. This results in a HALE for males at age 15 of 53.6 years, 5.7 years less than their life expectancy, and a HALE for females at age 15 of 57.9 years, 7.6 years less than their life expectancy. The relative HALE, expressed as a percentage of remaining life expectancy, is 90.4% for males and 88.4% for females. These provisional results, though not directly comparable, are consistent with Wilkins and Adams (1983)[10] in that females have a higher absolute HALE, but a smaller relative HALE; the difference between the males' and females' relative HALEs at age 15 is approximately 2%.

Survival and Health-Adjusted Survival Curves

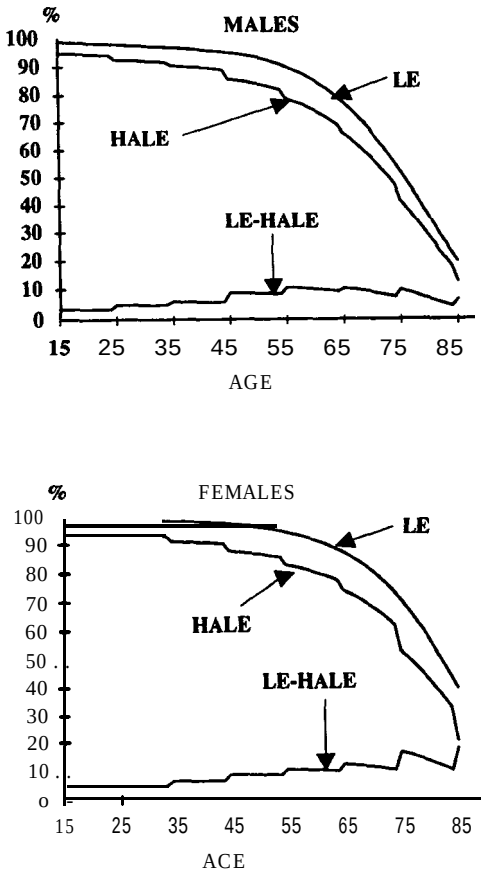


Fig. 6. 1985-1987 Survival Curve and Health-Adjusted Survival Curve for Ontario, by sex. The survival curve is **truncated** at age 85, since only aggregate estimates exist for ages over 85.

The reduction in health-related quality Of life (i.e., LE - HALE) was disaggregated into the 6 attributes, using a cause-deleted technique. For each individual in the sample, a given attribute is set to its fully healthy level, with the remaining attributes allowed to remain in their originally observed levels. Health status index scores are then re-computed individual by individual and re-averaged. This hypothetical "attribute-deleted" health status index calculation thus indicates the effect of each attribute on the overall HALE. Figure 7 shows that the pattern is relatively similar for males and females: the sensation attribute explains approximately 30% of the difference between life expectancy and HALE, followed by the pain attribute (25%) and the emotion and cognition attributes (12%). Given the non-linear character of the multi-attribute utility function that underlies the health status index, there is a residual -- that is, the difference between life expectancy and HALE not explained by the sum of the individual attribute-deleted calculations. This residual reflects the interaction of the level of one health status attribute with another. Since females generally have higher levels of disability than males, these interaction effects are more significant and the residual is therefore somewhat higher for females (14%) than for males (10%).

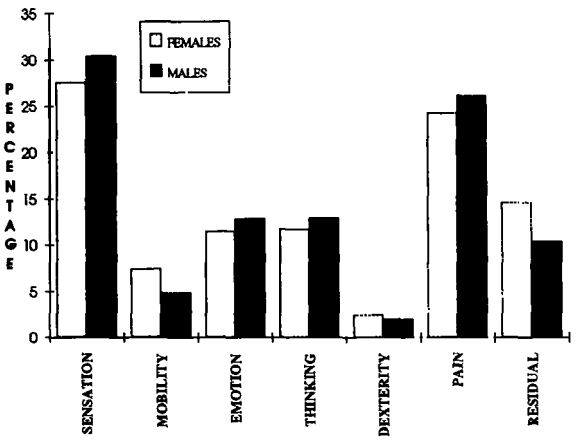


Fig. 7. Reduction **in** HALE by Attribute and sex, based on 1985-1987 life table for Ontario.

Discussion

The use of a multi-attribute health status measurement system in conjunction with a utility function provides a valuable framework for estimating Health-Adjusted Life Expectancy based on the Sullivan method. In our application of this approach, several important data limitations were encountered. One is temporary, due to the use of a provisional multi-attribute utility function for computing health status index scores for the Ontario Health Survey. This should be remedied by the forthcoming development of a utility function explicitly based on the OHS-CHSMS.

A second concern is the lack of appropriate health attribute information for the institutionalized population. It is clear that more information on the institutionalized population is required to improve this method. The measurement of the health status of the institutionalized population, using a health status measurement system like the CHSMS is required. The CHSMS used for the household population is not necessarily appropriate for the institutionalized population. For example, it is not clear if the emotion and cognition attributes capture properly the mental health states of persons with learning disabilities, mental handicaps and/or any physical conditions that resulted in an emotional or nervous disability. This could be an important factor for the institutionalized population, since 71 % of the persons aged 15 and over were diagnosed under this type of disability in the institutional portion of the 1986-1987 HALS. Research should also be initiated to adapt the CHSMS so that it can be used for both the institutionalized and the household populations. The measurement of the health status of the institutionalized population using a CHSMS like instrument is required in order to calculate

reliable health status index scores and, consequently, reliable HALEs for the elderly. In addition, the use of proxy responses during the data collection has to be evaluated. It is unlikely that proxy responses are appropriate for subjective attributes such as emotion, pain and cognition[19]. However, proxy responses may be the only means available to elicit responses for those with severe disabilities.

The availability of a health status measurement system, integrated within traditional health statistics reporting, is important to monitor the changes of the health-related quality of life of the Canadian population over time. In Canada, a new National Population Health Survey (NPHS) is under development and should be operational by 1994. It will be a longitudinal survey of around 22,000 households, and will be conducted every two years over a 20 year period. The goal of the National Population Health Survey is to provide more comprehensive information on current population health status[20]. In addition to providing the basis for monitoring health trends over time, the inclusion of a health status measurement system in the NPHS provides a baseline estimate for the distribution of health status in the general population. The same suite of questions can then also be included in clinical randomized and case-control studies to place such studies in a broader population context.

Finally, it is important to note the limitations in the Sullivan method implicit in the results reported here. This method, by using only cross-sectional prevalence data on health status, takes absolutely no account of the dynamics of health and disease processes. As a consequence, the "attribute-deleted" results presented above, while indicative of the relative importance of various health problems, probably do not represent realistic scenarios. Still, as a first estimate, the Sullivan approach is computationally straightforward, widely used, and not very data demanding[21].

Acknowledgments

The authors would like to acknowledge the contributions made by Carol Paul and Mary-Anne Pietrusiak of the Ontario Ministry of Health, Russell Wilkins, Gary Catlin, Adele Furrie, and Janet Pantalone of Statistics Canada, and George Torrance and William Furlong of McMaster University.

References

1 Nagnur, D. (1986): *Longevity and* historical life tables: *1921-1981* (abridged), Canada and the *provinces,* Statistics Canada, Catalog 89-506, Ottawa, Ontario.

2 Statistics Canada (1988): *Current Demographic Analysis: Report on the Demographic Situation in Canada 1988,* Statistics Canada Catalog 91-209E, Ottawa, Ontario.

3 McKinlay J.B., McKinlay S.M., Beaglehole R. (1989): A *review of the evidence concerning the impact of medical measures on recent mortality and morbidity in the United States.* International Journal of Health Services 19:181-208.

4 Mathers, Colin (1991): *Health expectancies in Australia 1981 and 1988.* Australian Institute of Health, Australian Government Publishing Services, Canberra, Australia.

5 Sullivan, D.F. (1971): *A single index of mortality and morbidity.* HSMHA, Health Reports 86:347-354.

6 Robine, J.-M., Michel, J.-P. (1992): *Towards international harmonization of health expectancy indices,* REVES paper No. 86. Presented to the 5th Meeting of the International Network on Health Expectancy, Ottawa, Ontario. February.

7 Wolfson M.C., (1992), *POHEM - A Framework for Understanding and Modelling the Health of Human Populations,* Analytical Study Branch Paper #34, Statistics Canada, Ottawa, Ontario.

8 Wolfson, M. C. and Manton, K. G. (1992): *A Review of Models of Population Health Expectancy: A Micro-Simulation Perspective,* REVES paper No. 96. Presented to the 5th Meeting of the International Network on Health Expectancy, Ottawa, Ontario. February.

9 Torrance G.W. ,(1976): *Health status index models: a unified mathematical view,* Management Science. 22 (3):129-136.

10 Wilkins, R. and Adams, 0. (1983), *Healthfulness of Life: A unified view of mortality,* institutionalization, and non-institutionalized *disability in Canada, 1978.* Institute for Research in Public Policy, Montreal.

11 Torrance G.W., Furlong W., Feeny D., and Boyle M.H. (1992a): Final *Report to Statistics Canada of Project No. 44400900187 "Provisional Health Index for the Ontario Health Survey",* McMaster University, Hamilton, Ontario.

12 Torrance G.W., Zhang, Y., Feeny D., Furlong W., and Barr, R. (1992b): *Multi-attribute preference* functions *for* a comprehensive *health* status *classification system",* McMaster University, Center for health economics and policy analysis, working paper no 92-18, Hamilton, Ontario.

13 Torrance, G.W., Boyle, M.H., and Horwood, S.P. (*1982): Application of Multi-Attribute Utility Theory to Measure Social Preferences for Health States,* Operations Research. 30 (6):1043-1069.

14 Adams O., Dowler J., Lafleur L., Jordan-Simpson D., and Wilkins R. (1991): *Profile of persons with disabilities residing in health care institutions in Canada,* Statistics Canada, Catalog 82-615, vol. 6, Ottawa, Ontario.

15 Statistics Canada (1990): *The health and activity limitation survey; microdata user's guide, adults in institutions,* Statistics Canada, Ottawa, Ontario.

16 Ontario Health Survey (1992): *The 1990 Ontario* health *survey -- documentation report,* Ontario Ministry of Health, Toronto, Ontario

17 Dupuy, H.J. (1978): *Self-representations of general psychological well-being of American adults.* Paper presented at American Public Health Association Meeting, Los Angeles, California.

18 Jamieson E., Reynolds, D, Bennett, K., Badley, E., Chambers, L., Goldsmith, C., Stock, S., Torrance, G., Tugwell, P. (1992) : *Modelling the population health impact of the rheumatic diseases, workshop* part *1:* Arthritis, Report submitted to *Statistics Canada.* McMaster University, Hamilton, Ontario.

19 Feeny, D., Barr, D.R. Furlong, W., Torrance, G. and Weitzman, S. (1991): *Quality of life of the treatment process in pediatric oncology: an approach to measurement,* in Effect of Cancer on Quality of Life, Chap 6:73:88, D. Osoba editor,CRC Press, Florida.

20 Catlin, G., Will, B.P. (1992): *The National Population Health Survey -- Highlights of Initial Developments,* Health Reports, Statistics Canada Catalog 82-003, Ottawa, Ontario, 4.,No.3:313-319.

21 Robine, J.M., and Mathers C.D., (1993): *Measuring the compression or expansion of morbidity through changes* in *health expectancy,* Colloque INSERM/John Libbey Eurotext Ltd., Vol. 226, pp. 269-286.

ACCOUNTING FOR DEATHS WHEN MEASURING HEALTH STATUS OVER TIME

Paula Diehr, University of Washington
Donald Patrick, David Grembowski, Joseph Picciano

Much research is underway to improve people's health, and thus increase their years of healthy life. Measures of health status over time are used extensively in such research. This paper addresses the frequently occurring problem of how to treat deaths in these studies. The connection between this problem and that of measuring or estimating years of healthy life is also discussed.

For illustration we use data from ongoing research evaluating the cost-utility of preventive services for Medicare enrollees in an HMO. This program, sponsored by the Health Care Financing Authority, is a randomized clinical trial of people over age 64 in which people in the treatment group are invited to special clinic visits and classes aimed at health promotion and disease prevention, while those in the control group have usual care [1]. The goal is to determine whether the treatment slows the decline in health status. For the evaluation, detailed survey information is collected at baseline (T0) and 2 and 4 years thereafter (T2 and T4). For this paper only the T0 and T2 data are used.

INCLUDING THE DEATHS

Various health status measures were collected for each person. Unfortunately, some subjects died before T2 and their health status information was not obtained.

For some health status measures deaths are not a problem, since death has a pre-assigned value. For example, the Quality of Well-Being scale was designed to give death a value of zero. For most health measures, however, deaths result in an unknown value. Researchers often report only the *complete* case, discarding the deaths. However, this has the effect of limiting the evaluation to a biased set of subjects, those in better health. Below we consider some approaches for including the deaths in the analysis. These include non-parametric analyses and strategies for obtaining a code for deaths so they can be included in the analysis.

There is some related literature dealing with missing data in longitudinal studies [2-4]. These approaches differ from our situation in two respects: first, they involve studies with frequent measurements over time, allowing missing data to be estimated from known data for that subject; and second, they treat deaths as equivalent to other types of missing data, when in fact knowledge that a person is dead provides a good deal of information about his health.

The health status measure discussed here is "would you say that your health is: excellent, very good, good, fair, or poor'? This measure was chosen because it is *commonly* used. Findings for this variable are not necessarily those of the complete study. People whose T2 measures are missing for other reasons than death were excluded, to make this example more clear.

Results

Table 1 shows the distribution of the people in the treatment and control groups at T0. Note that, despite randomization and the large study size, the treatment group has more in poor health than the control group, and fewer in excellent health. The control group is almost significantly healthier (p=.06).

Table 1. Health Status at T0

	Control	Tx
Poor	23	38
Fair	165	178
Good	467	460
Very good	417	423
Excellent	137	103
Total	1209	1202

Table 2 shows the distribution after 2 years, at T2. There were unfortunately 89 deaths (3.7%), and more unfortunately, more (38) were in the treatment group. The death rates in the two groups were statistically significant (p=.027). Note also that the number of people in excellent health is higher in the treatment group than in the control group, even though the reverse was true at T0.

Table 2. Health Status at T2

	Control	Tx
Dead	32	57
Poor	43	34
Fair	211	190
Good	492	468
Very good	353	362
Excellent	78	91
Total	1209	1202

A 2x6 chi-square test *is* almost significant (p=.054), but the direction of the difference is unclear, and the test for trend is not significant. If the deaths are discarded, the two groups are not significantly different. We thus need to know which group did 'better', and how to adjust for the T0 differences in the analysis.

Table 3 shows the results of several approaches for handling the deaths. For compactness we have coded the 5 categories on a scale from 1 (poor) to 5 (excellent), and will present means and use t-tests. At T0, as noted before, the controls are significantly healthier than the treated (line a of Table 3). At T2, analyzing only those with valid data, the treatment group is healthier (line c).

The health status of the treatment group thus declined considerably less than that of the control group. However, the T0 mean includes some people who were not alive at T2, and the comparison over time may be misleading. Eliminating the T2 deaths from the T0 mean (line b), we see first that the resulting group was healthier than the average of the whole group at T0, and second that the change scores (T2-T0) are significantly different (line f), with the treatment group declining significantly less than the control group. This finding, however, is made on a biased subset of the subjects, and so may itself be biased.

Table 3. Mean Health Status

	Contr	Tx	p
Health at TO			
a alive at TO	3.40	3.31	.025
b alive at T2	3.42	3.35	.094
Health at T2			
c alive at T2	3.18	3.25	.070
d death-0	3.10	3.10	.995
e death=-10	2.83	2.62	.052
T2-T0			
fve at T2	-.24	-.10	.000
g death-0	-.30	-.22	.027
h death=-10	-.57	-.69	.216

One frequent method of including deaths in such an analysis is to assign them an arbitrary value, such as zero. When this is done, the treatment and control means are almost identical at T2 (line d), which again results in a significant change score difference favoring the treatment group (line g). However, the choice of zero was arbitrary. If we had instead used -10 for deaths, the control mean would be considerably higher than the treatment mean at T2 (line e, p=.052), and the change scores would not be significantly different (line h).

All of the previous approaches involve arbitrary decisions. A more satisfactory approach is a non-parametric analysis, in which we treat the deaths as a group whose T2 health is lower than "poor", by an unspecified amount. The chi-square analysis mentioned above is appropriate, but it provided uninterpretable results. An analysis that takes into account the ordering of the categories, such as the Wilcoxon rank sum test [5] which examines the ordering of the data, is more interpretable. By this method, the controls were nearly significantly healthier at TO, and there was no significant difference at T2.

Neither the chi-square nor the Wilcoxon analysis adjusts for differences between the groups at TO. Such analyses can be modified by stratification on the TO category, and combining the results [4]. After stratification, both the chi-square and the Wilcoxon analyses showed significant results, with the Wilcoxon favoring the treatment group (Table 4, lines a and b).

The stratified Wilcoxon test has some advantages, but is not a common test, and does not provide a readily interpretable summary statistic. For these reasons we explored approaches that would recode the data, so that deaths could be used in the analysis. Two methods already discussed are to recode the 5 usual categories as 5 4 3 2 1 with deaths coded as either 0 or as -10. Another approach is to create a variable which takes on the value 1 if the person's T2 value is better than at TO, 2 for the same, and 3 for worse (deaths are put into the third category). The numbers of people in the better, same, worse categories were 425, 616, and 168 for the control group and 365, 626, and 211 for the treatment group. This analysis significantly favored the treatment group (p=.002).

Table 4. Comparison of Health Status Including the Deaths and Adjusting for TO Health Status

Method	Contr	Tx	P
a Chi-square stratified		?	.03
b Wilcoxon stratified		+	<.05
c B/S/W		+	.00
d 'healthy'		+	.11
e Pr(survival)	+		.08
f Pr("healthy")		+	.10
g death=0		+	.03
h death=-10	+		.22

Another logical way to recode variables is to recode excellent, very good and good as 1 and fair, poor and dead as 0. The resulting variable is interpretable as *healthy? yes/no'. It can be analyzed by chi-square or t-test, and the change score from TO can be calculated. In this case, the treatment was better (p=.11).

Another recoding, which maintains the distinct categories, is to recode the categories according to a standard. One possibility is to assign each category the percent probability of surviving two years, which can be estimated from our own data: 98% of those in excellent or very good health survived two years, as did 94% in good health, 88% in fair health, and 73% in poor health. Deaths, of course, are recoded as 0, since they will not be alive two years later. Standardizing a health status measure against mortality may seem inappropriate. A related approach is to assign to each category the percent probability of being "healthy' (in excellent, very good or good health) in two years, which can also be estimated from our data. The resulting codes for the 6 categories are, respectively, 96, 93, 76, 35, 18, and 0. The t-test on change-scores for the probability of survival favors the controls (p=.08), while that on the probability of remaining healthy favors the treatment group (p=.10). (Note that the probability of surviving is similar to coding death as "-10" above, since the code for deaths is far from the others; similarly, the probability of being "healthy" is more similar to coding deaths as 0, since all categories are reasonably evenly spaced).

We have suggested 8 approaches to testing the hypothesis that the change in health status is different for the treatment and control groups. Six of these favored the treatment group (4 significant) and 2 favored the control group (neither significant). The two that favored the control group were those that gave very high weights to the deaths.

We want a method that will: include deaths; be logical rather than arbitrary: provide a test of the null hypothesis; allow adjustment for baseline differences (although this is not required in the case of a randomized controlled trial); and provide unambiguous result as to which treatment is better. Our assessment of the various methods is that three of the methods: 'healthy*, probability of being *healthy", and better/same/worse meet more of the criteria than the other methods. We will probably use these three approaches in our evaluation.

"Healthy" is simple to understand, but has the disadvantage that a program that moved people from poor to fair health would receive no credit. One might also consider examining the cause of death and coding clearly accidental deaths as "missing" rather than as deaths for this purpose; this should be done with great caution to avoid creating or disguising unexpected program effects.

We should note that this particular measure of health status was chosen because of its interesting and ambiguous findings. Other health status variables, which will also be used in the evaluation, showed no significant difference between treatment and control using these different methods of including the deaths. Analysis of the data to explain the excess deaths in the treatment group is also underway.

YEARS OF HEALTHY LIFE (YHL)

There is some relation between the previous topic and years of healthy life. Suppose we frequently measured health status, and followed all people until their death. We could then plot the percent of people who were "healthy' at each point in time. The area under this curve would be the years of healthy life. If separate plots were made for the treatment and control groups, the difference in the areas would be the

additional years of healthy life due to the treatment.

Our evaluation will have only three measurements for each person, at TO, T2, and T4. These points will allow us to estimate the years of healthy life for each group durins the studv, but will not allow us to estimate the years after the study ends. Using only the TO and T2 data, we can estimate that the mean years of healthy life between TO and T2 in the treatment and control groups were 1.5865 and 1.6079 years, respectively, and that the two are not significantly different.

Recall that the control group was slightly healthier at TO, and that the two groups were similar at T2. Above, when we analyzed change scores, the initial TO difference gave credit to the treatment group, since its members had been slightly less healthy at the start. In the years of healthy life calculation, however, -that initial difference is credited to the control group, since that group has more area under the curve. Neither analysis found a significant difference in the two groups, but it is disturbing that the initial baseline difference is treated in different ways in the two analyses. Some method must be found to adjust for initial baseline differences in the years of healthy life approach. Weighting to make the TO distributions the same for the two groups would have this effect, but would make subsequent analysis difficult since most software does not handle this type of weighting correctly in calculating standard errors 161.

Years of healthy life after the vrograrn ends are more problematic. One strategy is to assign each individual his actual years of healthy life during the evaluation, and add to that the estimated remaining years of healthy life for those still alive at the end of the evaluation. One possibility is to use the life-table method described by Erickson [7]. The 'health' weights that we used for each age group (the Q's) were taken as the percent of "healthy" people in that age category, using T2 data for the treatment group. We might alternatively have decided to use TO data or to use different estimates for the treatment and control groups; here, these choices did not make much difference.

Once the healthy life table is constructed, the estimated remaining years of healthy life are estimated for each person based on his age at the end of the evaluation, by looking up the number in the table (people dead receive no extra years of life). Since the same table is used for treatment and control, the additional years of life will differ between the two groups only if the age distribution at the end of the evaluation is substantially different in the two groups. We estimated that the average total years of healthy life in the treatment and control groups were 10.4 and 10.5 years respectively, with a 95% confidence interval for the difference of (.0908, .149) years, a significant but small difference.

A second approach to estimate the remaining years of healthy life was attempted, using data on how many people moved from one health state to another during the evaluation. For example, 9 of 178 controls in poor or fair health at TO were in excellent or very good health (evg) at T2. We fit a polytomous logistic regression model of the form:

$$\ln[P(evg,T2)/P(dead,T2)] = a+b(age)+c(evg,T0)+d(g,T0)+e(fp,T0)=\alpha_{evg}$$

with similar expressions for good (g) and fair/poor (fp) versus death. Let $D = \exp(\alpha_{evg})+\exp(\alpha_g)+\exp(\alpha_{fp})$. Then $P(dead,T2)=1/D$, $P(evg,T2)= \exp(\alpha_{evg})/D$, and so on. The regression equations were solved using the conditional logistic regression analysis option of the EGRET package (Barbara McKnight, personal communication). Probabilities were estimated separately for treatment and control. These probabilities were then applied repeatedly for each person alive at the end of the intervention (T2).

Table 5 shows the estimates for a 70-year-old who was in excellent or very good health at T2 (100% probability of being E/VG). At T4 we estimate a 70% probability that his health is still E/VG, 24% probability of Good, 5% of Fair/Poor, and 1% probability of being Dead. Applying the equations to the estimated T4 data to predict T6, the person has now only a 50% probability of being E/VG, and so on until T16 at which time there is 100% probability of being dead. We kept track of the amount of time each person was expected to spend in each state, and used this to estimate years of healthy life for each person. The bottom line of Table 5 shows that this person spent 32% of 14 years in E/VG health and 29% in Good health, or 8.5 additional years of healthy life. If he was also in excellent health at TO, he would have a total of 10.5 years of healthy life.

Table 5. Percent probabilities of being in health states, over time (starting in excellent health, age 70 at T2).

Time	State			
	E/VG	Good	F/P	Dead
T2	100	0	0	0
T4	70	24	5	1
T6	50	38	19	3
T8	31	47	15	6
T10	16	48	29	16
T12	5	37	21	37
T14	0	10	9	81
T16	0	0	0	100
%	32	29	11	28

On average we estimated 7.55 and 6.71 years of healthy life between TO and T16 for the treatment and control groups, respectively. The 95% confidence interval for the difference in yhl was (.806,.874) or almost a year. (In fact, some controls were still alive at T16, so the difference would be even bigger).

Note that years of healthy life are considerably lower than estimated by the life table method, and that the difference between the two groups is considerably larger. Some of this discrepancy can be attributed to the poor regression results; we estimated only 10.6 years of life remaining in the control group, compared to 13.3 by the life table method. Since our study group is healthier than average (volunteer bias, analyses not shown), this is an unreasonable result. The regression method clearly kills people off too fast. A different model based on TO health but without age left half the people alive at T40, when all would be over 100 years old!

Clearly, the model can be improved. Transformations and interaction terms are possibilities. There are problems of inadequate data, since even with a sample as large as ours some of the transitions are poorly estimated; only 7 people who were in excellent or very good health at TO died, for example. In addition, the estimation method multiplies estimated probabilities times estimated probabilities, and propagation of error is another likely explanation for the bad fit. It is also unclear whether the transition probabilities that held from TO to T2 would be appropriate at later times, given that our sample of volunteers was healthier at TO (volunteer bias) and might be expected to regress to the mean at a later time.

All of these reasons suggest that the use of projected years of healthy life in randomized clinical trials may be a problem if future years need to be estimated rather than measured. It is

probably wiser at this time to remain within the range of the data that are known.

Bibliography

1. Durham M, Beresford S, Diehr P, Grembowski, et al. Participation of higher users in a randomized trial of Medicare reimbursement for preventive services. 1991. The Gerontologist 31:603-606.
2. Brown CH. Protecting against nonrandomly missing data in longitudinal series. Biometrics 46:143-155. 1990.
3. Zwinderman AH. Statistical analysis of longitudinal quality of life data with missing measurements. Quality of Life Research 1:219-224. 1992.
4. Dawson JD, Lagakos SW. analyzing laboratory marker changes in AIDS clinical trials. Journal of Acquired Immune Deficiency syndromes 4:667-616. 1991.
5. Hollander M, Wolfe D. Nonparametric statistical methods. 1973. John Wiley and Sons. New York.
6. Connell F, Diehr P, Hart G: The use of large data bases in health care studies. Annual Review of Public Health 8:51-74. 1987.
7. Erickson P: Calculating years of healthy life for Healthy People 2000. Presented at this conference.

Funding: The data used to illustrate this paper are from a larger project funded by the Health Care Financing Administration (HCFA) under Cooperative Agreement No. 95-C-99161. The total (direct and indirect) costs of the six-year project, which ends April 30, 1994, are estimated to be $2'318,479. HCFA has funded 75% ($1,731,422) of the total costs, with the balance provide by a grant from the W.K.Kellogg Foundation to Group Health Cooperative of Puget Sound.

Disclaimer: Interpretations of the data are the authors' own and do not necessarily represent the official opinion of the Health Care Financing Authority.

Session B

Infant Mortality

Toward the year

2000

Refining the measures

UNDERLYING CAUSES OF FETAL DEATH: EFFECT OF AN ADDITIONAL PANEL OF CHECK-BOXES ON THE WISCONSIN 1989 REVISED FETAL DEATH REPORT FORM

Russell S. Kirby, University of Arkansas for Medical Sciences
Michael Soref, Wisconsin Center for Health Statistics

Most states implemented the recent revision of the national standard certificates of live birth and fetal death in 1988 or 1989. The new standard certificates have been described elsewhere [Freedman, Gay, Brockert, et al. 1988; Taffel, Ventura, Gay 19891. The new certificates collect data on medical risk factors affecting the pregnancy, obstetric procedures, complications of labor and/or delivery, method of delivery, conditions of the newborn, congenital anomalies of the child, and maternal tobacco and alcohol use and weight gain during the pregnancy.

In the state of Wisconsin, the fetal death report form was further modified through the addition of checkboxes for certifiers to indicate that cause of death and/or congenital anomalies data were 'pending' assessment, and a panel of check-boxes was added to collect information on the types of assessment or evaluation done to determine the cause of fetal death. In this paper we report on results from the first three years of data using the 1989 revised fetal death report form, with comparisons to data from fetal death reports for the two years prior to 1989.

Background

Underlying causes of fetal death are poorly documented in the United States [Kirby 19931. An important proportion of fetal death certificates contain data that lead to nosological classification of causes that are a) implausible - codes that relate to gestation and/or describe the condition of the fetus or define the length of gestation rather than the underlying process which caused the death; b) extremely unlikely - (1) congenital anomalies not clearly incompatible with fetal survival, (2) conditions which should only be reported in newborns but which are unlikely to cause death even in liveborn infants, or (3) conditions primarily relating to respiratory function, with nosological codes which relate only to newborns but which are otherwise valid underlying causes of death; or c) unspecified - codes which impart so little information as to give almost no indication of the actual cause of death (for example: spontaneous abortion, intra-uterine fetal demise, unspecified congenital anomalies, termination of pregnancy, unspecified perinatal conditions, unknown cause). Among five states studied for 1985-1987, the proportion of underlying causes of fetal death which fell into these three categories ranged from 29.5 to 42.8 percent of fetal deaths in each state.

An earlier study compared underlying cause of fetal death and congenital anomaly reporting from Wisconsin fetal death reports with independent stillbirth assessments [Greb, Pauli, Kirby 19871. The Wisconsin Stillbirth Service Project from which the stillbirth assessment data were obtained is described by Pauli and Reiser (19841. There was limited concordance between reported underlying causes of fetal death and etiologic assessments when causes of death were grouped into fetal, placental/cord, maternal/environmental, and unknown. For specific causes reported on fetal death report forms, especially placental/cord complications and fetal anomalies, documentation to support the diagnosis was often lacking. Similar findings have been reported by Harter, Starzyk and Frost [1986] for Washington state and by Lammer et al. [1989] for Massachusetts. In part, as a result of this study, efforts were made to modify the fetal death report form for Wisconsin.

Data Sources

Data for this report were drawn from the computerized statistical files of fetal death reports maintained by the Wisconsin Center for Health Statistics. All valid fetal death reports of Wisconsin occurrence events for calendar years 1987-1991 were analyzed. A total of 900 fetal deaths for 1987-1988 were studied; for 1989-1991, 1,386 events were studied. The panel of seven check-boxes for type of evaluation of assessment added to the 1989 revision of the fetal death report form is shown in Figure 1. Underlying causes of fetal death were grouped into the following categories: infectious disease (ICD-9 001-139), cancer/benign neoplasms (140-239), congenital anomalies (740-759), maternal conditions (760-761), placenta, cord and membranes (762), fetal growth/birth weight (764-766), birth trauma, hypoxia, asphyxia (767-768), other respiratory conditions (770), perinatal infections (771), other and ill-defined perinatal conditions (779, 799), all other perinatal conditions (763, 769, 772-778), and all other conditions (balance of ICD-9). This classification is for purposes of data presentation only: readers are referred to studies by Baird et al. 1954; Georgsdóttir et al. 1989; Joshi et al. 1992; Keeling et al. 1989; McIlwaine et al. 1985; and Raghuveer 1992 for several proposals for classification of causes of fetal death. Other variables analyzed included birth weight, gestational age (estimated from date of last normal menses in completed weeks), timing of fetal death (before labor, during labor, during delivery, not stated), and plurality (singleton or multiple pregnancy).

Results

Almost all occurrence fetal deaths during 1989-1991 had one or more form of evaluation or assessment. However, there were marked differences in the proportion of events with each type of assessment. External clinical evaluation was reported most frequently (54.8% of events), followed by autopsy (51.2%), placental examination (44.1%) and photographs (38.0%). Less common were karyotyping or chromosome analysis (24.2%), cultures for infectious disease (16.4%) and X-rays (14.2%). The trend by year of event is shown in Figure 2. The rate of placental examination declined from 48.4% in 1989 to 39.4% in 1991: a more modest decline occurred for autopsy (52.6% declining to 49.3%); photographs showed the largest increase, rising from 26.1% in 1989 to 49.8% in 1991, while external clinical examinations increased from 51.6% to 64.1%.

Examining data on type of assessment by 500g birth weight strata (Figure 3), no specific conclusions can be drawn, although fetal deaths of 1000-1999g birth weight appear more likely to have autopsy, karyotyping and cultures. By birth weight, external clinical evaluation was most frequent in the smallest and largest fetal deaths, By estimated gestational age, fetal deaths of 37 or more weeks gestation were more likely to be autopsied, photographed, x-rayed, and to undergo placental examination than were fetal deaths with shorter gestations.

By timing of fetal death, 8.6% of the events studied were unknown or not stated, with 78.9%

occurring prior to labor, 10.1% during labor, and 2.5% during delivery. Among fetuses with death prior to labor, assessments by external clinical evaluation, autopsy and karyotyping occurred more frequently than among fetal deaths occurring during labor or during delivery (Figure 4). Surprisingly, cultures for infectious diseases and placental examinations were more common among fetuses with demise during labor, although autopsy and external clinical evaluation were less frequent than among fetal deaths occurring prior to labor or during delivery.

We anticipated more frequent assessments among multiple pregnancies involving fetal demise than among singleton deliveries. Multiple pregnancies were more likely than singletons to be evaluated by photographs and x-rays. However, the proportions karyotyped did not differ between singletons and multiple pregnancies. External clinical evaluation, autopsy, placental examination and cultures were done more frequently for singletons than for multiple pregnancies.

Type of evaluation or assessment was also examined across selected underlying cause of fetal death groups (Figure 5). Among events with an underlying cause of congenital anomalies (7.6% of all fetal deaths), autopsy was the more frequent form of assessment. These cases were more likely to be evaluated by photographs and x-rays, but less likely to have external clinical evaluation. Also, 39.0% of these events received karyotyping, compared to 24.2% of all events.

Approximately 30% of events had underlying causes attributed to complications of placenta, cord, and membranes. These events were more likely to have placental examination (54.8% compared to 44.1% for all events), but less likely to have any of the other forms of evaluation or assessment.

Fetal deaths with fetal growth or birth weight as the underlying cause of fetal death accounted for 7.1% of all events studied. These events were less likely to be evaluated by any method of assessment than other cases, with the exception of cultures (18.2% compared to 16.4%).

Birth trauma and/or hypoxia accounted for 5.3% of all events. These fetal deaths were more likely to be evaluated by external clinical evaluation and photographs, and received autopsy or placental examination at approximately the same rate as all events.

More than a third (37.7%) of all events were classified to other and ill-defined perinatal causes of death. These fetal deaths were slightly, but not remarkably, more likely to undergo external clinical evaluation, autopsy, photographs, x-rays, and karyotyping, but less likely to have placental examinations or cultures than were fetal death with specified causes.

Data for 1989-1991 were compared with data from 1987-1988, to determine whether the addition of the check-boxes had influenced the reporting of fetal cause of death information on the Wisconsin report form. Data on type of assessment were available only for autopsy in the earlier period. The rate of autopsy declined from 57.5% in 1987-1988 to 51.2% in 1989-1991. Autopsies were performed less frequently in each birth weight group except those less than 500g, and in each gestational age group, in the 1989-1991 period.

Figure 6 shows the distribution of the leading causes of fetal death among Wisconsin occurrence events. The same categories constitute the five leading causes in both time periods. There was a decline in the proportion of fetal deaths attributed to maternal conditions, and a modest increase in the proportion attributed to congenital anomalies from 1987-1988 to 1989-1991.

The proportion of Wisconsin occurrence fetal deaths with underlying causes classified as implausible, extremely unlikely, or unspecified for each calendar year is shown in Figure 7. Implausible causes have remained at essentially the same frequency over the period, while unspecified causes have generally been in the range of 35-40 percent, with erratic values for 1990 and 1991. The proportion of events with extremely unlikely causes of fetal death has risen dramatically, and is a source of some concern.

Discussion

The foregoing analyses examine a large series of fetal deaths from the perspective of the forms of evaluation and assessment undertaken to determine the cause of the fetal demise. These data extend our knowledge of how clinicians determine these causes, and offer a number of opportunities for enhancing both the quality of vital statistics data and the comprehensiveness of clinical care through more thoughtful interaction between clinicians, epidemiologists, and public health professionals.

Others have described the role of the autopsy in assessment of stillbirths and neonatal deaths [Carey 1987; Craft and Brazy 1986; Mueller et al. 1983; Winter et al. 19881. Comprehensive evaluation of the stillbirth requires assessments in several areas, including study of the placenta, ruling out various infectious disease processes, testing for chromosomal abnormalities or genetic defects, and dysmorphological assessment where birth defects syndromes are suspected, among others.

Inclusion of the series of check-boxes on type of evaluation or assessment was expected to decrease the proportion of fetal deaths with unspecified causes of death, and to lead, over time, to increases in the proportion of fetal deaths examined more comprehensively in the determination of the cause of death. To date, little change in the leading causes of fetal death has resulted from this innovation. This is primarily due to the limited involvement of clinical specialists or data analysts in the field training program, and to limited querying of reports with implausible, extremely unlikely or unspecified causes of death. The completeness and accuracy of reporting of types of evaluation or assessment is not known. This study reported only on statistical data in the Wisconsin computerized fetal death files. More extensive validation of these data is clearly needed, as evidenced by the fact that the instructions for completing the fetal death report form suggest that all dead fetuses will have undergone external clinical evaluation, whereas only in 1991 did as many as three out of five reports indicate this method of assessment.

The Wisconsin experience with inclusion of the panel of check-boxes on type of evaluation or assessment demonstrates that there exists an unrealized but potentially symbiotic relationship between clinical care, medical quality assurance, and statistical reporting requirements in the field of vital statistics. One example may serve to illustrate this relationship. Numerous clinicians have shown the value of placental examination in routine perinatal care [e.g. Winter et al. 19881. The findings of this study suggest that, far from becoming more common, the proportion of fetal deaths with placental examinations declined sharply from 1989 to 1991 among fetal deaths, a group of events which are far from routine and should merit extra consideration and clinical evaluation. Even among fetal deaths with Causes attributed to complications of the placenta, cord and membranes,

only about half of these cases had reported placental examinations. Multiple gestations, in which abnormal placentation is considerably more frequent, were less likely to have placental examinations, while fetal deaths occurring during labor were more likely to be examined than those occurring prior to labor or during delivery. These findings suggest a need for clinical updates and educational in-services for physicians and nurses involved in labor and delivery care. Field programs to improve compliance with vital statistics reporting requirements should integrate a clinical component which demonstrates the ability of state vital statisticians to evaluate appropriateness of care. The focus should not be solely on completeness of reporting, but rather on the relevance to clinical practice of the data items required on the vital statistics certificate.

The symbiotic link between state reporting requirements, quality assurance for facilities and clinicians, and clinical care must be recognized and integrated in public health statistics systems if we are to realize the full potential of enhancements to vital statistics certificates such as the panel of check-boxes on type of evaluation or assessment for cause of fetal death.

References

Baird D, Walker J, Thomson AM, The causes and prevention of stillbirths and first week deaths. Part III: A classification of deaths by clinical cause: The effect of age, parity and length of gestation on death rates by cause. *J Obstet Gynecol Brit Emp* 1954;n.s., 61:433-448.

Carey JC, Diagnostic evaluation of the stillborn infant. *Clin* Obstet *Gynecol* 1987;30:342-351.

Craft H, Brary JE, Autopsy: High yield in neonatal population. *AJDC* 1986;140:1260-1262.

Freedman MA, Gay GA, Brockert JE, Potrzebowski PW, Rothwell CJ, The *1989 revision* of the US standard certificates of live birth and death and the US standard report of fetal death. Am *J Public Health* 1988;78:168-172.

Georgsdóttir I, Geirsson RT, Jóhannsson H, Biering G, Snædal G, Classification of perinatal and late neonatal deaths in Iceland. A survey from a defined population. *Acta Obstet Gynecol Scand 1989;*68:101-108.

Greb AE, Pauli PM, Kirby RS, Accuracy of fetal death reports: Comparison with data from an independent stillbirth assessment program. Am *J Public Health* 1987;77:1202-1206.

Harter L, Starzyk P, Frost F, A comparative study of hospital fetal death records and Washington state fetal death certificates. *Am J Public Health* 1986;76:1333-1334.

Joshi AR, Daga SR, Daga AS, Perinatal audit through Wigglesworth's classification. *Ind Pediatr* 1988;25:525-529.

Keeling JW, MacGillvray I, Golding J, Wigglesworth J, Berry J, Dunn PM, Classification of perinatal death. *Arch Dis Child* 1989;64:1345-1351.

Kirby RS, The coding of underlying cause of death from fetal death certificates: issues and policy considerations. Am *J Public Health* 1993;83:1088-1091.

Lammer EJ, Brown LE, Anderka MT, Guyer B, Classification and analysis of fetal deaths in Massachusetts. *J A MA* 1989;261:1757-1762.

McIlwaine GM, Dunn FH, Howat RC, Smalls M, Wyllie MM, MacNaughton MC, A routine system for monitoring perinatal deaths in Scotland. *Brit J Obstet Gynaecol* 1985;92:9-13.

Mueller RF, Sybert VP, Johnson J, et al., Evaluation of a protocol for post-mortem examination of stillbirths. N *Engl J Med* 1983;309:586-590.

Pauli RM, Reiser CA, The Wisconsin Stillbirth Service Project: The "Wisconsin Idea" in practice. Wis *Acad Rev* 1984;December:62-64.

Raghuveer G, Perinatal deaths: Relevance of Wigglesworth's classification. *Paediatr Perinat Epidemiol* 1992;6:45-50.

Taffel SM, Ventura SJ, Gay G: Revised U.S. certificate of birth--New opportunities for research on birth outcome *Birth 16* (1989), *188-93.*

Winter RM, Knowles SAS, Bieber FR, Baraitser M. The pathological examination of the infant and examination of the placenta. In: Winter RM, ed. *The Malformed Fetus and Stillbirth.* Chicester, England: John Wiley and Sons, 1988:40-68.

FIGURE 1

STATE OF WISCONSIN
DEPARTMENT OF HEALTH AND SOCIAL SERVICES

REPORT OF FETAL DEATH

FOR STATISTICAL PURPOSES ONLY

43. ASSESSMENT (Check Each Done)
01A ☐ ..... External ClinicalEval.
02B ☐ ..... Autopsy
03C ☐ ..... Photographs
04D ☐ ..... X-Rays
05E ☐ ..... Chromosomal Studies
06F ☐ ..... Placental Exam
07G ☐ ..... Cultures (Infectious)

FIGURE 2

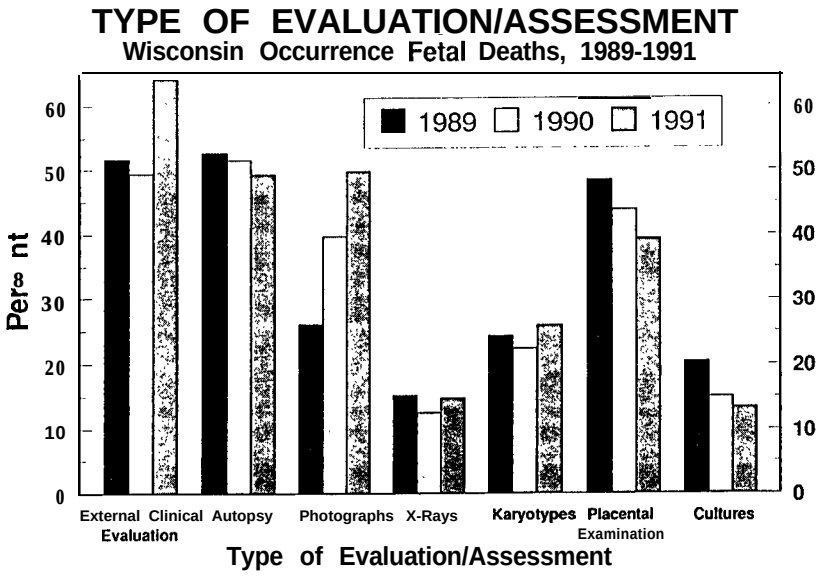


FIGURE 3

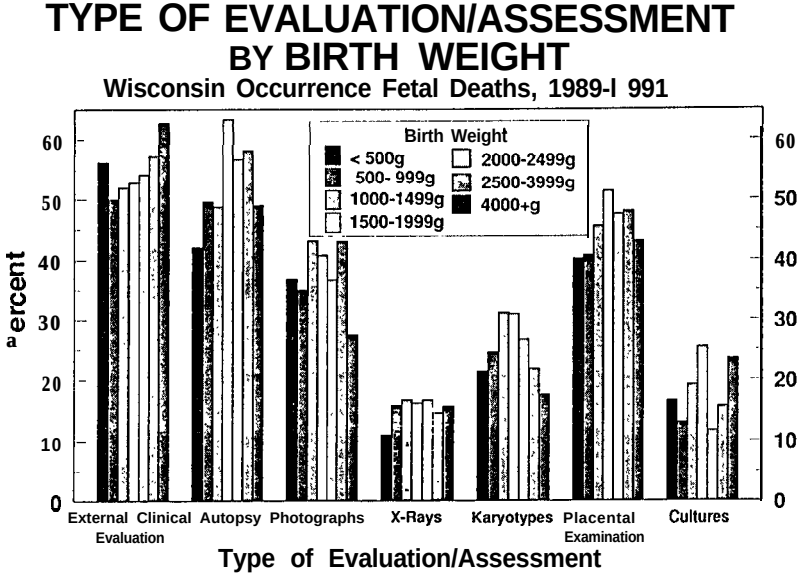


FIGURE 4

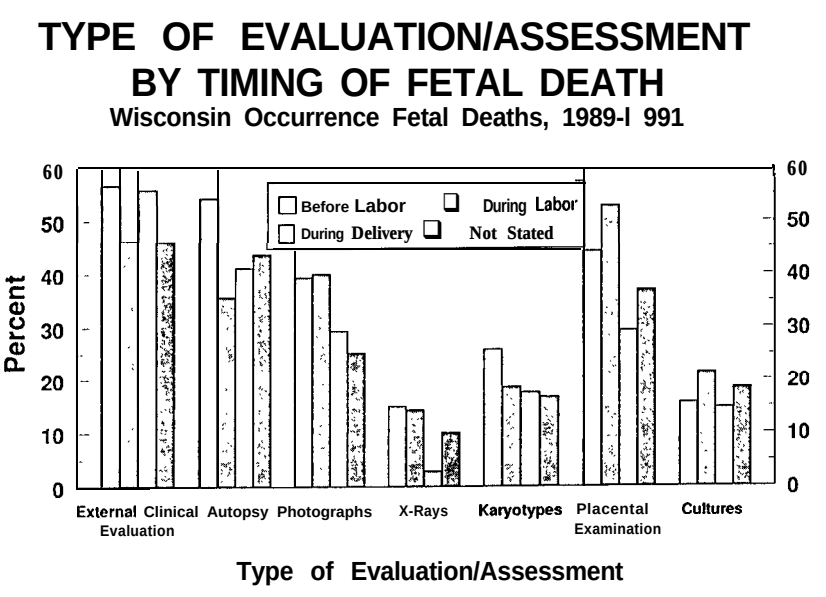


FIGURE 5

TYPE OF EVALUATION/ASSESSMENT FOR SELECTED UNDERLYING CAUSES OF FETAL DEATH

Wisconsin Occurrence Data, 1989-l 991

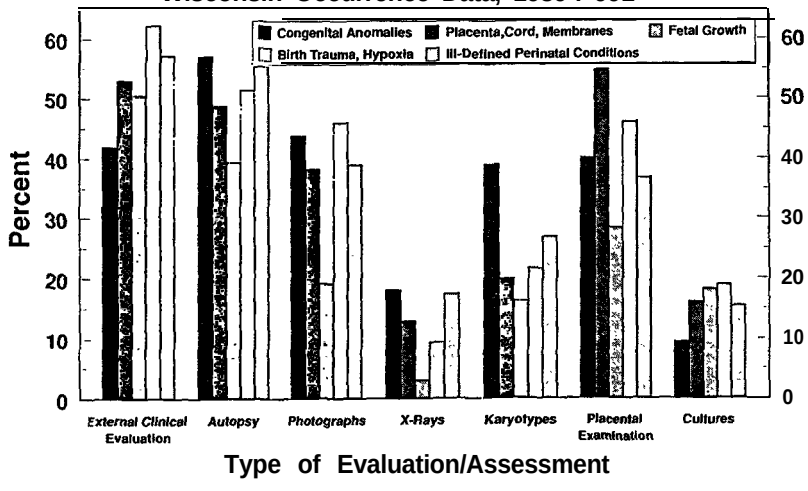


FIGURE 6

LEADING CAUSES OF FETAL DEATH WISCONSIN, 1987-88 AND 1989-91

	1987-1988			1989-1991		
	N	Percent	Rank	N	Percent	Rank
Other and Ill-Defined	337	37.4	1	523	37.7	1
Placenta, Cord, Membranes	272	30.2	2	418	30.2	2
Maternal Conditions	108	12.0	3	100	7.2	4
Fetal Growth/Birth Weight	71	7.9	4	99	7.1	5
Congenital Anomalies	57	6.3	5	105	7.6	3
Five Leading Causes	845	93.9		1,245	89.8	

FIGURE 7

IMPLAUSIBLE, EXTREMELY UNLIKELY, AND UNSPECIFIED UNDERLYING CAUSES OF FETAL DEATH, WISCONSIN, 1987-1991

	1987	1988	1989	1990	1991	Total 1987-1991
Implausible	0.9	0.5	1.2	0.4	0.5	0.7
Extremely Unlikely	2.0	1.4	0.0	5.9	8.6	4.6
Unspecified	39.7	36.1	39.1	31.5	44.4	38.2
Total	42.6	37.9	40.3	37.8	53.5	43.6

Status and Effectiveness of Perinatal Regionalization in Virginia

Sherry Allison-Cooke and Rachel M. Schwartz

Introduction

This study, funded by the Commonwealth of Virginia explored the effectiveness of perinatal regionalization from both a process and outcome perspective. The concept ofperinatal regionalization was formally promulgated in the mid 1970's through a document called "Toward Improving the Outcome of Pregnancy." The goal of the document, and those who developed it, was to promote uniformity nationwide concerning the perinatal care of high risk pregnancies and high risk newborns. Perinatal Centers or tertiary level III centers were to become the focal point of wide range of activities within designated regions providing risk assessment, emergency care, consultation, outreach/education, transportation, systems management, coordination of care, and specialty services to high risk mothers and newborns.

In some locations activities were formal in others they were not. In Virginia the efforts were more formalized in nature. In 1980, Virginia established a State Perinatal Services Advisory Board and in 1983 the Board developed a Statewide Perinatal Services Plan which recommended a formal coordinated regionalized system of care. Seven perinatal regions were develop, all except one, had a designated perinatal center devoted to high risk services. In the late 1980's the Perinatal Services Advisory Board (PSAB) requested an assessment of the effectiveness of regionalization and this paper presents the highlights of that assessment.

Method

The methodological approach was qualitative and quantitative in nature. The qualitative analysis focused on identifying current perceptions of gaps, strengths and weaknesses of each region as perceived by hospitals staff and physicians both within perinatal centers and outside. The quantitative analysis focused on measuring the regionalization process and quantifying its outcome. This paper describes the quantitative process and outcome analysis and findings.

The regionalization process analysis asks the question. Did regionalization occur? The most easily used measure would be whether or not high risk infants were more likely to be born in a tertiary center care center post regionalization. In Virginia the process of regionalization was initiated in the late 1970's so a trend analysis which examined the location of births over ten years for each region was used to identify if a clustering of high risk births in regional centers occurred. While there are probably numerous ways to measure high risk births using the birth and death certificate this study selected very low birthweight births (those between 500 and 1500 grams) as the least controversial measure. The study examined the change in the percentage of very low birthweight births in Perinatal Centers, and the proportion VLBW deliveries that occurred in tertiary settings beyond what would have been expected to occur based on the hospitals share of all births in its service region. The latter measure was called the "residual" because it is the proportion of the residual market which the perinatal center had obtained. This approach is a modified version of that published by Powers and McGill.[1] In Virginia calculating the proportion a Perinatal Center had of the "residual" was relatively easy because the perinatal regions were defined by the state and did not need to be created for the study. The calculation is relatively straight forward and is done by:

1. identifying the proportion of VLBW births at the perinatal center of all births in the region, or "market share".
2. subtracting those from the total; and
3. calculating the proportion of the remaining VLBW newborns that were born at the center. That is, how many more VLBW births occurred beyond what would be expected?

In this approach births in the center are credited to the center while neonatal transports are not. This is a tough measure of the regionalization process because it demands referral of the pregnant women either during prenatal care or in the early stages of labor. This approach reflects the preferred trend, that is, assuring that the high risk birth occurs in the tertiary center through a combination of patient and physician education and when necessary maternal transport.

To perform the process analysis the study documented the obstetrical inpatient care service delivery system in each region, i.e. which hospitals were perinatal centers; which were level two facilities, and whether these were similar across regions. A specially designed survey of all obstetrical hospitals in Virginia was used to score sophistication based on services and staffing. The residual marketshare analysis was then performed for each regional center in every region. Data were aggregated into five, two year intervals to reduce the problem of small numbers.

To examine the impact of regionalization on outcome the study used as its primary measures adjusted early neonatal mortality (AENM) by hospital within region for each of the five year two year intervals. The adjustment was done by applying the birthweight specific mortality rates of Virginia to the 1985 birthweight distribution for the U.S. This allowed for a trends analysis which controlled for changes in the birthweight distribution, thus any changes in biiweight distribution would not be an issue for this analysis. Finally, the analysis was performed on where biis occurred. Thus, biis in a region, were based on the location of birth not residence of mother. However the study did examine births by residence of mother to understand the out-of-state utilization and its impact on perinatal regionalization.

---

[1] Powers, WF McGill L, "Perinatal Market Penetration Rate. A Tool to Evaluate Regional Perinatal Programs" AMJ Perinatal; 4:24-28, 1987.

The Virginia birth-death certificate data were used for all the analyses. It should be noted that while it would be convenient to use national data files to allow comparison across states, national files do not include the hospital of birth or detailed enough address information on the mother. To perform this analysis as described both are needed. In addition, state hospital discharge records cannot be used because deaths can not attributed to the hospital of birth and birthweight is often not collected.

Findings

Process Analysis

Over all, in Virginia in 1987-1988 the analysis showed that 90.5% of all births occurred within region of residence. Of out-of-region births 3.7% occur in Virginia and 5.8% occur out-of-state. VLBW births out-of-state are more critical to understanding regionalization, 174 or 8.6% of all VLBW births occurred out of state, of these 143 or 82% were from only two regions and it is not known if they were delivered in a perinatal centers, but there was reason to believe from the qualitative component of the study that they did go to perinatal centers. The out-of-state issue is discussed further below.

All regions of the state experienced a trend toward regionalization as measured by the percentage of the residual market for each perinatal center. Of critical importance is the clear trend for the entire state – perinatal centers' share of the residual, increased by 90.3% over the period studied while all other hospitals' share dropped by 40%. While all regions increased there was variation: increases were 30.4, 40.5, 45.8, 56.9, 57.9.

The graph in Exhibit A shows the change between 1979 to 1988 for each region. Here it is very clear that some subsections of the state are more regionalized than others. There are three regions – A, C, and E where regionalization appears to be less active. The case study analysis and the hospital surveys revealed that in region E there is only one designated perinatal center, but there are three other NICU's in the region. When VLBW newborns born in any hospital with a NICU are combined it was found that 93% of Region E VLBW newborns are born in appropriate settings, as such analysis which uses only the residual market share of the perinatal center can be misleading. Here the centralized component of regionalization was not adhered to, yet most infants were born in an appropriate location. The outcome analysis examines if it matters whether multiple sites provide care.

Region A is quite different no hospital strictly qualified as a perinatal center, but there is a hospital with a special care nursery. In region A 50% of VLBW birth are born out-of-state. Although it is unknown where these infants are born it appears that they rely on a Perinatal Center in a nearby state. Nonetheless 2/3's of those VLBW infants born in the region A are not born in the special care nursery.

Region C has virtually no out-of-state VLBW births and only 12.7% of the residual occurs in the Perinatal Center. Analysis of where women in Region C deliver showed that 41.8% of the VLBW births occurred in a perinatal center in another region, thus 69% of all VLBW births to women residing in region C occurred in a perinatal center. However, like region A significant proportion of the VLBW births which occurred in the region not occur in a perinatal center.

The graph in exhibit B shows VLBW newborn percentages for perinatal centers. Only region A and C are noticeably lower than the others. That is 2/3's of VLBW births which occurred in region A and slightly more than half in region C occurred in a non-perinatal center. From perspective of where VLBW infants are born it appears that Virginia did very well with regard to regionalization except there is room for improvement in Region A and *C*.

Outcome

The next component of the analysis explored how regionalized high risk care impacted on newborn outcome. As already noted we examined birthweight adjusted early neonatal mortality (AENM) as the primary measure of outcome. Statewide we found that AENM dropped from 6.5 to 3.9 or 38% (see Exhibit C). AENM declined for all regions, but remained highest in the two regions A and C, those with the lowest percentage of VLBW infants born in tertiary settings. These two regions are 24% and 25% above the rate for the state. Interestingly region E, with the highest number of Neonatal Intensive Care Units has the lowest mortality.

Both perinatal center and non-perinatal center mortality have improved dramatically during the period studied. But the analysis also revealed that perinatal centers perform far better than non-perinatal centers even in regions where market share could be improved. For example in 1987-1988 the perinatal center in region C had AENM of 4.2 compare to 6.0 for the non-perinatal centers. In 1987 and 1988, the overall non-perinatal center adjusted early neonatal mortality rate was 27% higher than for perinatal centers. The message of these data is clear – the location of biih is important to the infant outcome.

A final issue addressed indirectly in this study, is the issue of race. The data were analyzed to assess whether minority infants benefited equally from regionalized care. Although this was not a focus of the study this issue deserved attention. The data revealed that the AENM for whites declined 43.8% while for non-whites it declined 31%. For whites it dropped for all regions, but for non-whites AENM actually rose in two regions, in region C it rose from 6.3 to 7.3 and in region D it rose from 3.8 to 4.2. In addition, in these two regions non-white AENM was lower than for whites in 1979-1980, but was higher than that of whites in 1987-1988. These findings would seem to indicate that the benefits of regionalized care accrued unevenly by race in these two regions.

The report recommended very specific actions to the Commonwealth of Virginia for each region all aimed at improving referral patterns, coordination and collaboration because the study identified very specific improvements that could be made. The study however demonstrated that regionalization not only was implemented, but had a dramatic effect on outcome. Graph D shows perinatal centers: increased their total share of VLBW births by 67.4% while non perinatal centers declined 32.7%; increased their residual market share by 90.3% while non-perinatal centers declined by 39.5% and reduced adjusted early neonatal mortality declined by 42.9% while non perinatal centers rates dropped by 33%. The most important result of perinatal regionalization was that adjusted early neonatal mortality dropped by 40% statewide.

Exhibit A

Percentage of Residual VLBW Births by Perinatal Center in Virginia 1979-1988

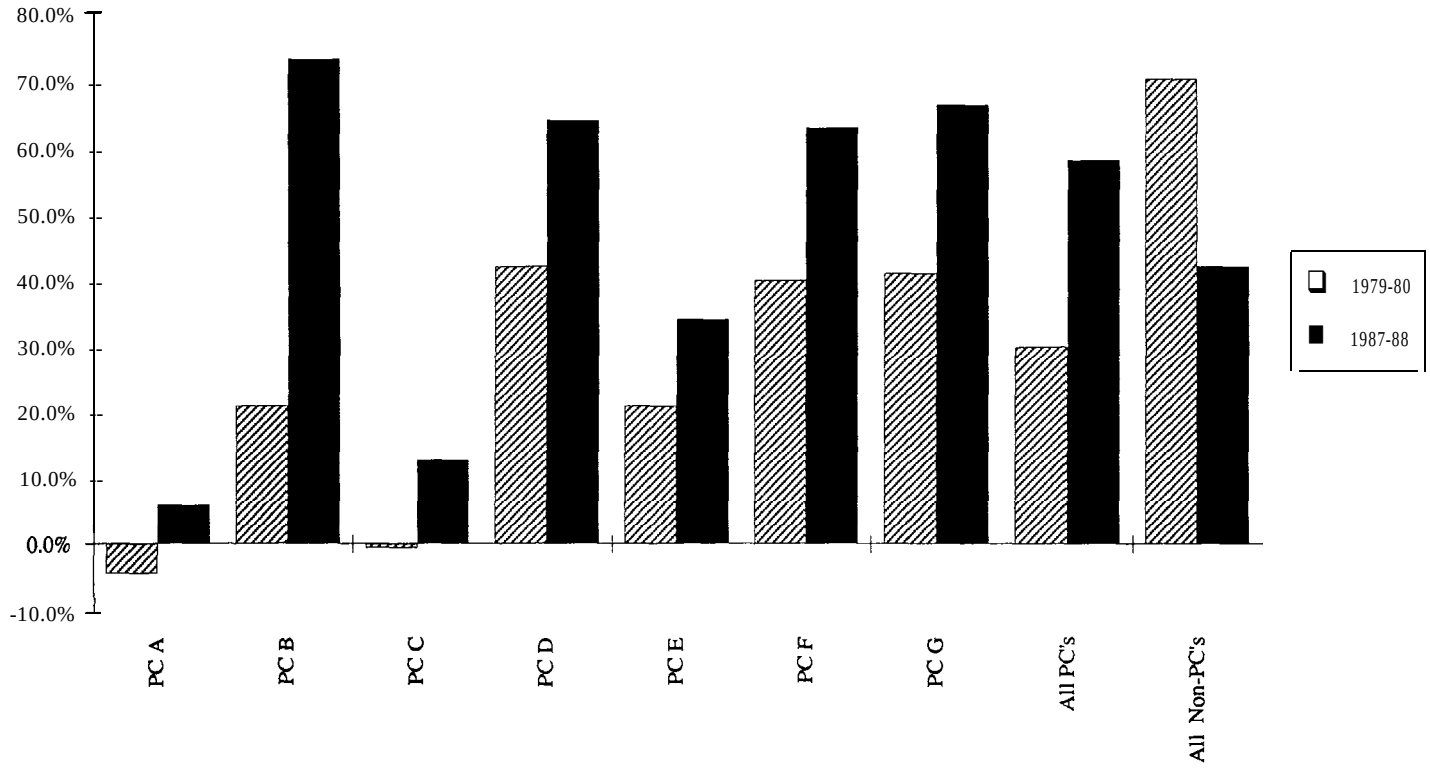
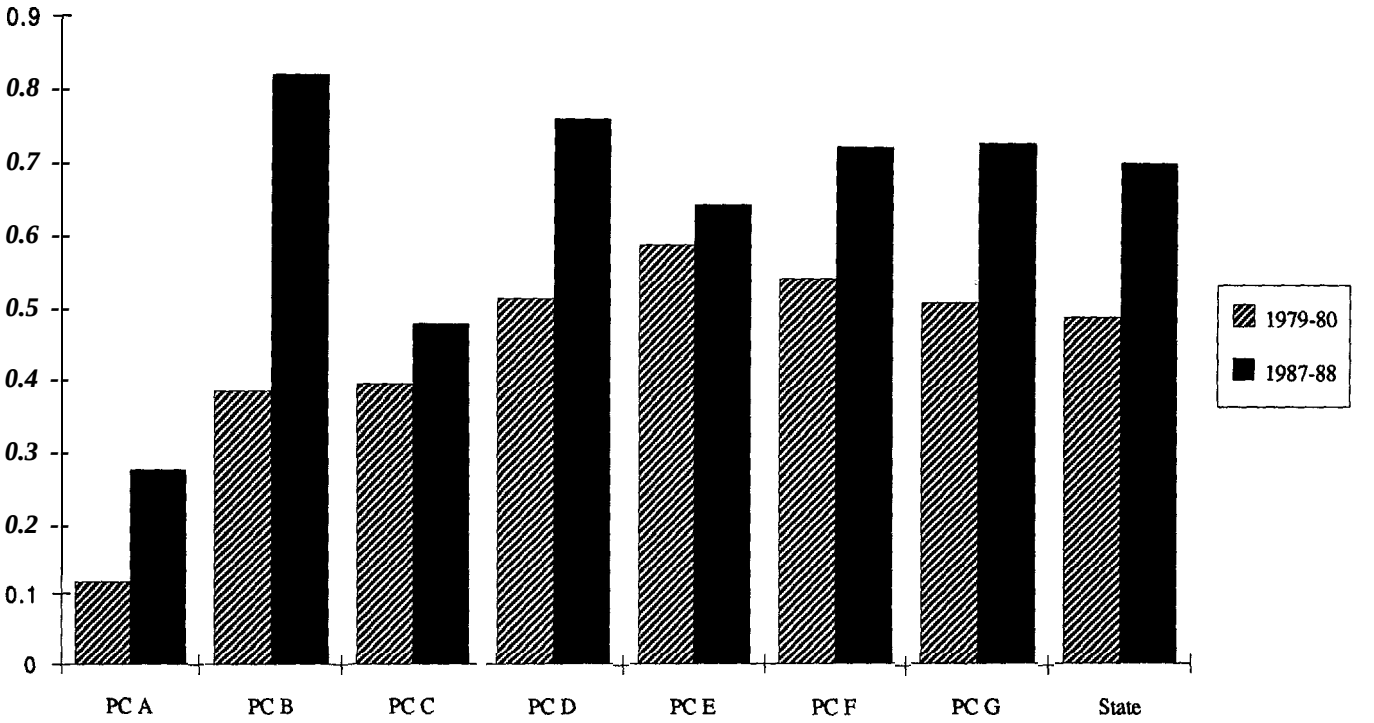
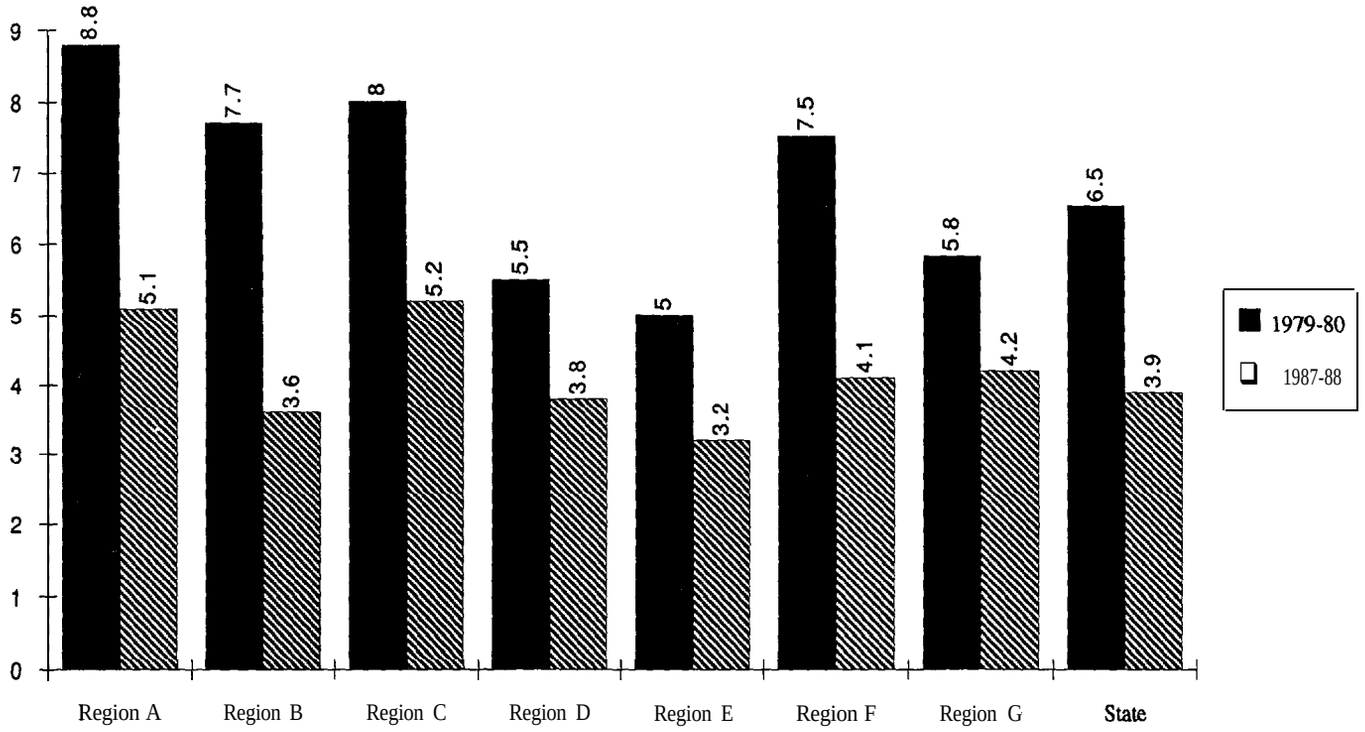


Exhibit B
PERCENT OF VLBW BY PC:
VA, **1979-80** TO **1987-88**



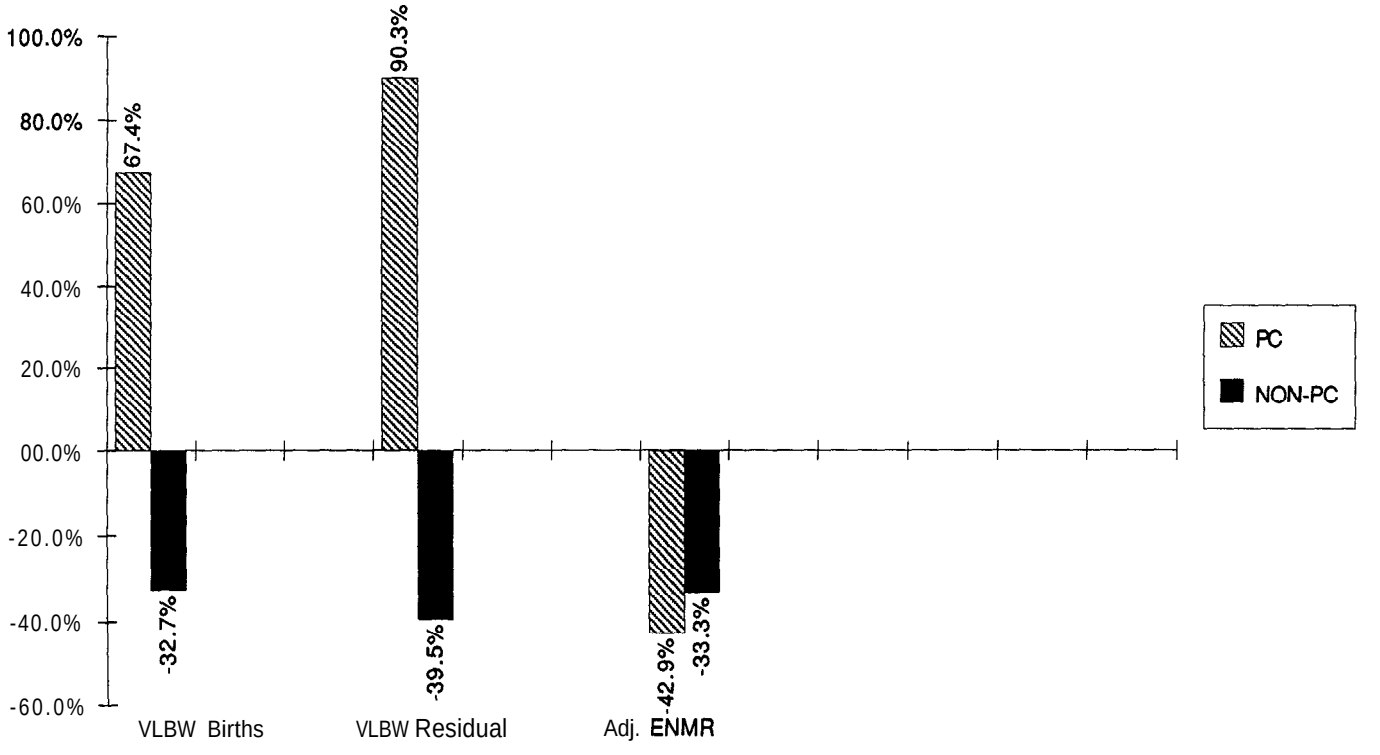
NPIC Evaluation of Perinatal Regionalization in Virginia

Exhibit C
EANMR BY REGION:
VA, **1979-80** TO **1987-88**



NPIC Evaluation of Perinatal Regionalization in Virginia

Graph D
PERCENT CHANGE IN SELECTED INDICATORS:
PC VS. NON-PC, VA, 1979-80 TO 1987-88



A METHODOLOGICAL APPROACH TO IMPROVING PREGNANCY OUTCOMES IN AFRICAN AMERICAN WOMEN

Cecile H. Edwards, College of Allied Health Sciences, Howard University
0. Jackson Cole, Ura Jean Oyemade, Enid M. Knight, Allan A. Johnson,
Ouida E. Westney, and Lennox S. Westney

I am honored to have been invited to present some aspects of research in the Howard University Program Project, **" Nutrition, Other Factors, and the Outcome of Pregnancy."**

Although there has been a steady decline in the infant mortality rate in this nation, the two foldhigher incidence of infant mortality in African Americans has continued to persist. The same pattern has existed for low birth weight; the incidenceof lowbirthweight (LBW) infants delivered to African American women is twice that of Caucasian or Hispanic women. But, because a larger number of low birth weight African American infants survive, the human suffering which attends the higher incidenceoflowbirthweightismagnifiedinthegreater numbers of black infants with learning difficulties, visual and hearing problems, autism, cerebral palsy, chronic lung problems, mental retardation, and birth defects. Although researchers have given intensive study to factors influencing infant mortality and low birth weight, little progress has been made in understanding these problems.

In our study, the incidence of LBW was reduced from 20.6%, the 1988 incidence in the Howard University Hospital (HUH) , to 8.9% during the three year period 1985-88. In women enrolled in the prenatal clinic who were eligible for the project, but not recruited, the incidence of LBW was 19.4%. In this paper we describe factors which may account for this reduction and make recommendations to decrease LBW in low income populations.

METHODS

This longitudinal, prospective observational study was designed to describe low income women during their pregnancies. Purposive sampling was employed in the recruitment of women who had been admitted to the prenatal clinics of the HUH or the Department of Human Services of the District of Columbia. The objective of the program project is to investigate nutritional, medical, biochemical, psychosocial, socioeconomic, lifestyle, and environmental factors that influenced the outcome of pregnancy in these woman.

The eligibility requirements of the project are as follows: participants were (a) nulliparous, (b) between the ages of 18 and 35 years, (c) free from abnormal hemoglobins (such as sickle cell disease, thalassemia, and hemoglobin C) and diabetes, and (d) not more than 26 weeks gestation at the time of enrollment.

After the women had given informed consent, the project- trained interviewers administered the Entry Questionnaire which provided demographic, lifestyle, dietary, and sociocultural information. Fasting blood samples were taken during each trimester of pregnancy at a nurse-scheduled regular clinic appointment. All other procedures were performed routinely in the HUH Prenatal **Clinic** or the clinics of the D.C. Department of Human Services. Women recruited in the latter prenatal clinics had to consent to be delivered at the HUH as a requirement for participation in the project.

Researchproject **interviewers** werecautionednot to give information on nutrition or health, as this was tobeprovidedaccordingtothe usual procedures of the clinicbyits staff. The Project Clinical Nutritionist, who met the participants at each prenatal clinic visit, mailed appointment reminders to each participant on Thursdays and telephoned them the day before their appointments. Soon after they had been recruited, participants telephoned the Clinical Nutritionist on Sundaystotalkabouttheirproblems, oftencollectively more than three hours each week.

Physical and behavioral assessments were performed within 48 hours afterbirth, andincludedthe Braselton for which the project infant examiners were specially trained.

Participants completed 24 hour food and lifestyle recallandreceivedaseriesofpsycholcgicalassessmants duringtheirregularmonthlyorbiweeklyvisitsscheduled by the clinic nurses. Analyses in blood were made by standard clinic procedures for 24 biochemical substances. The generic "stress" or "anxiety" construct was formulated from several validated instruments. In addition, a psychosocial interview was conducted which assessed body image and the social support network.

RESULTS

Demographic Characteristics

Eighty percent of the women were single and 17 percent were married. The mean age of the woman was 22 years and 38 percent had completed high school. Forty-one percent of the household incomes were below $11,000/ year.

Sources of Stress in African American and Caucasian Environments

Demographic and Economic Stressors

Any profile of demographic, social, andeconomic factors which characterize Caucasians and African Americans in 1990 in the United States is suggestive of sources of stress which disproportionately visit African Americans in this decade. Life in the central cities ofthenation, withitsexposuretogreaterenvironmental pollution, crime, and a higher level of unemployment, exacerbates the stress level already imposed on single parent households. In addition, public documents (1) over the years show that African American people have less income and wealth, liveinmoredilapidatedhousing, have access to a lower quality of social and public services, have fewer years of school attendance, and suffer a higher rate of incarceration than their Caucasian counterparts.

Psychosocial and Physiologic Stressors

Although crowding, excessive air and noise pollution, violence, and crime contribute heavily to stress from the external environment, African Americans are uniquely exposed to a higher incidence of specific stressors associatedwiththeintemalenvironmsnt: for example, lactoseintolerance, obesity, andhypertension. The continuing psychosocial stressor is racism. One of the consequences of racism is the loss of self esteem and the intensification of the feeling of hopelessness. Until recently, however, the possible impact of racism on low birth weight andpremature delivery has received little attention.

Nutrition, Stress, Social Support, and Pregnancy Outcome

It is widely recognizedthatmatemalnutrition, social support, stress, substance abuse, and early prenatal care are related to pregnancy outcome (2,3).

In the Nutrition Program Project, my colleagues have made significant observations relating psychosocial factors with pregnancy outcomes (4):

- Fewerpregnancy symptoms were associatedwith having a positive body image, heightened self-esteem, fewer stressful life events, relativelyloweranxietylevels, and higher concern about the delivery and birth.
- There was a significant positive associationbetween the frequencyofpositiveandnegativeinteractionswith social network members.
- Anxiety was negatively associated with the degree of self-esteem, the number of disagreements and the time spent together with the partner, as well as the rated degree of happiness in the relationship.
- Greater frequency of pregnancy symptoms was associated with reports of more stressful life events, higher anxiety levels, andraorematemalanxietyconcemingthe delivery and birth.

More recently, we observed that the size of the social support system was positively correlated with maternal serum vitamin E (n= 48, r= .35, P= 0.015) andfolate (n= 140, r= .24, P= 0.004) at delivery, and with maternal plasma zinc (n= 181, r= .20, P= 0.007) duringthethird trimester. Zinc is essential for immune function and vitamin E enhance3 immune response.

Social Support, Health, and Immune Function

Berkman and Syme (5) reported that the significantly lower mortality risk3 in their patients with most social ties was independent of physical health status; health behaviors such as alcoholandtobaccouse; physical activity; utilization of preventive health services; and socioeconomic status.

In epidemiologic studies, decreased mortality and morbidity are associatedwith strong social support systems. Bovard (6) hypothesizedthatstressfulstimuli initiate a complex series of neurcendocrine responses which lead to general catabolism and immunosuppression, whereas supportive relationships set in motion a competing response which modifies and helps to neutralize these harmful effects. Thomas, Goodwin, and Goodwin (7) suggest that social support systems may intervene between the stressful stimuli and the physiologic responses to those stimuli.

Oakley (3) concluded, after reviewing the recent literature on social support in pregnancy as a medium for decreasing the incidence of low birth weight, that there is considerable evidence to suggestthatprcgrams of intervention with the objective of improving social support in antenatal care are capable of influencing birth weight and other clinically acceptable measures of pregnancy outcome.

Psychosocial Factors and Pregnancy Outcomes

In Table 1 are shown correlations of other maternal psychosocial factors with infant gestational age and head circumference which have been observed in our research.

Although the mechanism for the positive effect of selfesteemongestationalage andheadcircumference cannotbeidentifiedfromourdata, we believe that these observations hold important implications for care delivery in the public health arena. The methodological approach to prenatal care shouldevolve in the hospital environment from one of "assembly line care" to one of empowering the low income women who come forhealthcare. We can empower low income women, who often have low self esteem and less than average support systems (4).

CONCLUSIONS

Intheprenatalclinicpopulationenrolledinthis study, we have hypothesized that the presence of an additional "support system," (provided unintentionally by the programproject as its Clinical Nutritionist and trainedinterviewers from 1985 through 1988, may explain the remarkable reduction in the incidence of low birth weight within an urban hospital. There are important methodological implications from this research to the reductionof low birth weight by practical and manageable means in the low income population.

Although review3 oftheeffectofmatemalsocial support onmatemal health consistently show a positive effect, considerably greater attention need3 to be focused on this variableinthe public healtharena. The following recommendations are made: (a) Training of nurses who work in the prenatal clinic to give culturally relevant psychosocial support to women who come for care; (b) Refocusing the special sensitivity of physicians assigned to the prenatal clinic on the likelihood of higher stress levels in low income women and the need to stimulate higher self esteem in them and decrease anxiety, where possible; (c) Addition to staff of prenatal clinics andsocialwelfare agencies of a social worker or other person who could serve a3 a listener or confidant to provide a valve for stress. This could reduce not only the higher rate of low birth weight in this nation among African Americans, but also the human suffering attendant to raising an infant with more than one indemnity at birth.

However, it may well be that those who do not come for prenatal care prior to delivery consider their personal problems too complex and incriminating to take a chance onincreasingtheir intensity by confiding them in nurses, doctors, and other health care personnel associated with the prenatal clinic. In this scenario, stress could decrease care seeking. Those who would attempt to bring them in earlier, will need to make the environment of interaction one of confidentialityandto advertise this characteristic in public messages to women who remainoutside ofthehealthcare systemduringtheir pregnancies; (d) A viable alternative method for reducing stress would be the installation of a "Hot Line" in prenatal clinics where the pregnant women could anonymously call into have questions answered, or just to talk with the person who is qualified to answer their question3 and willing to listen to them. This recommendation for the installationof a Hot Line can be implemented wherever low income pregnant women go...in offices which give assistance with Food Stamps, Aid to Families with Dependent Children, WIC, and employment, ... and in churches and schools a3 well. Thus, those who do not come to hospitals for prenatal care also can access this practical approach to stress reduction.

Cole and his coworker3 (4) have highlighted the centrality of the roles of self esteem, anxiety, stressful life events, social network size, and the quality of social network interactions inthepregnancy experience of African American women. It will be of interest to determine whether the provision of stress

valves during pregnancy in African American environments can reduce the disproportionately higher incidence of low birth weight in African American women.

TABLE1

CORRELATION OF PSYCHOSOCIAL VARIABLES WITH PREGNANCY OUTCOMES IN AFRICAN AMERICAN WOMEN

Psychosocial Variables and Pregnancy Outcomes	N	P value
GESTATIONAL AGE		
More Respect for Self	343	0.001
Do Things Same as Most People	187	0.042
Number of Social Network Persons	339	0.000
HEAD CIRCUMFERENCE		
I am a Person of Worth	191	0.011
TERM DELIVERY		
Positive Self Attitude	180	0.036

REFERENCES

1. U.S. Department of Commerce, Bureau of the Census. Summary Population and Housing Characteristics: 1990; Current Population Survey, March 1990; Money Income and Households, Families, and Persons in the U.S.: 1990; Poverty in the United States: 1990. Washington, D.C.: U.S. Government Printing Office. September 1991.
2. Thompson JE. Maternal Stress, Anxiety, and Social Support during Pregnancy: possible directions for prenatal intervention. In Merkatz IR et al., (eds) New Perspectives on Prenatal Care. 1990; 319-35.
3. Oakley A. Social Support in Pregnancy: the "soft way" to increase birthweight? Soc Sci Med 1985; 21: 1259-68.
4. LaPoint V, Alleyne SI, Cole OJ, Edwards CH, Westney OE. Ecological Perspectives on Critical Issues Confronting African American Families: In Sontag MS, Wright SD, Young GL, (eds) Human Ecology: Strategies for the Future. East Lansing: Society for Human Ecology 1991: 117-46.
5. Berkman LF, Syme SL. Social Network, Host Resistance, and Mortality: a nine-year follow-up study of Alameda County residents. Am J Epidemiology 1979; 109: 189-203.
6. Bovard EW. The Effects of Social Stimuli on the Response to Stress. Psycho1 Rev 1959; 66: 267-77.
7. Thomas PD, Goodwin JM, Goodwin JS. Effect of Social Support on Stress-related Changes in Cholesterol Level, Uric Acid level, and Immune Function in an Elderly Sample. Am J Psychiatry 1985; 142: 735-37.

INTERMEDIATE VARIABLES AS DETERMINANTS OF ADVERSE PREGNANCY OUTCOME IN HIGH RISK INNER CITY POPULATIONS

R.K. Sharma, University of Pittsburgh

Introduction - Delineation of the pathways through which social and economic factors operate to influence adverse pregnancy outcomeslhas been an important analytical advance. This study focused on the role of maternal medical risks and health behavior during pregnancy as intermediate variables through which social and economic factors influence preterm delivery and low birth weight - common endpoints for the expression of many social and biological processes that influence birth outcome.[2] Gestational age, as reflected in preterm delivery, and birthweight are causally related and are important predictors of perinatal mortality.[3]

Methods - The analysis uses data that was originally collected to assess baseline maternal and child health status for a federally funded initiative "Healthy Start" that is designed to decrease infant mortality in Pittsburgh, PA.[4] It was based on telephone interviews of 1004 women of childbearing ages (15 to 44 years), selected through random digit dialing procedure, who had a live birth within three years before the survey. The focus on live birth may lead to a selection bias because of exclusion of pregnancies that do not end in a live birth, resulting in the diminution of the adverse impact of socio-demographic factors. A potentially serious selection bias resulting from the use of random digit dialing procedures is that it ignores women from households without telephones who might be at high risk of adverse pregnancy outcomes. The interpretation of the data is also influenced by the fact that women in the sample are from geographic areas of Pittsburgh and vicinity that are predominantly African-American and most of the populations have low incomes and education. The survey instrument was constructed from selected questionnairs used in the "National Maternal and Infant Health Survey" developed by the National Center for Health Statistics and "Pregnancy Risk Assessment Monitoring System (PRAMS)" developed by Centers for Disease Control. The survey, which took place between March 19th and May 12th, 1992, targeted women living within six high risk geographic areas. Five areas were located in Pittsburgh, Pennsylvania, the sixth in the neighboring municipalities of Duquesne and Braddock. These areas had an average infant mortality rate of 19.8 per 1,000 during 1988-1990, compared to 10.4 per 1,000 for Allegheny County, Pennsylvania, the county that encompasses these communities. The nonwhite infant mortality in the study area was 26.3 per 1,000, about three times higher than the white rate of 9.2 per 1,000. Logistic regression is used to estimate the adjusted odds ratios for preterm delivery and low birth weight. SPSS PC+ V4.1, a statistical software package, was used to estimate the logistic regression coefficients.[5]

In this exploratory analysis, only the main effects are estimated. The original intent of the survey was to collect data on major risk factors related to adverse reproductive outcomes and not to test specific hypotheses. Therefore, given the sample size, the statistical power is insufficient to detect hypothesized effects that are small in size.

Results and Discussion - Analysis was based on the hypothesis that sociodemographic and economic factors influence preterm delivery and low birthweight through maternal health status, maternal health behavior and service utilization. Sociodemographic factors that elevate the risk of preterm delivery (defined here as live births of less than thirty-seven weeks of gestation) and low birth weight (defined here as live births of less than 2,500 grams) include; younger age (less than 20 years), marital status (single), education (less than 12 years) and race (black). Maternal age was included as a risk factor even though there is debate as to whether it is an independent determinant of pregnancy outcome.[6] Economic status was measured by household income. Low economic status, as a social class indicator, can be expected to increase the risk of adverse pregnancy outcome, despite the fact that existing research has not been able to determine the mechanism that connects economic status to adverse outcomes. Table 1 presents unadjusted odds ratios for low birth weight and preterm births. Being single, black or having relatively low income has a negative impact on pregnancy outcome. These factors tend to cluster since black mothers are likely to be mostly single and come from low income households. However, when other factors are taken into account (see Table 2) none of these variables were significant predictors of adverse pregnancy outcome in this study population. This was expected since sociodemographic and economic variables influence birth outcomes through intermediate variables which are controlled by the analysis.

Two indicators of maternal health status were used; (i) the number of health problems experienced during pregnancy, and (ii) prior fetal loss. Inner city residents without access to preventive health care and timely medical care are likely to have poor pregnancy health status and consequently are at a higher risk of developing health problems during pregnancy. Just under one third (31%) of the mothers reported one or more medical problems that could interfere with successful pregnancy outcome. As indicated by both, adjusted and unadjusted, odds ratios the risk of preterm delivery and low birth weight increased with number of medical problems experienced by mother during pregnancy. Women who developed two or more medical risks had about three and a half times the risk of preterm delivery and two and a half times the risk of low birth weight compared to those without such problems (see Table 2).

Prior fetal loss significantly increases the risk of both preterm delivery and low birth weight. About 22 percent of the women in the survey reported that they had suffered one or more fetal loss. According to adjusted odds

ratios in Table 2, there was a twofold increase in the risk of preterm delivery and low birth among women who had suffered prior reproductive loss.

Unintended pregnancy has been linked with the risk of low birth weight,[7,8] and it has been suggested that a woman whose pregnancy is intended is likely to engage in healthy behaviors, such as seeking timely prenatal care, avoiding or decreaseing smoking and alcohol consumption, gaining appropriate weight during pregnancy, etc. There are few studies that treat "unintended" pregnancy as an explanatory variable. Over half (54%) of the women in the survey reported that they either did not want the pregnancy or wanted it later. As shown in Table 2, unintended pregnancy was a significant predictor, resulting in one and one-half to twofold increases in the risk of low birth weight and preterm delivery, respectively. While there are lingering doubts about conceptual clarity and measurement of the wantedness of prengnacy,[9,10] it nevertheless focuses attention on the role of fpterventions that reduce unintended pregnancy.

Maternal health behaviors known to be associated with low birth weight include; gestational weight gain,,[12] maternal smoking[13] and maternal alcohol consumption. Gestational weight gain can be grouped into three categories: less than 25 lbs., 25 to 35 lbs., or more. Inadequate weight gain among women during pregnancy (less than 25 lbs.) increased the risk of preterm delivery by about 50 percent compared to those whose weight gain was between 25 and 35 lbs. (see Table 2).

Smoking was a significant risk factor for low birth weight but not preterm delivery among these women. Women who smoked during pregnancy were 1.6 times as likely to deliver a low birth weight baby as those who did not smoke.

Adequacy of prenatal care was measured using the Kessner index[14,15] which considers the number and timing of visits in relation to gestational age. About 74 percent of mothers in the survey had adequate prenatal care; 24 percent had intermediate level of prenatal care and only 3 percent reported inadequate prenatal care. Intermediate or inadequate level of prenatal care is significantly related to low birth weight but not to preterm births (see Table 1). When adjusted for other variables, adequacy of prenatal care is not a significant predictor of adverse pregnancy outcomes (see Table 2).

There is a lack of consensus on what the determinants of preterm delivery and low borth weight are, except perhaps for the immediately precipating events during pregnancy.[16] Most of significant risk factors in this study have also been noted in previous studies.[17] The high risk behaviors noted in this population may not be within the purview of individual choice or edcision making.[18] The risk-taking behavior of a substantial proportion of underclass, inner city residents is shaped by chronic intergenerational poverty, environmental deprivation, learned cultural beliefs and attitudes, social alienation and disadvantaged neighborhood effect.[19]

Maternal medical risks and health behavior are significant predictors of preterm delivery and low birth weight. Medical problems that develop during pregnancy should have been diagnosed during prenatal care and managed through appropriate follow-up. Prior reproductive loss is also a risk factor and therefore is included in initial comprehensive risk assessment at the first prenatal visit The fact that these high risk pregnancies resulted in a preterm delivery or a low birth weight infant raises considerable doubts as to the adequacy and efficacy of prenatal care as currently delivered to the inner city women in this survey.

1. MOsely WH and Chen LC. An analytical framework for the study of child survival in developing countries. Child Survival: Strategies for Research, Supplement to Population and Development Review. 1984; 10:25-45.
2. Kline, J., Stein, Z. and Susser, M. 1989. Conception to Birth: Epidemiology of Prenatal Development. New York: Oxford, pp. 191-207.
3. Ibid., p. 170.
4. Kramer M.S. Determinants of low birth weight: methodological assessment and meta-analysis. Bulletin of the World Health Organization. 1987; 65:663-737.
5. Kiely JL and Susser M. Preterm birth, Intrauterine growth retardation and perinatal mortality. AM J Public Health. 1992;83:341-345.
6. Allegheny County Health Department. Healthy Start Comprehensive Plan. Pittsburgh, PA: Allegheny County Health Department, 1992.
7. Norussis MJ. SPSS Advanced Statistics User's Guide. Chicago: SPSS Inc., 1992.
8. Geronimus A., Bound J. Black/White differences in women's reproductive-related health status: evidence from vital health statistics. Demography. 1990;27: 457-466.
9. Institute of Medicine (1985). Preventing Low Birthweight. Washington, D.C.: National Academy Press.
10. U.S. Dept. of Health and Human Services (1986). Report of the Secretary's Task Force on Black and Minority Health (Vol. 6): Infant Mortality and Low Birthweight. Washington, D.C.: U.S. Government Printing Office.
11. Klerman LV, Jekel JF. Unintended pregnancy. In: Bracken MB, ed. Perinatal Epidemiology. New York: Oxford; 1984:283-300.
12. Marsiglio w. and Mott FL. Does wanting to become pregnant with a first child affect subseauent maternal behaviors and infant birth weight. Journal of Marriage and Family. 1988; 50:1023-1036.
13. World Health Organization. Unintended pregnancy and infant mortality/morbidity. In: Amler Rw and Dull HB, ed. Closing the Gap: The Burden of Unnecessary Illness. New York: Oxford, 1987:130-142.

14. Institute of Medicine. Subcommittee on Nutritional Status and Weight Gain During Pregnancy. National Academy of Sciences. Washington, DC: National Academy of Sciences, 1990.
15. Kleinman JC, Maddans JH, The effects of maternal smoking, physical stature and educational attainment on the incidence of low birth weight. Am. J Epidemiol. 121:843-55. 1985.
16. Gortmaker SL. The effects of prenatal care upon the health of newborn. AJPH. 1987;69: 653-660.
17. Infant Death and Analysis by Maternal Risk and Health Care. Washington, D.C.: Institute of Medicine, 1973.
18. Eberstadt N. America's infant mortality puzzle. Public Interest, 1991:30-47.
19. Kliegman, RM. Perpetual poverty: child health and the underclass. Pediatrics. 1991; 89:710-713.

Table 1. Unadjusted Odds Ratios for Low Birth Weight and Preterm Births

	Low birth weight	Preterm births
Variables	Unadjusted Odds Ratio	Unadjusted Odds Ratios
Maternal age		
14-19	1.57	1.35
35-44	1.13	1.27
20-34	1.00	1.00
Marital status		
Single or previously married	2.04	1.16
Married	1.00	1.00
Years of school Completed		
<12	1.63	1.01
12	1.13	0.98
>12	1.00	1.00
Ethnic Background		
Black	1.98	1.28
White & Others	1.00	1.00
Household Income		
<$15,000	1.84	1.38
$15,000-$24,999	0.65	1.09
Not ascertained	1.56	1.23
$25,000	1.00	1.00
No. of medical problems during pregnanoy		
One	1.45	1.65
Two or more	2.61	3.89
None	1.00	1.00
Prior fetal loss		
One	1.99	1.61
Two or more	2.99	2.80
None	1.00	1.00
Unintended pregnancy		
Yes	2.02	1.52
No	1.00	1.00
Weight gain during pregnancy		
<25 lbs	2.39	1.73
>35 lbs.	0.95	0.88
25-35 lbs.	1.00	1.00
Smoke during pregnanoy		
Yes	1.78	1.27
No	1.00	1.00
Alcohol consumption during pregnancy		
Yes	1.14	1.01
No		1.00
Adequacy of prenatal care		
Inadequate or intermediate	2.24	1.49
Adequate	1.00	1.00

Table 2 Adjusted Odds Ratio for Low Birth Weight and Preterm Delivery based on Multiple logistic Regression Model

	Low Birth Weight	Preterm births
Variables	Adjusted Odds Ratio	Adjusted Odds Ratio
Maternal age		
14 to 19	1.22	1.29
35 to 44	1.00	1.10
20 to 34	1.00	1.00
Marital Status		
Single or previously married	1.59	0.91
Married	1.00	1.00
Years of School Completed		
<12 years	1.01	0.80
12	1.02	0.90
>12	1.00	1.00
Ethnic background		
Black	1.54	1.20
White and Others	1.00	1.00
Household income		
Under $15,000	0.72	1.14
$15,000 - $24,999	0.46	0.93
Not ascertained	0.46	0.85
$25,000 +	1.00	1.00
No. of Medical Problems during pregnancy		
One	1.59	1.77
Two or more	2.32	3.61
None	1.00	1.00
Prior fetal loss		
One	2.09	1.50
Two or more	2.07	2.24
None	1.00	1.00
Unintended pregnancy		
Yes	1.72	1.52
No	1.00	1.00
Weight gain during pregnancy		
<25 lbs.	1.99	1.51
>35 lbs.	0.88	0.79
25 to 35 lbs.	1.00	1.00
Smoking during pregnancy		
Yes	1.65	1.15
No	1.00	1.00
Alcohol consumption during pregnancy		
Yes	0.99	0.92
No	1.00	1.00
Adequacy of prenatal care		
Inadequate or intermediate	1.42	1.17
Adequate	1.00	1.00

INFANT MORTALITY

Discussant:

Michelle A. Williams
University of Washington

(Not available for publication)

Session C

Children and Youth

Toward the year

2000

Refining the measures

THE "BRIGHT SMILES, BRIGHT FUTURES" CURRICULUM EVALUATION: DESIGNING HEALTH EDUCATION EVALUATIONS WHICH ADDRESS THE METHODOLOGICAL CHALLENGES OF COLLECTING DATA FROM YOUNG CHILDREN

Meg Leavy, Ph.D., Macro International, Inc.
Robert S. Gold, Ph.D. and Maria Hernandez, M.S.

Introduction

The purpose of research is to advance knowledge and/or to guide decisions (Green, and Gordon 1982). Basch (1982) identifies assessing the comparative worth of programs, providing objective information which decision makers can use to make judgements for the improvement of policies and program approaches, and enhancing program efficiency and effectiveness as among the most important reasons for conducting evaluation research. In an applied discipline such as health education, evaluation research is needed to accomplish all of these important tasks. However, evaluating young students for whom reading and writing may not yet be possible poses challenges to this evaluation. Ball (1971) noted four problems in assessing attitudes of young children: (1) there is little stability on children's attitudes, (2) young children lack test-taking skills, (3) young children are eager to please adults and will try to answer according to what they think the adult wants, and (4) the child may answer all questions with the same response. To address these challenges Ball (1971) suggested the use of a control group to help reduce the error associated with the instability of children's attitudes, the administration of instruments to children individually, and a well-constructed test administered by a skilled interviewer. Finally Ball (1971) suggested the using pictures with measurement instrument items.

The evaluation of the *Bright Smiles, Bright Futures* program, an oral health curriculum developed by the Oral Health Improvement Program for Elementary Schools (OHIP), was faced with all of these challenges. The project was funded by the Colgate-Palmolive Corporation as part of their commitment to improving the oral health of U.S. children. The overall goal of the program is to provide a planned and sequential school-based oral health curriculum which is sensitive to the needs, interests, and cultural values of African American and Hispanic children. The specific objectives of the *Bright Smiles, Bright Futures* curriculum are as follows:

- to improve the health and oral health knowledge of African American and Hispanic and other high risk students in the areas of personal health (the shape, function and care of different body systems), nutrition (the relationship between eating patterns, foods, and oral health), and safety (high risk behaviors and other environmental risks to oral health);
- to affect African American and Hispanic students' self-esteem and attitudes toward oral health preventive care and the value of such preventive actions by encouraging them to accept personal responsibility for such activities; to maintain a healthy body; and to appreciate the role of others in helping them maintain their health;
- to positively influence African American and Hispanic students' health behaviors by teaching good decision-making and risk-evaluation skills, providing opportunities to promote and maintain personal health, and encouraging adaptation to varied social situations.

Purpose

The purpose of this study was to examine the impact of the *Bright Smiles, Bright Futures* curriculum on three first grade classrooms in the District of Columbia. Specifically, the study was designed to determine if exposure to the curriculum under controlled conditions was associated with increases in oral health knowledge, better attitudes toward oral health, better oral health skills, and more positive oral health behaviors. In order to accomplish this primary purpose, it was necessary to: (1) develop measurement instruments which were capable of assessing changes in students' knowledge of oral health issues, attitudes toward oral health, attitudes toward themselves, and toothbrushing skills; (2) to assess the validity and reliability of the methods used to determine program impact; and (3) to assess the level of parental involvement in the program. An independent implementation evaluation of the curriculum was conducted in Oakland, CA, and Philadelphia, PA. The goals of the implementation evaluation were to assess how the curriculum was used and/or modified when implemented under normal classroom conditions as opposed to a controlled environment.

Hypotheses

The hvootheses to be tested related to changes in students' oral health knowledge and behaviors, their attitudes toward oral health, and their toothbrushing skill. The specific hypotheses to be tested were:

When initial differences are controlled for, students exposed to the oral health curriculum will have an increase in knowledge about oral health and will have more positive attitudes and behaviors toward oral health than their counterparts who did not receive the curriculum.

When initial differences are controlled for, students who are exposed to the oral health curriculum will be more skilled at brushing their teeth than their counterparts who were not exposed to the curriculum.

Methods

A non-equivalent comparison group with pretest and posttest measures was used. Criterion referenced measurement instruments were developed to assess changes in students' knowledge, and attitudes. A skill check analysis was used to assess if the students who received the curriculum were more skilled at brushing their teeth than those students who did not receive the curriculum. In addition, a survey to assess parental involvement was developed and administered.

Student Outcome Instrument

In order to test the hypotheses it was necessary to develop a student outcome instrument to measure changes in knowledge, attitudes and self reported behaviors, and a skill check analysis to assess students' tooth-brushing skills. Criterion referenced knowledge items were developed to assess if the students had attained the specific knowledge objectives identified in the curriculum scope and sequence. Similarly, criterion referenced attitude items were developed to assess if the students had attained the specific attitude objectives identified in the curriculum scope and sequence. The student instrument was designed with graphic elements to enhance each respondent's active participation and understanding.

Picture usage has been evaluated as effective in the measurement of health knowledge of first grade children (Solleder, 1979) and preschool children (Jubb, 1982). Researchers in the Community and Family Medicine Department at the Georgetown University School of Medicine also support the use of pictures. Interviewing first grade students using pictures as well as words seems to generate the most reliable data (Bush and Iannotti, personal communication). Consequently, pictures were developed to accompany each item. The pictures were developed by an artist and resembled images presented in the curriculum's ancillary materials. Bush and Iannotti also support the use of comparison groups, and the individual administration of measurement instruments.

To determine if the students who received the curriculum were better at brushing their teeth than students who did not receive the curriculum, a tooth brushing skill check protocol was developed.

The final student outcome instrument had thirty-five items, including the nutrition and safety items, and a skill check analysis which was administered at posttest only. The items were grouped by content (general health, importance of healthy oral structures, tooth loss, preventive measures, dental visits, nutrition, safety, and sealants). Items were grouped in this fashion to maintain continuity. Items whose correct response was "yes" were interspersed with items whose correct response was "no". Each item (both text and picture) was placed on a separate page. All thirty-five pages were placed in a three ring lose leaf binder.

Quality control measures were employed to keep the measurement error at a minimum and increase inter-rater reliability. With criterion referenced measurements, systematic error affects the correspondence between observed scores and the criterion. In other words, data with systematic error will fail to reflect the criterion of interest. Specifically, data collectors received extensive training. In addition each data collector field tested the instrument with approximately fifteen first grade students enrolled in two District of Columbia public schools which were not participating in the evaluation.

Measurement Validity and Reliability

The following efforts, which were executed during instrument development and field testing, were made to maximize the supporting evidence for validity. These efforts focused on reducing measurement error and on providing strategies for the appropriate use of the resultant scores.

To increase the utility of the scores, criterion referenced items were used. Criterion referenced, as opposed to norm referenced tests, provide absolute interpretations (Glaser, 1963). Popham (1988) suggests that criterion referenced tests are more useful than norm-referenced tests in determining the effectiveness of a curriculum. The measurement instruments which were developed determined whether the student had achieved a specific objective or

not, and were useful in deciding what aspects of the curriculum enhanced student learning.

The following steps were taken to provide supporting evidence for the validity of the instruments. The student outcome instrument was reviewed by an expert panel to determine if the instrument adequately addressed the content of the specific objectives. Educational, dental, and developmental specialists reviewed the instrument to determine not only if it covered the content, but if it was developmentally appropriate for the students as well. The instrument was field tested in the Washington, D.C. area. During field testing data concerning the clarity of the instrument items, the difficulty of the instrument items, and the ordering of the instrument items were collected. Warm-up questions were administered prior to the criterion- referenced items. These warm-up items assisted the interviewer in determining if the student understood the test taking procedures. The items also helped the children feel comfortable with the test-taking situation. Developmental specialists were consulted to help develop appropriate the warm-up items.

For those objectives which require more than one item, a measure of internal consistency will be calculated. An alpha coefficient was calculated to determine the reliability of all scales.

Student Outcomes

Four statistical strategies were used to analyze the student outcome data; repeated measures analyses of variance, a one-way analysis of variance, a discriminant analysis, and follow-up Chi-square analyses. Four repeated measures analyses of variance were conducted to determine changes in knowledge. A one-way analysis variance was conducted to determine changes in skill. A discriminant analysis was conducted to determine which variables, if any, discriminated between students from the experimental and comparison schools. Finally, chi-square analyses were conducted in a post hoc examination of the findings of the discriminant analysis. Descriptive statistics were calculated, where appropriate, for both distributions to provide evidence for normality and variance homogeneity.

Although students were not randomly assigned to treatment and control groups, the data suggest that there were no serious pre-existing differences between groups which would threaten the internal validity of the findings. Only eight cases were lost to student attrition and/or the inability to match pre and posttest scores.

Knowledge

There were no significant differences between pretest scores for the experimental and the comparison schools, indicating that the students from all four classrooms had similar oral health knowledge, attitudes, and behaviors. However, at posttest, students from the experimental school had a significant increase in knowledge about the importance of healthy oral structures, and had significantly more knowledge about the importance of healthy oral structures than their counterparts who did not receive the curriculum. Similarly, students from the experimental school had a significant increase in knowledge about preventive oral health measures, and, at posttest, had significantly more knowledge about the preventive measures than their counterparts who did not receive the curriculum.

There were no significant differences between the experimental and the comparison schools for knowledge of dental sealants. All of the students in the study knew more about dental sealants at posttest, however the sample mean for the five point scale was low, indicating that neither the experimental nor the comparison group had a substantial amount of knowledge about dental sealants. Similarly, all of the students knew more about tooth loss and safety at posttest time. However, the change in knowledge about tooth loss, although statistically significant, was small.

Students from the experimental school had a significant increase in knowledge about foods and drinks which were good for their teeth, and had significantly more knowledge about foods and drinks which were good for their teeth than their counterparts who were not exposed to the curriculum.

Attitudes and Behaviors

Students from the experimental school reported at higher frequencies (1) visiting the dentist, and (2) brushing their teeth in the morning and at night. These two behaviors discriminated between students who received the curriculum and those who did not. No other attitude or behavior items discriminated between the groups.

Skill

Students from the experimental school were better skilled at brushing their teeth than those students who did not receive the curriculum.

Conclusions and Recommendations

Results from this study indicate that students who received the *Bright Smiles, Bricrht Futures* curriculum successfully met seven of the twelve curriculum objectives, and obtained

requisite knowledge for the attainment of the nutrition and safety behavioral objectives. However, more importantly, results form the evaluation of the Bright Smiles, Bright Futures curriculum suggest that valid and reliable data can be collected from young children. Specific recommendations are as follows:

1. Have data collection instruments reviewed for both content and for developmental appropriateness.
2. Use pictures to enhance individual student's participation and understanding.
3. Intersperse correct yes and no, or a, b, c responses.
4. Use warm-up items to verify that students understand the procedures and to help the students become comfortable with the process.
5. Administer instruments individually to help students focus.

All in all, health education is faced with numerous challenges, one of which is collecting data from young children. The objectives of health education curricula and adequate measurement techniques need to be well conceived before program implementation. Appropriately addressing research challenges will enhance the quality of data produced by health education evaluations, which hopefully will lead to demonstrated program effectiveness.

References

Ball, S. (1971). Assessing the attitudes of young children towards school. Journal of experimental Child Psycholosy, 31, 3-5.

Basch, C.E. Evaluation of a School Health Education Curriculum: Implementation Analysis and Outcome Assessment. Published Doctoral Dissertation, Southern Illinois University at Carbondale, Carbondale, IL, 1982.

Glaser, R. (1963). Instructional technology and the measurement of learning outcomes. American Psycholosist, 18, 519-522.

Jubb, W.H. The Development of an Instrument to Assess Health Knowledge of Children Prior to First Grade. Unpublished Doctoral Dissertation, University of Alabama, Tuscaloosa, AL, 1982.

Popham, W.J. (1988). Educational Evaluation. Edgwood Cliffs, NJ: Prentice Hall.

Solleder, M.K. (1979). Evaluation Instruments in Health Education. Fourth Education. Reston, VA: Association for the Advancement of Health Education, American Alliance for Health, Physical Education, Recreation and Dance.

EFFECTS OF MATERNAL EDUCATION AND ACCESS TO MEDICAL CARE ON ANALYSIS OF INJURY RISK ACCORDING TO PATTERNS OF CHILD CARE

Mary D. Overpeck, National Institute of Child Health and Human Development
Jonathan B. Kotch, Ann C. Trumble

An analysis of injury risk to all U.S. children ages 0-5 demonstrated that the relationship of socioeconomic patterns of child care use and access to medical care were similar.' Such similarity may lead to bias in assessment of medically-attended injury risks associated with non-guardian child care of all types. Without stratifying by maternal education and adjusting for having either health insurance or Medicaid, having a place to go for sick or injured care, and other demographic variables, the findings on injury risks associated with child care use would be reversed. The following analysis demonstrates the effects of the interaction of patterns of use of child care and access to medical care coverage on both the reporting and analysis of medically-attended injury risks associated with non-guardian care in centers, family day care homes, and in the child's own home.

Otherpopulation-based survey studies have shown lower incidence of reported injuries and chronic conditions associated with low socioeconomic status or no health insurance and between racial categories [2,3,4] The need to consider the effects of access to medical care coverage was found in a previous analysis of these data which showed that children <18 years old without either health insurance or Medicaid were significantly less likely to report medically-attended injuries than those with any medical care covera e even for injuries defined as serious.[4] The analysis showed those with Medicaid were as likely to report medically-attended injuries as those with health insurance, for both total and serious injuries, after adjustment for having a place for sick or injured care, maternal education, age, race and sex. Therefore, having either health insurance or Medicaid is used as the definition of 'having medical care coverage' in the following results.

METHODS. Data on medically-attended injuries and also on regular sources and hours of child care for children under six years old were collected as part of the nationally representative sample of all U.S. children from the 1988 Child Health Supplement (CHS) to the National Health Interview Survey (NHIS). The NHIS is a multistage probability sample household interview survey of the U.S. civilian noninstitutionalized population.[6] For the CHS, one sample child <18 years old was selected from each family with children according to a predetermined probability of birth order yielding an overall response rate for the NHIS-CHS of 91 percent. The sample included 6,209 children under age six representing 22 million children in the U.S. after final weighting with the poststratification variables of age, race, and sex. Final weights adjusted for nonresponse are applied to all data in the results section.

Analysis is based on linear regression modelling of the predicted risk of injury associated with an increasing number of hours in each type of child care after adjusting for confounders (predicted injury risk = hours in each type of care + confounders). Both injury risk and the number of hours of care are continuous variables. A backward stepwise approach using rescaled weights to assess significant covariates in the model is being used as recommended for large complex surveys.[7] Confidence intervals are based on standard errors produced by SUDAAN which adjusts variance estimates to account for the complex multistage sample design of the survey.'

Injury questions included details on the nature, cause, place of occurrence and severity of all medically-attended injuries occurring in the previous 12 months (including doctor visits, telephone calls, or emergency room visits). Predicted probability of injury risk is a continuous outcome variable based on number of injuries occurring during the previous 12 months divided by number of children in the stratum. Injuries resulting from complications of medical treatment (E-codes 929-949, 870-879) are excluded.

Injuries are analyzed according to severity levels. Medically-attended injuries serious enough to have some effect on the child show more stability with increasing time from the injury event to the time of reporting due to the effects of recall and differences in care-seeking patterns.' Serious injuries are defined by the following effects on the child: (1) caused him/ her to stay in bed more than half of a day; (2) had any limitations or was prevented from doing usual childhood activities; (3) had one or more nights in the hospital; (4) had surgery, including bone setting or stitches; (5) had pain often or all of the time; or (6) bothered him/her.

Child care questions were asked about any regular types of non-guardian child care arrangement during the previous four weeks, any changes in care during the past year, 'main' type of child care used in the previous year prior to the last four weeks, place of each type of care, and hours per week in each place of care.

Number of hours per week during the past four weeks in any type of regular child care were asked for up to three places of care in addition to attendance in any nursery school, preschool, or school setting (including extended day care before and after school). Number of times the place of care had changed in the past 12 months was also determined.

Types of regular non-guardian child care are classified as follows:

- center-based care - day care centers, nursery schools, preschools, or school settings with or without extended day care, and day camps;
- family day care home - baby sitter's home, including the home of relatives; or
- own home - babysitter or relative in child's own home.

If a child had no care in the past four weeks but had received care in the past 12 months, hours per week in the main type of care received in the past 12 months are used. After deleting records where either number of hours or type of care was unknown for either the past four weeks or during the past 12 months, the analytic sample size is 6,071 (99.4% of the total sample). Data were edited to account for changes in the care pattern across the previous four weeks and past 12 months.

Significant covariates retained in all models include having health insurance or Medicaid, having a place for sick or injured care, years of maternal education, multiple places of child care, age, sex, and race. Maternal education is categorized as <12 years, 12 years, and >12 years. Additional, but generally nonsignificant, covariates assessed are changes in the child care place during the year, number of adults and children in the household, hours worked by the mother, income level, urban/rural place of residence, and number of times child had moved.

RESULTS. Bivariate unadjusted medically-attended injury risks are lowest for both total and serious injuries for children without medical care coverage (either health insurance or Medicaid) or whose mothers have lower educational levels (Figure 1). Children without coverage are 30% less like to have a medically-attended injury reported than those with coverage (RR=.70, CI=.54 to .91). Injuries defined as serious also are about 50% less likely to be reported for children without coverage compared to those with (RR=.47, CI=.31 to .71). A similar pattern is shown when comparing medically-attended injury risk

Figure 1. Unadjusted Injury Risk by Medical Care Coverage and Maternal Education

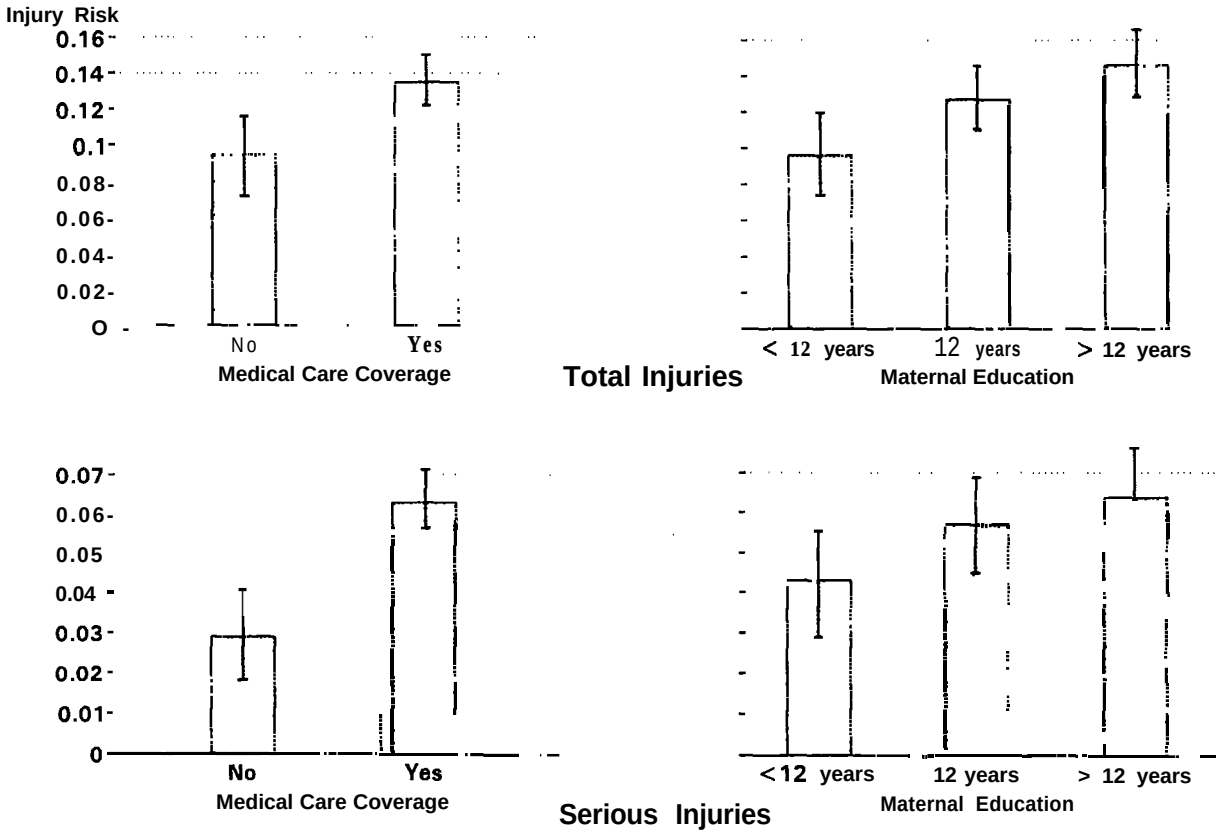
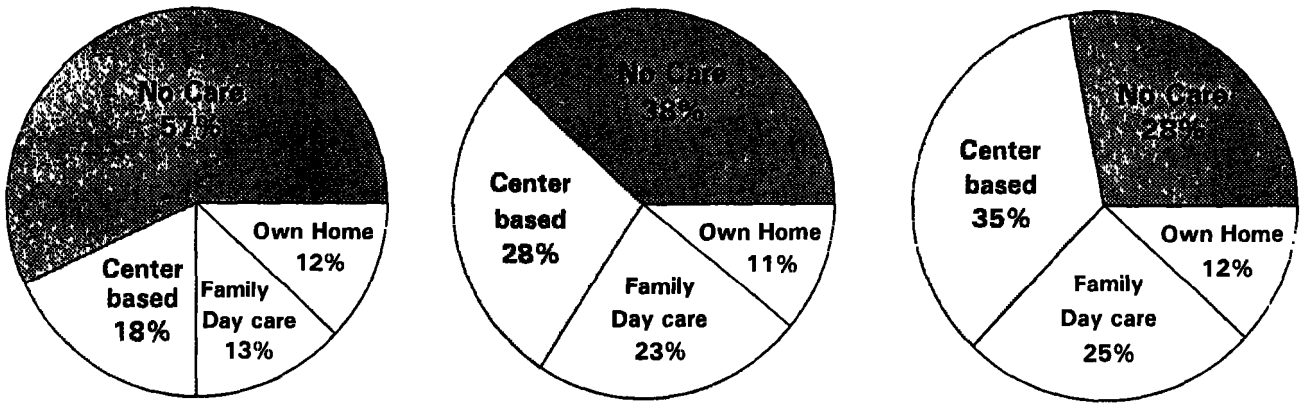


Figure 2: Use of Non-guardian Child Care by Maternal Education Levels



< 12 Yrs. Education **12 Yrs. Education** **> 12 Yrs. Education**

by maternal education levels. Total injury risk for children of mothers with <12 years of school is about 30% less than for those at the highest education level (RR=.66, CI=.50 to .87). The latter effect is significant even when controlling for medical care coverage within maternal education levels.

Figure 2 shows that use of non-guardian child care differs by maternal education levels. Fifty-seven percent of children with mothers who had <12 years of education are not in any child care compared to 38% with 12 years and 28% at the highest education level. Mothers with the highest educations are most likely to use center-based care (at 35%) compared to only 18% at cl2 years of school.

Table 1. Percent all U.S. children in each type of care according to maternal education by hours in care for each age, 1988

	Percent of children in each type of care*							
			Center-based		Family home		Own home	
Maternal Education for each age	Total all types	No Care	≤35 hours	>35 hours	≤35 hours	>35 hours	≤35 hours	>35 hours
All ages	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
< 12 years**	20.0	28.4	13.2	8.6	10.2	12.7	15.0	26.0
12 years	41.1	42.1	38.4	42.1	42.0	43.8	38.8	41.0
> 12 years	38.9	29.6	48.3	49.4	47.8	43.5	46.2	33.0
< l-l year	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
< 12 years**	21.2	27.0	13.5	10.2	11.3	10.2	13.5	18.3
12 years	38.3	39.0	28.5	37.2	40.1	40.3	30.7	39.1
> 12 years	40.5	34.0	57.9	52.5	48.6	49.6	55.7	42.5
2-3 years	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
< 12 years**	20.0	28.6	7.4	8.9	9.1	13.9	14.8	23.3
12 years	41.1	44.4	34.3	41.7	42.5	45.4	42.6	43.3
> 12 years	38.9	27.0	58.3	49.4	48.4	40.7	42.5	33.4
4-5 years	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
< 12 years**	21.2	31.8	15.3	7.7	10.3	15.1	16.4	41.2
12 years	38.3	45.8	41.0	44.0	43.2	47.1	42.1	40.8
> 12 years	40.5	22.4	43.7	48.3	46.5	37.8	41.5	18.0

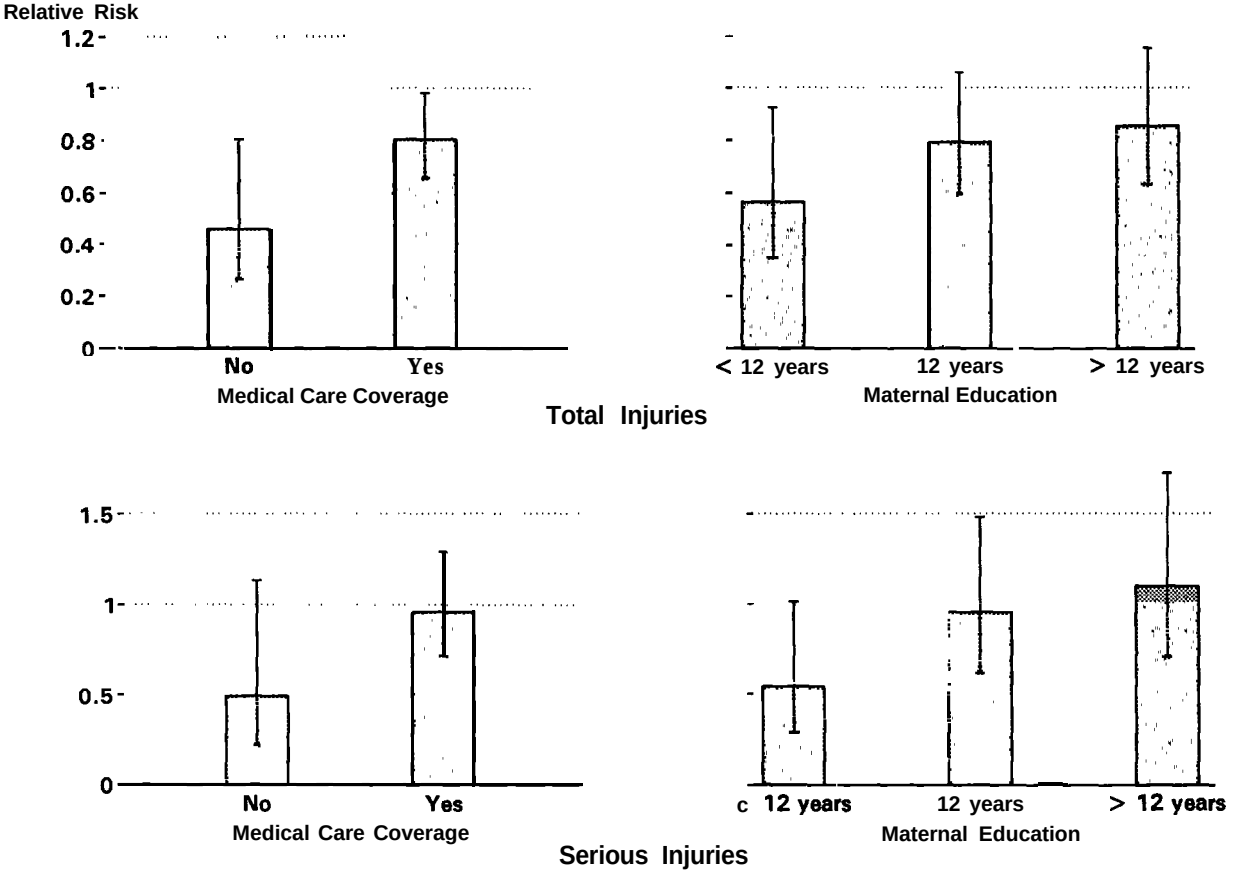
*The number of hours is based on hours in each type used during the previous twelve months. A child may be in more than one type of care.
**Includes 2 percent of all mothers for whom education was unknown because the mother did not live in the household or respondent did not know the number of years completed.

The proportion of children in each type of child care, or no care, also differs by maternal education levels: 20% of U.S. mothers have <12 years of school but their children represent 28% of all children with no child care (Table 1, columns 1 & 2). Almost 40% of all mothers have 212 years of school while their children comprise only 30% of children without child care.

Table 1 also shows that as the child ages, the types of care used and the hours of exposure to each type of care change. The difference in distribution of children by maternal education level in each type of care increases with age of the child. Children of mothers with the lowest education represent 27% of ages ≤1 in no care. While the percent increases only slightly to 31.8% of those not in care at ages 4-5, the proportion of children of mothers with the highest education decreases from 34% of those with no care at ages ≤1 to only 22.4% at ages 4-5. Similar reversing of proportions occurs for children cared for in their own homes with differences by number of hours of care. Mothers with the highest education level have a disproportionately high percent of all children in center-based care at any age. However, they represent almost 60% of children in center-based care for ≤35 hours a week until ages 4-5.

The crude relative risk of injury comparing no child care to any care by medical care coverage and maternal education status demonstrates the potential effect from the association between no medical care coverage and low maternal education (Figure 3). Those with no medical care coverage consistently report lower injury risks for children with no child care compared to children with care. This is shown by a 95% confidence interval of the relative risk that is below the value 1. The pattern is the same for the lowest maternal education level, injury risk, and use of child care. At higher maternal education levels, the relative risk for those with no child care is not significantly different from those using

Figure 3. Unadjusted Relative Risk of Injury comparing No Child Care to Any Care by Medical Care Coverage and Maternal Education



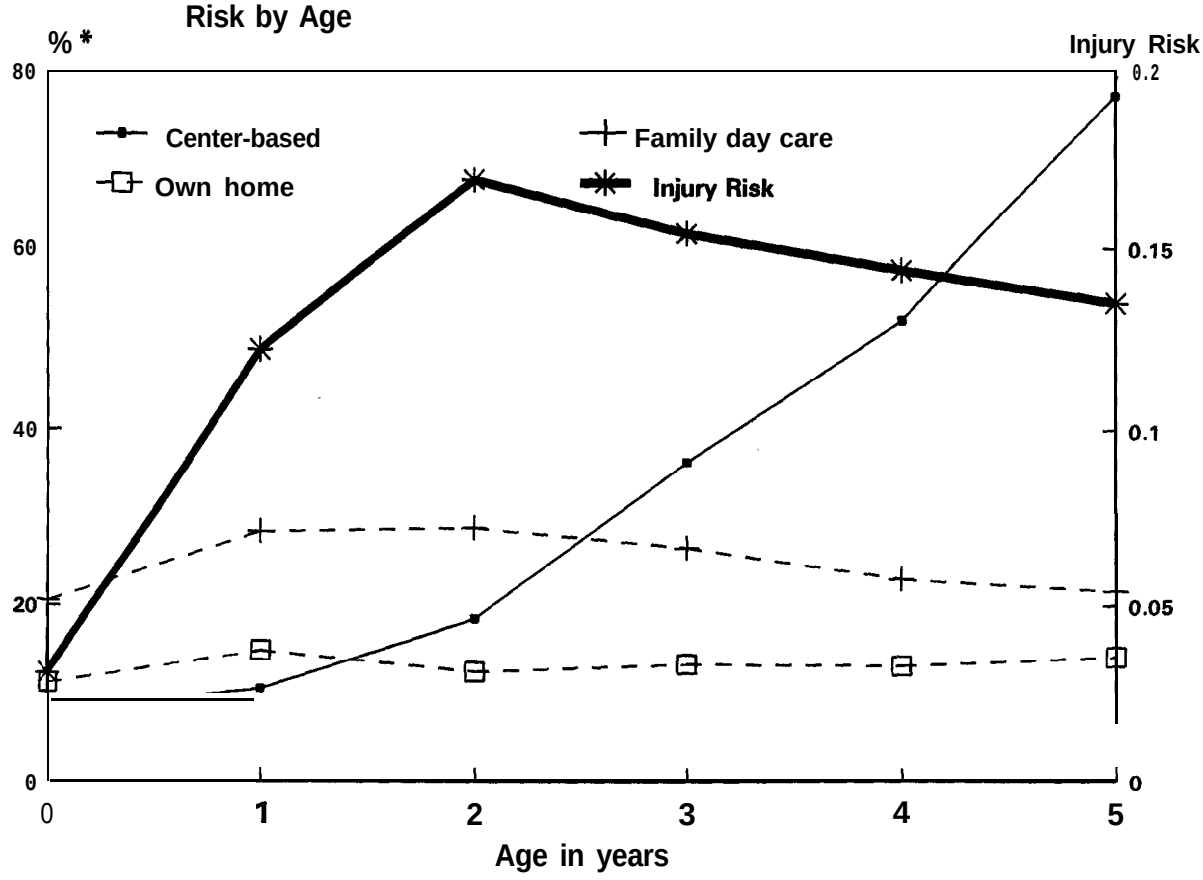
care. The same interactions occur for serious injuries.

However, the simple comparison of the relative risk of injury for those in no child care versus those in any care is complicated by changes in the pattern of use of child care and in occurrence of injuries according to age of the child (Fig. 4). Use of child care in the child's own home remains fairly constant with age. Use of family day care homes decreases slightly after age two. The predominate child care shift with age is from no child care to center-based care. At the same time, risk of injury increases with age, consistent with developmental stages of children and their wider injury exposures. The result is an apparent correlation between increasing use of child care and injury risk, particularly for center-based care.

This interaction between patterns of child care use by maternal education and injury risks with age requires a stratified data analysis by both maternal education level and age. The predicted risk of injury associated with each type of child care is modelled within the three levels of education for three two-year age groups and for both total and serious injuries. The resulting models are then adjusted for medical care coverage and other covariates. The large sample size is robust enough to allow tests of significance in all models as long as both injury risk and hours of child care exposure were kept as continuous variables. The final results generally show no increased injury risk with increasing hours of child care at any age in most care types. However, one of the findings which would have been reversed without accounting for maternal education and medical care coverage is a decreased risk of serious injury with increasing hours in center-based care at ages 4-5 ($p<0.05$).

DISCUSSION. This analysis assesses the effects of the interaction of patterns of use of child care and medical care coverage on medically-attended injury risks associated with non-guardian child care. Both total and serious medically-attended injury risks are lower for children without medical care coverage or those whose mothers have lower education levels. The least educated are less likely to use any child care, and particularly center-based care, while those with the highest education levels are most likely to use center-

Figure 4. Percent of Children in each Type of Non-guardian Care and Injury Risk by Age



* Percents add to more than 100 because a child may be in more than one type of care

based care. The distribution of children by hours in each type of care also differs with maternal education level and age as more children move from no child care to center-based care. The apparent effect resulting from interaction between patterns of use of child care and medical care coverage is lower injury risk assessments for children with no child care compared to children using care, and particularly center-based care. Those using center-based care may renort higher medically-attended injury risks because they are more likely to have medical care coverage.

Therefore, assessment of risks of injury associated with child care may be biased without measuring and accounting for socioeconomic factors. This study demonstrates the potential bias toward findings of lower injury risks in low socioeconomic groups that can result when using records frommedical care sources or from self-reports of medically-attended injury. However, the opposite bias could result if parents are less likely to seek medical care for injuries occurring at centers than at home as suggested by Kotch et al.[10]

Without the additional **information** on socioeconomic factors of families reporting injuries that was available for adjusting results of this analysis, children with no child care would have appeared to have less injury risk. Assessment of relative risks of injury according to patterns of use of child care is biased without accounting for interactions among use of medical care, maternal education, and types and hours of child care.

Recent studies have addressed safety and health in non-guardian child care.[11,12,13] Measures of the quality of a child care setting such as licensure, the staff/child ratio and care-giver education were incompletely reported in these data.[14] They should be included as modifiers in other studies of injury risk by type of care. At a minimum, studies using medical **treatment** sources for **injury** analysis should control for maternal education or income and distribution of types of child care used in their patient population.

REFERENCES

1. Overpeck MD, Kotch JB, Trumble AC. Injuries in preschool children according to patterns of child care. Abstracts from the International Conference on Child Day Care Health: Science, Prevention and Practice. Atlanta, Ga: US Center for Disease Control, June 15-17, 1992.

2. Ries P. Health of black and white Americans, 1985-87. Hyattsville, Md; National Center for Health Statistics, Vital Health Stat 10(171). 1990.

3. Lieu TA, Newacheck PW, McManus MA. Race, ethnicity, and access to ambulatory care among US adolescents. Am J Public Health. 1993;83:960-965.

4. Wagener DK, Winn DW. Injuries in working populations: Black-white differences. Am J Public Health. 1991;81:1408-1414.

5. Hare1 Y, Scheidt PC, Over-peck MD, Trumble A, Bijur PE, Jones D. Childhood and adolescent injury morbidity in the United States: implications for research and prevention. Abstracts from the Second World Conference on Injury Control. Atlanta, Ga: US Center for Disease Control, May 20-23, 1993.

6. National Center for Health Statistics. Current estimates from the National Health Interview Survey: United States, 1988. Vital Health Statistics Series 10(173), 1989.

7. Lee ES, Forthofer RN, Lorimor RJ. Analysis of complex sample survey data: problems and strategies. Social Meth and Research 1986; 15:69-101.

8. Research Triangle Institute. Software for the Survey Data Analysis (SUDAAN) Version 5.30. Research Triangle Park, NC: Research Triangle Institute; 1991.

9. Hare1 Y, Overpeck MD, Jones DH, Scheidt PC, Bijur PE, Trumble AC, Anderson, J. Effects of recall on estimating annual non-fatal injury rates for children and youth. Am. J. Public Health (in press).

10. Kotch JB, Harms T, Clifford RM, McMurray MP, Loda F. Parents' reports of out-of-home injuries among children attending child care centers. Early Child Development and Care. (In press).

11. Kotch JB, Bryant D. Effects of day care on the health and development of children. Current Opinion in Pediatrics 1990; 2:883-894.

12. Panel on Child Care Policy. Who Cares for America's Children? Washington, D.C: National Academy Press, 1990; 113-114.

13. Runyan CW, Gray DE, Kotch JB, Kreuter, MW. Analysis of US child care safety regulations. Am J Public Health 1991;81 (8):981-985.

14. American Public Health Association and American Academy of Pediatrics. Caring for Our Children: National health and safety performance standards for out-of-home child care programs. Washington, DC: American Public Health Association, 1992, pgs 1-4.

DRUG AND ALCOHOL FINDINGS:

NEW YORK STATE STUDENTS IN FIFTH AND SIXTH GRADES

Blanche Frank, New York State Office of Alcoholism and Substance Abuse Services
Gregory Rainone, Rozanne Marel, Zahra Aryan, Norman Williams, Torrington Watkins

INTRODUCTION

Periodically, the Bureau of Applied Studies of the New York State Office of Alcoholism and Substance Abuse Services conducts a major survey of substance use among students in public and private schools throughout the State. In the winter of 1989-1990, the Bureau conducted its most recent survey, which sampled an especially large segment of the youthful population. Students in grades seven through 12 usually constitute the survey sample. This time the study included fifth and sixth graders as well, since findings from earlier school surveys suggested that many students began using drugs before entering the seventh grade. A sample of 10,656 fifth and sixth graders was included in the study, representing approximately one-half million students in these grades in public and private schools throughout the State. This paper describes findings for these younger students, most of whom are 10 and 11 years of age.

Methodologically, a self-administered questionnaire was designed specifically for fifth and sixth graders. The instrument maintained students' anonymity, and was available in both English and Spanish. Specific questions tapped the use of and exposure to alcohol, tobacco, and a variety of illicit substances. The following section focuses on the youngsters' own substance use and exposure to substance use on the part of others. The second section focuses on the important lines of communication to the children.

1. Substance Use and Exposure to Use

Overall. fifth and sixth qraders show very low rates of substance use. Nevertheless, it becomes evident from the survey's findings that many youngsters live in environments where their close friends use, where they have been offered a variety of substances, and where they have actually witnessed others using illicit substances. The highlights that follow discuss use rates among students and student exposure to substance use.

A. Substance Use

Of the substances queried by the survey, those most frequently used were alcohol and, to a lesser extent, cigarettes. Nearly one-third (31 percent or 131,000) Of the students in grades five and six reported that they had had a drink of alcohol in their lifetime--"not just a sip or taste"--and 15 percent or 62,000 students reported that they have smoked cigarettes. Regarding use of specific alcoholic beverages, wine or wine coolers were the most popular, followed closely by beer. Over one-quarter (28 percent or 113,000) of the students reported that they had had wine or wine cooler, while one-quarter (101,000) reported that they had had beer in their lifetime. The number of students who reported that they had had a drink of liquor in their lifetime was much lower (12 percent or 49,000).

The rate of use of illicit substances was, as one would expect, quite low when compared to licit substances such as alcohol and cigarettes. Some illicit 'substances, however, were more prominently reported than others. The rates of use of marijuana (two percent or 7,000) and inhalants (also two percent), while extremely low in absolute terms, were nonetheless approximately twice as high as the rates reported for other illicit substances. In addition to lifetime drug use, the survey also asked about recent (in the past six months) and current (in the past 30 days) use. Given the low rates of use in this population, the analysis of findings on substance use will be restricted to lifetime use.

B. Substance Use of Close Friends

As noted in the introduction, the survey asked about the extent of student's exposure to substance use as well as the extent to which students perceived that their friends engaged in substance use. The value of exploring these factors is that both exposure to substance use and perceptions surrounding such usage may influence one's own propensity to use substances. Thus, three facets of exposure were queried:

- substance use of close friends;
- witnessing actual episodes of use; and
- being offered licit and illicit substances.

The findings regarding substance use of close friends, as expected, parallelled findings regarding students' own usage, but almost consistently exceeded students' use. The substances most prominently reported by students as being used by close friends were cigarettes (28 percent or 117,000) and alcohol (23 percent or 94,000). Regarding close friends' use of illicit drugs, marijuana was the most frequently reported substance (seven percent or 27,000), followed by the use of inhalants, and some form of cocaine (both at four percent).

Although the proportion of students reporting use of specific substances by close friends usually exceeded the proportion of students who reported that they had actually used these substances, alcohol use was an exception. While only 23 percent of students reported that their close friends used alcohol, 31 percent reported that they themselves had actually

had a drink of alcohol in their lifetime. It should be noted that some discrepancies may be accounted for by the fact that respondents always knew about their own use but may not have known about their friends' use. In addition, a student may have many close friends and these may also be the close friends of other respondents.

C. Students' Exposure to Actual Use

In addition to questions on student's perceptions of friends' substance use, the survey also included questions on students' exposure to episodes of actual use. Students were asked about instances of actual use "in real life (not just on TV or in movies)" and being offered substances. In many instances, the findings were particularly dramatic.

Regarding the extent to which students witnessed instances of actual use of illicit substances, marijuana was the most frequently reported substance (42 percent or 174,000), followed by the use of cocaine in any form (27 percent or 113,000), crack or freebase (20 percent or 79,000), and, most unexpectedly, heroin (13 percent or 55,000). Data on ever seeing alcohol and cigarettes are not presented since these are legal substances that are widely used in the adult population.

Interestingly, instances of inhalant use (glue and "laughing gas" or nitrous oxide), while more widespread than the use of cocaine, crack, or heroin within this population, were, nonetheless, less likely to have been witnessed by students. Instances of glue sniffing were witnessed by 11 percent or 44,000 students and instances of the use of nitrous oxide were witnessed by six percent or 25,000 students.

D. Students Being Offered Substances

Regarding the extent to which students were offered both licit and illicit substances, alcohol and cigarettes were the most frequently reported. Nearly one-quarter (24 percent) of the students reported being offered each of these substances. As far as illicit substances are concerned, marijuana was the most frequently offered substance (seven percent or 30,000), followed by cocaine in any form (five percent or 19,000). Exhibit 1 presents for many of the substances a comparison of findings for "ever seen use," "ever offered," and "ever used." As expected, the likelihood of having ever seen a particular substance used is considerably greater than having been offered the substance, which is, in turn, usually greater than the students' use of the substance. Findings for having been offered alcohol, however, remain an exception. As with use of alcohol by close friends, the proportion and number of students who reported having been offered alcohol (24 percent or 99,000) was exceeded by the proportion and number of students who reported having had a drink of alcohol (31 percent or 131,000). One explanation may be that for some students, the early use of alcohol may occur on a solitary basis.

Two other findings are also noteworthy. First, the number of students who reported that they were offered cocaine in any form exceeded the number who reported being offered glue (three percent or 13,000). Second, the number of students who reported that they were offered crack (three percent or 12,000 students) was virtually the same as the number who were offered glue and, at the same time, somewhat higher than the number who were offered "laughing gas" or nitrous oxide (two percent or 7,000). Yet Exhibit 1 suggests that, despite this higher level of exposure, students were still more likely to have used inhalants than cocaine. This may be because of the higher level of danger and also social stigma that is attached to cocaine use vis-a-vis inhalant use for this age group.

II. Influences and Lines of Communication

Although these young students have had little experience in using the substances queried, they surely have been exposed to the use of these substances on the part of others. Consequently, it is important to understand additional influences on these young students that may be useful in the prevent ion of substance use. The questionnaire probed a variety of links to the students, including the perception of parental attitudes, recent sources of information in school, and students' preferences for seeking help for a drug or alcohol problem.

A. Parental Attitudes and Substance Use

As a way of discerning parental attitudes regarding substance use, students were asked two questions. One question asked how students' parents felt about children their age drinking beer; the second question asked how their parents felt about children their age using marijuana. Generally, students' perceptions regarding parental attitudes on marijuana use were clearer than those regarding parental attitudes on alcohol use. While slightly over one-third of the students (35 percent or 138,000) either did not know how their parents felt regarding alcohol use or felt that they neither approved nor disapproved, less than one-quarter of the students (22 percent or 87,000) made similar claims about parental attitudes regarding marijuana use.

Moreover, students were more likely to report that their parents disapproved of marijuana use, While parental disapproval of beer drinking was reported by b3 percent (252,000) of students, parental disapproval of marijuana smoking was reported by 75 percent (305,000).

B. School Sources of Information

The students were asked to identify school sources for drug or alcohol information that they may have received "since school began in September." Several school sources were listed, and students were asked to indicate "No" or "Yes" for each choice. "Health class" and "assembly program" were the most frequently reported sources of information (40 percent for each). The second most frequently identified was "drug counselor" (36 percent or 140,000), followed by "science class" (27 percent or 106,000), "guidance counselor" (19 percent or 72,000), and, finally, "social science class" (16 percent or 62,000). While the predominance of health class in disseminating information is not very surprising, the finding that over one-third of students have received information from drug counselors is interesting for two reasons. First, it may be an indication of the widespread use of school-based prevention services. Second, it is an illustration of how drug abuse prevention counseling provides information as well as intervention.

C. Student Preferences for Help

Finally, the survev included a question on where students would go if they "were using drugs or alcohol so much" that they needed help (see Exhibit 2). The question provided students with a list of sources of help to respond "Yes" or "No" to as many sources as they wanted. The answers give some clue to the resources that young students value.

Clearly, the most popular source of help chosen by students was "parents" (74 percent or 291,000), while the least popular was "your brother or sister" (44 percent or 168,000). "Friends" was only slightly *more* popular than siblings as a source of help (47 percent or 179,000). "Program or counselor in school," chosen by 54 percent (211,000) of students as a source of help, was virtually tied with "program or counselor outside of school" (53 percent or 206,000). Furthermore, both choices were less likely to be selected as sources of assistance than either "medical doctor" (59 percent or 228,000) or "another person," such as grandparents, other relative, or clergyman (57 percent or 219,000 students). These findings apparently suggest that fifth and sixth grade students tend to turn to particular adults outside the school setting, and that parents hold a very special trust in this regard.

D. Findings for New York City

In analyzing the findings in depth, New York City students are of special interest. Substance use rates for these fifth and sixth graders were extremely low and were identical to the State average. What differed was the students' report of witnessing drug use by others "in real life (not just on TV or in movies)." New York City youngsters were much more likely to have seen illicit drugs used than were the students in the State as a whole. For instance, 45 percent of New York City students had witnessed the use of cocaine, but only 27 percent of New York State students had been similarly exposed. A total of 22 percent of the city's fifth and sixth graders had seen heroin used, compared with 13 percent of students in the State as a whole.

The school, however, may be a mediating force for New York City students. Various lines of communication to the students were explored. New York City students were more likely than other students in the State to indicate that a school drug counselor was a source for drug or alcohol information and that a school counselor was a preference for help with drug or alcohol problems.

Exhibit 1
Extent of Exposure to and Use of Selected Substances
Among 427,000 Students Enrolled in Grades5 and 6 in New York State
Winter 1989190

Type of Substance	%	N
Alcohol[a]		
Ever Offered	24	99,000
Ever Used	31	131,000
Cigarettes[a]		
Ever Offered	24	100,000
Ever Used	15	62,000
Any Inhalants[b]		
Ever Seen	14	59,000
Ever Offered	4	17,000
Ever Used	2	9,000
Marijuana		
Ever Seen	42	174,000
Ever Offered	7	30,000
Ever Used	2	7,000
Crack		
Ever Seen	20	79,000
Ever Offered	3	12,000
Ever Used	1	3,000
Any Cocaine		
Ever Seen	27	113,000
Ever Offered	5	19,000
Ever Used	1	4,000
Heroin[c]		
Ever Seen	13	55,000
TOTAL NUMBER OF STUDENTS		427,000

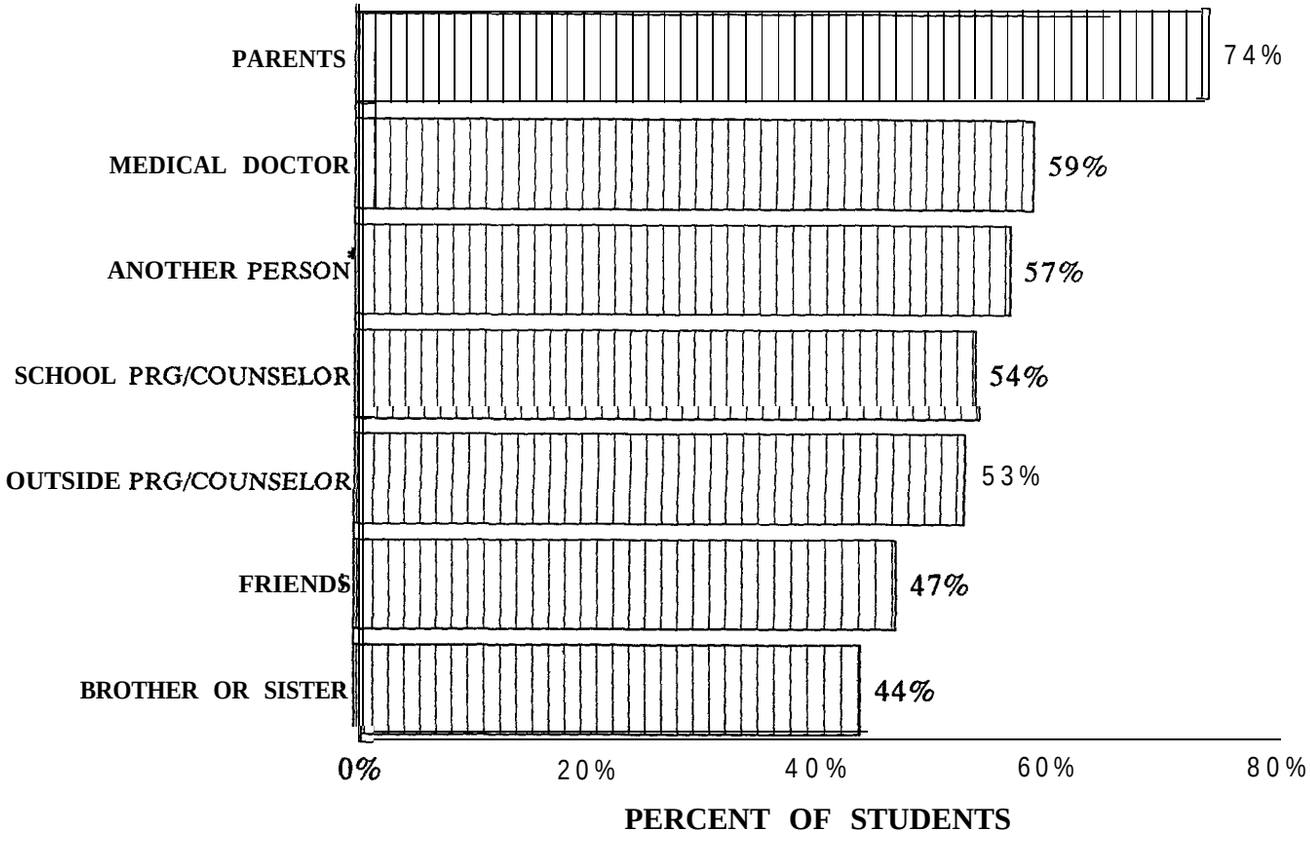
Example of How to Read the Table: Among fifth and sixth graders, an estimated 42 percent have seen the use of marijuana "in real life (not just on TV or in movies)," seven percent have been offered marijuana, and two percent have actually used marijuana.

[a]Since alcohol and cigarettes are legal substances that are widely used, "Ever Seen" is not presented.

[b]Inhalants include glue, sprays, and "laughing gas."

'Questions specifically asking "Ever Offered" or "Ever Used" heroin were not included.

EXHIBIT 2: PREFERENCES FOR HELP WITH A DRUG OR ALCOHOL PROBLEM
AMONG 5TH AND 6TH GRADERS IN NEW YORK STATE, 1990



*SUCH AS GRANDPARENTS, OTHER RELATIVE, MINISTER, PRIEST, RABBI

HOMELESS STREET YOUTH: STRATEGIES AND OBSTACLES IN GATHERING DATA ON ANTECEDENTS TO HOMELESSNESS

Therese van Houten, Macro International Inc.
Gail H. Shur

I. INTRODUCTION

Few adolescents are as at risk as homeless or street youth, and few are as difficult to systematically survey regarding the health and related risks to which they are exposed, e.g., alcohol and other drug use, poor nutrition, depression, physical assault, sexual exploitation, and suicidal ideation. Main methodological concerns include: defining the homeless youth population, obtaining a broadly representative sample, and getting these youth to agree to participate in a lengthy and in-depth survey interview covering personal, and perhaps painful, aspects of their lives.

This paper presents what was learned about strategies for and obstacles to gathering data on antecedents and consequences of homelessness in interviews with 298 homeless adolescents between the ages of 12 and 20 (194 were age 17 and younger) in seven cities in the continental United States.'

II. RESEARCH DESIGN AND SAMPLING PLAN

This section discusses the premises and assumptions underlying the research design, sampling and data collection methods. The next section discusses what was learned during the actual implementation.

1. Defining the study population

Since the study was conducted in response to a mandate from the Stewart B. McKinney Homeless Act of 1987 which required the Department of Health and Human Services (DHHS) to examine the causes of homelessness among youth, the contract asked that, with a few minor modifications, the operational definition of a homeless youth be based on the McKinney Act's definition of a homeless person. This definition, however, does not distinguish between adults and youth. It defines a homeless person as follows:

> "An individual who lacks a fixed, regular, and adequate night time residence: and an individual who has a primary night time residence that is (a) a supervised publicly or privately operated shelter designed to provide temporary living accommodations (including welfare hotels, congregate shelters, and transitional housing for the mentally ill): (b) an institution that provides a temporary residence for individuals intended to be institutionalized: or (c) a public or private place not designated for or ordinarily used as regular sleeping accommodations for human beings." (Public Law 100-77, July 22, 1987).

The definition that was eventually developed for purposes of this study was one that took into account the above definition, as well as existing definitions of homeless youth used in federal legislation related to runaway and homeless youth.' To be included in the study a youth had to:

- Be age 20 or younger
- Not in the care of a parent or guardian
- Either be without a regular fixed address or living in:
  - shelter not designed for human beings
  - shelter obtained illegally or in exchange for sex or drugs
  - shelter for runaway and homeless youth following an episode of inappropriate or illegal beha-vior
  - transitional shelter for homeless youth, homeless adults or battered women.

This definition did not allow an easy way to determine whether a youth belonged to the intended study population. Not all street youth are homeless; many of the unemployed and out-of-school youth who spend most of their days and part of their nights on the street do have a regular fixed address. This had the following implications for the sample selection and interviewer recruitment:

- Sample selection. How does one accurately sample when it is not always immediately clear who belongs to the study population? To what extent should one oversample if the final decision about who belongs in the study population cannot be made until after the interview has been started, or in some instances, concluded?
- Interviewer recruitment. The interviewers had to be individuals who were so extremely knowledgeable about the local street youth culture that they would almost intuitively know which youth to interview.

2. The sampling plan

The contract called for interviews with youth in urban settings (preferably 50 youth per site). This required a two-phase sampling plan. It was necessary first to sample cities and then to sample homeless youth within each sampled city.

The city sample. Youth who are homeless, like homeless adults, tend to live in areas where their presence is at least tacitly accepted by the community, and where, legally or otherwise, they can meet basic survival needs for food, shelter, and sanitary facilities.

Considering the limited resources available for this study, it was not possible to randomly select the cities nationwide and then assume that it would be possible to find a sample of homeless youth in each city. Instead, the approach taken was to create a list of cities known to have large homeless youth populations, and to select a geographically representative sample. There were drawbacks to this approach, however, as the selection of a city with a reportedly high number of homeless youth did not in and of itself guarantee that the field interviewers would be able to gain access to the youth. We used the following selection criteria to sample cities: the willingness of local youth providers to help recruit interviewers and to let it be known "on the street" that this was a legitimate study: geographic diversity within the continental United States: and ethnic diversity. It was hoped that the geographic location and ethnic composition of the selected cities would ensure inclusion of diverse racial and ethnic groups in the youth sample.

This approach would not have been suitable if one of the study's intended outcomes was to estimate the size of the homeless youth population. Since the study was essentially an exploratory one, it was deemed appropriate to use a purposive sampling approach.

The homeless youth sample. Limiting the study population to homeless youth known to local service providers would have allowed a random selection of youth. However, this would have resulted in an unacceptable bias, as youth who were not known to local providers would have been excluded from the study. Instead, we selected a purposive quota sample with the three quota being the youths' 1) level of contact with service providers, 2) age, and 3) gender.

- Level of contact with service provider. A distinction was made between homeless youth who were in contact with service providers and those who were not. Low-contact youth were defined as youth who are only receiving services designed to meet basic survival needs: food, shelter, laundry., etc. High-contact youth were defined as those youth receiving services designed to terminate their homelessness. At least one-third of the youth were to be low-contact youth.
- Age. The definition of the study population set the maximum age limit at 20. No minimum age limit was set. We decided on two age categories: youth under age 18, and youth ages 18, 19, and 20. At least two-thirds of the youth were to be under age 18; the remainder could be ages 18, 19 or 20.
- Gender. It was initially planned to ask local service providers to estimate the male-female ratio of their homeless youth population, and to use these figures to set a gender quota for each city. However, in the pilot conducted in spring 1991 in Boston, this proved to be an unnecessarily cumbersome and expensive task. Instead, we sought a 50-50 male-female ratio.

The contract did not call for the use of a comparison group of non-homeless youth.

3. Data collection plan

As discussed above, the selection of cities was based in part on the proposed data collection plan: to recruit and train local youth workers to conduct in-person interviews with homeless youth. This plan presupposed that qualified interviewers would be available and that youth would be willing to participate in the interviews.

Recruiting youth interviewers. In each selected city, a director of a homeless youth service agency agreed to assist in interviewer recruitment. Applicants were screened and hired as temporary part-time staff by Macro, and were then trained on-site in the administration of the data collection instrument and research study procedures. Interviewers were added incentive to compensate for the extra time needed to locate the low-contact youth and to interest them in participating in the study.

Developing the interview instrument. As no existing instrumentation was sensitive to the complex risk behaviors exhibited by these youth, a modular, structured questionnaire that contained primarily close-ended questions was developed. It was based in part on a questionnaire used in an earlier ACF follow-up study on youth who had stayed in ACF-funded runaway and shelters.[3] Each module asked about a different aspect of the youths' lives. Questions were used that would allow comparisons with other national data sets, notably the Youth Risk Behavior Survey conducted by the Centers for Disease Control. The questionnaire covered the following topics:

- Youth demographics: age, gender, race and ethnicity, citizenship, marital status, and adolescent parenthood.
- Antecedents of youth homelessness: characteristics of family origin and of mother and father figure(s) who reared the youth, history of foster home placements, incidence of physical and sexual abuse, parental drug and alcohol abuse, and other stressors affecting parents and other household members. Other factors covered in the questionnaire were youth behaviors and characteristics (e.g., self-reported suicide attempts, prior episodes of running away, alcohol and other drug abuse, and school failure), as well as

characteristics of the neighborhoods in which the youth were reared.

- Consequences of youth homelessness: The youth's current living arrangements and psychosocial functioning and wellbeing, as well as services received by the youth in the past 30 days, obstacles to receiving services, and services desired.

The questionnaire was pretested in Boston and Washington, D.C., with nine homeless youth, and piloted with 54 homeless youth in Boston. A major concern during the pre- and pilot test was to determine whether the youth would be willing to respond to the questions and to sit through the interview.

Engaging the youth in participating in the interview. To compensate the youth for their time, and to demonstrate the importance of the study, each youth was to be paid $15.00 per completed interview. In addition, the interviewers could spend up to $5.00 per interview on providing a fast-food meal or snack to the youth during or after the interview.

III. IMPLEMENTATION OF TEE RESEARCH PLAN

This section discusses the implementation of the sampling and data collection plans: what worked and what did not, and how implementation problems were resolved.

1. Implementing the sampling plan

The city sample. The following seven cities were selected according to the predetermined criteria (the probability of access to the youth and geographic representativeness): Boston, Chicago, Kansas city, Seattle, San Diego, Tucson, and Ft. Lauderdale. (Because of budgetary considerations, the size of the sample was reduced from eight to seven cities.) Purposely excluded were San Francisco, Los Angeles, and New York because of the possibility that the problems faced by homeless youth in these cities would be viewed as worse than those in other urban centers and therefore nonrepresentative.

The youth sample. The number of interviews per site was intended to be 50. The number of completed interviews per site ranged from 22 to 57. Fifty or more interviews were conducted in Boston, Ft. Lauderdale, and Seattle. Sites with fewer than 50 interviews were: Tucson (44 interviews), Kansas City (30 interviews), San Diego and Chicago (each with 22 interviews). The number of youth interviewed per site appears to have been a function of 1) the amount of time local interviewers (most of whom worked fulltime in youth service agencies) were able to devote to the interviewing task, and 2) the fact that we had to halt data collection due to lack of resources (both funding and time). We do not believe it to be a reflection of the size of the homeless youth population in each sampled city.

There were two methodological concerns about the youth sample: 1) the feasibility of determining which youth fit the study definition and, 2) meeting the sampling quota. The inability to prescreen which youth were homeless meant that the decision to include a youth in the study had to be made prior to the interview by the interviewer, and then be confirmed after the interview based on the youth's responses to questions regarding his current living situation. In all, 318 interviews were conducted. Five interviews were immediately rejected as clearly inappropriate: two interviews with recent runaways staying in a youth shelter, one interview with a street youth who actually had a home to return to, and two duplicate interviews conducted by different interviewers. The interviewers were not paid for these interviews. Of the remaining 313 interviews, an additional 15 were later rejected as not strictly fitting the definition of homeless youth. Most were older youth who had spent the past 30 days living with friends. Thus, 298(94%) of the 318 interviewed youth fit the study population.

The final sample size (298) represented 85% of the proposed sample size of 350 (50 interviews per city). There were indications that the youth who did not meet the study definition of youth homelessness were "doubling up" and did not have a regular fixed address. However, this question was not asked. Responses to the following five questions were used to ascertain a youth's homeless status: 1) places of night time shelter during the 30 days preceding the interview; 2) place where youth spent the night prior to the interview: 3) the length of time the youth had been away from parent or guardian; (4) the place where youth usually ate meals; and 5) the youth's perception of him/herself as homeless. In addition, for a few youth, we were able to obtain clarification by reviewing the youth's responses to several of the open-ended questions, as well as the interviewers' comments.

- Night-time shelter during the preceding 30 days. In the 30 days preceding the interview, the youth used the following places of night- time shelter: home of friends or relatives, 60%; youth shelter, 54%; "on the street," 45%; hotel, 11%; correctional facility, 5%; general hospital, 3%; and psychiatric hospital, 1%.
- Shelter during the night prior to the interview. Fifty-four percent of the youth stayed in a youth shelter, and 3 percent in an adult shelter: 25% stayed in a house or apartment: 16% spent the night "on the street," 2% were in a hotel, and one youth (<1%) had been in jail. There were significant differences across sites (chi square=82.07; df=30; p<.000). In Ft. Lauderdale and Chicago, over 80% of the youth were in a youth shelter, In San Diego, 46% spend the night on the

street, and in Tucson, 40% had been in a house or apartment usually with friends.

- Lensth of time awav from parents. On the average, both youth reported being away from parent or guardian for 18 months. Across sites, these differences were significant (F=5.6; p<.000). The Kansas City youth being away from home for the shortest period of time (6 months), and the Seattle youth the longest period of time (32 months).

  There were some differences across sites between low- and high-contact youth, but in no city were these statistically significant. In Boston, the mean length of time that low-contact youth had been away from home was 20 months; for high-contact youth it was 13 months. (t=1.57; df=34;p<.125) In San Diego, low-contact youth had been away an average of 16 months versus 38 months for the high- contact youth. (t=-1.79; df=ll; p <.101).

  Place where vouth usuallv eats meals. Of 120 youth who had not stayed in a shelter during the past month, 18 percent reported eating in drop-in centers and shelter day programs.

- Whether youth considers- self to be homeless. About half of the youth who fit the study's definition of homelessness (40%) did not label themselves as homeless: They explained that they were not "bums," or that to consider oneself homeless would be "pitiful," or that local shelters provided a roof over their head and a place to sleep.

Meeting the sampling quota. The sampling plan called for the assignment of quota to the interviewers in each city. It soon became clear that certain quota were easier to fill in some cities than others. For instance, older high-contact youth were more readily recruited for interviews in Boston, whereas younger low-contact youth were more quickly found in Seattle. Because of time constraints, the decision was made to assign the quota to the entire sample rather than to each city. The sample met the three designated quota requirements.

- Distribution bv age. Sixty-five percent of all sampled youth (n=298) were between the ages of 12 and 17; 35% were ages 18, 19, or 20. The mean and median age was 17. In Boston, Kansas City, and Ft. Lauderdale, the sample distribution met the quota: two-thirds of the youth were age 18 or younger. In Chicago, only 50% of the 22 youth interviewed were under age 18. On the other hand, in Tucson and San Diego, over 75% of the interviewees were under age 18. These age differences by site were not statistically significant (chi square=7.8, df=6, sign.=.256). See Table 1.

- Distribution bv level of contact with service nrovider. The sample was evenly divided between low-contact (49%) and high-contact youth (51%). of the males, somewhat fewer youth (45%) were classified as high contact. The reverse was true for the females--52% were classified as high contact. However, this difference was not statistically significant (chi square= 6.51, df=6, sign.=.366).

- Distribution by gender. An approximately 50-50 male-female ratio was achieved in Boston, Ft. Lauderdale, and Seattle. However, in San Diego and Chicago, there was a far greater percentage of males (86% and 78%, respectively). In Kansas City, only 28% of the respondents were male. Across sites, this difference was significant (chi square=21.05, df=6, sign.=.002).

Distribution by race and ethnicity may have been a function of site location. Overall, 49% of the youth were white, 28% were African American, 17% were Hispanic, 5% were Native American and 1% were Asian. There were some differences across sites: In Boston, 72% of the youth were white, while in Chicago, 75% were African American. San Diego had the largest percentage of Hispanic youth (46%), and Seattle the largest percentage of Native American youth (12.5%).

2. Implementation of the data collection procedures

Start-up of data collection. As the study was commissioned through a Federal contract, no data could be collected until the contracting agency received clearance from the Office of Management and Budget (OMB). The OMB clearance process took longer than anticipated. This delay had the both negative and positive effects on the data collection activities:

- Recruitment of kev service nrovider agencies in participating cities. Cities with supportive key service providers had to be selected prior to the request for OMB clearance. Turnover in agency directors meant that when clearance was obtained, arrangements had to be made "from scratch" with the new project director.
- Recruitment of interviewers. For the same reason, interviewer recruitment had to be delayed until clearance was obtained. In most agencies, interviewers who were interested in the study were no longer available when we were ready collect data.
- Interviewer training. The delay had a positive effect on interviewer training. We had planned to recruit one interviewer per city, and to train all interviewers at one central location. To overcome problems caused by the delay, we recruited three to five interviewers in each city and conducted the training on-site as soon as the

Table 1

Percentage distribution of respondents by age, gender, and level of contact with service provider by city

City	Gender		Level of contact with service provider		Age	
	Male	Female	LOW	High	Under age 18	Age 18, 19 or 20
Boston (n=53)	49%	51%	51%	49%	66%	33%
Ft. Lauderdale (n=57)	47%	53%	49%	51%	61%	39%
Seattle (n=56)	54%	46%	55%	45%	64%	36%
Tucson (n=44)	48%	52%	64%	36%	77%	23%
San Diego (n=22)	86%	14%	50%	50%	78%	22%
Kansas City (n=30)	23%	77%	50%	50%	70%	30%
Chicago (n=22)	78%	22%	59%	41%	50%	50%
Total	49% n=146	51% n=152	51% n=153	49% n=145	65% n=195	35% n=103
Chi square	21.05		6.51		7.76	
Degree of freedom	6		6		6	
Level of significance	.002		.366		.256	

interviewer recruitment process for that city was completed. In retrospect, this approach was far better than the original plan.

- Completing sampling auota. The delay and the additional tasks meant that fewer resources were available to implement the data collection. It is for these reasons that the quota requirements were modified to encompass the seven-city sample, and that interviews were discontinued in four of the seven cities before they attained the planned sample size.

The use of local interviewers. After a thorough screening process, 28 individuals from the 7 designated cities were selected to conduct the interviews. Three-fourths of the interviewers were white. In three cities (Chicago, Boston, and Seattle) all interviewers were white. In general, there was little correspondence between the race and ethnicity of the interviewers and the interviewed youth. For instance, in Chicago where 75% of the youth were African American, all 3 interviewers were white. On the other hand, in San Diego, where 46% of the interviewed youth were Hispanic, 2 of the 4 interviewers were also Hispanic. Although most of the Hispanic origin interviewers were bilingual, only a few interviews in Tucson and San Diego were conducted partially in Spanish.

Fifty-seven percent of the interviewers were female. Their ages ranged from 22 to 40. The mean age was 30 (std. dev.=5.47); the median age was 27. The mean age per site ranged from 26 in Tucson to 34 in Ft. Lauderdale.

The number of interviews per interviewer varied significantly. Three interviewers conducted no interviews at all: one received an unexpected job transfer: and the other two misjudged the amount of free time they would have available. A fourth interviewer was only able to conduct one interview because of a lack of free time due to a demanding full-time job and a family. We found no pattern that would help predict, in a future study, what type of interviewer would be more likely to complete a large number of interviews. In a few instances, the financial incentive drove interviewers. to conduct a larger than average number of interviews, but this was not always the case. In general, those interviewers who failed to meet their quota misjudged the amount of time that they would be able to give to the project.

Characteristics of the interview.

- Lensth of the interview. For 16 interviews, the interview was conducted in two sessions. Only three of these occurred in the cities with more than 50 interviews. The length of the interview varied greatly. It ranged from 40 minutes to one 4-hour 2-session interview. The overall mean duration for all interviews (both single session and a-session interviews) was 88 minutes (std. dev.=27). The mean duration of a one-session interview was 90 minutes (std. dev.=23). The mean duration of the first part of a 2-session interview was 60 minutes (std. dev.=32). The range was from 15 minutes to 2 hours.

On the average, the second session lasted 54 minutes (std. dev.=29).

- The locale of the interview. The location of the interview was left up to the interviewer and the youth: Sixty-eight percent of the interviews were conducted in the offices of local shelters, drop-in centers, and other social service programs: 16% took place in fast food restaurants: 8% in parks; 3% each in malls and schools; and 2% in the homes of friends, (note: one youth [<1%] was interviewed in a library). This flexibility appears to have contributed to the interviewers' and interviewees' sense of ownership of the interviewing process.

  Response rates. The interviewers did not record the number of refusals. Informal feedback indicates that refusals were not a problem. We believe that less than 1 percent of the youth approached about the interviews refused to participate. Youth were also told several times throughout the interview that they had the right to respond to specific questions. Yet, all questionnaires that were started were completed. Missing responses did occur: however, these seem to have been the result of interviewer error rather than youth refusal to answer certain questions.
- The use of signed consent forms. Participation in the interviews was entirely voluntary. The youth were assured prior to the interview that not only was participation voluntary, but that they could also choose not to answer any specific questions. Prior to the interviews, the youth were asked to sign a written consent form. Contrary to our expectations, signing was not a problem, although a few youth chose to use aliases rather than their legal names.
- The use of incentives to ensure participation. After a few youth interviewed in the pilot test-misused the $15 cash incentive, interviewers were given the option of paying the youth in cash or in gift certificates. Approximately 10% of the youth interviewed in Boston and all youth interviewed in Tucson were paid in gift certificates to fast-food restaurants or local stores. In Tucson, the interviewers chose to pay the youth with $5 cash and a $10 K-Mart certificate.
- The provision of a light snack. Snacks ranged from a meal at a fast-food restaurant to a can of soda. In Ft. Lauderdale, one of the interviewers used too broad a definition of "snack," and the interviewer (who was a smoker), bought the youth a pack of cigarettes. Thereafter, we clarified the definition of "snack" in the training to ensure that there would be no further misunderstandings.

IV. SUMMARY

Findings indicate that the following procedures are key to successful recruitment and interviewing of high-risk populations of this nature: 1) careful selection and training of interviewers knowledgeable about, and comfortable with, the high-risk population being studied: 2) the allotment of sufficient time for interviewing as well as for finding respondents; 3) the use of an instrument that is sensitive to the specific culture of the population being studied: 4) the availability of funds to provide incentives to both interviewers and interviewees; and 5) flexibility in the procedures and adaptation to the local setting.

---

1. The study was sponsored by the Administration of Children, Youth and Families, Department of Health and Human Services, and conducted by Macro International Inc., Silver Spring, Maryland. The proposed methodology was reviewed by an advisory panel of representatives from Federal agencies (including the Centers for Disease Control, the Office for Juvenile Justice and Delinquency Prevention, and agencies in the former Alcohol, Drug and Mental Health Administration).

2. Title III of the Juvenile Justice and Delinquency Prevention Act, 1974, Section 302, P.L. 93.415. Amended 1988 to become the Anti-Drug Abuse Act, Section 321, P.L. 100-690.

3. Cohen, B. and T. van Houten. 1991. Followup Study of Youth Using Runaway and Homeless Youth Services. Washington, D.C.: The Urban Institute.

Session D

Mental Health Indicators-From Epidemiology to Service Delivery

Toward the year

2000

Refining the measures

THE SOCIAL COSTS OF ANXIETY DISORDERS

Andrew C. Leon, Ph.D., Cornell University Medical College
Myrna M. Weissman, Ph.D.

In 1977 the President's Commission on Mental Health convened. First Lady Rosalynn Carter served as the Honorary Chair of that Commission. Included in their investigation were three fundamental questions: 1) How many people are mentally ill? 2) Who are they? 3) What services do they receive?

These seemed to be simple questions, yet at the time there were no definitive answers. This highlighted the need for a comprehensive survey of mental disorders in the community. The Commission recommended that an epidemiologic study be conducted to learn about rates of mental disorders and corresponding service utilization. This recommendation led to the development of the National Institute of Mental Health Epidemiologic Catchment Area (ECA) Program. We will discuss the design of the ECA in a few minutes.

Let's move ahead 16 years. In 1993 the Task Force on Health Care Reform was convened and chaired by Hilary Rodham Clinton. In those hearings, rates of mental disorders and the corresponding impairment and costs were discussed. However, unlike in 1977, comprehensive data was now available. This was a result of the research need identified by the President's Commission on Mental Health (1978). The ECA data was used in the Task Force hearings this year. It provided empirical evidence of the prevalence, symptomatology and associated impairment of those suffering from mental disorders.

We employ the ECA data to evaluate the social costs of anxiety disorders. In doing so, we illustrate how a well-conceived data base can be used to answer these and other questions -- including questions that developed after the survey has been completed.

In this time of limited health care resources, policy makers ask if reimbursement for treatment of mental disorders can be justified? We use these data to evaluate the social costs of anxiety disorders. We examine the number of people afflicted with anxiety disorders and evaluate the social morbidity that is associated with the anxiety disorders.

We will show that anxiety disorders afflict a substantial proportion of the population in the US. The symptomatology can be painful and distressing and the resulting help-seeking places a great demand on the health care system. First, we describe the epidemiologic study.

The Epidemiologic Catchment Area Program. Included among the objectives of the ECA were: 1) To provide estimates of rates of mental disorders in the community (treated & untreated); 2) Identify subgroups at high risk for mental illness; 3) Examine the concordance of community needs and mental health services.

The ECA was a survey of mental disorders conducted in the United States in the late 1970s and early 1980s (Robins et al., 1984; Regier et al, 1984; Robins and Regier, 1991). The ECA sample consisted of over 18,000 non-institutionalized adults from five US sites: New Haven, Connecticut; Baltimore, Maryland; Durham, North Carolina; St. Louis, Missouri; Los Angeles, California. The data used in our analyses come from Wave I of that study.

Fifty-nine percent of the sample is female. Two-thirds of the sample is white. About half of those surveyed in the ECA study were married. Each adult age group was well represented in the study.

The NIMH Diagnostic Interview Schedule (DIS) is a highly structured interview. It was designed to be conducted by well-trained lay interviewers. The DIS ascertains information regarding symptomatology for each of the diagnoses. The symptom data include information on duration, frequency, age of onset, and associated impairment.

The DIS incorporates the DSM-III criteria (APA, 1980). Diagnoses were determined for different time frames (two weeks, one month, six months, one year, and lifetime). It also elicited information on demographics, and health service utilization. The psychiatric diagnoses that were determined include the affective disorders, schizophrenia, alcohol and substance abuse and dependence, anxiety disorders and somatization. In DSM-III there were several anxiety disorders. The anxiety disorders that we will discuss include panic disorder, the phobic disorders and obsessive compulsive disorder (OCD).

In the NIMH ECA Program, it was estimated that over 16% of the general population had suffered from a DSM-III anxiety disorder at some point in their lives -- 12% of males and 20% of females. More specifically, the lifetime rates of panic disorder are 1% for males and 2% for females; for OCD the rates are 2% for males and 3% for females; the rates of phobias were most prevalent: about 8% among males and 18% for females. The most common phobia was simple phobia. Females consistently have about twice the rates of each anxiety disorder. Put in the context of the US population, this represents over thirty million Americans having suffered from an anxiety disorder at one point in their lives.

The rates that we have mentioned refer only to those individuals meeting DSM-III criteria for an anxiety disorder. About 3% of the general population had panic attacks, yet did not meet DSM-III criteria for panic disorder. We will show that there is considerable social morbidity even among those panickers who were below diagnostic threshold for panic disorder.

These rates are not unique to the US, but look similar cross-culturally. For instance in the Munich Follow-up study the rates were comparable with those of the ECA, except for simple and social phobias.

There are effective treatments for anxiety disorders. However, the burden of anxiety disorders extends beyond the direct costs of treatment to the indirect costs of impaired social functioning. We will present evidence of financial dependence, help seeking and drug and alcohol comorbidity among each of the diagnostic groups. Those with anxiety disorders suffer from considerably more social morbidity than those with no psychiatric disorder.

FINANCIAL DEPENDENCE is elevated among individuals with anxiety disorders and this is shown in Table 1. For instance, those with an anxiety disorder, in the past six months, are more likely to be unemployed. In the ECA, about 20% of males without Axis I disorders were not currently employed. In contrast, about 36% of males with phobias, 45% with OCD, and 60% males with panic disorder were unemployed. That is a threefold increase over those with no Axis I disorder. About 30% of males with subclinical panic attacks were not currently employed. For females there is an elevation in rates associated with anxiety disorders, but it is less striking.

There is also a greater chance of chronic unemployment among those with anxiety disorders. "Chronic unemployment" refers to those not employed for at least five years. In Table 2 the differences are most striking in males. Although all anxiety disorders are associated with increased rates of unemployment for males, panic disorder has the greatest elevation. There are no such differences for females.

If a substantial proportion of those with anxiety disorders do not work, how do they pay for food, clothing and shelter? They are more likely to receive financial assistance than those who do not suffer from a psychiatric disorder as is demonstrated in Table 3. Males with panic disorder are about six times as likely to receive disability as those with no Axis I diagnosis; the rates are slightly lower for OCD and phobias. For females the rates of disability payments are lower.

Receipt of welfare payments is also elevated among those with anxiety disorders. This includes Aid for Dependent Children, thus the rates are higher among females. As has been seen earlier, the elevation is greatest in those with panic disorder. It also elevated for the other anxiety disorders.

Taken as a whole, those with anxiety disorders received considerably more financial assistance than those with no mental disorder. The gender differences disappear when the forms of financial assistance are combined.

Now help-seeking among those with anxiety disorders is discussed. There are safe and effective psychotherapeutic and psychopharmacologic treatments of each anxiety disorder. If the disorder is detected by a physician, proper treatment can reduce social morbidity. However because of the prevalence of anxiety disorders, and the inadequate screening procedures, the help-seeking behavior places a tremendous burden on the health care system. Rates of help-seeking for emotional, drug or alcohol problems in the six months prior to the interview are presented in Table 4 and compared across diagnoses.

SPECIALIZED MENTAL HEALTH SERVICES. Not surprisingly the rates of seeking specialized mental health services (such as outpatient visits to a mental health specialist) are magnified for those with anxiety disorders compared to those with no Axis I diagnosis. The rates are rather similar for males and females. Over one-third of those with panic

disorder seek such help; compared with about 16% of those with panic attacks or OCD; and less than 2% for those with no Axis I diagnosis. Indeed the rates of help-seeking are elevated in the ill. However, if these rates were higher the social morbidity might be reduced. What is the plight of the 60% of panic disorder patients who sought no specialized mental health services during the past six months?

GENERAL MEDICAL SERVICES. Those with anxiety disorders not only go to specialized mental health service providers, but also seek help in emergency rooms and from primary care practitioners. For that reason several efforts are underway to train physicians to look for panic disorder, for example, in the general medical setting. Consider that less than 4% of individuals without an Axis I diagnosis sought help from the general medical system for emotional problems in the prior six months. The rates are much higher for those with anxiety disorders -- especially panic disorder. This might result from the physiological symptoms of a panic attack such as palpitations and shortness of breath -- which mimic cardiac problems.

EMERGENCY ROOMS. Those with anxiety disorders are also far more likely to seek help for emotional problems from emergency rooms. This is primarily limited to those who have panic attacks or panic disorder. Although the absolute rates are low, there are large relative differences.

ANY HELP-SEEKING. Taken as a whole, the help-seeking for emotional problems among anxiety disorder patients is very high. Over half of the patients with panic disorder seek help. Nearly 30% of those with other anxiety disorders, yet only 10% of those with no diagnosis seek help for emotional problems.

OTHER COSTS of anxiety disorders include alcohol and substance abuse. People suffering from panic disorder have very high rates of alcohol and substance abuse. These data are presented in Table 4. This might reflect attempts at self-medication. The rates are highest for males with panic disorder or OCD. The magnitude of the rates are lower for females, but the association between anxiety and alcohol remains strong.

Likewise there are also elevated rates of substance abuse and dependence among those suffering from anxiety disorders. The substance rates are elevated from two to over fivefold in those with anxiety disorders. Again, females' rates of drug abuse and dependence are lower than males, but still elevated in females with anxiety disorders.

IN CONCLUSION, the social costs of anxiety disorders are very high. The costs include financial dependence, service utilization and alcohol and substance abuse. Based on the results of clinical trials (e.g., panic disorder: NIH Consensus Development Conference Statement, 1991; OCD: The Clomipramine Collaborative Study Group, 1991; phobias: Barlow, 1988; Gelernter et al., 1991), we believe that these costs would be reduced if the disorders were detected and treated by a physician. Short screens are currently being evaluated for use in primary care settings and in emergency rooms (Apfeldorf et al., in press; Broadhead et al., submitted).

This evidence of the prevalence and social morbidity of anxiety disorders is available because of the President's Commission on Mental Health. We acknowledge their foresight. It serves to illustrate progress that has been made in mental health research. In 1993 survey data is used to make policy decisions regarding mental health services.

TABLE1
OCCUPATIONAL FUNCTIONING (RATES/100)
AGES 18 - 64

NOT CURRENTLY EMPLOYED

DIAGNOSIS (past 6 months)	MALES	FEMALES
Panic Disorder	60.00	68.57
Obsessive Compulsive	44.78	65.25
Phobia	35.69	54.05
No Axis I Diagnosis	21.05	45.92

TABLE 2
OCCUPATIONAL FUNCTIONING (RATES/100)
AGES 18 - 64

NOT EMPLOYED IN THE PAST 5 YEARS

DIAGNOSIS (past 6 months)	MALES	FEMALES
Panic Disorder	25.00	28.57
Obsessive Compulsive	17.91	28.81
Phobia	12.72	28.02
No Axis I Diagnosis	4.63	23.72

TABLE 3
CURRENT FINANCIAL ASSISTANCE **(RATE/100)**

MALES: AGES **18-64**

DIAGNOSIS (Fast 6 months)	DISABILITY	WELFARE	UNEMPLOYMENT COMPENSATION	ANY ASSISTANCE
Panic Disorder	33.33	6.67	0.00	36.67
Obsessive Compulsive	22.50	2.50	2.50	25.00
Phobia	19.38	2.18	4.67	24.69
No Axis I Diagnosis	5.43	1.23	3.01	9.21
		FEMALES: AGES **18-64**		
Panic Disorder	14.58	30.93	2.06	41.67
Obsessive Compulsive	12.93	13.61	3.40	29.25
Phobia	7.63	16.53	2.69	24.71
No Axis I Diagnosis	4.24	9.59	2.19	15.13

TABLE 4
HELP-SEEKING for EMOTIONAL, DRUG
or ALCOHOL PROBLEMS **(RATE/100)**
(past 6 months)

MALES

DIAGNOSIS (Fast 6 months)	GENERAL MEDICAL SYSTEM	SPECIALIZED MENTAL HEALTH SYSTEM	HUMAN SERVICES	EMERGENCY ROOM	ANY HELP-SEEKING
Panic Disorder	33.33	36.67	13.33	3.33	63.33
Obsessive Compulsive	7.07	17.17	4.04	0.00	24.24
Phobia	8.33	8.33	3.92	1.23	17.40
No Axis I Diagnosis	1.93	1.70	1.05	0.07	4.35
			FEMALES		
Panic Disorder	25.93	31.48	11.11	3.70	56.48
Obsessive Compulsive	11.96	16.30	4.35	0.00	27.72
Phobia	10.67	7.53	4.71	0.78	20.00
No Axis I Diagnosis	3.80	2.13	1.94	0.18	7.19

REFERENCES

American Psychiatric Association, Committee on Nomenclature and Statistics (1980). Diagnostic and Statistical Manual of Mental Disorders, Third Edition. Washington, DC: American Psychiatric Association

Apfeldorf WA, Shear MK, Leon AC, Portera L. A brief screen for panic disorder. (in press) Panic attacks in the community: Social morbidity and health care. Journal of Anxiety Disorders

Barlow, D (1988). Anxiety and its disorders: The nature and treatment. New York: Guilford Press

Broadhead WE, Weissman MM, Leon AC, Barrett JE, Blacklow R, Gilbert T, Keller MB, Higgins E, Olfson M. The SDDS-PC screening questionnaire for multiple mental disorders in primary care: A validation Study. submitted for publication

The Clomipramine Collaborative Study Group (1991). Clomipramine in the treatment of patients with obsessive-compulsive disorder. Archives of General Psychiatry, 48:730-738

Gelernter CS, uhde TW, Cimbolic P, Arnkoff DB, Vittone BJ, Tancer ME, Bartko JJ. (1991) Cognitive-behavioral and pharmacologic treatments of social phobia. Archives of General Psychiatry, 48:938-945

NIH Consensus Development Conference Statement (1991). Treatment of Panic Disorder. NIH Consensus Development Conference, September 25-27, 1991; 9(2)

President's Commission on Mental Health (1978). Report to the President from the President's Commission on Mental Health (Stock No. 040-000-00390-8, Vol. 1). Washington, DC: US Government Printing Office

Regier DA, Myers JK, Kramer M, Robins LN, Blazer DH, Hough RL, Eaton WW, Locke BZ (1984) The NIMH Epidemiological Catchment Area Program: historical context, major obstacles, and study population characteristics. Archives of General Psychiatry 41:934-941

Regier DA, Boyd JH, Burke JD, Rae DS, Myers JK, Kramer M, Robins LN, George LK, Karno M, Locke BZ (1988) One-month prevalence of Mental Disorders in the United States: Based on five Epidemiologic Catchment Area sites. Archives of General Psychiatry 45:977-986

Robins LN, Helzer JE, Weissman MM, Orvaschel H, Gruenberg E, Burke JD, Regier D (1984) Lifetime prevalence of specific psychiatric disorders in three sites. Archives of General Psychiatry 41:949-958

Robins LN, Helzer JE, Croughan J and Ratcliff KS. (1981) National Institute of Mental Health Diagnostic Interview Schedule: Its history, characteristics, and Validity. Archives of General Psychiatry, 38:381-389

REFLECTIONS ON APPROPRIATE MENTAL HEALTH QUESTIONS FOR THE HEALTH INTERVIEW SURVEY

James S. Larson, University of Arkansas at Little Rock

A musical friend of mine, who directs a choir, told me the following joke: What's the difference between a terrorist and a first soprano? You can negotiate with a terrorist. That is a little like the difference between the physical and mental health. One is relatively easy to understand, while the other is hard to negotiate with, hard to understand.

Mental health is an important component in a person's overall health. We are all familiar with the WHO definition of health as "complete physical, mental, and social well-being," but perhaps we are not as familiar with the fact that most of the well-known instruments measuring overall health have a significant mental health component.

(SEE TABLE 1)

The Sickness Impact Profile measures physical, mental, and social health, and devotes 17% of its questions to mental health. The Nottingham Health Profile also measures the three dimensions, and devotes 33% of its questions to mental health. Looking at the rest of the table, all instruments except one have questions for physical, mental and social health.

At the bottom of the table, all of the instruments together average 54% of their questions for physical health, 29% for mental health, and 17% for social health. This table is a reminder that mental health is an important component of overall health, and should be included in any serious assessment of "health" in a population. It is also a reminder that social health should be included, but this paper is devoted to the subject of mental health.

What are the components of mental health?

(SEE FIGURE 1)

In a paper given at this conference two years ago (Larson, 1991), I presented this figure. Mental health consists of both the absence of mental disorders and, on the positive side, mental well-being. Mental well-being is subdivided into the absence of psychological distress and positive well-being, which includes satisfaction with life. Psychological distress includes anxiety, depression, and other factors.

This figure was presented to psychologists in the form of a survey which I conducted last year (Larson, 1992). They were asked to assign weights to these various components, and to add any additional components of their own. The sample was taken randomly from members of the American Psychological Association, and 113 responded.

(SEE TABLE 2)

As the table indicates, the psychologists believed that 40% of mental health in a population is the absence of mental disorders, 30% is the absence of psychological distress, and 30% is positive well-being. Psychological distress was subdivided into 10% anxiety, 10% depression, and 10% self-control and coping. The results of this survey are notconclusive, but they provide preliminary information on the weighting of these components. When combined with other information, they help to suggest the number of questions which should be devoted to each component.

Another way of weighting the different areas is to look at instruments measuring various aspects of mental health.

(SEE TABLE 3)

At the top of the table, three specialized mental health instruments are found to have four major components: anxiety, depression, well-being, and control. Each of the dimensions is weighted about equally, with slightly less weighting given to depression. The rest of the table looks at the general health instruments, but they include only items on anxiety and depression, with depression receiving the major emphasis. From this table, we can conclude that probably anxiety, depression, and control should be about equally weighted. We might give depression slightly more weighting by including it in questions relating to disorders.

What about disorders and their measurement? Which instruments are best?

(SEE TABLE 4)

In the survey of psychologists mentioned earlier, they were asked to rank the best instruments for measuring mental disorder. Their first choice was the MMPI, followed by the DIS-III (the Diagnostic Interview Schedule, Version III), the Rorschach Test, the Thematic Apperception Test (TAT), and the General Health Questionnaire (GHQ) of Goldberg. The Rorschach Test and TAT are not really appropriate for surveys of the general population, so the remaining instruments are analyzed in Table 5.

(SEE TABLE 5)

At the top of the table, four instruments are listed: MMPI-2, the latest version of the MMPI, DIS-III, DSM-III-R, which is the Diagnostic and Statistical Manual, Version III Revised, and the GHQ. DSM-III-R is the American Psychiatric Association's classification scheme for mental disorders, a major source for classifying disorders.

As the table indicates, depression, schizophrenia, hypomania or mania, and somatization are the major components mentioned most often. Core questions measuring mental disorder in the Health Interview Survey might focus on these more commonly mentioned disorders, but an extended supplement might include all of these dimensions.

The measures of mental well-being are summarized in the next table, once again from the survey of psychologists.

(SEE TABLE 6)

They include measures of anxiety, depression, well-being, and quality of life. As the table indicates, the first choice for measuring anxiety is the State-Trait Anxiety Inventory, the first choice for depression is the Beck Depression Inventory, the first choice for well-being is the Life Satisfaction Index, and the first choice for quality of life is also the

Life Satisfaction Index. The final list of questions which are chosen for the Health Interview Survey should include questions from these instruments.

In the appendix of this paper, the questions which are recommended for the Health Interview Survey are listed. The numbers of questions in each category reflect the weighting suggested from the survey of psychologists and the analysis of prominent instruments. The wording of the questions has been altered to fit the format of the new instrument, so it will be necessary to test these new questions for validity and reliability. These questions are presented as suggestions for possible use in the Health Interview Survey, and are not the final word in questionnaire development.

REFERENCES

Larson, J. (1992). Defining and measuring mental well-being: A survey of psychologists. Presented at the American Public Health meeting, 1992.

Larson, J. (1991). Survey methodology and mental health: Recent developments. Proceedings of the 3991 Public Health Conference on Records and Statistics. DHHS Pub. No. 92-1214.

TABLE1

SUMMARY OF GENERAL HEALTH INSTRUMENTS

Instruments	Physical	Mental	Social
1. Sickness Impact Profile	64%	17%	19%
2. Nottingham Health Profile	58%	33%	9%
3. MOS Short Form General Health Survey	53%	29%	18%
4. McMaster Health Index Questionnaire	32%	34%	34%
5. Functional Status Questionnaire	30%	15%	41%
6. Duke-UNC Health Profile	24%	62%	14%
7. OARS Multidimensional Functional Assessment Questionnaire	55%	33%	12%
8. Arthritis Impact Measurement Scales (AIMS)	64%	21%	9%
9. Physical and Mental Improvement of Function	42%	41%	17%
10. Functional Assessment Inventory	48%	35%	17%
11. Multilevel Assessment Instrument	54%	18%	28%
12. CORE-CARE	71%	17%	12%
13. Quality of Well-Being Scale	100%	0%	0%
MEDIAN VALUE	54%	29%	17%

SOURCES: Ian McDowell and Claire Newell, MEASURING HEALTH: A GUIDE TO RATING SCALES AND QUESTIONNAIRES. (Oxford: Oxford University press, 1987). David Wilkin, Lesley Hallam, and Marie-Ann Doggett, MEASURES OF NEED AND OUTCOME FOR PRIMARY HEALTH CARE, (Oxford: Oxford University Press, 1992).

FIGURE 1

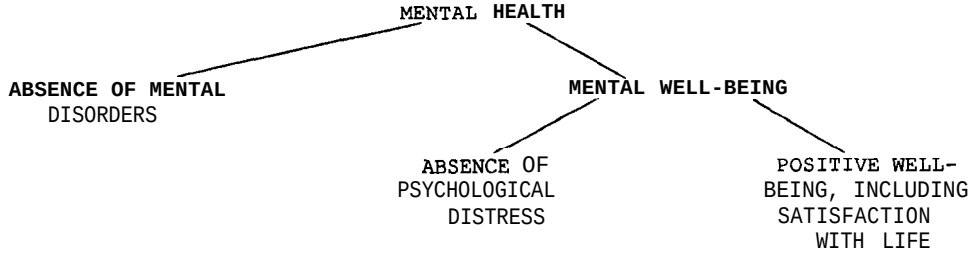


TABLE 2

WEIGHTING OF *COMPONENTS FROM* SURVEY OF PSYCHOLOGISTS

Absence of mental disorders		40%
Absence of psychological distress		30%
Absence of anxiety	10%	
Absence of depression	10%	
Self control and ooping	10%	
Positive well-being		30%
Total.....,...		100%

TABLE 3

WEIGHTING OF COMPONENTS FROM SURVEY *OF* INSTRUMENTS

Specialized Instruments	Anxiety	Depression	Well-being	Control
1. General Well-Being Schedule	22%	17%	17%	28%
2. Mental Health Inventory	26%	13%	29%	24%
3. General Health Questionnaire (GHQ-28)	25%	25%	0%	0%
MEDIAN VALUE	25%	17%	23%	21%

General Health Instruments	Anxiety	Depression	Other
1. Sickness Impact Profile	43%	43%	14%
2. Nottingham Health Profile	27%	64%	9%
3. MOS Short Form	40%	60%	0%
4. AIMS	50%	50%	0%
5. Physical and Mental Impairment of Function	42%	58%	0%
6. OARS	33%	67%	0%
7. CORE-CARE	0%	100%	0%
8. Multilevel Assessment Instrument	14%	43%	43%
MEDIAN VALUE	37%	59%	0%

SOURCES: McDowell and Newell, Wilkin et al.

TABLE 4

MEASURES OF DISORDERS FROM SURVEY OF PSYCHOLOGISTS

First Choice:	Minnesota Multiphasic Personality Inventory (MMPI)
Second Choice:	Diagnostic Interview Schedule Version III (DIS-III)
Third Choice:	Thematic Apperception Test (TAT)
Fourth Choice:	General Health Questionnaire (GHQ) (Goldberg)

TABLE 5

FOUR MAJOR MEASURES OF DISORDER AND THEIR DIMENSIONS

Dimensions	MMPI-2	DIS-III	DSM-III-R	GHQ
Depression	X	X	X	X
Schizophrenia	X	X	X	
Hypomania (Mania)	X	X	X	
Somatization		X	x	x
Hypochondriasis	x			x
Anxiety	x			x
Social Introversion	X			x
Panic Disorder		x	x	
Phobias		x	x	
Anorexia		x	x	
Alcohol Abuse		x	x	
Drug Abuse		x	x	
Obsessions and Compulsions		x	x	
Antisocial Behavior		x	x	
Cognitive Impairment (Measured by Mini-Mental State Exam)		x	x	

TABLE 6

MEASURES OF DISTRESS, WELL-BEING AND QUALITY OF LIFE
FROM THE SURVEY OF PSYCHOLOGISTS

MEASURES OF ANXIETY

First Choice: State-Trait Anxiety Inventory (Spielberger et al.)

Second Choice: Taylor Manifest Anxiety Scale

Third Choice: Zung Self-Rating Scale

MEASURES OF DEPRESSION

First Choice: Beck Depression Inventory

Second Choice: Hamilton Depression Rating Scale

Third Choice: Center for Epidemiological Studies Depression Scale

MEASURES OF WELL-BEING

First Choice: Life Satisfaction Index

Second Choice: Quality of Well-Being Scale

Third Choice: Mental Health Inventory

MEASURES OF QUALITY OF LIFE

First Choice: Life Satisfaction Index

Second Choice: Quality of Life Index

Third Choice: General Well-Being Schedule (Dupuy)

APPENDIX

MENTAL HEALTH QUESTIONS FOR THE HEALTH INTERVIEW SURVEY

Response Categories: Please describe yourself in relation to each of the following statements, ranking yourself on a scale of 1 to 5. Give yourself a 1 if the statement doesn't describe you at all, and a 5 if the statement describes you exactly. (Adapted from the Duke-UNC Health Profile)

DISORDERS

1. I have talked to a medical doctor about problems with my memory.
2. I don't get real satisfaction out of anything anymore.
3. Recently I have been able to sleep as well as usual.
4. I get annoyed or irritated more easily than I used to.
5. I have never considered myself to be a nervous person.
6. I have occasionally had a strong fear of something even though I knew there was no real danger.
1. There have been periods of a week or more when I was so happy or excited that my family or friends worried about me.
8. There have been times when I felt people were watching me or spying on me.
9. More than once I have heard things other people couldn't hear, such as a voice.
10. I get all the sympathy I should.
11. I am sure I'm being talked about.
12. At times I feel like smashing things.

SOURCES: MMPI-2, Diagnostic Interview Schedule-III, Beck Depression Inventory.

ANXIETY AND DEPRESSION

1. I often feel nervous.
2. I often feel calm.
3. During the past month, I have often felt rattled, upset, or frustrated.
4. During the past month, I have felt downhearted or blue.
5. During the past month, I have not been moody or brooded about things.
6. I do not feel like a failure in life.

SOURCES: State-Trait Anxiety Inventory, Beck Depression Inventory, Mental Health Inventory.

EMOTIONAL/BEHAVIORAL CONTROL/COPING

1. During the past month, I have cried a number of times.
2. I feel that things are going my way.
3. I get upset when someone criticizes me.

SOURCES: Mental Health Inventory, Affect Balance Scale.

POSITIVE WELL-BEING AND QUALITY OF LIFE

1. During the past month, I have felt that the future looks hopeful and promising.
2. During the past month, I have felt relaxed and free of tension much of the time.
3. During the past month, I have not really enjoyed the things that I have done.

4. During the past month, I have awakened feeling fresh and rested.

5. I am not happy or satisfied with my personal life.

6. As I look back on my life, I am fairly well satisfied.

7. This is the dreariest time of my life.

8. During the past month, I have often felt cheerful.

9. During the past month, I have been worried or concerned about my health.

10. During the past month, I have been feeling full of energy.

SOURCES: Life Satisfaction Index, General Well-Being Schedule, Mental Health Inventory.

SPECIALTY MENTAL HEALTH SERVICE UTILIZATION IN THE U.S.

Marilyn J. Henderson, Center for Mental Health Services

In his introductory remarks, Dr. Manderscheid presented a chart detailing the annual epidemiology of mental disorder. In this presentation, I will focus in more detail on the 23 million persons who receive treatment in a year, present some characteristics of persons seen in the specialty mental health sector, and discuss possible future directions in the collection of data on mental disorder.

Data from the Epidemiologic Catchment Area (ECA) show that persons receive mental health treatment in a variety of settings.[1] The ECA was completed with five research sites in of the National Institute of Mental Health to collect community based estimates of prevalence/incidence as well as service use for different mental disorders from 1980-1985. When comparing data on the treatment provided in four sectors: the Specialty Mental Health and Substance Abuse Sector, the General Medical Sector, the Human Services Sector, and the Voluntary Support Network, it is apparent that people are seen throughout all these sectors, not just in the traditional specialty sector. Overall, when overlap is accounted for, 43 percent of persons treated in a year received mental health services within the General Medical Sector, with 32 percent receiving care exclusively in this setting; 40 percent received services within the Specialty Sector; 20 percent received treatment within the Human Services Sector, and 28 percent received care within the Voluntary Support Network. Fifteen percent of persons received their mental health care exclusively within the voluntary network.

From these ECA data, it is clear that the mental health service system encompasses a wide variety of settings. Extremely limited information exists about mental health services in many of these settings. Our most comprehensive, detailed information about clients and services is collected through the National Reporting Program within the Center for Mental Health Services (CMHS) which has focused over time primarily on the traditional specialty mental health care setting, a subset of the Specialty Mental Health and Substance Abuse Sector. Although this setting no longer provides the bulk of care, it does provide fairly intensive services to those who receive treatment there.

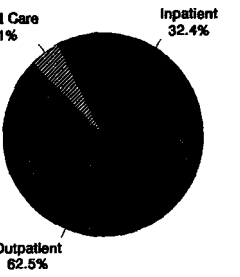
Now I will focus further on a few characteristics of the approximately 5.7 million persons seen within the organized specialty mental health sector. Data are from the CMHS 1986 Client/Patient Sample Survey of Inpatient, Outpatient, and Partial care programs. Programs included in this survey were all those that provided inpatient, outpatient, or partial care services within specialty mental health organizations identified through our biennial Inventory of Mental Health Organizations and General Hospital Psychiatric Services. Not included were services provided by private practitioners, specialty alcohol and drug abuse services, the general medical sector, human services sector, and the voluntary support network. We are looking at a small, specific piece of the pie.

In the 1986 survey, we obtained estimates of persons continuing care at a point in time (beginning of survey) as well as for those admitted during the year. Added together, these estimates provide us with an estimate of the total persons seen within each program type. Note that throughout this discussion, persons have not been unduplicated between the settings; persons receiving care in more than one setting are included in the totals for each setting.

If we look at the relative distribution of persons who received services by setting (chart 1), we see that outpatient programs provided care to the largest number of persons (over 62%). About one-third of persons were seen within inpatient programs, and a very small group, only 5 percent were seen within partial care programs. Over time, the specialty mental health system has increasingly moved toward care in ambulatory settings.

CHART 1

PERCENT OF PERSONS SEEN IN INPATIENT, OUTPATIENT, AND PARTIAL CARE PROGRAMS
SPECIALTY MENTAL HEALTH ORGANIZATIONS, 1986

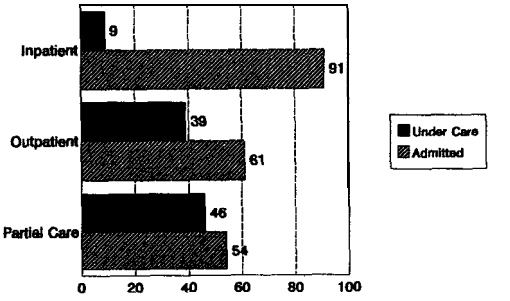


Source: CMHS Client/Patient Sample Survey of Inpatient, Outpatient, and Partial Care Programs

Inpatient, outpatient and partial care programs differed with respect to the percentages of their caseloads who were admitted during the year versus those in active treatment at the beginning of the year (chart 2). Whereas almost all of the persons seen in inpatient settings were admitted during the year (91%), in outpatient programs only 61 percent were admitted during the year, and in partial care programs, persons admitted and persons under care at the beginning of the year were about equally represented. With the increased emphasis on short inpatient stays within psychiatric hospitals, it is not surprising to find a smaller number of persons resident at any one time, and a larger number of persons moving in and out of the system. In ambulatory programs, caseloads are more balanced between new persons entering the program and ongoing clients.

CHART 2

PERCENT OF PERSONS UNDER CARE AND ADMITTED
SPECIALTY MENTAL HEALTH INPATIENT, OUTPATIENT, AND PARTIAL CARE PROGRAMS - 1986

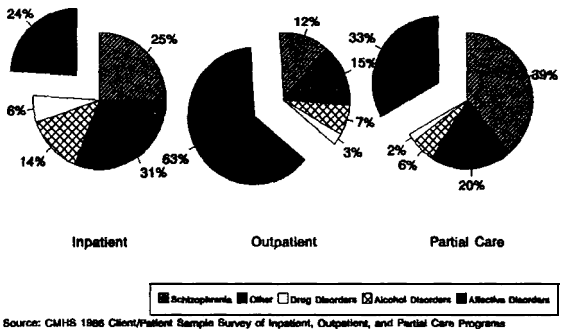


Source: CMHS 1986 Client/Patient Sample Survey of Inpatient, Outpatient, and Partial Care Programs

As could be expected, differences also existed in the diagnostic distributions for persons seen within the three service settings (chart 3). To illustrate differences among the settings, I have grouped principal diagnoses into four very broad diagnostic groupings: schizophrenia and related disorders, major affective disorders, alcohol related disorders, and drug related disorders. When comparing these broad groupings, differences among the settings are dramatic. Over three-fourths of the persons in inpatient programs received one of these diagnoses as their principal diagnosis; about two-thirds of those in partial care received these diagnoses; and only slightly over one-third received these diagnoses within outpatient settings. In outpatient settings, more focus was on less severe disorders such as adjustment disorders and social conditions. This points to very different groups of people being seen in each setting with respect to diagnosis.

CHART 3

PERCENT OF PERSONS WITH SELECTED PRINCIPAL DIAGNOSES SPECIALTY MENTAL HEALTH INPATIENT, OUTPATIENT, AND PARTIAL CARE PROGRAMS - 1986



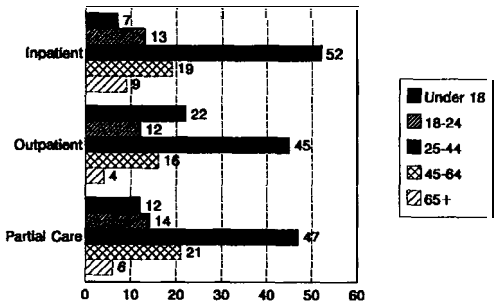
Source: CMHS 1986 Client/Patient Sample Survey of Inpatient, Outpatient, and Partial Care Programs

It is of interest to note the percentages of persons with principal diagnoses of alcohol and drug related disorders within the specialty mental health sector, which does not include separate programs specifically focused on substance abuse disorders. Alcohol and drug related disorders accounted for approximately 20 percent of all principal diagnoses within inpatient settings, 10 percent within outpatient settings, and 8 percent within partial care settings. Since we did not collect secondary or dual diagnoses in the 1986 survey, these estimates do not include clients who had co-occurring diagnoses of alcohol and drug related disorders. Hence, the total number of persons being treated with alcohol or drug related disorders in the specialty mental health sector represents a fairly substantial group, underscoring the need for emphasis on integrated services for persons with co-occurring disorders.

When comparing the relative age distributions within these settings (chart 4), we see that children and youth under 18 represented higher percentages of persons in ambulatory care settings than in inpatient settings, a finding consistent with our emphasis on providing services in the least restrictive settings, particularly for children and youth. Looking solely at these overall counts, however, obscures differences in the populations of clients seen by different types of facilities.

CHART 4

PERCENT OF PERSONS BY AGE GROUPINGS SPECIALTY MENTAL HEALTH INPATIENT, OUTPATIENT, AND PARTIAL CARE PROGRAMS - 1986



Source: CMHS 1986 Client/Patient Sample Survey of Inpatient, Outpatient, and Partial Care Programs

If we further look at age and diagnostic information by type of organization, we see a marked difference in the percentage of children and youth that were seen in private psychiatric hospitals when compared with the State and county public hospitals and nonFederal general hospitals with separate psychiatric services (table 1). This held for both inpatient and outpatient settings. Similarly, organization types differed considerably with respect to the percentages of persons receiving principal diagnoses of schizophrenia or affective disorders. Much higher percentages of persons were diagnosed with schizophrenia within State and county mental hospitals when compared with the other two facility types, whereas major affective disorders were much more prevalent in private psychiatric hospitals and nonFederal general hospitals.

TABLE 1

PERCENT OF PERSONS WITH SELECTED CHARACTERISTICS BY TYPE OF ORGANIZATION SPECIALTY MENTAL HEALTH INPATIENT AND OUTPATIENT PROGRAMS - 1986

Characteristic	State and County	Private	General Hospital
INPATIENT			
% Under 18	5	21	6
% Schizophrenia	42	12	19
% Affective	15	49	38
OUTPATIENT			
% Under 18	14	33	20
% Schizophrenia	29	6	13
% Affective	18	23	23

Source: CMHS 1986 Client/Patient Sample Survey of Inpatient, Outpatient, and Partial Care Programs

These are just a few of the types of comparisons that may be made with organization-based data from the National Reporting Program. We could go on to look further at other sociodemographic variables, at length of stay, and types of treatment from this survey. As noted earlier, however, these data cover only a specialized part of the mental health system. When trying to construct a total picture of what is happening within the entire mental health system, it is extremely difficult to apply a building block approach, pulling pieces of available information from the large range of organization based data systems. Many gaps exist in our data collection, with very little information available about the human service system, voluntary support network, or the criminal

justice and education systems.

Within the National Reporting Program of CMHS, we are trying to gain a better understanding of the mental health services provided in these nontraditional settings, by conducting surveys of organizations outside the traditional specialty sector and also outside of organized settings. In 1988, we conducted a survey of mental health services in State Correctional Institutions, and are currently going into the field with a survey of mental health services in local jails. This year we will be completing feasibility studies for conducting surveys of the juvenile justice system and also of organized consumer run self-help groups. These additional surveys will enable us, for the first time, to gain a better understanding of mental health service availability in these settings.

In order to follow a person's course of care throughout the entire system, however, we need to have the capacity to track the care of persons with mental disorder *across* all sectors where they receive care. As we have already seen, persons with mental disorder seek mental health care in a variety of settings. Our data needs span both epidemiology and service patterns.

Within the Center for Mental Health Services, we have been supporting a system called the Mental Health Statistics Improvement Program (MHSIP), a cooperative effort with the States, one purpose of which is to define and collect standard data elements for comparable data across the country. Data standards are in place for client, organization, and event reporting, and we have been supporting grants to the States to implement these standards, and to make use of data for oolicv decisions. Our effort has been to increase participation to include providers of mental health services other than States, and to include consumer input into our data collection.efforts.

Currently, the MHSIP data standards are organization-based. Over time, we have increasingly recognized the need to augment our current organization-based reporting system with a person-based system. It is clear that mental health services will continue to move out and encompass new settings, and that people will seek care within a varietv of innovative services in overlapping sectors: The future of data collection in the area of mental health should include both person-based and organization-based reporting. We are exploring the collection of person-based information so that we may look at patterns of care. We also see a need to continue the collection of specific organization-based data that provide program management and system-level information. Both of these together would increase our capacity to document the types of care persons with mental disorder receive regardless of where they obtain these services.

Reference

[1] Regier, D.A.; Narrow, W.E.; Rae, D.S.; Manderscheid, R.W.; Locke, B.Z.; and Goodwin, F.K. The de facto US mental and addictive disorder service system: Epidemiologic Catchment Area prospective 1-year prevalence rates of disorders and services. Archives of General Psychiatry 50:85-94, 1993.

Session **E**

Family and Health

Toward the year

2000

Refining the measures

PARENTAL FACTORS INFLUENCING PATTERNS OF PRENATAL CARE UTILIZATION

Peter T. D'Ascoli, University of Minnesota
Greg R. Alexander, Donna J Petersen, Michael D. Kogan

Introduction:

In our society, prenatal care is provided as the major preventative measure to assure healthy pregnancy outcomes. The utilization of prenatal care has been associated with several health benefits, including diminished low birth weight (LBW) percentages. While the content of prenatal care and the recommended number of prenatal care visits are currently being debated and are in an ongoing state of development, it is generally accepted that obtaining prenatal care is one of the health behaviors necessary to obtain a healthy pregnancy outcome. Further, for many primigravid young women, prenatal care may represent their first adult contact with the health care system and their satisfaction with and knowledge gained from that interaction may influence their and their families' subsequent use of preventive health care services.

Obtaining prenatal care can be viewed as an individual health-seeking behavior. In a systems approach, this view may be too narrow. There is evidence that geography, knowledge, attitudes economics, and health insurance, which is based on employment, all play a role in obtaining prenatal care. Family structure and function may also influence these relationships as well as having a direct effect on a woman's prenatal care seeking behavior. While the role single marital status plays in prenatal care acquisition has been studied, the role of the father in prenatal care and pregnancy outcome has had limited review.

In this paper, we will look at the influence that socio-demographic characteristics of the father and mother have on utilizing prenatal care in Minnesota. Our specific area of inquiry is the impact that marital status and the mother's and father's education have on both the initiation and adequate use of prenatal care services. Recognizingthat differences in family makeup and educational level may represent variations in income, insurance, social support and attitudes related to family planning and health care seeking behavior, we postulate that these socio-economic and demographic indicators will be positively related to earlier and more adequate utilization of prenatal care.

Data:

The data set utilized for this investigation was obtained from the live birth file of the State of Minnesota for the years 1990 through 1991. Single live births to resident mothers were selected: 131,845 births were available for analysis.

Educational attainment of mothers and fathers was defined as follows: For adults (18+ years of age): low (<12 years of education), average (12 years), high (>12 years); for adolescents (<18 years of age): low (2+ years below expected grade level), average (within 1 year of expected grade level), high (2+ years above expected grade level).

High parity for age was defined as 1 or more previous births for adolescents (<18 years), 3 or more previous births for mothers 18-21 years, 4 or more previous births for mothers 22-24 years and 5 or more previous births for all other mothers. Adequacy of prenatal care utilization was defined according to the index proposed by Alexander and Cornely. Reported race of mother was used to define ethnicity (Alexander and Cornely, Am J Prevent Med 1987; 3:243-253.)

Multiple logistic regression was employed to calculate odds ratios for the independent effects of parental factors on various measure of prenatal care initiation and adequacy.

TABLE 1

SOCIO-DEMOGRAPHIC CHARACTERISTICS
1990-1991 SINGLE LIVE BIRTHS
TO MINNESOTA RESIDENT MOTHERS
N=131,845

MATERNAL CHARACTERISTICS	%	(#)
RACE OF MOTHER		
WHITE	90.5	(119,308)
BLACK	4.0	(5,215)
AMERICAN INDIAN	2.0	(2,620)
OTHER	3.6	(4,702)
MARITAL STATUS		
MARRIED	78.5	(103,256)
NOT MARRIED	21.7	(28,589)
AGE OF MOTHER		
<18 YRS	2.7	(3,550)
18-29 YRS	62.1	(81,795)
30+ YRS	35.3	(46,474)
EDUCATIONAL ATTAINMENT*		
LOW	9.0	(11,331)
AVERAGE	38.3	(48,367)
HIGH	52.7	(66,607)
PARITY		
PRIMIPARA	61.2	(80,031)
MULTIPARA	38.2	(50,818)
HIGH PARITY FOR AGE**	2.7	(3,623)

Findings:

Socio-demographic characteristics of the study population are displayed in Table 1. Infants of white mothers

Being over 30, a primipara and having high educational attainment are positively associated with early prenatal care initiation. Single teens and adults, women with a high parity or low educational attainment, as well as the identified minority groups, are at a greater risk of initiating prenatal care late. As a note of caution the logistic regression detailed in columns one and two cannot be directly compared to the remaining regression in column three.

TABLE 4

Odds Ratios and Confidence Intervals from Logistic Regression Analysis of Receipt and Adequacy of Prenatal Care Indicators

1990-1991 Single Live Births to Minnesota Resident Mothers

Maternal Characteristics	1st Trimester <Adequate vs Adequate	2nd Trimester Inadequate vs Intermediate	No Care vs Any Prenatal Care
Teen (<18 yrs), Not Married	1.59 (1.34-1.89)	1.80 (1.33-2.44)	9.58 (6.60-13.90)
Teen (<18 yrs), Married	2.40 (1.56-3.69)	1.60 (0.76-3.40)	1.13 (0.15-8.27)
Adult (18+ yrs), Not Married	1.57 (1.47-1.68)	1.63 (1.39-1.91)	9.92 (7.86-12.52)
Mother 30+ Yrs of Age	0.92 (0.88-0.97)	0.88 (0.74-1.04)	1.06 (0.85-1.32)
Primipara	0.65 (0.62-0.69)	0.52 (0.44-0.61)	0.62 (0.51-0.76)
High Parity for Age	1.97 (1.69-2.29)	1.57 (1.26-1.96)	3.37 (2.62-4.34)
Low Educational Attainment	1.37 (1.26-1.50)	1.63 (1.40-1.91)	1.45 (1.20-1.76)
High Educational Attainment	0.80 (0.76-0.84)	0.75 (0.63-0.90)	0.73 (0.58-0.93)
Black	1.65 (1.45-1.87)	2.37 (1.95-2.88)	2.33 (1.87-2.90)
American Indian	2.78 (2.40-3.22)	3.63 (2.94-4.50)	1.74 (1.30-2.32)
Other Race	1.85 (1.60-2.15)	2.52 (1.99-3.19)	1.84 (1.24-2.74)
Number	88,080	14,631	115,737

Table 4 presents the results of the analysis of adequacy of care within trimesters that prenatal care began against the same maternal characteristics as those considered in Table 3. Among those who initiated prenatal care in the first trimester, there is an increased risk of less than adequate care use for not married, teenaged, high parity, low education and minority mothers. These relationships were also found for those who started care in the second trimester and received inadequate prenatal care. Non-significant odds ratios are shaded.

Also provided in Table 4 is the results of the analysis of the predictors of receiving no care versus any prenatal care. Not being married was found to have a markedly high risk of receiving no prenatal care, regardless of age.

Parental Characteristics:

In order to examine the impact of parental characteristics on prenatal care, a subset of singleton births to white mothers 21 years of age or older, were selected. Other ethnic groups were excluded due to a lack of sufficient cases for meaningful analysis. The results of the logistic regressions, focusing on initiating and adequacy of prenatal care use, are presented in Table 5. The

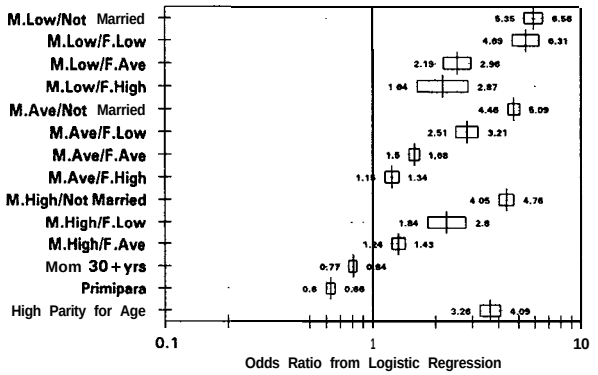
Odds Ratios and Confidence Intervals from Logistic Regression Analysis of Timing of Initiation and Adequacy of Prenatal Care 1990-1991 Single Live Births to Minnesota Resident White Mothers, 21+ Years of Age

Parental Characteristics	No Care vs Any Prenatal Care	3rd vs 2nd vs 1st Trimester Initiation	1st Trimester, <Adequate vs Adequate
Not Married; Educ: Mom Low	13.63 (8.03-23.12)	5.93 (5.35-6.56)	2.55 (2.14-3.05)
Married; Educ: Mom Low, Dad Low	3.19 (1.11-9.19)	5.44 (4.69-6.31)	2.19 (1.67-2.89)
Married; Educ: Mom Low, Dad Ave	3.07 (1.07-8.74)	2.55 (2.19-2.96)	1.57 (1.24-1.99)
Married; Educ: Mom Low, Dad Hi	2.96 (0.40-21.75)	2.17 (1.64-2.87)	1.33 (0.85-2.07)
Not Married; Educ: Mom Average	15.88 (10.57-23.84)	4.76 (4.46-5.09)	1.97 (1.76-2.20)
Married; Educ: Mom Ave, Dad Low	1.80 (0.55-5.89)	2.84 (2.51-3.21)	1.72 (1.42-2.08)
Married; Educ: Mom Ave, Dad Ave	1.64 (1.00-2.69)	1.59 (1.50-1.68)	1.32 (1.23-1.42)
Married; Educ: Mom Ave, Dad Hi	1.40 (0.71-2.75)	1.24 (1.15-1.34)	1.15 (1.04-1.27)
Not Married; Educ: Mom Hi	11.95 (7.38-19.35)	4.39 (4.05-4.76)	1.56 (1.35-1.81)
Married; Educ: Mom Hi, Dad Low	1.86 (0.25-13.64)	2.26 (1.84-2.80)	1.49 (1.08-2.06)
Married; Educ: Mom Hi, Dad Ave	1.34 (0.69-2.64)	1.33 (1.24-1.43)	1.23 (1.12-1.35)
Mom 30+ Years of Age	1.21 (0.91-1.61)	0.81 (0.77-0.84)	0.96 (0.91-1.02)
Primipara	0.62 (0.45-0.85)	0.63 (0.60-0.66)	0.66 (0.62-0.70)
High Parity for Age+	7.85 (5.24-11.76)	3.65 (3.26-4.09)	1.91 (1.56-2.33)

reference group was married, multiparous women with average parity for age, 21-29 years of age, with both mother and father having 13 years or more completed education.

High parity for age women, and not married of any educational level were at markedly higher risk of receiving no prenatal care. Married women of low educational attainment with a husband with low or average educational attainment were also at risk of receiving no prenatal care, but at lower odds ratio. The receipt of no prenatal care is relatively rare in this population, evidenced by the broad confidence intervals of the odds ratios.

Odds Ratios and Confidence Intervals
Trimester Care Began (3rd vs 2nd vs 1st)
1990-91 Minnesota Single Live Births



White Mothers, 21+ Years of Age

represent over 90% of the study group. Over three-fourths of the infants were to married mothers. Only 2.7% of the mothers were adolescents (<18 years of age), and over one-third were 30 years of age or older.

The level of educational attainment among mothers was relatively high with more than half having some college education. Almost 40% of the births were to primiparous women and nearly 3% were to women with a relatively high number of previous births for their age.

Table 2 depicts the number and percentage of births by prenatal care use categories. These categories indicate the trimester prenatal care began and the adequacy of the utilization. The adequacy of utilization takes into account the number of prenatal care visits adn the gestational age at delivery. Over 60% of the women started care in the first trimester and had adequate utilizaton of prenatal care. Less than one percent received no prenatal care and for 16% of the study population, prenatal care could not be categorized due to missing data.

The overall LBW percentage of this study population was 4.19%. There is a progression of LBW from 3% in the adequate group, to 5% in the intermediate group, to 6% in the inadequate group, to better than 21% in those who received no prenatal care. Those with missing data had 7% LBW. (not shown)

TABLE 2

PERCENT LW BIRTH WEIGHT BY TRIMESTER PRENATAL CARE BEGAN AND ADEQUACY OF UTILIZATION*
1990-1991 SINGLE LIVE BIRTHS
TO MINNESOTA RESIDENT MOTHERS

TRIMESTER OF INITIATION AND ADEQUACY OF UTILIZATION	% LOW BIRTH WEIGHT	DISTRIBUTION % (#)
1ST TRIMESTER PRENATAL CARE		
ADEQUATE	2.96	62.93 (82,900)
INTERMEDIATE	6.95	5.65 (7,439)
INADEQUATE	7.75	0.44
2ND TRIMESTER PRENATAL CARE		
INTERMEDIATE	4.08	11.60 (15,286)
INADEQUATE	11.70	0.88 (1,154)
3RD TRIMESTER PRENATAL CARE		
INADEQUATE	3.78	2.43 (3,204)
NO PRENATAL CARE	21.43	0.49 (644)
MISSING	7.20	15.58 (20,530)

* ADEQUACY OF PRENATAL CARE UTILIZATION BASED ON INDEX BY ALEXANDER AND CORNELY, AM J PREVENT HEO 1987; 3:243-253.

Table 2 also provides low birth weight (LBW) percentages by prenatal care categories. Among those who started prenatal care in the first trimester, there was 2.96% LBW among those who were in the adequate group. However, for those who received intermediate or inadequate care, the percentage of LBW was 6.95 for intermediate and 7.75 for inadequate, more than twice that of the adequate group. Among those who began prenatal care in the second trimester the LBW percentage of the inadequate group (11.70%) was nearly three times that of the intermediate group (4.08%). (By definition in this index, starting prenatal care in the second trimester can not be adequate.)

It is noteworthy that the LBW percentage for women initiating prenatal care in the third trimester is 3.78, which is below the Year 2000 Objective of 5%.

Mothers who start care in the first or second trimester, but get less than adequate care, as well as the no prenatal care and the missing data categories, represent the groups at highest risk of LBW. The "less than adequate" groups that initiated care in the first trimester are of particular interest as they entered the health care system early but failed to maintain regular contact.

TABLE 3

ODDS RATIOS AND CONFIDENCE INTERVALS FROM LOGISTIC REGRESSION ANALYSIS OF INITIATION OF PRENATAL CARE INDICATORS

1990-1991 SINGLE LIVE BIRTHS TO MINNESOTA RESIDENT MOTHERS

MATERNAL CHARACTERISTICS	2ND VS 1ST TRIMESTER	3RD VS 2ND TRIMESTER	3RD VS 2ND VS 1ST TRIMESTER
TEEN (<18 YRS), NOT MARRIED	3.47 (3.15-3.81)	1.52 (1.29-1.80)	3.61 (3.31-3.93)
TEEN (<18 YRS), MARRIED	2.18 (1.64-2.91)	0.91 (0.54-1.51)	1.99 (1.53-2.58)
ADULT (18+ YRS), NOT MARRIED	2.75 (2.63-2.87)	1.65 (1.51-1.81)	2.98 (2.86-3.10)
MOTHER 30+ YRS OF AGE	0.76 (0.73-0.79)	0.80 (0.72-0.88)	0.74 (0.71-0.76)
PRIMIPARA	0.71 (0.69-0.74)	0.89 (0.81-0.97)	0.70 (0.68-0.73)
HIGH PARITY FOR AGE	2.49 (2.25-2.74)	1.81 (1.58-2.08)	2.88 (2.65-3.13)
LOW EDUCATIONAL ATTAINMENT	1.41 (1.33-1.49)	1.26 (1.15-1.39)	1.50 (1.43-1.57)
HIGH EDUCATIONAL ATTAINMENT	0.68 (0.65-0.71)	0.88 (0.80-0.97)	0.67 (0.65-0.69)
BLACK	1.73 (1.60-1.86)	1.62 (1.43-1.82)	1.97 (1.84-2.11)
AMERICAN INDIAN	1.57 (1.41-1.73)	1.87 (1.62-2.16)	1.95 (1.79-2.13)
OTHER RACE	2.53 (2.31-2.77)	2.02 (1.75-2.33)	3.00 (2.76-3.25)
NUMBER	111,718	20,837	115,146

Maternal Characteristics:

Logistic regression was used to examine the association of maternal characteristics on three measures of initiation of prenatal care. This is seen in Table 3. The comparison group is white, adult (18-29 years), multiparous mothers with average parity for age and average educational attainment (high school graduates). The first column is a comparison between initiating care in the second versus the first trimester. The second column 3rd versus 2nd trimester initiation. The last column is a trichotamous analysis of all three trimesters, which compare the 3rd to the 2nd and the 1st trimesters and the 1st to the 2nd and 3rd.

Using the previously employed trichotomous outcome indicator of trimester of prenatal care initiation, a distinct pattern to the odds ratio emerges. For each maternal education strata, an increased risk of delayed initiation of prenatal care is observed for lower paternal educational attainment. Within each maternal education strata not married mothers demonstrate the highest risk of late prenatal care initiation. This pattern can be more clearly observed in Figure 1. Among the marital status and parental education groups, not married women of low educational attainment exhibit the highest risk of delayed care. Holding mother's education constant, the risk of delayed initiation of prenatal care decreases for married women as fathers education increases. A similar association with marital status and parental education was found in the analysis of adequacy of the prenatal care, given a first trimester initiation of prenatal. This is again seen in Table 5.

Discussion:

As has been previously noted, variations in the utilization of prenatal care are associated with differences in infant birth weight. In this study, those women receiving adequate care exhibited the lowest proportion of low birth weight infants while those receiving no prenatal care experienced the highest percentage. Further, this study demonstrates that maternal characteristics affectthetiming of initiation and adequacy of prenatal care utilization. Clearly, the highest risk for receiving no prenatal care was among unmarried women while the lowest risk for inadequate prenatal care was found among primiparous women and those with a higher level of educational attainment.

The examination of the effects of father's education among married women on the adequacy and initiation of prenatal care provides an important starting point for discussion, particularly when compared with the same data for unmarried women. Among white adult Minnesota resident mothers, higher educational attainment of the fathers as well as the mothers seems important for the early initiation of care and attendance to a recommended schedule of visits. As a father's education declines, so too does utilization of prenatal care. In the absence of a marital partner, prenatal care utilization becomes lower still.

The reasons for this phenomenon can only be speculated upon here. Possible considerations include:
1) a better educated husband may command a higher income and have access to more comprehensive health insurance benefits: 2) a better educated husband may be more knowledgeable about the need for ongoing medical care and may encourage its use: 3) a better educated husband may have participated in the planning of the pregnancy and may have other resources generated out of personally motivated interest to devote to assuring a healthy pregnancy: 4) a better educated husband may be a more supportive partner in general, providing the social support considered important in promoting a positive pregnancy outcome: 5) single women may well experience a disproportionate number of unplanned or unwanted pregnancies resulting in a delay in initiation of care because the pregnancy was not recognized or because of the ambivalence surrounding the pregnancy.

An area for further investigation is the role of paternal education in the situation of cohabitation out of wedlock. Additionally, as this study's findings are based on largely a white population, further investigation of these relationships are needed in other ethnic groups.

FAMILY AND CULTURAL FACTORS IN MENTAL AND PHYSICAL COMORBIDITY IN CHILDREN

Ronald J. Angel, The University of Texas at Austin
Jacqueline L. Angel

A growing body of epidemiological data based on structured instruments such as the Diagnostic Interview Schedule for Children (DISC) and the Child Behavior Checklist (CBCL) reveals relatively high rates of childhood psychopathology in community samples.[1,2,3,4,5,6] In a review of the findings of recent studies, Costello reports rates of significant psychopathology ranging from 17.6% to 228.' Of particular importance to our present purpose is that fact that the evidence also shows that rates of childhood psychopathology are influenced by various social class factors.[8,9] Because of the fact that minority group children are often exposed to serious psychological stresses associated with poverty, they are at elevated risk of emotional and behavioral problems. The evidence suggests, therefore, that there is a great deal of unmet need for mental health care services among children in general, and among poor minority group children in particular.

Although we do not have accurate estimates of mental health care use by poor minority children, data for the general population suggest that most children in need of services never see a mental health professional. As is the case for adults, most children who receive any care at all for mental health problems are seen exclusively bv general practitioners, primarily pediatricians.[10,11,12,13] Extrapolating from data for the general population presented by Regier et al.[14] Costello estimates that among children with a mental disorder only one in five receives specialist treatment, while three receive care from a pediatrician, and one receives no treatment at all.[10] Offord et al.[15] come to a similar conclusion. These figures dramatically emphasize the fact that, as is the case for adults, general medical care providers are the only contact that the majority of children with mental health problems have with the medical care system. For poor, minority group children, among whom rates of psychopathology are likely to be high, even this contact may be rare.

Because they are often poor, and consequently more likely to be exposed to the health risks of poverty, black and Hispanic children in the United States are at particularly high risk of poor health. To make matters worse, the children of the poor often do not receive the same quality of health care as the children of the middle class." In addition to poverty black and Hispanic children are more likely to live in single-parent families.[17] A growing body of research' indicates that children who grow up without a father often experience serious developmental problems, including disturbed sex-role identities and antisocial behavior.[18,19,20,21,22]

Children also experience major health problems as a consequence of divorce itself.[23,24] Clearly, most single-parent families are not dysfunctional nor do all children suffer irreparable damage as the result of divorce.[25] Nonetheless, the economic and social stresses that many female-headed families experience can place a child at increased risk of both mental and physical illness.[22,26] The children of poor single mothers face a double jeopardy in terms of the risk of emotional and behavioral problems.

In this Paper we examine patterns of co-occurrence of physical and mental illness among children and adolescents. In addition to physical and mental health problems we examine the use of both specialty mental health and general medical care because of the possibility that children with sub-clinical, but nonetheless significant emotional and behavioral problems, are more likely than children without such problems to be taken to the doctor. As is the case for adults, many children who are taken to the doctor for physical problems such as recurrent abdominal pain, may be suffering more from emotional than physical problems. Our major focus is on the impact of race and Hispanic ethnicity, as well as family structure, on patterns of mental and physical illness and medical care use.

Single Motherhood and Childhood Psychopathology

The identification of mental health problems among children is a difficult task primarily since there is little agreement as to what psychopathology means in childhood. Gutterman, O'Brien, and Young note that "[m]any diagnostic categories for children have a provisional quality, given rudimentary knowledge regarding the prognosis, prevalence, and etiologies of these diagnoses ."[27(p.625)] Happily, substantial progress in the development of structured diagnostic instruments for the estimate of the prevalence of specific emotional and behavioral problems in the community has been made in recent years [7,28,29]

As of yet, however, there are no large scale studies of the role of family structure on the mental health of children in different racial and ethnic groups that use these new highly structured diagnostic instruments. As a consequence, the little that we know of the role of family structure, race, and ethnicity on the mental health of children and adolescents is based primarily on specialized studies that employ symptom checklists dealing with developmental, behavioral, and emotional problems.[30,31,32,33] Although such data do not allow us to make estimates of the prevalence of specific mental health problems in the community they do provide important information on the role of family structure in childhood psychopathology.

In the following analyses we rely on parents' reports of emotional and behavioral problems in their children. We justify the use of such non-clinical assessments on the grounds that maladaptive behaviors and emotional problems that are not serious enough to qualify as full-fledged clinical entities can still cause serious life-long problems for a child. Relatively few children suffer from the most serious mental illnesses, and psychological problems in childhood and adolescence often manifest themselves as emotional and conduct disorders.[34] Difficulties with parents, teachers, and other social authorities can have serious long-term consequences. Children whose behavioral and emotional problems interfere with their education and social development are at high risk of delinquency and of the long-term consequences of underachievement. Data on the adult consequences of childhood maladjustment show that, although most troubled adolescents grow up to be normal adults, the majority of adults with serious adjustment problems were troubled adolescents.[35]

Before proceeding we should note that the phenomenon we are dealing with is related to what in the clinical literature is referred to as somatization disorder, in which emotional problems are manifested physically.[36] However, we are interested in a far more general phenomenon than the formal diagnostic category of somatization disorder. Although full-blown

somatization disorder is rather rare, the expression of psychological distress as physical symptomatology, or as generally diminished physical well-being, is much more common. This is because individuals do not experience their mental and physical selves separately; subjectively, distress is expressed through multiple channels and problems in any one area of life can easily manifest themselves in others.[37]

Since we rely on parental reports of children's mental and physical health status we should also caution the reader that one of the major problems with the use of third party reports to assess psychopathology or any other aspect of an individual's functioning is that such reports are subject to error from various sources. For example, it has been shown that a mother's perception of her child's physical health is influenced by her own psychological state." Mechanic" found that mothers who were under stress reported more illness for themselves and for their children than mothers who were not under stress.

Many single mothers are under great economic and emotional stress since they must often raise their children with little help. The need to work, at the same time that she must be both mother and father to her children, can lead to what has been termed "role overload."" It is easy to see how such a situation might influence a mother's perception of her child's health. Of course, the same stress that affects a mother's perceptions can lead to very real decrements in her child's health so that it is very difficult to disentangle the mother's own psychological state from the child's actual state on the basis of the mother's reports alone. Nonetheless, parental reports have been shown to be largely reliable sources of data on a child's mental health status and because they are widely used in the literature we rely on them here.[28]

As we stated earlier, one of our major concerns in this paper is to determine the impact of a mother's marital status on her children's mental health. However, single mothers may well be more different than they are similar. They differ for example in age, education, and earnings capacity.[40] One other aspect in which they differ, previous marital status, has important implications for children's health. There is a substantial body of literature that shows that children who experience their parent's divorce are particularly vulnerable to behavioral and emotional problems.[18,19,21,22,23] Evidently, the trauma and family turmoil that precedes the separation, in conjunction with the dramatic changes that accompany their parent's divorce cause children great emotional trauma.

On the other hand, the children of never-married mothers, many of whom are young and have few job skills, are often subjected to poverty and a disorganized social environment throughout their young lives. Since black, Hispanic, and non-Hispanic white single mothers differ greatly in previous marital status their children also differ in the sorts of physical and mental health risks to which they are exposed. Although we cannot definitively answer the question of the impact of previous marital status on children's mental health we examine the issue to some extent in the following analyses.

Data

Our analysis is based on the 1988 child health supplement to the National Health Interview Survey (NHIS), a nationally representative sample of the civilian, noninstitutionalized population under 18 years old.[9] In the child health (NHIS-CH) supplement, one child was randomly selected from each family contacted as part of the basic survey. The total NHIS-CH sample includes 17,110 children. We employ a subset of 15,259 black, Hispanic, and non-Hispanic white children that lived with both parents or their mother only. Because of the small number of Puerto Ricans and Other Hispanics our findings concerning these groups are only tentative. Because of the absolute lack of information on the role of family structure on the health of these children, however, we retain them in those parts of the analysis in which their numbers are sufficient to justify an examination of the net impact of Hispanic ethnicity.

Information on the child's health was provided by the mother or some other household member. The respondent provides information on the child's health, behavioral problems, health service utilization, and family characteristics. Race and Hispanic ethnicity are determined on the basis of the respondents self-report. In the following analyses we examine four outcomes: (1) whether the child has ever seen a psychiatrist, doctor, or counselor about any emotional, mental, or behavior problem (this question is asked only for children three years and older); (2) whether the child has ever had an emotional or behavioral problem that lasted three months or more; (3) whether the child has seen a doctor at least once during the past 12 months; and (4) whether or not the child has manifested one of a number of behavioral problems during the previous three months.[41]

Our definition of family type is based on the marital status of the child's mother and we compare two-parent families to mother only families. These families can contain any number of other relatives. In our multivariate analyses, therefore, we control for total family size.

Results

Preliminary analyses of the data (not shown) reveal three distinct dimensions of economic disadvantage: race, specific Hispanic ethnicity, and headship status. Regardless of race or ethnicity, a much larger proportion of female-headed than two-parent households have low incomes. Among blacks, Puerto Ricans, and Mexican Americans fewer than half of single mothers are employed. Previous research shows that a large fraction of these families depend on welfare." Even when they are employed, though, female-headed families are at a distinct economic disadvantage for two reasons: women earn less on average than men and in couple-headed households there are two potential earners.

Table 1 shows the proportion of children who have ever had a mental health visit broken down by their mothers present or previous marital status for Mexican American, black, and non-Hispanic white children. The number of Puerto Rican and Other Hispanic children is insufficient for a detailed breakdown. These data reveal some rather large differences in the proportion of children with mental health visits for each racial and ethnic group. In all three groups, children whose parents have divorced are more likely to have had a mental health visit than children in intact homes or those whose mothers have never married. For Mexican American and non-Hispanic white children this is also the case for those whose parents have separated.

These data, then, suggest that the children of divorced and separated parents are also more likely to have emotional problems. These findings corroborate more qualitative research that shows that the family conflict and emotional strain that surround divorce cause mental health problems in children.[27] These data also indicate that there are substantial group differences in the probabilities of mental health visits and emotional problems as a function of the mother's previous marital status and race and Hispanic ethnicity. The impact of previous marital status on psychopathology and mental health care utilization of children in different racial and ethnic groups clearly warrants further study. In

the United States race and Hispanic ethnicity are correlated with socioeconomic status and differences in material resources, as well as differences in culture and social class may translate into different mental health risks for children.

In Table 2 we examine patterns of physical comorbidity among young school-aged children with behavioral problems. In this table those children who often experience behavior problems are compared to those who experience them only sometimes or never. Preliminary analyses revealed that, as one might expect, a large fraction (three-quarters) of children occasionally act up, but that far fewer do so often. It is clearly those who manifest serious behavior problems and who do so frequently who are at highest risk of long-term maladjustment.

This table shows that, as one would expect, a larger proportion of those children who often experience behavior problems have seen a psychotherapist compared to children who never or only sometimes act up. Additionally, a higher proportion of non-Hispanic white children who frequently manifest behavior problems have seen a physician in the last year. At least for non-Hispanic whites, then, children with behavior problems are more likely than those without such problems to see a general practitioner. The last row of Table 2 shows that for all three groups a larger proportion of children with behavior problems also have emotional problems that have lasted 3 months or more. This table also reveals large differences between non-Hispanic whites, blacks, and Mexican Americans. A larger fraction of non-Hispanic white than black or Mexican American children who often manifest behavioral problems have either seen a psychotherapist or a general medical doctor.

The information on physical comorbidity in Table 2 offers some insights. It appears, in general, that children who often manifest behavior problems are more likely than children who do not to experience those physical illnesses that can have a psychogenic etiology. For example, among children who often manifest behavior problems a slightly higher fraction have had asthma. This difference is statistically significant only for non-Hispanic whites, however. There is also a tendency for children with behavior problems to have experienced food and respiratory allergies. Finally, for all groups a higher proportion of children with behavior problems have had ear infections, a diagnosis that is frequently made in pediatric practice. In addition, as we might expect, a larger proportion of children with behavior problems have experienced headaches and enuresis, both physical conditions with clear emotional components. Again, there are clear differences in rats of these illnesses and conditions between the three groups, perhaps reflecting group differences in the tendency to seek medical care. we might note that, by and large, the same patterns are evident among the younger children in Table 2 emerge among adolescents.

These results suggest that there is a great deal of emotional and behavioral comorbidity associated with physical illness among children and adolescents. Although our data do not allow us to examine associations between particular mental health problems and specific physical conditions, they suggest that children with behavioral problems also experience physical illnesses that can have a substantial psychogenic etiology, such as asthma, headaches, and enuresis. Many of these children are no doubt brought to the doctor for the physical problem and if they receive specialty mental health care do so as the result of a referral. The literature we reviewed earlier, though, suggests that most receive no treatment for their emotional problems or receive it only from the general practitioner.

In Table 3 we examine the association between family structure and the various physical, emotional, and behavioral problems we have been discussing. Since we cannot continue our disaggregation, in this table we simply present the proportions of children with each problem by family headship for the various racial and ethnic groups. This table reveals that a larger proportion of children in female-headed households often suffer emotional and behavioral problems. As in the previous tables there are large differences between groups that may be partially the result of differing sample sizes. Although single motherhood is stressful for all racial and ethnic groups it interacts with other factors, such as income and minority group status, to lead to different emotional and behavioral outcomes.

Table 3 also shows that, in general, a larger Proportion of children in female-headed families have seen a psychotherapist. On the other hand, except for Mexican Americans, there are few differences in the proportion of children who have seen a general practitioner during the

Table 1: Percent of Children with a Mental Health Visit or Emotional Problem by Race/Ethnicity and Mother's Marital Status
(Weighted Percentages; Weighted N's in Parentheses)

Race/Hispanic Ethnicity	Mental Health Visit		Emotional Problem	
Mexican American				
Never Married	0.0	(30)	8.0	(36)
Divorced	12.9	(49)	13.2	(57)
Separated		(25)	1.5	(27)
Widowed	2.0	(10)	0.0	(10)
Married		(438)	4.7	(493)
Black				
Never Married	3.5	(441)	6.0	(498)
Divorced	6.3	(263)	4.9	(301)
Separated	3.8	(164)	3.7	(179)
Widowed	3.2	(44)	3.2	(47)
Married	3.4	(721)	3.0	(796)
Non-Hispanic White				
Never Married	5.7	(131)	4.4	(149)
Divorced	14.6	(706)	14.6	(918)
Separated	16.1	(202)	14.2	(253)
Widowed	5.3	(57)	14.6	(71)
Married	5.1	(6,139)	4.9	(7,046)

past year. Although the previous tables revealed differences in physical morbidity among children with and without behavioral problems, there are no striking differences in physical morbidity between children in two-parent and female-headed families. Evidently, then, although a larger proportion of children in female-headed families have behavioral and emotional problems than children in two-parent families, these differences are not large enough to result in discernable differences in physical comorbidity in the aggregate.

The data in these tables, then, suggest that among children and adolescents emotional and physical problems tend to occur together. There are several possible explanations for this phenomenon. Since we are relying on proxy responses, those of the parent, reports concerning the child's behavior and physical health may be influenced by the parent's own emotional state. On the other hand, those parents who seek specialty mental health care for their child may also be more likely to use general medical care. Finally, some children with emotional and behavioral problems may be referred from the general medical care provider to a specialty mental health care provider.

Finally, we sought to examine the multivariate correlates of general medical care use and of specialty mental health care use (table not shown). These analyses had two objectives, first, to determine whether race, ethnicity, and single motherhood have independent effects on the probability of seeking care and, the second, to determine which of the physical conditions listed in Table 2 *are* associated with physical or mental health visits. Since single motherhood, race ethnicity, and even physical conditions are associated with other economic and health-related factors these models control for such health and utilization-related factors as family income, health insurance, mother's education, the child's birth order, and family size. They also control for the mother's marital status and whether or not the child manifests behavioral problems often.

The results of these analyses revealed that the children of single mothers have more medical visits per year than the children of married mothers. They also showed that Black and Mexican American children have fewer doctors visits per year than non-Hispanic white children. As one would expect, most of the physical morbidity factors are highly significant predictors of physician's visits. What is of particular interest, though, is that behavior problems are highly significant predictors of general physician visits even after all other factors are taken into account.

Our analysis of the probability of a mental health visit showed that children in female-headed households were much more likely than children in two-parent households to have had a mental health visit. Once again, black and Mexican American children were less likely than non-Hispanic white children to have had a mental health visit. Again, as one would expect, having behavior problems greatly increases the probability of a mental health visit, but in addition, having respiratory problems, ear infections, or enuresis also increases the probability of a mental health visit.

Discussion

Clearlv. we would prefer to be able to examine the &-occurrence of specific diagnostic categories and particular medical conditions among these various racial and ethnic groups using instruments such as the DISC. As of yet, there are no large scale surveys that allow this sort of detailed examination for the children of single parents among different racial and ethnic groups. Hopefully this situation will be remedied in the near future. Our findings, therefore are only suggestive, but they add to a growing body of evidence that mental and physical health problems tend to occur together for children as well as for adults. They also reveal that the children of single mothers are at elevated risk of emotional and behavioral

Table 2: Childhood Physical and Mental Comorbidity and Mental Health Service Utilization by Race and Hispanic Ethnicity

	6 Months - 11 Years					
	Behavioral Problems					
	Non-Hispanic White		Black		Mexican American	
	Often[1]	Sometimes/ Never[2]	Often[1]	Sometimes! Never'	often'	Sometimes/ Never[2]
Ever Seen-Psychotherapist'	8.1"	3.7	4.4'	2.6	4.5	3.4
Past 12 Months-Doctor Visit	83.8'	11.4	73.5	67.0	66.6	66.4
Asthma	9.5"	5.6	9.1	6.9	8.0	5.0
Ear Infections	36.8"	29.8	28.1**	12.6	31.4'	19.7
Food Allergy	9.3"	6.9	5.1	2.6	4.6	3.5
Respiratory Allergy	16.1'	13.1	9.0	6.1	5.4	5.0
Pneumonia	9.7	1.8	8.2'	4.5	5.1	4.4
Headaches	7.3"	3.7	8.3"	3.0	3.4	2.5
Enuresis	11.2"	6.4	11.3"	3.6	12.3'	6.6
Ever Had- Behavior or Emotional Problems[3]	15.2"	2.3	6.5"	1.1	9.7"	1.8
Unweighted N	(1,255)	(2,857)				

reports at least one behavioral problem occurs
problem sometimes or
[3]Amongand older.

* $p \leq .05$
** $p \leq .01$

problems and that black and Hispanic are less likely than non-Hispanic white children to see either a general or specialty mental health care provider.

These findings have important implications for research into childhood psychopathology as well as for pediatric practice. As with adults, the general medical care provider is likely to be the first, and often the only, contact that a child with mental health problems has with the medical care system. Since they are on the front lines in the battle against childhood mental illness, it is important that pediatricians be aware of the substantial emotional comorbidity that is likely to exist among their patients who present with strictly medical problems.

The assessment of emotional and behavioral problems in pediatric practice, however, is a difficult task especially when there are large cultural and social class differences between the physician and the patient. Cultural and social class distance between the physician and the patient can decrease the probability that problem behaviors will be recognized and referred for appropriate care. On the other hand, cultural difference between patient and provider may increase the likelihood that a physician will label relatively normal behavior as problematic.

By now there is a large literature that shows that culture and social class influence both the individual's expression of distress and the responses of those that make up his or her immediate social network. Although it is likely that childhood psychopathology has certain universal core aspects, the language and behaviors that different cultures and social classes provide a child influences the expression of psychological distress. Children from different social groups may differ in the extent to which they manifest emotional problems as physical illness at the same time that parents from different social groups may be differentially sensitive to behavioral, emotional, or physical problems.

It is imperative, therefore, that we begin to understand how race and Hispanic ethnicity, in conjunction with social class influence both the occurrence and progression of mental health problems in children and how they influence the child's contact with the general medical care as well as the specialty mental health care system. Those social class and economic factors that influence the probability that a child sees a pediatrician also influence the probability that he or she receives help for mental health problems. Although having to deal with such factors as race and etbnicity, family structure, and social class complicates the study of children's mental health, the failure to do so runs the risk of seriously misrepresenting the needs of children at high risk of psychopathology.

Reference6

1. Anderson J, Williams, S., McGee R. Silva, P: DSM-III disorders in preadolescent children. Arch. Gen Psychiatry 1987; 44:69-76.
2. Bird, ER, Gould, MS, Yager, T, Staghezza B, Canino, G: Risk factors for maladjustment in Puerto Rican children. J. Am. Acad. Child Adolesc. Psychiatry 1989; 28:847-50.
3. Boyle, MB, Offord, DR, Hoffman, HG et al.: Ontario child health study I, Methodology. Arch. Gen. Psychiatry 1987; 44:826-31.
4. Cohen, P, Brook J: Family factors related to the persistence of psychopathology in childhood and adolescence. Psychiatry 1987; 50:332-45.
5. Costello, EJ, Costello, AJ, Edelbrock C, et al: Psychiatric disorders in pediatric primary care: prevalence and risk factors. Arch Gen. Psychiatry 1988; 45:1107-16.
6. Offord, DR, Boyle, MB, Racine, Y: Ontario child health study: correlates of disorder. J. Am. Acad. Child Adolesc. Psychiatry 1989; 28:856-60.
7. Costello, EJ: Developments in child psychiatric epidemiology. J. Am. Acad. Child Adolesc. Psychiatry 1989; 28:836-41.
M, Wunsch-Hitzigld R , Dohrenwend B: Estimating the prevalence of childhood Acad 20:462-476.

Table 3: Physical and Mental Comorbidity and Medical Care Utilization of Children in Two-Parent and Female-headed Families by Race and Hispanic Ethnicity

	Non-Hispanic White		Black		Mexican American	
	Two Parent	Female-Headed	Two Parent	Female-Headed	Two Parent	Female-Headed
Problems[1]	5.4"	13.6	3.0"	5.3	4.9	5.8
Often Experience- At Least One Behavioral Problem	28.1"	38.2	32.3	40.2	35.4	37.3
Ever Seen- Psychotherapist'	5.4"	13.3	3.4	4.6	2.4'	6.0
Past 12 Months- Doctor Visit	82.3	80.7	72.0	73.3	67.4*	80.1
Asthma	5.1"	7.9	1.6	7.9	4.2	5.0
Other Respiratory	12.8	14.9	7.1'	6.2	5.6	4.5
Food Allergy	7.8	6.6	3.1'	3.9	3.8	3.9
Pneumonia	7.1	8.0	3.8*	5.8	3.6	6.1
Headaches	4.8"	6.5	4.0	4.1	2.5'	6.3
Earaches	21.9	28.3	16.7	15.2	19.7	20.0
Enuresis	5.0	4.3	4.2	4.3	4.2	4.9
Unweighted N	(9,886)	(1,445)	(1,207)	(1,235)	(684)	(139)

years and
years and

*p ≤ .05
"p ≤ .01

9. Zill N, Schoenborn CA: Developmental, learning, and emotional problems: health of our nation's children, United States, 1988. Advanced Data from the Vital and Health Statistics; no 190. Hyattsville, MD: National Center for Health Statistics, 1990.
10. Costello EJ: Primary care pediatrics and child psychopathology: A review of diagnostic, treatment, and referral practices. Ped 1986; 78:1044-1051.
11. Costello EJ, Burns BJ, Costello AJ, et al.: Service utilization and psychiatric diagnosis in pediatric primary care: The role of the gatekeeper. Ped 1988; 82:435-441.
12. Starfield B, Hankin J, Steinwachs D, et al.: Utilization and morbidity: Random or tandem? Ped 1985; 75:241-47.
13. Kelleher K, Starfield B: Health care use by children receiving mental health services. Ped 1990;85:114-118.
14. Regier DA, Goldberg ID, Taube CA: The de facto mental health services system. Arch Gen Psych 1978; 35:685-693.
15. Gfford DR, Boyle MB, Szatmari P et al.: Ontario Child Health Study II: Six-month prevalence of disorder and rates of service utilization. Arch Gen Psych 1987; 44:832-836.
16. Wolfe BL: Children's utilization of medical care. Med Care 1980; 18:1196-1207.
17. Angel R, Worobey JL: Acculturation and maternal reports of children's health: Evidence from the Hispanic Health and Nutrition Examination Survey 1988; 69:707-721.
18. Biller HB: The mother-child relationship and the father-absent boy's personality development. Mer Palmer Quart 1971; 17:227-41.
19. Hetherington EM, Cox BM, Cox R: Long-term effects of divorce and remarriage on the adjustment of children. J of Amer Acad Psych 1985; 24:518-30.
20. Hetherington EM, Cox EM, Cox R: Play and social interaction in children following divorce. J Soc Issues 1985; 35:26-49.
21. Kellam SG, Ensminger MB, Turner FJ: Family structure and the mental health of children. Arch Gen Psych 1977; 34:1012-22.
22. Wallerstein JS, Blakeslee S: Second Chances: Men, Women, and Children a Decade after Divorce. New York:Ticknor and Fields, 1989.
23. Guidubaldi J, Cleminshaw H: Divorce, family health, and child adjustment. Fam Relations 1985; 34:35-41.
24. Dawson DA: Family Structure and Children's Health: United States, 1988. National Center for Health Statistics. Vital Health Statistics 1991; 10(178). Hyattsville, MD.
25. Munroe, HB, Boyle, MH, Offord, DR: Single-parent families:Child psychiatric disorder and school performance. J Child Adol Psy 1988; 27:214-219.
26. Angel R, Worobey JL: Single motherhood and children's health. J Health Soc Beh 1988; 29:38-52.
27. Gutterman EM, O'Brien JD, Young JG: Structured diagnostic interviews for children and adolescents: Current status and future directions. J Amer Acad Child Adol Psychiat 1987; 26:621-30.
28. Rutter, M: Isle of white revisited: Twenty-five years of child psychiatric epidemiology. J. Am. Acad. Child & Adolesc. Psychiatry 1989; 28:633-53.
29. Cantwell DP: Classification, Pp. 133-47 in Solnit AJ, Cohen DJ, and Schowalter JE (eds.) Child Psychiatry. New York:Basic, 1986.
30. Achenbach T: Psychopathology of childhood: Research problems and issues. J. Consult. Clin. Psych. 1978;46:759-76.
31. Rutter M, Tuma AH: Diagnosis and classification: Some outstanding issues. In: Rutter M, Tuma AH, Lann, IS (eds.): Assessment and Diagnosis in Child Psychopathology (pp. 437-52); New York:Guildford, 1988.
32. Achenbach, R. Behavioral problems and competencies reported by parents of normal and disturbed children aged 4 through 16. Monographs of the Society for Research in Child Development 1981; 46:#188.
33. Quay HC, JS Werry: Psychopathological Disorders of Childhood (3rd edition) :New York:John Wiley, 1986.
34. Rutter M: Changing youth in a changing society: Patterns of adolescent development and disorder. Cambridge, MA:Harvard Univ. Press, 1980.
35. Robins LN, Rutter M (eds.): Straight and devious pathways from childhood to adulthood. New York:Cambridge Univ. Press, 1990.
36. Angel R, Guarnaccia P: Mind, body, and culture: Somatization among Hispanics. Social Science and Medicine 1986; 23:1-10.
37. Idler E, Angel R: Age, chronic pain, and subjective assessments of health. Pp. 131-152 In: Albrecht GL(ed.): Advances in Medical Sociology, Vol 1. Greenwich, CT:JAI Press Inc., 1990.
38. Mechanic DC. The influence of mothers on their children's health, attitudes, and behavior. Ped 1964; 33:444-454.
39. Goode W: A theory of role strain. Amer Soc Rev 1960; 98:483-96.
40. Garfinkel I, McLanahan S: Single mothers and their children: A new American dilemma. Washington, DC:Urban Institute, 1986.
41. Behavior problems for young children (4 to 11 years) included: changes in mood, complaints about love, high strung, cheats or tells lies, fearful or anxious, argues too much, can't concentrate, easily confused, bullies, cruel, or mean, disobedient at home, misbehaving, trouble getting along with children, trouble getting along with teachers, acts without thinking, feels worthless, not liked by others, one track mind, can't sit still, stubborn, sullen, or irritable, very strong temper, unhappy, sad, or depressed, withdrawn, breaks things, clings to adults, cries too much, demands attention, dependent on others. For adolescents (12 to 17 years) the questions about breaking things, clinging to adults, crying too much, demanding attention, and depending on other were replaced with four questions associated with adolescent behavior: others out to get him/her, hangs with trouble makers, secretive, worries too much. Parents responded that the behavioral problems occurred often, sometimes, or never.
42. Angel R, Worobey J: Intra-group differences in the health of Hispanic children. Soc sci Quarterly 1991; 72:361-378.

ASSESSING FAMILY STRENGTHS AND YOUTH BEHAVIOR: ANALYSES OF THREE NATIONAL DATABASES ON CHILDREN AND FAMILIES

Barbara W. Sugland, Child Trends, Inc.
Kristin A. Moore, Donna Ruane Morrison, and Brett Brown

Introduction

Over the past 20 years dramatic changes have occurred in the United States and in the make-up of the American family. While two-parent families are still the norm, divorced, never-married single-parent families, step-families and cohabiting-parent families now represent a sizeable proportion of families with children?

The changing face of America's families and shifts in the context of family life have prompted numerous discussions about the impact of such changes on the well-being of children. Many researchers have empirically investigated the impact of different family situations on child and adolescent development,' and others have begun to investigate differences in family processes in order to learn more about how family life affects the development and the behavior of children.[3]

Despite these efforts, we still know relatively little about family processes and their impact on children. This paper explores three research questions: 1) What is the utility of national-level data to adequately assess measures of family processes?; 2) What is the appropriateness of family strengths constructs for varied sub-groups, such as single and two-parent families, blacks and whites?; 3) Are family strengths associated with positive outcomes for children, and do associations hold after controlling for other family characteristics?

Data and Methodology

This study summarizes selected findings from analyses of the National Survey of Families and Households (NSFH), the National Longitudinal Survey of Youth, Child Supplement (NLSY-CS), and the National Survey of Children (NSC). Details of the full analyses are available.[4]

Data

The NSFH is nationally representative sample of 13,000 U.S. Households in 1987, 2,300 of which contain adolescents between twelve and eighteen years of age. A fairly rich array of family process data, such as parental involvement in youth organizations and time parents and children spend together, was obtained from an adult respondent. Data for the 1987 survey are used for this study.

The NLSY-CS contains assessments of children born to female respondents in the NLSY. Child development data from 1986, 1988, and 1990, are linked with socioeconomic, family background, and marital history data reported by the NLSY respondents in 1979. The NLSY has a limited number of family process measures, but surveys are obtained every year, permitting prospective analyses. Also information is collected from both the mother and the child.

The NSC is a nationally representative household survey of children aged 7-11 in 1976, designed to examine the social, physical, and psychological characteristics of U.S. children. Three waves of the NSC have been conducted. The first in 1976, the second in 1981, and the third in 1987. In each wave, both the parent and child were' interviewed. Although the purpose of the NSC was not to assess family processes, it has a rich set of family process measures, including parent-child communication, parent-child activities, social networks, and family adaptability.

Measures

The family strengths measures used in this study are: parent-child communication, parental affection, frequency of parent-child activities; importance of marriage and marital fidelity, and parental discipline including parent's attitude toward harsh punishment and use of harsh discipline.

Two behavior outcomes from each data base are selected. In the NSFH, we use a summary measure of behavior problems comprised of difficult behavior, personality characteristics, such as a being irritable or sad, fearful, and bullying (Behavioral Problems I), and a measure of more serious behavior problems, such as being suspended or expelled, running away, or being in trouble with the police (Behavioral Problems II).

In the NLSY-CS, we explore the Behavioral Problems Index (BPI), reported by the adult respondent. Developed by Zill and Peterson: from prior indices, the BPI measures acting out behaviors, depressed and withdrawn behaviors, as well as distractable or hyperactive behavior.

The second behavior measure in the NLSY-CS includes the number of times youth stayed out late, hurt someone badly enough to require medical attention, lied to parent(s), stole something or skipped school.

In the NSC, we examine the BPI and a youth-reported scale measuring delinquent behaviors in the previous 12 months.

Principal components analyses, bivar-

iate correlations and ordinary least squares regression are the techniques employed.

Results

Table 1 summarizes the range of family strengths constructs available across the three data sets. All three data sets have questions that assess religious orientation and attendance, extended family networks, consistent roles and expectations for children, and time spent together with children. There are also items in each to assess parental conflict.

Some constructs cannot be assessed. The NSFH has insufficient measures for positive communication, appreciation, and parental discipline. The NLSY-CS does not contain measures of commitment to marriage or family, or family adaptability. None of the three data sets have items with which to measure encouragement and support of family members.

There are also differences in the breadth of the substantive detail that one is able to measure among available items. For instance, social connectedness in the NSC is measured via frequency of parental contact with friends and the number of close friends in near geographic proximity. In the NSFH, social networks can be measured by the total number of adult kin outside the home, the total number of' family members living near the respondent, and the number of days per year the adult socializes with friends, neighbors or co-workers. The NLSY-CS has a single item about the frequency of family visits with friends/relatives.

We explore the psychometric properties of available items. Factor analysis suggests that the family process measures have limited psychometric properties. In the NSFH, two of the three items (family activities and social networks) have reliabilities of .60 or more, but the social networks measure produces four factors; The first factor explains close to 40% and remaining factors explain less variance.

In the NLSY-CS, two constructs demonstrate satisfactory properties (parent-child communication and appreciation). Both produce a single factor on which all items load at least .30; Both have acceptable levels of reliability (.75 and .84).

The NSC has the greatest number of family process items, but presents the weakest set of psychometric properties. Three of the six constructs have reliabilities of .60 or more -- parent-child communication, appreciation, and parental discipline. Only parent-child communication produces a single factor. While the reliability of the remaining two are acceptable (Chronbach's alpha equals .75 and .71 respectively), they both produce three factors. In each case, the first factor explains less than 40 percent of the variance in the items.

While measures are available in these three data sets, more relevant items are needed to construct better summary measures of family processes. In most cases, only a few items are available to construct scales, reducing scale reliability, and some constructs could not be assessed at all, particularly in the NLSY-CS.

The presence of family strengths is also examined separately by sample subgroups. Few substantial differences in the distributions of mean scores on indicators of family strengths across family type were observed for all three data sets. This suggests that family strengths are common among all types of families, irrespective of family structure or race/ethnicity. Notable differences occur among white families and two-parent families, where family strength characteristics show a more positive ranking.

Bivariate Analyses

The most central question in this study is whether family strengths affect the incidence of problem behaviors in children and youth. We first explore this question using bivariate correlations controlling for family type. Correlations indicate that family strengths is associated with fewer behavior problems among children and youth, irrespective of family structure. Measures of harsh or strong punishment, on the other hand, predict to later problems. Measures of internal family functioning, such as parent-child communication, appreciation, and strong punishment, are more predictive of outcomes than external measures, such as social networks; Child-reported measures are more predictive than parent-reported measures. The magnitude of significant associations varies in some instances, with Pearson's correlations ranging from .10 to .20. Similar effects on child behavior are observed across race/ethnic sub-groups as well.

Multivariate Analyses

Table 2 summarizes the influence of family strengths on the selected outcomes based on findings from final regression models. Controlling for socioeconomic variables, such as parental education, income, race, and family structure, tends to diminish but not erase the effects of family process variables. In the NSC, parent-child communication demonstrates a significant influence on youth outcomes, leading to lower scores on the BPI and fewer. delinquent behaviors. Commitment to family reduces delinquent behavior, but has no significant influence on BPI scores.

In the NLSY-CS none of the family strength measures consistently affects children's behavior, though appreciation is associated with fewer child-reported

behavior problems, and harsh discipline is associated with more behavior problems. 'The lack of effects of family strengths on child outcomes in the NLSY-CS may reflect the paucity of strong measures of family processes or the limited variability found in the disadvantaged sample of NLSY-CS mothers with school-age children. Since a goal of examining family strengths, however, is to identify family processes that represent a positive resource for families regardless of their socioeconomic assets, the minimal effects in this sample are important to recognize.

In the NSFH, the variables found to be the most important were those that tap the internal family processes, including parent-child time together and parental commitment to the family. The availability of social networks was not found to predict directly to child outcomes in the multivariate models.

Measures of parental discipline, marital conflict, parental depression and marital disruption have negative effects on children's development, net of background factors and other measures of family strengths. In the NLSY-CS, the use of strong punishment (spanking) by the parent did predict to subsequent behavior problems. Strong punishment was shown to increase negative behaviors in youth in the NSC as well. In the NSFH, both parental depression and marital conflict were associated with higher scores on the index for Behavioral Problems I.

Discussion

The results from this study suggest that including family strengths constructs in large-scale national surveys is promising. Measures of family strengths predict to later behavior problems even within family structure and race/ethnic subgroups. Associations also hold after controlling for social and economic variables, although the variance explained by family strength variables is generally modest.

The minimal associations found here may reflect the lack of a conceptual framework for the family strengths measures. The constructs employed were developed by researchers and practitioners who tended to first identify successful families and then to identify the characteristics of those families. This strategy yielded an intuitively meaningful set of family strengths that were not theoretical linked with other disciplines (e.g., child development, family sociology). A stronger linkage between the successful families literature and the theoretical perspectives of other traditions would strengthen the development and operationalization of family strengths measures.

Limited associations may also be due a lack of adequate measurement in existing national surveys. Typically, few resources are devoted to developing scales appropriate for survey administration, relative to such tasks as sample selection. As a result, the range and quality of available items in large scale surveys is often limited; Entire constructs are frequently measured with a single item; The validity of items and scales and constructs in different sub-populations is not typically assessed. This has important implications for exploring and understanding subgroup differences. To understand the role of family processes apart from family socioeconomic resources will require an investment in measure development.

One methodological issue to which the answer appears clear is the need for multiple respondents. In particular, obtaining the perspective of the child or youth on family processes and on their own behavior seems to be important. Family strengths, as reported by the child-respondent, were more predictive of child outcomes than measures reported by the adult respondent.

The value of this study is that it has systematically taken promising constructs developed in one literature and examined them with stringent multivariate methods. These analyses indicate that most family strength constructs do affect the development of children and adolescents net of socioeconomic variables across varied social groups. At the same time, the results indicate a need for theory-driven measures and more reliable survey items that are appropriate for various types of families and race/ethnic groups.

References

'U.S. Bureau of the Census. Department of Commerce Economics and Statistics Administration (1992). "Housing of Single-Parent Families. Statistical Brief"; U.S. Bureau of the Census, Current Population Reports, Series P-20, No. 458, "Household and Family Characteristics". U.S. Government Printing Office, Washington, D.C.

[2]Dawson, D.A. (1991). Family Structure and Children's Health and Well-being: Data From the 1988 National Health Interview Study. *Journal of Marriage and the Family*, 53:573-584; Jayakody, R. Chatters, L.M., and Taylor, R.J. (1993). "Family Support to Single and Married African AmericanMothers:The Provision of Financial, Emotional, and Child Care Assistance". *Journal of Marriage and the Family.* 55(May):261-276; Kurdek and Fine, M.A. (1993). "The Relation Between Family Structure and Young Adolescents' Appraisals of Family Climate and Parenting Behavior". *Journal of Family Issues.* 14(2):279-290;

[3]Capaldi, D.M. and Patterson, G.R. (1991). "Relation of Parental Transitions to Boys' Adjustment Problems: I. A Linear Hypothesis. II. Mothers at Risk for Transitions and Unskilled Parenting. *Developmental Psychology.* 27:489-504; Dishion,

T.J., Patterson, G.R., Stoolmiller, M., and Skinner, M.L. (1991). "Family, School, and Behavioral Antecedents to Early Adolescent Involvement With Anti-social Peers". *Developmental Psychology.* 27:172-180.

[4]Brown, B. (1993). "Family Functioning and Adolescent Behavior Problems: An Analysis of the National Survey of Families and Households". Washington, D.C.: Child Trends, Inc; Morrison, D.R. and Glei, D. (1993). "Assessing Family Strengths in the National Longitudinal Survey of Youth - Child Supplement". Washington, D.C.:Child Trends; Sugland, B.W. (1993). "The Effect of Family Strengths on Youth Behavior: An Analysis of the National Survey of Children. Washington, D.C.:Child Trends, Inc.

[5]Zill, N. (1990). Behavior Problems Index Based on Parent Report. Washington, D.C.: Child Trends, Inc.

Table 1. Family Strengths/Family Process Measures Across the NSFH, NLSY-CS, and the NSC

	NSFH	NLSY-CS	NSC
Family Strengths Constructs			
Communication	---	√	√
Encouragement of	- -	---	---
Individuals	√	.	√
Commitment to Filmily	√	√	√
Religious Orientation/ Training/Attendance	√	√	√
Social Connectedness/Social Networks	-	---	√
Family Adaptability	-	√	√
Expressing Appreciation	√	√	√
Clear Roles	√	√	√
Time Together/Family Activities			
Other Family Process Constructs			
Strong Punishment/Spanking	---	√	√
Mother-Partner Relationship Satisfaction	- -	√	√
Parental/Marital Conflict	√	√	√
Parental Agreement about Child		√	√
Parental Depression	√	√	---

Key: "√" = family strengths/family process variable(s) are available.

Table 2: Summary of Results of Multiple Regression Analyses: Family Strengths, Background Characteristics, and Youth Behavior in the NSFH (1987), NLSY-CS, (1986, 1988, and 1990), and the NSC (1981 and 1987)

	NSFH		NLSY-CS		NSC	
Independent Variables	Behavior Problems I	Behavior Problems II	BPI (Parent Reported)	Behav Problems (Child Report)	BPI (Parent Reported)	Delinquent Behav (Child Reported)
Family Strengths						
Communication	na	na	0	0	+	+
Appreciation	na	na	0	+	0	0
Family Activities	+	+	na	na	0	0
Commitment to Marriage & Family	0	0	na	na	0	+
Social Networks	0	0	0	0	0	0
Discipline measures						
Use of Strong Punishment	na	na			-	-
Mother/Partner Measures						
Parental Depression		0	na	na	na	na
Marital Conflict		0		0	na	na
Marital Disruption	na	na	na	na	0	0
Parent-Parent Communication	na	na	0	0		

Key: "0" = no statistically significant association, "+" is positive influence (i.e., an increase in family strengths characteristics leads to less negative behavior outcomes); "-" is a negative influence (i.e., an increase in the family characteristic leads to more negative behavior outcomes); "na"=variable not included in the final model. Statistical significance is at the $p \leq 0.05$ level.

Note: Table values are based on weighted data.

Source: Child Trends, Inc. Tabulations of the National Survey of Families and Households, 1987; the National Longitudinal Survey of Youth-Child Supplement, 1986, 1988, and 1990 waves and; the National Survey of Children, Waves II and III (1981 and 1987).

FAMILY AND HEALTH

Discussant:

Martha Farnsworth Riche
Population Reference Bureau

(not present)

FAMILY CHARACTERISTICS AND HEALTH

Karen smith Thiel, Health Resources and Serviaes Administration

Bach of these three very interesting papers utiliae existing data sets to try to examine the relationship between family characteristics and some aspect of physical or emotional health. Ascoli, et al. use the Live Birth File for the State of Minnesota 1990-1991. The Angels use the 1988 Child Health Supplement to the National Health Interview Survey and Sugland, et al. use these data sets: the 1987 National Survey of Families and Households: the 1986, 1988, and 1990 National Longitudinal survey of Youth-Child Supplement; and Waves II and III of the National Survey of Children.

Rather than address the findings of these papers, I'd like to focus my comments on the limitations of the data used in each paper.

Ascoli et al's paper is limited in its generalizability by the population studied. The 1990-1991 Minnesota Live Birth File is limited in terms of the range of ethnic, marital, and age characteristics of this population. Over 90% are White, 79% are married and 97% are age 18 and over.

Secondly, although the introduction to the paper refers to the influence of both family structure and family function, in fact only family structure variables can be captured with this data set. For example, while the paper's strength is to note that the role of the father in prenatal care and pregnancy outcome has had limited review, in fact "role" can be measured as a set of process variables and is little captured in the data available to these researchers.

In terms of the outcomes studied, this is a population in whiah over 60% of the mothers had first trimester prenatal care and adequate prenatal care. As the authors note, less than 10% had no prenatal care. The overall low birthweight incidence was approximately 4.2%. In the communities in which I've worked in the past and certainly in the 15 Healthy Start Communities in which I work now, these oertainly would be dream statistics.

Ascoli et al. provide good documentation of the relationship between prenatal care and birth outcome - noting the positive relationship between the reoeipt of adequate levels of prenatal care, with low birthweight incidence progressing from 3% in the adequate care group to 5% in the intermediate care group, to 6% in the inadequate care group, to 21% of those who had no prenatal care. The authors' foaus on prenatal care is doubly important because the prenatal period provides some window of opportunity to address other health issues (e.g. substance abuse and parenting skills).

The paper also allows some profiling of groups at risk of poor birth outcome. The authors note that mothers who receive care early but not adequately are at high risk, and that adolesaent, unmarried, high parity, low eduaation minority mothers are at increased risk for low birth weight even if they enter care early.

By way of interpretation, I would note that in the evaluation of the State of California's Adolescent Pregnancy Program we *found* early entry into prenatal care often meant a pregnancy test, but actual entry into care often did not occur until the seoond trimester. This phenomenon was so much the case with the adolescent mothers we studied, that we decided to use total number of prenatal case visits instead of trimester of entry into care as the variable in our models.

Not being married appears to put a mother at risk for reaeiving no prenatal care, regardless of age or educational level. Aside from marriage, however, the only father variable added to the analysis is father's educational level. As these authors are aware, this hardly oaptures dimensions of the concept "father's role".

Interestingly, most of the interpretations of findings offered by the authors about the relationship between a father's education, maternal oharaoteristics, and prenatal care outcomes are process interpretations relying on process measures unobtainable in the data set utilized in this analysis. The question of what fathers do in terms of their provision of access to care, their knowledge or motivational input, or the social and emotional support they supply to their partner remains unanswered.

I empathire with these authors and their inability to extract proaess data from vital record data. Unfortunately, our maternal and child health bias of women and children first results in a lifeboat in which fathers' characteristics and roles don't exist. Putting the family unit in the lifeboat by including fathers *in* services as well as data collection aertainly would enrich our understanding of the impact of *prenatal care on* birth outcomes.

The Angels paper is also limited by its data. Relying on mothers' reports of

children's emotional and behavioral problems may either over report, due to maternal characteristics like perception, stress, education, and income or under report, due to mothers' not knowing the full range of their children's behaviors (e.g. smoking, drinking, drug use, school fights, skipping school, etc.). This is particularly the case with the 12-17 year old children depicted in the tables.

The Angels paper does, however, go beyond Ascoli et al. in breaking out the category "not married" into never married, divorced, separated and widowed for a portion of its analysis. The paper also presents data in Table 1 showing that non-Hispanic (white) female headed households are eight times more likely to be low income than are two parent families and that Black and Hispanic female headed families are each five times more likely to be low income than two parent families in this data set. Moreover, non-white Hispanic, white and Black female-headed families in this data set have almost one year less education than two parent families. Hispanic female headed families have almost two years less education than two parent families. These disparities in income and education by marital status and race are important for this analysis.

With respect with the conceptual basis of their analysis, I question the way in which the authors have operationalised two of their key concepts. The paper offers a hypothesis to explain the greater number of emotional and behavioral problems of Black and Hispanic children as well as the children of single mothers. However, the term "psychopathology", which I consider a clinical term, is used interchangeably with the terms "behavioral problems" and "emotional problems". The use of the term psychopathology, even in the discussion section of the paper, gives these mothers' reports a clinical connotation that the data do not reflect.

Secondly, the authors test a hypothesis about the receipt of mental health care, but then proceed to ask mothers whether their children have seen a psychiatrist, doctor, or counselor. I would argue that most children's emotional and/ *or* behavioral problems don't end up in any of these offices, but instead are handled by parents and other family member? at home, by teachers, school principles, school guidance counselors, or by the juvenile justice system. Added to this is the fact that access to a pediatrioian or a psychotherapist both involve more general "access to care" issues.

Finally, I would argue that one statement made in the paper is completely unsupported by the methodological design of this paper. After the authors note the differenaes between often and sometimes/never responses to the category "ever had behavior or emotional problems" in Table 3, they state... "it is olearly those who manifest serious behavior problems and who do so frequently who are at risk of long term maladjustment." First, there is no "degree of seriousness" measure to the behavior reported by these mothers in this data set. Secondly, this is not a longitudinal study. These mother's reports of children's and adolescents' behavioral and emotional problems are not linked to any data on long term behavioral outcomes or long term maladjustments.

Like Ascoti et al's paper, the Angel's paper also is limited by the limited family process data in the Child Supplement to the Health Interview Survey. It would be interesting to know how mothers respond to children's behavioral and/or emotional problems beyond appointments with physicians or psychotherapists, whether problems pervade ohildhood, and how serious these problems are perceived to be by families.

Given the limitations in these data, I would have preferred that Angel et al offered their findings as hypotheses to test in fuller, more clinically oriented data sets, rather than as results of the impact of these limited sets of variables on child behavioral problems.

Finally, Sugland et al. attempt to literally scour three existing data sets to derive family process measures. The paper is a clear response to the call for these efforts made in the National Center for Health Statistics/Centers for Disease Control and Prevention sponsored March, 1993 Workshop on *Family* Data and Family Health Policy.

These authors have the advantage of working with comparatively rich data sets. Two of the data sets used in their analysis are longitudinal or as the authors state, "employ a prospective study design", and two data sets incorporate children's responses as well as mothers' reports of children's behavior.

More importantly, the paper takes a "cup half full approach" and examines the impact of measures of family strengths on children's behavior. The paper uses theoretical constructs and attempts to analyee three domains and their

dimensions, and selects variables or sets of variables with which to measure these domains: parent/child interactions, including parent/child communication; feelings about marriage and family, and use of soaial networks; and discipline and harsh punishment. The study also uses a specific and expansive set of child behavioral problems captured in the Behavioral Problems Index.

The correlational findings of the relationship between the presenae of family strengths and fewer behavioral problems among children and adolesaeats are impressive, even if modest. The findings that family strengths aorrelate with behaviors regardless of race or gender of child are also important.

The findings with respect to the laak of association between family strengths and child outcomes in the NLSY, once socioeconomic status variables are controlled, points to an interesting area for further *research. Given* that these respondents tended to be younger mothers, a portion of whom gave birth as adolesaent mothers, we need to learn about these families' strengths and what f actors mediate children's behaviors in these households.

With respect to the soaial network findings, social network measures don't predict to either of the behavior problem measures. But, only one data set, the *NSFH,* includes emotional ties as a social network measure.

The Sugland et al. paper is an interesting paper made even more so by its level of self-critique, specificallyitsdisaussioa of the ways in which constructs, data sets, and measures can be improved.

clearly, all three papers struggle to aaalyae existing data sets with implications for important family policy issues. These papers refleot some of the key points made at the NCHS/CDC March 1993 workshop on Family Data and Family Health Policy. Researchers and poliay analysts concluded that there was a need to: measure family struotures and process; use the family as the unit of analysis or at least aolleat data on multiple family measures; include the male partner; study social connections, social networks, and parent/child communication: develop standard indicators of family struoture and process across Federal data sets, and build consensus in the field to strive to collect these data in other researah; and focus on studying childbearing in a biosoaial context.

Both the Sugland et al. paper and a recent paper on risk and resilienay among youth in single-parent families by Peter Benson and Eugene Roehlkepartan of the Search Institute note demographia ahanges in U. S. soaiety where approximately one in four families are single headed and estimates are that 50-60% of children born today will experience at least a portion of their ohildhood or adolesoeaoe in single-headed families. In faat, some *experience in a* single-headed family *for* children in our soaiety may beaome normative. To understand how single-headedaess, partiaularly female-headedaess, iateraats with reduaed income or even poverty and family fuaatiooiag measured by family strengths as Well as family defiaits is imperative if we are to design health, mental health, and eduaation interventions to meet changing family needs.

Session F

Nutrition

Toward the year

2000

Refining the measures

ASSESSING FRUIT AND VEGETABLE INTAKES: TOWARD THE YBAR 2000

Susan Krebs-Smith
National Cancer Institute

(Not available for publication)

METHODOLOGICAL ISSUES IN ANALYZING SCHOOL MENUS

Patricia McKinney, Food and Nutrition Service

U.S.D.A.'s Food and Nutrition Service (FNS) recently completed a nationally representative study of schools in which menus were collected from 515. The objective of the FNS study was to

> "determine the average nutrient content of USDA meals as offered to children in schools and compare USDA-reimbursable breakfast and lunch meals to dietary standards defined on the basis of the *Dietary Guidelines* and the *Recommended Dietary Allowances."*

The study also will supply baseline data for monitoring The *Healthy People 2000* Objective for School Food Services which reads:

> "Increase to at least 90 percent the proportion of school lunch and breakfast services and child care food services with menus that are consistent with the nutrition principles in the Dietary Guidelines for Americans."

Methodological issues first resulted from differing needs of the two objectives, interpretation of the *Healthy People* objective and the need to quantify the Dietary Guidelines. Most of the Guidelines are given as serving sizes with ranges [e.g., 6-11 servings of breads, cereals, rice and pasta.] with guidance that "the content of the total diet over the day or more is what counts" and "most people should have at least the lower number of servings". The FNS study did not address the Guidelines for "Variety" or "Plenty of Vegetables, Fruits and Grains.". The meal pattern requirements for USDA's school lunch and breakfast are given as servings. We looked at variety within each food group but did not develop a meal-based standard for the number of servings. These two Guidelines are emphasized especially for their complex carbohydrates, dietary fiber and nutrients. It was felt that the nutrient data base for complex carbohydrate was inadequate so percent of calories from carbohydrate was used as a crude indicator of complex carbohydrates. The Guidelines for cholesterol and salt are not quantified. The NRC's *Diet and Health* recommendations were used as the basis of comparison. Again, because of an inadequate data base and no quantification of the amount, the sugar Guideline was not analyzed.

Additional methodological issues had to be addressed:

- Guidelines designed to be applied over a total day had to be applied to a single meal.

USDA regulations requirethatmeals should over time provide one-third of childrens' daily nutrient needs, whereas breakfast regulations do not include specific RDA target goals. Based on USDA recommendations, twenty-five percent was used as the target. The same percentages were used for quantifying the Guidelines, i.e. 33% for lunch and 25% for breakfast.

- Methods had to be developed to deal with the large number of choices available at each meal, and to assign amounts for self-serve options such as salad bars for which there are no standard servings.

Based on information supplied by food service managers (description, amounts served, recipes, and labels) for one week's menus, each food was assigned a USDA meal component code, and a code to link items served together such as salad with salad dressing or hamburger with a bun. The assumption was made that each meal as offered contained the numbers and types of foods required under the USDA meal pattern, plus any non-creditable items such as dessert or condiments. Production counts were not used to weight the foods, based on the premise that each student has an equal opportunity to select each food item that is offered.

All entrees offered, all additional bread/bread alternates, and all fruits and vegetables were calculated separately and an average nutrient value for meal component was obtained.

For salad and other food bars, quantities were assigned based on the assumption that students in elementary schools are offered the minimum meal pattern components of 1 serving of bread, 2 02 of meat/meat alternates and 3/4 cup fruit/vegetable [assuming that 1/2 of the vegetable is lettuce], plus 1 serving of salad dressing and 3 "toppings" such as bacon bits, croutons, olives, etc; middle and high school students were offered somewhat larger portions.

To calculate the average nutrients in meals offered each day of the week, all of the foods offered were considered:

- Foods on the "offered each day" list were included with the foods served each day of the week and coded using the same rules.
- Items from salad (and other self serve) bars were included as discussed above; and
- All types of milk offered were weighted equally.

The average nutrient content of each lunch was thus calculated as the sum of nutrients in one entree, one bread, two fruit/vegetables, one milk and one dessert and/or condiment (if offered).

Average nutrients offered in school lunches during the week were then calculated by totalling the average nutrients offered each day and then dividing by the number of days lunch was served.

MONITORING PROGRESS TOWARD MEETING YEAR 2000 OBJECTIVES: NHANES III, 1988-91

Clifford L. Johnson, National Center for Health Statistics
Ronette Briefel

The importance of nutrition in the Year 2000 objectives is obvious. Not only is there a full chapter on nutrition objectives, but there are a number of nutrition-related health objectives throughout the National Health Promotion and Disease Prevention Objectives [1]. In this paper, data from selected National Health and Nutrition Examination Surveys (NHANES) that provide baseline data and current or future data with which to monitor progress towards the goals of the Year 2000 nutrition-related objectives are discussed.

Year 2000 dietary-related objectives for which the NHANES surveys provide information for measuring progress toward meeting the objectives include those on dietary fat intake, calcium intake, consumption of complex carbohydrate and fiber-containing foods, consumption of alcohol, and infant feeding practices related to dental health. Nutrition-related Year 2000 health topics for which NHANES data are used to monitor progress include serum cholesterol, blood pressure, overweight, iron deficiency, and blood lead. These are not an exhaustive list of Year 2000 objectives relevant to nutrition or for which NHANES data will be used to monitor progress toward the respective goals. However, it does provide a view of the broad public health relevance of the NHANES data.

Over the last thirty years, the National Center for Health Statistics has conducted a number of surveys to provide information on the health and nutritional status of the U.S. population by means of standardized interviews and physical examinations. Three National Health Examination Surveys (NHES) were conducted for selected age groups during the 1960's [2-4]. Beginning in 1971, a nutrition component was added and the survey became known as NHANES. Two such national surveys, NHANES I and NHANES II, were conducted in 1971-75 and 1976-80 respectively [5-6]. During 1982-84 a special survey of selected Hispanic subgroups, known as the Hispanic Health and Nutrition Examination Survey (HHANES) was conducted [7]. The latest national survey, NHANES III, began in 1988 [8].

In particular, NHANES III was designed as a six-year survey and consists of two consecutive, independent, nationally representative samples of the U.S. population. The first phase of NHANES III was conducted from 1988-1991. The second phase is currently ongoing and will be completed in late 1994. Results presented in this paper represent some of the first findings based on analyses of data from the first phase of NHANES III. Upon completion of phase 2 of the survey, data from that phase alone as well as the combined data from both phases will be available for further monitoring of the health and nutritional status of the U.S. population and the Healthy People 2000 objectives mentioned previously.

There are a number of important aspects of the NHANES survey design that are worth discussing. First, NHANES III is a nationally representative sample of the noninstitutionalized population of the entire United States. Secondly, it covers

Finally, in order to be able to address a number of research and analytic objectives including population subgroups of special interest, NHANES III includes oversampling of young children, adults ages 60 years

Americans. As in all NHANES, information in NHANES III is collected during an in-person interview in the home and in a

[9]. For comparison purposes, the data were age-adjusted using the direct method to the total adult resident population for 1980.

The Year 2000 goal for one cardiovascular objective (Heart Disease and Stroke Objective 15.6) is to reduce the mean serum cholesterol level to no more than 200 mg/dL in adult men and women. The NHANES data shown in Table 1 demonstrate that there has been a significant decline in mean serum total cholesterol levels in adult men and women over the past 30 years. The mean cholesterol level in men has declined from 220 mg/dL to 205 mg/dL. For women, the decline in mean cholesterol level has been from 225 mg/dL to 205 mg/dL. It appears that the Year 2000 goal for serum cholesterol (200 mg) is attainable.

Another Year 2000 cardiovascular objective (Heart Disease and Stroke Objective 15.7), is to reduce the percent of adults with "high" blood cholesterol levels to no more than 20 percent of adults. As shown in Table 1, NHANES data indicate that this Year 2000 objective has essentially been met. The percent of adult men and women with high blood cholesterol is about 20%.

The ability to examine trends in cholesterol levels depends on the knowledge of the comparability of the measurements. For cholesterol, laboratory methods changed several times over the course of the 30 years covered by the NHANES surveys. However, detailed internal and external quality control procedures in each survey plus the existence of a lipid standardization program at the Centers for Disease Control and Prevention in Atlanta made it possible to evaluate the extent of differences in laboratory methods on the observed decline in cholesterol levels. Only a small part of the observed decline in cholesterol levels could have been due to laboratory methods. This was critical to the interpretation of these recently published findings in the Journal of the American Medical Association [9].

Other risk reduction Year 2000 objectives (in the Heart Disease and Stroke chapter) of the Preventive Services section of Healthy People 2000, relate to control of blood pressure and action to control high blood pressure. Again, it is possible to use NHANES survey data to track these objectives. Preliminary analyses of NHANES III phase 1 data indicate that the prevalence of hypertension has decreased [10]. The prevalence of hypertension has decreased from 30 percent in NHANES II (1976-80) to 26 percent in NHANES III (1988-91). The percent of persons who are aware, treated, or controlled has increased between NHANES II and NHANES III. The Year 2000 objectives 15.1 and 15.2 specifically deal with control and "action" to control high blood pressure. More detailed analyses of these data are currently in process and will be available in the near future.

Blood pressuremeasurements have been made in each of the NHANES surveys, although the number of measurements, equipment, and procedures have changed to some extent over time. In order to more fully evaluate the interpretation of the above findings, a number of analyses are currently being conducted to evaluate potential methodologic effects on the results. Although these methodologic differences are not likely to be major, there is no doubt that ensuring comparability for blood pressure measurements over time is difficult. In addition, definitions for hypertension have also changed during the 30 year time period which requires reanalysis of previous survey data.

A year 2000 nutrition objective which is cross-referenced to other health objectives is the prevalence of overweight (objective 2.3). The goal is to reduce the prevalence of overweight in adults to no more than 20 percent. The baseline prevalence estimate was 26 percent based on NHANES II data. A few provisional estimates of the prevalence of overweight in U.S. adults have just recently been calculated using NHANES III phase 1 data and indicate that the prevalence of overweight has increased by approximately 8 percent for men and women between 1976-80 and 1988-91. Further analyses are ongoing to examine these results in more detail. If the final data confirm these preliminary results, then this will be an objective where progress has not been made, in fact, the situation has worsened. Since comparable equipment and procedures have been used over time in the NHANES surveys, it is unlikely that methodologic differences could have resulted in any of the differences observed in the overweight results.

In summary, the ability to examine trends in these risk factors and to monitor progress for the Year 2000 objectives specific to these variables is not only dependent on the collection of national survey data, but on the quality of the data and the ability to document that the data are comparable over the entire time period. If this is not true, then the ability to interpret change over time in the findings is compromised. Although the examples have all been physical measurements from NHANES surveys, the same concerns apply equally to questionnaire data and to all surveys and surveillance systems.

NHANES data are a valuable resource and these interpretive factors have been continually stressed in the collection and analyses of the data. It is critical for the interpretation of changes observed over time and the evaluation of dietary and nutrition-related health objectives, such as those for the Year 2000. The importance of this concept is apparent from the fact that there are entire sessions at this conference devoted to these topics. It is critical that long-term cross-sectional measures at the national, state, and local levels be collected and analyzed with these considerations in mind.

Table 1. Mean serum cholesterol levels and percent of persons with high levels (≥ 240 mg/dL) for men and women ages 20-74 years.

	1960-62	1971-74	1976-80	1988-91
Mean				
Men	220	214	211	205
Women	225	217	215	205
Percent				
Men	30.7	26.8	24.9	19.0
Women	36.3	29.6	28.5	20.3

References

1. Dept of Health and Human Servies. Healthy People 2000: National Health Promotion and Diseae Prevention Objectives for the Nation. Washington, DC: Public Health Service; 1991.

2. National Center for Health Statistics. Plan and initial program of the Health Examination *Survey.* National Center for Health Statistics. Vital health Stat l(4). 1965.

3. National Center for Health Statistics. Plan, operation and response results of a program of children's examinations. National Center for Health Statistics. Vital Health Stat l(5). 1967.

4. National Center for Health Statistics. Plan and operation of a Health Examination Survey of U.S. youths 12-17 years of age. National Center for Health Statistics. Vital Health Stat l(8). 1969.

5. National Center for Health Statistics. Plan and operation of the Health and Nutrition Examination Survey, United States 1971-73. National Center for Health Statistics. Vital Health Stat **1(10a)** 1973 and **1(10b)** 1977.

6. McDowell A, **Engel** A, Massey JT, Maurer K. Plan and operation of Second National Health and Nutrition Examination Survey, 1976-80. National Center for Health Statistics. **Vital Health** Stat **1(19).** 1981.

7. National Center for Heatlh Statistics. Plan and operation of the Hispanic Health and Nutrition Examination Survey, 1982-84. National Center for Health Statistics. Vital Health Stat **1(19).** 1985.

8. **Ezzati TM,** Masses JT, Waksberg J, Chu A, Maurer KR. Sample Design: third National Health and Nutrition Examinaiton Survey. Vital Health Stat **2(113).** 1992.

9. Johnson CL, Rifkind BM, Sempos CT, Carroll MD, Bachorik PS, Briefel RR, Gordon DJ, Burt VL, Brown CD, Lippel K, Cleeman JI. Declining serum total cholesterol levels among US adults. The National Health and Nutrition Examination Surveys. **JAMA 269(23).** 1993.

10. National High Blood Pressure Education Program. Working Group Report on Primary Prevention of Hypertension. National Institutes of Health. National Heart, Lung, and Blood Institute. October 30, 1992.

MONITORING THE NATIONS' HEALTH: THE 1991 HEALTH SURVEY FOR ENGLAND - RESULTS AND METHODOLOGY

Amanda White, Office of Population Censuses and Surveys

1. Introduction

The 1991 Health Survey for England[1] was commissioned by the Department of Health from the Office of Population Censuses and Surveys (OPCS). It is the first in a new series of annual health surveys designed to monitor trends in the nation's health. This series will form part of a wider programme of surveys related to health. The overall aim of the 1991 Health Survey was to obtain information on important aspects of health relevant to cardiovascular-disease (CVD) and nutrition including data which are currently unavailable - at least on a regular national basis - or available only from diverse sources and so cannot be used to examine combinations of risk factors and their effects.

This paper focuses on anthropometry and nutrition related health measures including blood cholesterol, blood pressure and nutritional biochemistries (haemoglobin and ferritin). It reports on the methodologies used to obtain these measurements and presents some of the main survey findings.

2. Survey design

The 1991 Health Survey for England was carried out between September and December 1991. The Survey aimed to select a nationally representative sample of adults (aged 16 or over) living in private households throughout England. This was the first official survey of this type to include people aged 65 and over and provides important new information on the health of elderly people.

Fieldwork was divided into two elements: an initial visit to the household by an interviewer and a second visit by a qualified nurse. At the first visit, each adult in the household was asked to participate in a health and socio-economic interview, have measurements .of their height and weight taken, and agree to a nurse visiting them at home. At this visit adults were asked to give details of any prescribed medicines they were taking, and to have their blood pressure, demi-span, and waist and hip circumferences measured. Those aged 18 or over were also asked to provide a blood sample for the analysis of total cholesterol, haemoglobin and ferritin.

The sample was selected using a multi-stage random probability design. A total of 3242 adults completed a full interview. This represented 81% of the eligible sample. Height and weight measurements were made for 78% and 77% of the sample. Demi-span measurements were made for 71%, waist-hip ratios for 70%, blood pressure measurements for 70% and 57% gave a sample of blood.

Comparison of the characteristics of the responding sample with those of the total population and other large surveys showed good general agreement, although there were indications that the 16-24 age group may have been under represented.

3. Anthropometric measurements

3.1 Methodology

An important aim of the survey was to measure obesity. Thus the following anthropometric measurements were included as part of the survey: height and weight, demi-span, and waist and hip circumferences (for the calculation of waist-hip ratio).

Height and weight were measured in order to calculate body mass index (BMI), weight(kg)/height(m)^2, a widely used measure of obesity.

Although it was possible to obtain height measurements for the majority of those aged 65 or over there were some problems in the pilot study with height measurements of those who were unsteady on their feet or who had physical disabilities. Our attention was drawn to an alternative measure of skeletal size which is less affected by such problems, the demi-span measure [2,3,4]. The demi-span is the distance between the sternal notch and the finger roots with the arm out-stretched laterally. The measure is relatively easy to perform, can be taken with the informant standing, sitting or lying down, and does not cause unnecessary discomfort or distress. Demi-span was included in the survey as an alternative measure of skeletal size, particularly for the elderly, and its use as an alternative to height in the assessment of obesity was examined.

The waist to hip ratio is an indicator of the distribution of body fat (adipose tissue) and especially of intra-abdominal fat deposits [5,6]. It is calculated as waist girth divided by hip girth, and on the Health Survey waist and hip circumferences were measured over light clothing.

It was decided that all measurements were to be taken in informants' homes. Therefore it was important to ensure that they were practicable under such circumstances. Additionally it had to be possible to train interviewers and nurses to obtain reliable and valid measurements. The following points were therefore essential to enable accurate, representative, national data to be collected. The equipment needed to be portable: for example a specialist contractor developed a portable stadiometer for measuring height. The

equipment also needed to be robust to enable transportation to and from the office and between fieldworkers. It was also important that the equipment was accurate to the required level. Scales and stadiometers were recalibrated before and during fieldwork to check their accuracy. Strict procedures and protocols had to be developed to ensure all measurements were taken in a standardised way. Widely used protocols, or protocols used in previous surveys, were adopted wherever possible to ensure comparability with other data sets and over time. For example, the WHO protocol for waist and hip circumference measurements was adopted, modified to be carried out over light clothing. Full details were collected by fieldworkers of any deviation from the protocol during measurement taking in order to assess the accuracy and validity of the measurements obtained. Training sessions were held for all measurement takers. Estimates were made of inter and intra measurer variability.

3.2 Results

Height and weight

The mean height of men in the 1991 survey was 174.2 cm (5ft 8½in), and of women 161.0 cm (5ft 3½in). The mean weight of men was 78.0 kg (12st 4lbs), and of women 65.9 kg (10st 6lbs).

Comparison with a previous survey, *(The heights and weights of adults in Great Britain'* carried out in 1980) shows evidence of a significant increase in the mean height of men and women aged 16-64 over the past decade. Among men in this age group mean height increased by 0.8 cm, from 174.1 in 1980 to 174.9 cm in 1991. Among women mean height increased by 1.0 cm, from 161.0 cm in 1980 to 162.0 cm in 1991. However, mean weight increased by more than height over this period, from 73.7 kg to 78.3 kg for men, and from 62.2 kg to 66.2 kg for women (Table 1).

Body mass index (BMI)

Mean BMI was 25.6 in men and 25.4 in women. More than half (53%) of men and 44% of women had a BMI of more than 25, and so were considered to be overweight or obese: 13% of men and 16% of women had a BMI of more than 30, described as obese.

There has been a significant increase in mean BMI in the 16-64 age group for both men and women since 1980. Mean BMI increased over this period from 24.3 to 25.5 for men, and from 24.0 to 25.2 for women (Table 1). Comparison with *The dietary and nutritional survey of British adults'* carried out in 1986/87 shows that over the last 4-5 years the proportion of adults aged 16-64 with a BMI of more than 30 has increased significantly - from 7% to 13% for men, and from 12% to 15% for women.

Demi-span

The analysis of the relationship between height and demi-span suggested that demi-span could be a suitable alternative to height as an indicator of skeletal size in the elderly, for whom it can be difficult to measure height. The mean demi-span of men (82.4 cm) was 7.5 cm greater than that of women (74.9 cm), and similar associations between demi-span and various population characteristics were found as for height.

waist-hip ratio

As expected mean waist-hip ratio was greater for men (0.901) than for women (0.794) and this difference was evident in all age groups. Among both men and women, waist-hip ratio increased with age. Among men it increased up until age 45-54 and then levelled off somewhat: among women there was a consistent increase through the whole age range (Figure 1).

4. Blood pressure

4.1 Methodology

Blood pressure was measured using the Dinamap 8100 monitor which is a portable, automatic monitor based on the oscillatory principle. The monitor measures systolic blood pressure, diastolic blood pressure, mean arterial pressure and pulse rate automatically at pre-selected time intervals.

Observervariance, movementartifacts and background noise are the main difficulties when conducting home examinations for a large population. Monitors based on the auscultatory principle detect Korotkov sounds by placing a stethoscope over the brachial artery whereas monitors based on the oscillatoryprinciplemeasureoscillations (i.e. vibrations) in the artery. Unlike monitors based on the auscultatory principle, monitors based on the oscillatory principle produce readings which are not affected by background noise. A fully automatic monitor allows, to a large degree, the standardisation of the procedure and minimises observer variance. It was therefore decided that a fully automatic monitor based on the oscillatory principle was more suitable than a manual or automated device based on the auscultatory principle for measuring blood pressure in this survey.

Three blood pressure readings were taken from each informant and the results presented are based on the mean of the

second and third measurements.

4.2 Results

Among all adults (including those on medications which might have affected their blood pressure) 16% of men and 17% of women were classified as hypertensive, that is their systolic pressure was 160 mmHg or more, and/or their diastolic pressure was 95 mmHg or more as measured on the survey. This included 13% of men and 11% of women classified as hypertensive untreated, that is they had raised blood pressure and were not currently taking any medication to control it: the remainder were taking such medication (Table 2).

5. Blood sample

5.1 Methodology

Blood sampling was carried out to measure three analytes,: total cholesterol, haemoglobin and ferritin. Cholesterol was measured because it is a risk factor for cardiovascular disease. Haemoglobin and ferritin were measured as indicators of iron status. A sample of 9.5ml of blood was taken from all adults aged 18 or over. Written consent was sought for the blood sample and if informants agreed their results were sent to their GPs. All procedures associated with obtaining the blood sample were tested rigorously prior to the main survey to ensure that they were safe and acceptable to informants, nurses and medical profession. Local Research Ethical Committee approval was sought and obtained from each of the areas involved in the survey and from the ethical committee of the British Medical Association.

The blood sample was taken in informants' homes by a nurse -using the vacutainer system. Strict procedures were specified to the nurses to ensure that all specimens were obtained in a standardised way.

The blood samples were sent to one central laboratory for analysis, and had to arrive within 48 hours of being taken. To have arranged local handling and storage facilities for the blood would have been difficult and expensive. It was decided that in a population survey of this type it was impractical to ask people not to eat or drink for several hours prior to the blood sample being taken. For these reasons the number of analytes was limited to the three listed above. For the 1992 and 1993 health surveys the blood is also being tested for fibrinogen, glycosylated haemoglobin, gamma GT and serum cotinine.

5.2 Results

***Serum** total cholesterol*

Mean serum total cholesterol concentration was 5.8 mmol/l in men and 5.9 mmol/l in women. Overall 69% of men and 70% of women had serum total cholesterol concentrations above the desirable level, that is 5.2 mmol/l. Four per cent of men and 9% of women had severely elevated cholesterol levels, that is 7.8 mmol/l or greater (Table 3).

***Iron** status: ferritin and haemoglobin*

Serum concentrations of ferritin were measured to assess iron stores. A value below 25 µg/l is generally considered to indicate low iron stores. Overall, 9% of men and 34% of women had serum' ferritin concentrations less than 25 µg/l. Among men this proportion did not vary much with age but among women it was much higher among the younger age groups: 47% for those under 45, falling to 31% among the 45-54 age group, and to 17% for those aged 55 or over. Mean ferritin concentration in men (121.1 µg/l) was more than twice as high as that for women (55.8 µg/l). Ferritin levels for both men and women rose steadily with age. Women over the age of 45, many of whom would have reached the menopause, had much higher mean ferritin levels (means between 59.6 µg/l and 91.4 µg/l) than those below this age (means between 29.7 µg/l and 37.8 µg/l) (Figure 2).

The normal range of haemoglobin concentration is 12.0 g/dl-17.0 g/dl for men and 11.0 g/dl-16.0 g/dl for women. Almost all men (97%) fell within the normal range for haemoglobin concentration: only 2% had levels below 12.0 g/dl and less than 1% had levels above the normal range limit. Among women, 96% fell within the normal range: 4% had haemoglobin levels below 11.0 g/dl and less than 1% had levels above the normal range limit. The mean haemoglobin level for men was 14.3 g/dl and for women was 12.8 g/dl. There was no relationship between haemoglobin concentration and age in either sex.

Table 1 Comparison of 1991 Health Survey results for height, weight and body mass index (BMI) with other surveys

England only: all aged 16-64

Measurement and survey	Men				Women			
	Mean	Median	Standard error	*Base*	Mean	Median	Standard error	*Base*
Height (cm)								
1991 Health Survey	174.9	174.9	0.21	1147	162.0	161.9	0.18	1318
Dietary and Nutritional Survey (1986/1987)	174.7	174.5	0.22	*1000*	161.8	161.8	0.20	996
Heights and Weights Survey (1980)	174.1	n/a	0.14 *	4042	161.0	n/a	0.11 *	4428
Weight (kg)								
1991 Health Survey	78.3	77.0	0.42	1159	66.2	63.0	0.42	1279
Dietary and Nutritional Survey (1986/1987)	76.2	75.2	0.36	1032	64.6	62.2	0.43	1021
Heights and Weights Survey (1980)	73.7	n/a	n/a	3935	62.2	n/a	n/a	4289
BMI								
1991 Health Survey	25.5	25.1	0.12	1141	25.2	24.0	0.16	1264
Dietary and Nutritional Survey (1986/1987)	25.0	24.6	0.11	998	24.7	23.6	0.17	995
Heights and Weights Survey (1980)	24.3	n/a	0.05 *	3884	24.0	n/a	0.08 *	4236

* Standard error for full sample, including Scotland and Wales.

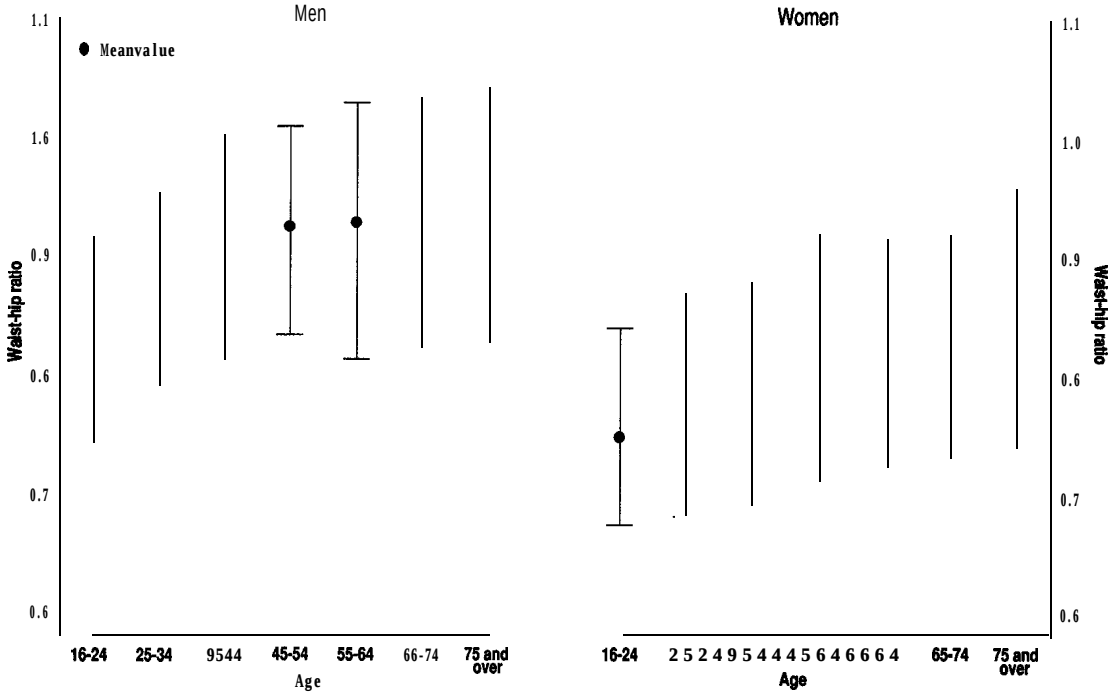
Table 2 Blood pressure level by age and sex

Blood pressure level	Age							Total
	16-24	25-34	35-44	45-54	55-64	65-74	75 and over	
	%	%	%	%	%	%	%	%
Men								
Normotensive (untreated)	98	95	91	79	58	45	29	76
Normotensive (treated)	1	1	2	6	13	18	24	8
Hypertensive (treated)	–			2	8	11	16	4
Hypertensive (untreated)	2	4	7	13	20	26	31	13
Base	163	245	230	189	171	152	*100*	1250
Women								
Normotensive (untreated)	100	97	93	86	58	39	22	76
Normotensive (treated)		1	3	6	14	16	24	8
Hypertensive (treated)	–	–		2	10	18	21	6
Hypertensive (untreated)	–	2	4	6	19	27	33	11
Base	176	266	268	207	205	175	126	1423

Table 3 Percentage distribution of serum total cholesterol by age and sex

Total cholesterol (mmol/l)	Age							Total
	18-24	25-34	35-44	45-54	55-64	65-74	75 and over	
	%	%	%	%	%	%	%	%
Men								
Less than 5.2	66	48	28	17	19	16	23	31
5.2 - less than 6.5	30	38	49	46	46	43	41	43
6.5 - less than 7.8	2	13	19	29	29	36	36	23
7.8 or more	3	1	4	7	6	5	–	4
Base	108	219	210	177	148	129	86	1077
Mean	4.8	5.3	5.7	6.2	6.1	6.2	5.9	5.8
Standard area of mean	0.10	0.07	0.07	0.08	0.09	0.10	0.11	0.03
	%	%	%	%	%	%	%	%
Women								
Less than 5.2	70	51	43	17	7	4	7	30
5.2 - less than 6.5	2A	40	43	50	38	32	37	39
6.5 - less than 7.8	6	8	13	26	41	35	42	23
7.8 or more		1	2	8	14	28	14	9
Base	*101*	236	229	187	*164*	145	102	1164
Mean	4.8	5.2	5.4	6.2	6.7	7.0	6.7	5.9
Standard error of the mean	0.09	0.06	0.06	0.09	0.09	0.11	0.13	0.04

Figure 1 Waist-hip ratio by age and **sex:** mean values and 5th and **95th** percentiles *



*The vertical lines indicate the range between the 5th and 95th percentiles

Figure 2 Mean **ferritin** by age and sax

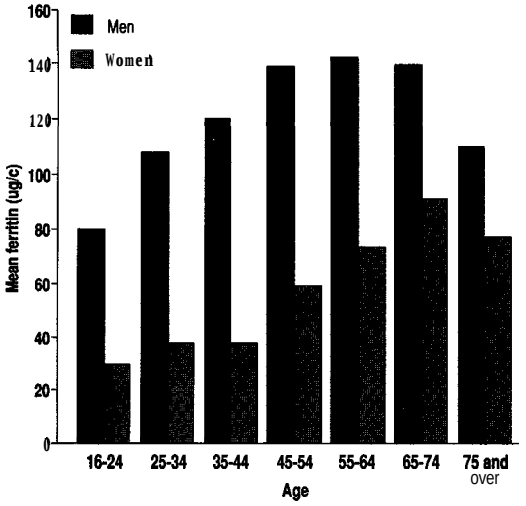
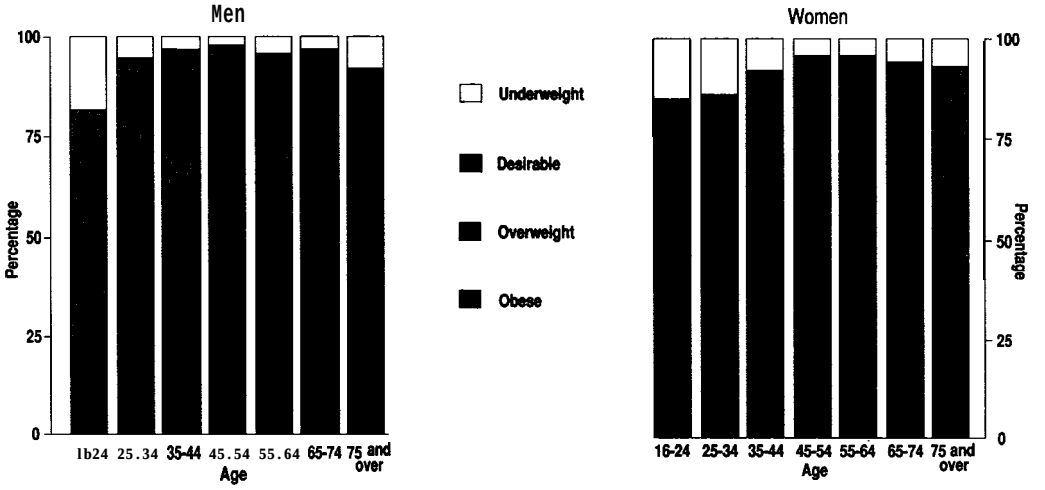


Figure3 Categories of body mass index (BMI) by age and sex



References

1. White A, Nicolaas G, Foster K, Browne F, Carey S. Health *Survey for England* 1991. London. HMSO, 1993.

2. Brown OT, Wigzell FW. The significance of span as a clinical measurement. In: *Current achievements in geriatrics.* Anderson WF (eds). London: Cassell. *See* pages 246-251.

3. Mitchell CO, Lipshitz DA. Arm length measurement as an alternative to height in nutritional assessment of the elderly. *Journal of Parenteral and Enteral Nutrition* 1982; 6 (3): 226-229.

4. Lehmann AB, Bassey EJ, Morgan K and Dallosso HM. Normal values for weight, skeletal size and body mass indices in 890 men and women aged over 65 years. *Clinical Nutrition* 1991;10.

5. Larsson B, Bjorntorp P, Suardsudd K, Wellin L, Wilhelmsen L, Tibbin G. Abdominal adipose tissue distribution, obesity and risk of cardiovascular disease and death: 13 year follow up of participants in the study of men born in 1913. *British Medical Journal* 1984; 288: 1401-8.

6. Lapidus L, Bengtsson C, Larsson B. Distribution of adipose tissue and risk of cardiovascular disease and death: a 12 year follow up of participants in the population study of women in Gothenburg, Sweden. *British Medical Journal* 1984; 289: 1257-61.

7. Knight I. *The heights and weights of adults in Great Britain.* London. HMSO, 1984.

8. Gregory J, Foster K. Tyler H, Wiseman M. *The dietary and nutritional survey of British adults.* London. HMSO, 1990.

Session G

Injuries and Violence

Toward the year

2000

Refining the measures

REPORTING EXTERNAL CAUSE OF INJURY AND POISONING CODES IN MEDICARE HOSPITAL DISCHARGE DATA

Harry L. Savitt and Alma McMillan, Health Care Financing Administration

INTRODUCTION

The National Committee on Vital and Health Statistics stated in its 1991 annual report that it is estimated that each year one out of every four persons in the United States is injured (1). About 100,000 persons die annually from unintentional injuries, which is the fourth leading cause of death, which accounts for 2.3 million years of life lost. An additional 45,000 persons die from suicide or homicide, which ranks the United States first among industrialized nations in violent death rates. For every death due to an injury, 16 persons are hospitalized and another 380 persons seek medical attention or have restricted activity due to an injury. The total lifetime cost of all injuries sustained in 1985 was estimated to be $200 billion. Productivity loses due to injuries are far greater than those for cardiovascular diseases and cancer combined. Trauma patients account for almost 7 percent of all health expenditures, and trauma care makes up approximately 12 percent, or $1 billion, of the overall cost of uncompensated care.

This paper reports on the frequency of E-code reporting nationally and by State. It will attempt to show the distribution of the more commonly reported causes of injury and poisoning using the E-code data recorded in the Medicare inpatient hospital file for calendar year 1990. Distributions of grouped E-codes are reported by categories and State. The billing form contains space for reporting a maximum of five (5) diagnoses. E-codes are reported on the billing form in diagnosis spaces 2, 3, 4, or 5, since the principal diagnisis code for the hospital stay must be in space 1. The number of diagnoses per bill, including the E-codes, are given. This paper also identifies the most frequently reported 3-digit and I-digit E-codes. The E-codes were analyzed to evaluate the association of the external causes of injury and poisoning with the principal diagnosis.

DATA AND METHODS

The data used in this paper are derived from a 20 percent sample of the 1990 Medicare Provider Analysis and Review (MEDPAR) file. This is an annual file of hospital discharges containing detailed accommodation and departmental charge data, days of care, diagnostic/surgical information, and beneficiary demographics for the aged and disabled. The sample of 2,128,147 million hospital discharge records contained 72,357 records with at least one E-code. Some records, about 7 percent, had more than one E-code reported, which resulted in 77,951 codes recorded. The sample number of E-codes was multiplied by 5 to estimate the universe of E-code reporting on Medicare inpatient hospital discharges recorded in the MEDPAR file in 1990.

External causes of injury and poisoning and their corresponding codes are listed in a supplementary section in the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) labeled "Supplementary Classification of External Causes of Injury and Poisoning" (2). The codes range from E800 through E999 and are referred to as "E-codes". This supplementary section in ICD-9-CM states that the E-codes "permit the classification of environmental events, circumstances and conditions as the cause of injury, poisoning and other adverse effects. Where a code from this section is applicable, it is intended that it shall be used in addition to a code from one of the main chapters of ICD-9-CM, indicating the nature of the condition." A small sample of records, with at least one E-code present (175), was randomly selected to ascertain the association of the E-code with a code 'from one of the main chapters of ICD-9-CM.

E-Code system

The external causes of injury and poisoning classification system consists of 193 3-digit categories of which 138 can be expanded with specified fourth digits, which produces 996 4-digit codes. The remaining 3-digit categories cannot be expanded, which produces 55 3-digit codes. Therefore, the E-code classification system contains 1,051 codes. For those codes with a fourth digit subdivision, the subdivision identifies:

1. the injured person (driver, passenger, pedestrian, etc.);
2. the place of occurrence (home, farm, residential institution, etc.);
3. the poisoning agent (drugs, iiquid substances, gases, etc.);
4. accidental falls (stairs, ladder, hole, etc.);
5. other accidents (types of fire, firearms, machinery, etc.);
6. adverse effects of drugs (antibiotics, sedatives, anti-depressants, etc.);
7. suicide and self-inflicted injury (drugs, gases, hanging, etc.);
8. homicide and injury purposely inflicted (rape, strangulation, submersion, etc.); and
9. injury due to war operations (bullets, fragments, gases, etc.).

An example of a 3-digit E-code is E882 - "fall from or out of building or other structure" - which does not have a fourth digit subdividion, and E883 - "fall into hole or other opening in surface" - which does have a fourth digit subdivision, such as, E883.1 - "accidental fall into well." In this sample, 608 of the 996 4-digit E-codes and 39 of the 55 3-digit E-codes were reported.

Eight of the 3-digit categories (E849, E929, E959, E969, E977, E989, E998, and E999) are designed to be used with other E-codes to denote late effects or place of occurrence. E-code grouped categories include motor vehicle, railway, water, and air and space accidents; accidental poisoning and falls; accidents caused by fire and submersion, suffocation and foreign bodies; adverse effects caused by drugs, medicinal and biological substances; suicide and self-inflicted injury; misadventures to patients during surgical and medical care; and homicide. Place of occurrence codes are included to denote the place where the accident or poisoning occurred.

E-code reporting

There is no national requirement to record E-codes and many hospitals do not routinely code external causes of injury on their billing forms or discharge abstracts. In addition, E-codes are not part of and do not effect the current Medicare reimbursement system. This lack of reporting may be attributed, in part, to the fact that the uniform hospital bill (UB-82) has room for only five diagnosis codes. This makes it less accommodating for recording E-codes which occupy space that might go to a reimbursable diagnosis. The UB-92, which will be implemented in October 1993, has room for nine diagnosis codes and an unique space for E-code reporting. Other reasons for not reporting E-codes include: incomplete medical records; inadequacies in the external causes of injury and poisoning classification system; lack of adequate rules, national guidelines and instructional materials; costs; potential liability issues; and inadequate index for coders.

Because the reporting of E-codes,is erratic, States are beginning to mandate the reporting of external causes of injury and poisoning for inpatient hospital discharges. Five states have mandated recording E-codes in their hospital discharge data systems: Rhode Island and Washington in 1989, Vermont, New York, and California in 1990. Arizona has mandated recording E-codes in hospital records. Colorado, Massachusetts, Virginia, and Wisconsin have encouraged voluntary reporting of E-codes.

Position of E-Codes recorded on UB-82 Billing Form

The billing form used by hospitals to bill for inpatient hospital services is referred to as the Medicare bill form 1450, also known as the UB-82. This billing form has spaces for five diagnosis codes. The first space is labeled for the principal diagnosis code and the other four spaces are labeled for other diagnosis codes. The principal diagnosis code is always entered in the space so labeled, since the Medicare reimbursement is determined by the principal diagnosis and specific comorbidities and complicating conditions and surgical procedures, if they occur. Specific comorbidities and complicating conditions, if they exist, are recorded in the four spaces labeled other diagnosis codes. Hospitals are aware which of the patients' comorbidities and complicating conditions affect the reimbursement and probably make every effort to report them accurately.

FINDINGS

It can be expected that the Medicare population, when hospitalized, would generally have comorbidities and complicating conditions. A review of 1,000 records with E-codes in the 1990 Medicare inpatient hospital discharge file showed the following:

5.0 percent had 2 diagnoses
5.0 percent had 3 diagnoses
14.5 percent had 4 diagnoses
75.5 percent had 5 diagnoses

Four or five diagnoses were reported in 90 percent of the records. Since E-codes are not required to be reported and they do not impact reimbursement, it can be expected that E-codes would not be routinely reported. The above referenced discharge file had 3.4 percent of the discharges with E-codes reported. The percent distribution of E-codes shown on hospital bills in spaces two through five are as follows:

24.4 percent had an E-code in space 2
25.9 percent had an E-code in space 3
25.4 percent had an E-code in space 4
24.3 percent had an E-code in space 5

The distribution of E-codes in spaces two through five is remarkably homogeneous. Since an E-code cannot be reported as a principal diagnosis, it can only appear in any of the four spaces on the bill that are reserved for other diagnoses. E-codes appear in spaces 2 through 5 equally, about 25 percent of the time in each.

Distribution of grouped E-codes

Almost half (46.1 percent) of the E-codes reported were associated with adverse effects of drugs, medicinal and biological substances. Accidental falls

accounted for over one-quarter (26.7 percent) of E-codes reported and abnormal reactions or later complications of surgical and medical procedures represented about one-tenth (9.5 percent) of the codes. "All other external causes of injury and poisoning" (8.6 percent) included accidents caused by fire, natural and environmental factors, submersion, suffocation, and foreign bodies (4,610); homicide and injury purposely inflicted by others (2,020); injury undetermined whether accidentally or purposely inflicted (2,010); and late effects of accidental injury (1,885). Place of occurrence, which is usually reported with accidents and accidental poisoning (E850-E869 and E880-E928), was reported 19,415 times.

Distribution of E-Codes by State

Every State and Puerto Rico reported some E-codes, ranging from 43,850 in New York to 210 in Alaska). Fifteen States accounted for two-thirds of all E-codes reported and generally included those States that require or encourage E-code reporting and have larger populations, hence more hospital admissions. The rate of E-codes reported by State varied from a high of 88.5 per 1,000 discharges for Vermont to a low of 10.1 for Minnesota, as compared to a national rate of 37.0. Rates for 8 of the IO States identified earlier, that require or encourage E-code reporting, were greater than the national rate, with Colorado and Rhode Island slightly below.

Distribution of E-Codes by categories

The 18 most frequently reported 3-digit categories represent over 80 percent of all the E-codes reported. Ten of the eighteen most frequently reported 3-digit categories are associated with adverse effects of drugs and medicinals and four with falls. The 10 categories of adverse effects of drugs and medicinals (E942, E933, E944, E935, E932, E939, E930, E936, E947, and E945) accounted for over 40 percent of all E-codes reported and the four categories of falls (E888, E885, E884, and E880) accounted for over 25 percent. Almost 10 percent of the categories reflect abnormal reaction of the patient, or later complication, caused by surgical operations or other procedures without mention of misadventure at the time of operation or procedure.

Distribution of E-Codes by 4-digit codes

The 25 most frequently reported 4-digit E-codes accounted for almost 50 percent of the 4-digit E-codes reported. The most frequently reported 4-digit E-codes are subsets of the 3-digit categories and basically follow the same order. Seventeen of the 25 most frequently reported 4-digit E-codes were associated with adverse effects of drugs and medicinals. (The fourth digit identifies the type of drug.)

E-Codes not reported

Of the 193 3-digit E-code categories listed in ICD-9-CM, all but 23 were reported in the 1990 MEDPAR file. Those that were not reported seem to be logically missing due to the types of causes of injury and poisoning they represent which are not expected in the Medicare population. Seven were related to legal intervention; 5 were associated with railway accidents; 4 were for injuries resulting from war operations; 3 were associated with homicide and injury purposely inflicted by other persons; 2 were related to accidents due to natural and environmental factors; and one each for water transport accidents and vehicle accidents not elsewhere classified.

Analysis of a small sample of records

In order to compare the association of the external causes of injury and poisoning with the principal diagnoses, a small sample of records from the 1990 MEDPAR file that contained at least one E-code was written off to include the state in which the hospital stay occurred, the principal diagnosis, and other diagnoses (including E-codes). The principal diagnosis is defined as the condition established after study to be chiefly responsible for occasioning the admission of the patient to the hospital for care. Other diagnoses are defined as all conditions that coexist at the time of admission or that developed subsequently, that affect the treatment received and/or 'the length of stay. Diagnoses that relate to an earlier episode which have no bearing on the current hospital stay are excluded.

A review of the principal diagnoses and the reported E-codes for a small sample of records showed that 17 states were represented in the sample. Seventy-five of the principal diagnoses were allocated to Injury and Poisoning, 50 of which were fractures; 40 to Diseases of the Circulatory System; 15 to Symptoms, Signs, and Ill-Defined Conditions;. and the remaining principal diagnoses distributed within respiratory (10), musculoskeletal (10), endocrine, nutritional and metabolic (10), genitourinary (5), and nervous system (5) diseases, and mental disorders (5).

DISCUSSION

The voluntary reporting by hospitals of external causes of injury and poisoning (E-codes) on approximately 3.4 percent of the 1990 Medicare inpatient hospital discharge file is encouraging. One would expect that E-code reporting would occur only in those few states that now require or encourage such reporting on all hospital discharge records. However, a review of those records with E-codes reveals that each of the 50 States, the District of Columbia, and Puerto Rico, reported some E-codes. The largest number of discharges reporting E-codes came from New York,

California, and Pennsylvania; the smallest number came from Alaska, Wyoming, and the District of Columbia. The rates per 1,000 discharges of E-codes reported were highest in Vermont, Maine, and Washington and lowest in Minnesota, Puerto Rico and Kansas.

Considering that the 1990 Medicare population consisted of approximately 30.9 million aged persons and 3.3 million disabled persons, it is not surprising that the largest numbers of E-codes were related to adverse effects of drugs, medicinal and biological substances in therapeutic use and accidental falls. Both of these causes are related to the aging process and disability problems. All of the grouped E-code categories were represented. As can be expected, certain grouped categories had small numbers reported such as: water transport accidents (155); air and space transport accidents (155); injury resulting from operations of war (60); railway accidents (25); and legal intervention (5).

The most frequently reported 3-digit categories and 4-digit detailed codes follow the expected external causes of injury and poisoning, i.e., adverse effects of drugs, medicinal and biological substances in therapeutic use and accidental falls. Other categories with significant numbers were surgical operations and procedures as the cause of abnormal reactions of patient, or of later complication, without mention of misadventure at the time of operation or procedure and suicide and self-inflicted poisoning by solid or liquid substances.

Almost 12 percent (23) of the 193 3-digit E-code categories were not reported in the 1990 MEDPAR file. It is no surprise that most of the unreported categories would not be encountered frequently. These were unusual types of railway accidents, legal interventions, and war operations.

E-codes were compared to the principal diagnoses in a small sample of records to ascertain if there was an association between the E-code and the principal diagnosis. In most cases there was a direct relationship between the E-code and the principal diagnosis. Fractures were associated with accidental falls and adverse effects of drugs were associated with principal diagnoses that would be treated by the types of drugs that were identified by the E-codes.

CONCLUSION

One can conclude the following:

- E-codes are being voluntarily reported on Medicare hospital bills;
- the E-codes that are being reported are valid, that is, in most cases the reported E-codes are associated with the principal diagnosis;
- almost 50 percent of the E-codes are due to adverse effects of drugs and over 25 percent are due to accidental falls;
- the large number of cases related to adverse effects of drugs suggests that follow-up studies should be carried out to determine how best to deal with drug reactions, interactions, and overdoses; and
- the large proportion of accidental falls also requires that greater effort should be directed at protecting those that are most vulnerable from these accidents.

REFERENCES

1. U.S. Department of Health and Human Services, The National Committee on Vital and Health Statistics, 1991. National Center for Health Statistics, DHHS Publication No. (PHS) 92-1205. Hyattsville, Maryland. May 1992.

2. U.S. Department of Health and Human Services, The International Classification of Diseases, 9th Revision, Clinical Modification. Health Care Financing Administration, DHHS Publication NO. (PHS) 80-1260, Vol. 1, Second Edition, September 1980.

EMERGENCY DEPARTMENT WEAPON-RELATED INJURY SURVEILLANCE

Victoria V. Ozonoff, Mass. Department of Public Health
Catherine W. Barber

INTRODUCTION

In recent years the field of public health has taken on the challenge of lowering the high rates of violent injury and homicide in this country. Many of the recently developed violence prevention programs, such as peer leadership programs, gun buybacks, and school-based mediation, are locally-based. Under the public health model, primary and secondary prevention programs like these ideally are driven in part by data. Data is used to describe the problem, to target risk groups, and to measure the impact of interventions. Many of us attending this conference are in the business of collecting and providing data for disease prevention. But where do we find sources of local injury data, and how can we develop these sources given the reduction in surveillance resources which many of us are facing?

Public health practitioners have often been frustrated by lack of adequate data. Mortality data are reasonably well documented, but we know that most violence does not result in death. Yet, data on non-fatal injuries are severely limited. Police data'on assaults represent only cases that come to the attention of law enforcement. Hospital discharge data typically lack E-codes, making it impossible to identify cause of injury; and outpatient data are rarely automated, making routine analysis impractical. Emergency department data has the potential for providing timely, population-based information on intentional injury at the local level if ongoing surveillance systems could be implemented.

At the Massachusetts Department of Public Health we at the Weapon-Related Injury Surveillance System have been exploring the use of emergency department data for ongoing injury surveillance. In this presentation, two injury reporting mechanisms which we have developed will be described. Both employ existing reporting requirements in order to limit expense and minimize burden on health care providers. The first is a paper reporting system for victims of gunshot wounds and stabbings, and the second is a computerized emergency department log which captures injury variables.

This presentation will describe both approaches, briefly give examples of data from both systems, discuss data quality and uses of the data, and end by reviewing implications for replication in other states.

DEVELOPMENT OF THE PAPER REPORTING SYSTEM

Massachusetts, along with at least 40 other states, has a mandatory reporting law for gunshot and stabbing injuries. It was this reporting requirement that provided the foundation for the first surveillance application. Medical providers are required to report all gunshot wounds and stabbings to the Massachusetts Department of Public Safety using short report forms. Reports were designed to assist in police investigations--not injury surveillance--and so lacked what public health practitioners might consider essential variables; In addition, the data were not automated, not routinely analyzed, and compliance was poor--about 30% on average in four hospitals at which we reviewed records.

Conversion of this reporting system into a weapon-related injury surveillance system required modifying the existing form to include information on injury intent, victim/offender relationship, location of injury, race and sex of victim, and precipitating circumstance. A copy of the modified form is sent to the Department of Public Health (MDPH), and the original continues to be sent to Public Safety.

Hospitals were actively recruited to participate in the new system: although reporting to police is mandatory, the expanded public health reporting is voluntary. Participation and compliance have been encouraged through active feedback to providers, staff training, workshops, and newsletters. Currently, forty-five hospitals--almost half of the EDs in Massachusetts--are now participating, and we anticipate a statewide system by the end of next year. Despite the conventional wisdom that the ED is no place for injury surveillance, no hospital to date has refused to participate.

The reporting forms are sent to the MDPH on a monthly basis and entered into a computer data base, permitting timely analysis of the data. Our success in recruiting hospitals is attributable to a few factors. Hospital personnel were aware of the reporting laws and did not object to replacing an old form with a more useful form. The additional variables are extremely limited. From the beginning, we felt that we could not reasonably expect high compliance from any report requiring more than a minute or two to complete. In addition, daily news reports of violence in our communities has made 'the topic compelling: providers were interested. Finally, we tried to "give something back" to EDs; generally this has meant staff training or data. In essence, we transformed a passive criminal justice reporting requirement into an active public health intentional injury surveillance system.

DEVELOPMENT OF THE COMPUTERIZED SYSTEM

Earlier I said that no hospital has refused to participate. That was a little misleading. One of the first hospitals we tried to recruit said, "We'd love to, but..." This was a mid-size community hospital whose ED director felt they were drowning in paperwork. He told us that if we could automate the ED log and integrate mandated reports into the system, the hospital would be happy to provide us with injury data. In this case, the existing data system was the ED log: the hospital was required to maintain a log according to the standards of the Joint Commission for the Accreditation of Healthcare Organizations. But because it was handwritten, the hospital could not use the log to answer basic questions about ED quality, services, management, or reporting.

We accepted the challenge and developed an automated log, now in use in two hospitals serving one geographic area. LOGBOOK, as the program is called, captures standard log information for all ED visits; for injuries, variables captured include injury cause, diagnosis, intent, job-relatedness,'and location. Cause is coded into collapsed E-code categories, and mandatory reports (gunshot wounds, dogbites,

and burns--for example) are integrated into the program. As was the case with the paper reporting system described earlier, we made the decision to keep injury surveillance variables to a minimum. In order to maximize compliance, the program is designed largely to serve ED management needs and patient care needs. Our assumption was that the simpler and more useful LOGBOOK was to ED managers --whose interest was not necessarily injury surveillance--the more effective it would be as an injury surveillance tool.

RESULTS: BOSTON

I will first report results from the database on gunshot wounds and stabbings. Almost half of the cases reported in 1992 were from the City of Boston. All eleven general emergency departments at Boston's acute care hospitals participated in the system. We were therefore able to approximate population-based rates for the city. A few examples of results from Boston are provided to show the kind of information available from the paper reporting system.

Table 1: 1992 Reported Cases of Gunshot Wounds and Stabbings per 10,000 Population, Treated by Boston Emergency Departments (City residents only)

	Cases 1990 Reported	Census Population	Crude Rate/10k
Boston residents, total	636	574283	14.6
Age Groups			
O-14	20	94361	2.1
15-19	la5	42714	43.3
20-34	460	208048	22.1
35-64	157	163966	9.6
65+	2	65152	0.3
Sex			
Male	698	275923	25.3
Female	113	296360	3.8
Race/Ethnicity			
Hispanic	76	61963	12.3
While	141	341142	4.1
Black	509	139692	36.4
Asian	a	29939	2.7
Race/Ethnicity and Sex			
Hispanic male	66	30964	21.3
Hispanic female	9	30999	2.9
White non-Hispanic male	118	163762	7.2
White non-Hispanic female	la	177390	1.0
Black non-Hispanic male	420	65236	64.4
Black non-Hispanic female	73	74456	9.8
Race/Sex and Age Group			
Black male, 15-24 yrs	213	12315	173.0
Non-black males, 15-24 yrs	78	46466	16.8

Age is missing in 12 cases, sex in 25 cases, and race/ethnicity in 95 cases. These cases are included in the Boston residents total. Sex is missing in 1 case for Hispanics, 5 cases for White non-Hispanics, and 16 cases for Black non-Hispanics. Cases falling under "Other" for race are not listed separately, but are included in the Boston residents total.

CASES: The eleven EDs in Boston reported a total of 1150 cases in 1992. Of these, 68% were stabbing injuries and 32% were gunshot wounds.

ARRIVAL AND DISPOSITION: About half of the cases reported to us arrived at the hospital by ambulance, and only about half were admitted. Thus, if we relied on fatalities, hospital discharge abstracts, and ambulance trip reports as data sources, we would miss a substantial proportion of the cases.

AGE/SEX/RACE: Most of the victims of gunshot wounds and stabbings were male--about 84%--and most were young. Seventy-eight percent were under the age of 35. The frequency of cases rose sharply in the mid-teens, and began to drop after the age of 25.

For victims who were residents of the city of Boston, we examined rates of injury by age group, as summarized in Table 1. Teens aged 15-19 were at highest risk with a rate of 43 cases per 10,000 teens, almost twice the rate for adults aged 20-34.

Males had victimization rates almost seven times higher than females, and this held true among Hispanics, non-Hispanic whites, and non-Hispanic blacks. However, black females were at slightly higher risk than white males, with a rate of 9.8/10,000 compared to 7.2 for white males.

The age/race/sex group at highest risk was young black males aged 15-24 years. Their injury rate of 173 stabbings and shootings per 10,000 youths was over ten times higher than that for other young males (16.8/10,000). In fact, this group made up only 2% of the Boston population but supplied 25% of the injury cases.

NEIGHBORHOOD: One of the strengths of the WRISS project has been our ability to provide local data for use by violence prevention programs. This will become more evident as the statewide system is developed: we will then be able to provide population-based rates for any local community. For the city of Boston, we have calculated injury rates by victim's neighborhood. We have used sip code to define neighborhood which for Boston roughly corresponds to the "popular" definitions. Neighborhood-specific maps such as that in Figure 1 were designed to aid violence prevention programs in identifying high risk areas.

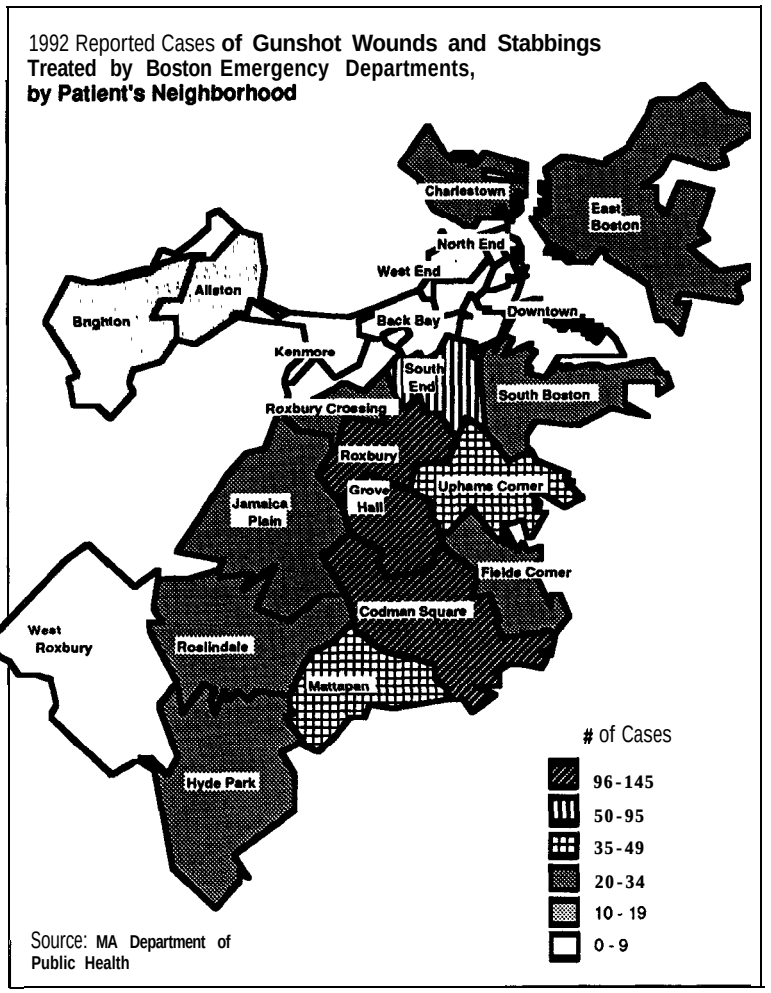
RESULTS: LOGBOOK

Results from Boston are based on the paper reporting system and therefore include only gunshot and stabbing injuries. LOGBOOK, on the other hand, provides information on all ED visits. It has been in use since 1991 at a community hospital which serves a low-income city of 70,000 residents and receives 40,000 ED visits annually. Among those visits for the year ending June 30, 1992, 27%--or 11,011--were injury visits. Injuries resulting from interpersonal violence accounted for about 7% of all injury visits, as indicated in Table 2.

Although gunshot wounds and stabbings may have the most serious outcomes, the vast majority of violent injuries were the result of bodily assaults (73%). Stabbings and other sharp instrument injuries were the second most common, accounting for about 10% of violent injuries.. Less than 1% were the result of gunshots.

SEX: In the Boston data, and elsewhere in the state, victims of stabbings and shootings

Flg . 1: Sample Neighborhood-specific Data Presentation Developed for Community Violence Prevention Group



1992 Gunshot Wounds and Stabbings Treated by Boston Emergency Departments, by Patient's Neighborhood

Neighborhood/Zip	# of cases	rate/ 10,000
Roxbury (02119)	117	46
Grove Hall (02121)	98	38
Codman Sq. (02124)	145	30
S End (02118)	61	28
Roxbury Crossing (02120)	29	20
Mattapan (02126)	46	17
Charlestown (02129)	22	15
Fields Corner 02122	29	14
Uphams Corner (02125)	41	13
Roslindale (02131)	33	10
S Boston (02127)	29	10
E Boston (02128)	31	9
Jamaica Plain (02130)	34	9
Hyde Park (02136)	22	9
W End (02114)	6	6
Allston (02134)	12	5
Bostondwnwn (02108-11, 02210)	6	5
N End (02113)	3	4
Brighton (02135)	15	4
Back Bay (02115-6)	16	4
W Roxbury (02132)	9	3
Kenmore (02215)	2	1
Boston, no zip	30	NA
Other MA city/town	202	NA
Out of State	19	NA
Homeless	15	NA
Unknown/Missing	78	NA

Neighborhood is according lo the patient's zip code. Rates/10,000 are based on the US Census estimates of the 1990 population by zip code area. Rates are not age-adjusted.

were overwhelmingly male: only 12-16% were female. However, the LOGBOOK data demonstrated that when we look at the wider range of non-weapon assaults, the percent of female victims increased to 41%.

AGE: The LOGBOOK data also points out that the age distribution of victims varies with the type of attack, In the age distribution for shootings and stabbings, there was a sharp peak between the ages of 15-24. The age distribution for sexual assaults (illustrated in Figure 2) shows an even younger victim group: 62% were under the age of 20. In fact, almost half were age 14 or younger, 'and, perhaps most troubling, 19% were age 4 or younger.

The MDPH sponsors an Adolescent Violence Prevention Coalition in the LOGBOOK community. We provide the coalition with local data to better understand the nature and scope of the problems they are addressing. Over time, the data will be vital in evaluating whether the Coalition has had an impact in lowering violent injury rates in the community.

COMPLIANCE AND- DATA ACCURACY

The paper and automated systems are monitored to assess data quality, accuracy and completeness. For the paper reporting system, compliance is measured by comparing reports filed with MDPH against actual cases identified during periodic record reviews by WRISS staff at each participating hospital. System compliance averages about 10%. Reported and unreported cases are entered in a database and statistically compared for systematic reporting bias. There has been no significant difference between reported and unreported cases on age or sex: 'however, gunshot wounds are more likely to be reported, and cases explicitly diagnosed as "gunshot wound" or "stab wound" are more likely to be reported than cases diagnosed as "lacerations."

For LOGBOOK, all ED cases are logged by the ED secretary, and data is transmitted electronically to the MDPH. The potential reporting problem is therefore not "missing cases" but incorrect coding of injuries. WRISS staff periodically compare a sample of the ED's injury medical records to the codes assigned in the LOGBOOK data. Coding accuracy exceeds 95% for each injury variable. We estimate that fewer than 10% of intentional injury cases are missed by the system. No apparent reporting bias is obvious although a more extensive analysis of data quality is planned.

DATA USES

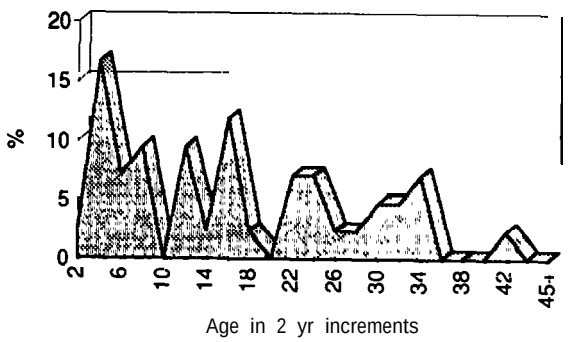
Our interest in developing ED surveillance was to create an ongoing, low cost system that could be used on the community level to develop violence prevention efforts. To that end, we have made the data available to local groups, providers, and public health violence prevention

Table 2: Injury Intent for Emergency Department Patients at LOGBOOK Site by Sex of Victim, July 1, 1991 - June 30, 1992

	Total		Females		Males	
	n	%	n	%	n	%
Unintentional	9716	88.2	3849	88.7	5840	88.0
Assault	731	6.6	250	5.8	478	7.2
Self-inflicted	143	1.3	72	1.7	71	1.1
Other	203	1.8	91	2.1	112	1.7
Unknown/Missing	218	2.0	78	1.8	139	2.1
TOTAL	11011	100	4340	100	6640	100

Total includes 31 victims for whom sex was not noted

Fig. 2: Sexual Assault-Related ED Visits to the LOGBOOK Sie by Age of Victim, 7/91-6/92



practitioners through timely newsletters, presentations, staff training, and workshops. We have been pleased to see several instances in which the data collected were directly used for program development. A few examples include:

-Boston City Hospital has used the data to assist in obtaining funding for a victim services program. Additionally, the data have been used to identify target prevention groups, estimate caseload, and schedule services.

-The Boston Violence Prevention Project has used the data to develop a portrait of Hispanic victims for service planning.

-In Lawrence, the data have been used by an adolescent violence prevention initiative to describe the victims of adolescent violence and to target programs.

-At least two Boston hospitals have used the data, and the data collection system, as a tool in resident training. Completion of the reporting form has been used to encourage residents to ask about intentional injury--questions they may have skipped in the past.

-Policy makers have also found the data useful: requests for data have come from the mayor's office in both Boston and Springfield. The governor's office has been provided with summary data, and injury control programs within the MDPH have used the data extensively.

DISCUSSION/CONCLUSIONS

What we have learned in Massachusetts is that violence-related injury surveillance at the emergency department level is feasible if existing reporting systems are adapted. Such systems can generate relatively accurate, population-based, local data. This has implications for other states trying to develop surveillance systems.

Most states already have laws requiring physicians to report all gunshot injuries (and in some states, stabbings and other attacks as well). In Massachusetts, we had the advantage that physicians were already required to fill out a short report form, whereas in most other states physicians notify police with a telephone call. Although modifying a form and improving compliance is doubtlessly easier than introducing an entirely new reporting form, the existence of mandated reporting itself is an advantage. Most ED managers were willing to cooperate with the system because they felt that weapon-related injuries were sufficiently grave to warrant special reporting. A less compelling issue may have met with less success. Additionally, reporting that is tied to action--in this case both feedback to providers and the development of violence prevention programs--is received with more enthusiasm than reporting that is seen as simply "bureaucratic" or "regulatory."

Although weapon injury reporting can provide extremely valuable community-level data, there are some disadvantages to this approach that should be mentioned. First, most violent injuries do not involve weapons and thirefore are not included in the system. Second, a paper reporting system, such as the one described here, relies on the memory and motivation of ED providers. An automated system, such as LOGBOOK, addresses these issues by directly connecting injury reporting to the ED log, a record which all EDs must maintain. Any hospital currently using a computerized log or planning to convert to one, should consider--and be encouraged by injury control advocates--to add injury variables that capture standardized cause-of-injury data in a retrievable format.

Despite the advantages of obtaining information on all injuries treated at the ED level, automation is years away for many hospitals. In the interim, a paper reporting system can provide extremely valuable local data, particularly if the system is simple, provides timely feedbadk, and has clear application to injury control projects. The data available on non-fatal violent injuries are so limited that we must take advantage of existing opportunities to develop surveillance systems at all levels.

SURVEILLANCE FOR NONFATAL FIREARM INJURIES

Roberta K. Lee, UTMB School of Nursing at Galveston
Robert G. Froehlke, Michigan State University College of Human Medicine
Mary J. Harris, UTMB School of Nursing at Galveston

Public Health Importance

Firearms are the second leading cause of injury mortality in the United States. In 1989, more than 34,000 Americans died due to the discharge of firearms.[1] In the United States, a child dies daily from handgun injuries. In fact, firearm injuries have become the leading cause of injury death in several states. In 1990, for example, firearms surpassed motor vehicles as the leading cause of injury mortality in Texas.[2] This trend is predicted to continue and firearms are expected to become the leading cause of injury mortality by the end of this decade.'

However, little is known about the contribution of firearms to nonfatal injury. This lack of information can be largely attributed to an absence of, or lack of uniformity in, firearm injury reporting and data collection. In order to estimate firearm injury incidence, it is necessary to include firearm injuries due to unintentional circumstances, as well as firearm injuries which are self inflicted (attempted suicide and suicide), other inflicted injuries (such as assault and homicide), firearm injuries inflicted during conduct of police work (legal intervention), and incidents where the circumstance can not be determined.

To measure the contribution of firearms to injury morbidity, we established a surveillance system for firearm injuries in Galveston County, Texas,[4] and, in several other locations, the National Electronic Injury Surveillance System (NEISS), was piloted as a potential surveillance system for firearm injuries. These firearm injury surveillance systems were then evaluated relative to established standards for surveillance systems.5

The Surveillance Systems

Galveston, Texas. To determine the incidence of medically treated or legally investigated firearm injuries, Lee[6] established a population-based surveillance system for firearm morbidity and mortality in Galveston, Texas. The case definition for this study was any resident who was injured due to the discharge of a firearm, with the injury being verified in the detailed reports of one of the following: police, hospital, vital registrar, or medical examiner. Cases where a firearm was used as a threat or a club were excluded.

This system was based on existing sources of data, gathered from medical, law enforcement, and judicial sources. (Figure 1) These data sources were the records of emergency medical services, emergency department records, hospital records, medical examiner records, vital records, and police records. Data were obtained retrospectively in order to exclude records being used in active criminal proceedings.

Data elements were chosen based on Haddon's framework of host, agent, and environmental factors applied to this injury problem. (Figure 2) We sought to develop a system which would address incidence rates and provide accurate, useful information to public policy-makers.

Prior to beginning the surveillance, the confidentiality and use of police and court records was discussed with the Galveston County District Attorney. He noted that the initial police record, which becomes part of the Uniform Crime Report of the Federal Bureau of Investigation, is a public document which is not protected by confidentiality laws. The detailed police reports, however, are confidential. After approval was obtained from the institutional review board, the police records were made available to us.

NEISS is a system developed by the Consumer Product Safety Commission (CPSC) to monitor injuries related 'to consumer products. The system is based upon a national, stratified probability sample of hospital emergency departments. From these sample hospitals, information is collected on persons who are treated in emergency departments for injuries related to consumer products. National estimates are then derived of the number and severity of these injuries. In 1990, The National Centers for Disease Control conducted a feasibility study to determine whether NEISS could be used to collect information about firearm injuries. An initial feasibility study was conducted at four hospitals-one from each of the four sampling strata of hospitals used by NEISS.'

The case definition for the NEISS study was any injury sustained as a result of the discharge of a firearm and initially treated in one of the four site hospital's emergency departments during the study period of July 1, 1990 to Sept. 30, 1990. Cases where a firearm was used as a threat or a club were excluded.

System Evaluation Procedures

Establishment of a "Gold" Standard. To evaluate the sensitivity of the surveillance systems at Galveston and the four NEISS sites, we needed a standard con-

FIGURE 1

GUNSHOT INJURY DATA FLOW CHART
GALVESTON COUNTY, TX

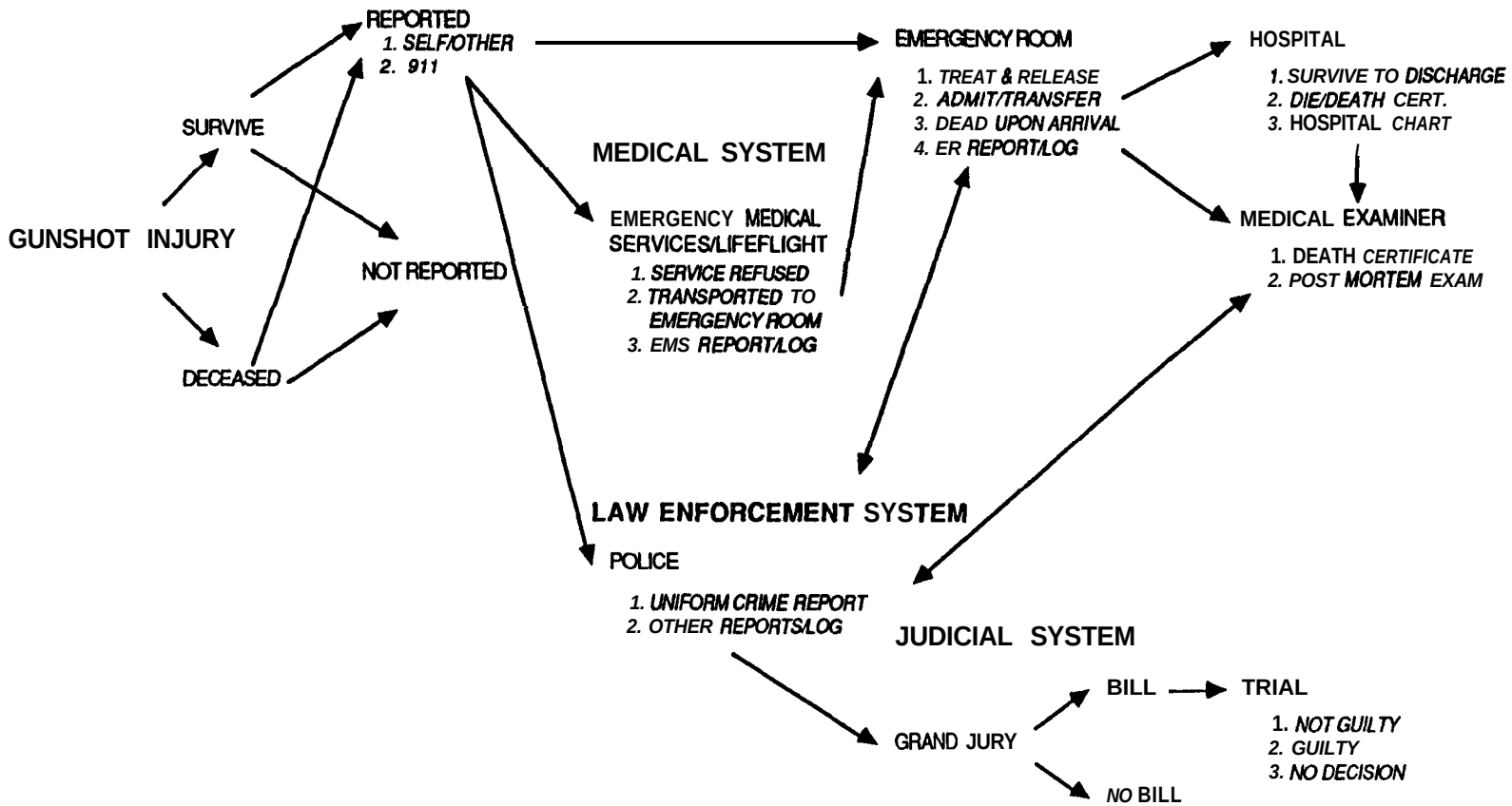
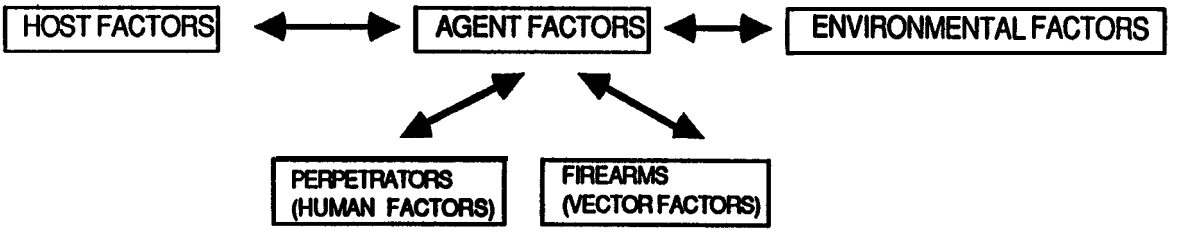


FIGURE 2

HOST, AGENT, AND ENVIRONMENTAL FACTORS IN GUNSHOT INJURY



PRE-INJURY	PSYCHIATRIC HISTORY TRAUMA HISTORY GUNSHOT HISTORY ALCOHOL POLICE HISTORY	PSYCHIATRIC HISTORY GUNSHOT HISTORY A L - POLICE HISTORY	WEAPON TYPE STORAGE LOCATION WHERE PURCHASED WEAPON HISTORY	TIME GEOGRAPHIC LOCATION RELATIONSHIP TO OTHER CRIME CIRCUMSTANCES OF INJURY
INJURY	BODY PART INJURED	WEAPON USED CALIBER #SHOTS FIRED	PROXIMITY TO HOST MALFUNCTION?	NUMBER PERSONS INVOLVED
POST-INJURY	INJURY SEVERITY SURVIVAL TIME	ARRESTED?		GEOGRAPHIC HOTSPOTS PLACE OF DEATH EMERGENCY RESPONSE POLICE RESPONSE
RESULTS	MEDICAL OUTCOME BY CAUSE LEGAL DECISION			

taining all of the true gunshot wound cases or, what is referred to as a "gold" standard. Since no such standard existed, one had to be established at each site. The method used to establish the standard was the same for each site and followed these steps:

1. Meeting and interviewing each site's emergency department administrators and staff in order to understand patient flow to and from that hospital's emergency department. Specific information was obtained about the management of patients with gunshot wounds.
2. Identifying all potential data sources and systems for patients treated in the emergency department.
3. Determining whether patients with a gunshot wound could be identified in these potential data sources.
4. Establishing the "gold" standard, based upon the compilation of all the data sources at each site which could identify a patient treated for a gunshot wound in the emergency department.

Different information and data sources existed at each of the five sites; thus each site's "gold" standard was specific to that site. Types of data and information systems used to establish the gold standard included:

1. Emergency department logs
2. Trauma registries
3. Billing information
4. Hospital admissions logs
5. Medical records searches
6. Outpatient clinic logs
7. Emergency department medical records reviews
8. Telephone consultations with neighboring hospitals.

Once the gold standard had been determined for each site, the standard was matched to the list of cases generated through. the surveillance system. Cases were considered to match if the medical record numbers matched and at least three of the following variables matched: 1. treatment dates (+/- one day), 2. date of birth, 3. race, 4. sex.

The medical records of cases that matched on all four criteria, but differed on medical record number, were reviewed to determine whether they were in fact a match. The medical records of all non-matches were reviewed to determine whether there had been any coding errors, and if in fact they were matches.

Results

Galveston, Texas. In Galveston, the sensitivity of the firearm surveillance

system varied according to the gunshot wound outcome (fatal or nonfatal). For fatal firearm injuries, we found that state vital records found 100% of firearm injuries, while both police and medical examiner records yielded 98% case ascertainment. For nonfatal firearm injuries, police records were the best identifier of firearm iniuries. with a sensitivity of 98%, follow&d by emergency room records. with a sensitivitv of 82%. The sensitivity of specific data elements also varied, with medical record sources being relatively complete for data such as age, race, and sex, but incomplete for information about the circumstance of the injury. Police records gave more complete information about the circumstance of the injury, including information about the perpetrator in assault and homicide cases. Table 1 shows the percent of records with information gathered from all sources--the EMS, hospital, police, medical examiner, and vital registry for the Galveston study. As can be seen, information about the injured person's age and race, the place the injury occurred, and the injury severity score were ascertained for almost all fatal and nonfatal cases. However, marital status, type of weapon, alcohol level and circumstances prior to the injury were hard to find for nonfatal injuries. No data source was very complete for data regarding the perpetrator in assault cases. Table 2 shows the percent. of records with information about the assailant in assaults and homicides. Again, even in fatal cases, we were unable to determine the extent to which alcohol contributed to firearm injuries.

TABLE 1. PERCENT OF GALVESTON GUNSHOT WOUND RECORDS WITH INFORMATION ABOUT THE PATIENT.

Information	Percent of Nonfatal Cases	Percent of Fatal 'Cases
Race/Ethnicity	100	100
Marital Status	52	89
Age	98	100
Occupation	57	97
Type of Weapon	81	93
Alcohol (Levels or Described)		
Police Record	7	76
Hospital Record	0	0
ME Record		76
Place Injury Occurred	94	98
Injury Severity Score	97	99
Circumstances Prior to Injury	71	90

TABLE 2. PERCENT OF GALVESTON GUNSHOT WOUND RECORDS WITH INFORMATION ABOUT THE ASSAILANT

For Assault/ Homicides:	Percent of Nonfatal Cases	Percent of Fatal Cases
Sex	92	93
Race/Ethnicity	87	93
Relationship-to Gunshot Case	86	88
Alcohol - Related	4	5

NEISS. During the study period at the 4 NEISS hospitals, a total of 538 gunshot wound cases were detected in the NEISS surveillance system. In our evaluation of the NEISS sites, we found 538 gunshot wound cases (our gold standard) during the same period, of which 520 were also in the NEISS system. Thus, the overall sensitivity of the NEISS surveillance system to detect gunshot wound cases was 96.7%. The sensitivity of NEISS varied by hospital, with one sensitivity value of 83%, and the remaining above 96%.

Discussion

The Galveston and NEISS studies indicate that several sources of data are potentially useful for the surveillance of firearm injuries. Vital records are relatively complete for cases having a fatal outcome. However, vital records may be miscoded if death certificates are not amended following the medical examiner's final description of the cases, or if the death occurred in a location not served by a qualified medical examiner's office.

In locations with a trauma registry, the trauma registry identified only those cases who survived long enough to be taken to a hospital and cases whose injuries were severe enough to warrant trauma care or hospitalization. Trauma registries excluded people whose gunshot injuries were immediately fatal and those with less serious wounds.

Police records were very good at identifying gunshot wound cases and included detailed information about the circumstances of these injuries. Medical records, especially emergency department records, were also very sensitive in identifying cases and describing the details regarding the pathology of these injuries; they were less detailed in describing the circumstances of the injury event. No source was sufficiently complete in providing information about the perpetrators of these injuries.

Recommendations

Based on this work, we recommend that public health professionals:

1. Work for mandatory reporting of all firearm injuries in all states. In forty states and the District of

TABLE 3. **STATES WITH GUNSHOT WOUND REPORTING LAWS**

STATE	REFERENCE	WHO REPORTS			TO WHOM
		MD	RN	HOSP	
AL	--------				
AK	ST §08.64.369	Y	Y	Y	Police
AZ	ST §13-3806	Y	Y	Y	Police
AR	ST § 12-12-602	Y		Y	Law Enf.
CA	ST §no 1, sect Penal 11160	Y		Y	Police
CO	ST § 12-36-135	Y			Police
CN	--------				
DE	ST TI 24 x 1762	Y		Y	Police
DC	ST §2-13613, 1362, 1363	Y		Y	Police
FL	ST § 790.24	Y	Y	Y	Police
GA	ST §31-7-11	Y	Y	Y	Police
HA	ST § 453-14	Y		Y	Police
ID	--------				
IL	ST §38§206-3.2	Y	Y	Y	Police
IN	ST § 35-47-7-1	Y		Y	Law Enf.
IA	ST §147.111	Y	Y	Y	Sheriff
KS	ST § 21-4213	Y			Police
KY	--------				
LA	--------				
ME	ST T. 17-A § 512	Y			Law Enf.
MD	CODE 1957, art. 27	Y	Y	Y	Police
MA	ST 112 § 12A	Y		Y	Police
MI	ST §750.411	Y	Y	Y	Police
MN	ST 626.52	Y	Y	Y	Police
MS	ST §45-9-31	Y	Y	Y	Police
MO	ST 578.350	Y	Y		Police
MT	ST § 37-2-302	Y	Y		Police
NB	ST §28-902	Y	Y	Y	Police
NV	ST §629.041	Y	Y	Y	Police
NH	ST §631.6	Y	Y	Y	Police
NJ	ST 2C:58-9	Y		Y	Police
NM	--------				
NY	PENAL § 265.25	Y		Y	Police
NC	ST § 90-21.20	Y		Y	Police
ND	§43-17-41	Y	Y	Y	Sheriff
OH	ST §2921.22	Y	Y		Police
OK	--------				
OR	ST §146.710-146.780	Y	Y	Y	Med Ex.
PA	ST §18 § 5106	Y		Y	Police
RI	ST § 11-47-48	Y	Y	Y	Police
SC	--------				
SD	ST 23-13-10	Y	Y	Y	Law Enf.
TN	ST § 38-1-101	Y	Y	Y	Law Enf.
TX	CIV ST Art. 447p	Y		Y	Police
UT	ST §26-23a	Y	Y	Y	Law Enf.
VE	ST §13.4012	Y		Y	Police
VA	ST §54.1-2967	Y			Police
WA	--------				
WV	Hospital Regulation			Y	Police
WI	ST §146.995	Y	Y	Y	Police
WY	--------				

Columbia, firearm injuries are specifically identified in existing state law as reportable events. The report usually is made to law enforcement personnel, rather than to public health officials. The states are listed in Table 3.

2. Work for mandatory use of ICD E-codes in hospital discharge data. The report Injury Control in the 1990s: A National Plan for Action recommends that E-codes be included in hospital discharge data and also recommends that federal and private health insurance systems require E-codes for reimbursement'.
3. Use multiple existing sources of data to ascertain cases of firearm injury. At a minimum, these should include vital, emergency department, and police records.
4. Keep the surveillance system "lean and mean." Do not collect unnecessary data. However, do include factors useful for planning intervention and prevention strategies and facilitating evaluation.
5. NEISS should be adapted to collect data on firearm injuries at all sites.
6. The development of creative programs to finance surveillance activities should be pursued. These programs might include taxes on guns and bullets, or fees for gun licenses.
7. The development of local surveillance systems for firearm injuries should be encouraged and facilitated.

Firearm injuries are affecting the lives of millions of Americans. Public health has the responsibility to monitor this problem, so that appropriate prevention and intervention strategies can be developed.

REFERENCES

1. U.S. National Center for Health Statistics. United States Detailed Mortality Data, 1989 (Public use data tape).
2. Zane DF, Preece MJ, Patterson PJ, Svenkerud EK. Firearm-related mortality in Texas (1985-1990). Texas Med 1991:87(11):63-5.
3. Cosby A. Presentation at second world conference on injury control, Atlanta, GA, May, 1993.
4. Lee RK. Incidence of fatal gunshot injuries occurring to Galveston County residents between 1979-1981. Dissertation Abstracts International 1989;50:04-B1352J(50004B).
5. Klaucke DN, Buehler JW, Thacker SB, et al. Guidelines for evaluating surveillance systems. MMWR 1988;37(S-5):1-18.
6. Lee RK, Waxweiler RJ, Dobbins JG, Paschetag T. Incidence rates of firearm injuries in Galveston, Texas, 1979-1981. A J Epidemiol 1991;134: 511-21.
7. Froehlke RG, Annest L, Conn J, Blank S, Mercy J, Lee RK. NEISS Firearm injury surveillance sensitivity: A report from CDC Division of Injury Control to CPSC. Atlanta, GA., 1992.
8. Waxweiler RJ, Rosenberg ML, Fenley MA. Injury control in the 1990s: A national plan for action. Report to the second world conference on injury control. Atlanta, GA: AAAM, May, 1993.

DEVELOPMENTS IN THE SURVEILLANCE OF CENTRAL NERVOUS SYSTEM INJURIES

D.J. Thurman, J.E. Sniezek, and D. Johnson
National Center for Injury Prevention and Control
Centers for Disease Control and Prevention

Abstract

Injuries to the central nervous system--the brain and spinal cord--frequently result in death or permanent disability. In the United States, traumatic head injuries are estimated to result in 500,000 hospitalizations and 75,000 to 100,000 deaths each year; 10,000 people may sustain traumatic spinal cord injuries. At present, national data are not adequate to define mortality, morbidity, and disability related to injuries of the central nervous system. The few available population-based studies on the incidence of head and spinal cord injury are difficult to compare because they differ in case definitions and types of data collected. Generalizing from these studies is difficult because most were conducted in limited geographic areas.

surveillance can provide the data needed to measure the impact of central nervous system injuries and to develop more effective methods for their prevention. The objectives of the Centers for Disease Control and Prevention (CDC) for surveillance of central nervous system injury address mortality, morbidity and disability. They are 1) to define incidence; 2) to identify groups at high risk; 3) to describe causes, circumstances, and mechanisms of injury; and 4) to develop and monitor prevention programs. To achieve these objectives, it is essential to establish standard case definitions and a standard minimum data set; these guidelines will allow for the compilation and comparison of surveillance information across jurisdictions.

A Drafting Group of CDC's Advisory Committee for Injury Prevention and Control has recommended modifications to a proposed case definition and minimum data set for traumatic head injury. A surveillance case definition for spinal cord injury has also been developed. The revised draft definitions and minimum data set are now being field-tested in three state and local health departments funded through cooperative agreements with CDC. After modifications based on these field tests are incorporated, final guidelines will be distributed to federal, state, and local agencies in 1994.

Backsround

Central nervous system injuries are the most likely type of injury to result in death or permanent disability. Only recently have we recognized that injuries to the central nervous system--the brain and spinal cord--are a major public health problem. Using the criteria by which we judge the public health importance of a health condition--the incidence and prevalence of the condition, its severity, its cost, and the degree to which it is preventable--we can discover the following facts about head and spinal cord injuries in the United States:

> Estimates of incidence indicate that head injuries may result in 500,000 hospitalizations and 75,000 to 100,000 deaths each year'. Over 10,000 people may be hospitalized with spinal cord injuries each year[3]; reliable estimates of the numbers of spinal cord injury deaths have not been published.
>
> The severity of the nonfatal injuries is shown by estimates that each year 70,000 to 90,000 people sustain head injuries resulting in permanent disability'. Spinal cord injuries, though fewer in number, are even more likely to be severe: among those who survive, most have irreversible neurologic impairments.
>
> The costs of central nervous system injuries--acute care, rehabilitation, chronic care, and indirect costs--are unknown but certainly enormous. One estimate suggests that head injuries impose an annual economic burden of $25 billion in direct and indirect costs[7]. The aggregate annual costs of spinal cord injuries exceed $6 billion'. There are no adequate measures of the personal and social costs of pain and suffering to injured persons and their families.
>
> These injuries are in large measure preventable. The leading causes of central nervous system injuries are motor-vehicle crashes, firearms, and falls. Some methods of prevention exist for greatly reducing the occurrence or severity of these injuries; others are yet to be developed.

Clearly, all of these criteria--incidence, severity, cost and preventability--indicate that central nervous system injury is a serious public health problem. Recognizing this, in 1989 the federal Interagency Head Injury Task Force recommended that traumatic brain injury be established as a category for public health surveillance?. The task force also recommended designating a lead federal

agency to plan and coordinate prevention programs. The Centers for Disease Control and Prevention (CDC) has since been assigned to this role. The surveillance of central nervous system injury is a CDC priority.

Surveillance, which is the foundation of effective prevention programs in public health, is defined as the ongoing systematic collection, analysis, and interpretation of health data necessary for planning, implementing, and evaluating public health programs. Linkage to injury prevention and control efforts is an essential part of injury surveillance. Surveillance is limited to obtaining basic information concerning the health events in question. It is not synonymous with epidemiologic research, which involves more focussed study with extensive data collection for a limited period and is designed to test specific hypotheses concerning etiology. Surveillance is important in order to

- determine the magnitude of injury mortality and morbidity (incidence and severity),
- identify populations at higher risk,
- identify risk factors, causes, and mechanisms of injury occurrence, and
- guide the development and monitor the effectiveness of injury control programs.

Cur understanding of the current incidence, severity, risk factors and causes of central nervous system injuries is limited. Recent national data have been published concerning the incidence and causes of mortality due to head injuries[5,6] but not spinal cord injuries. Published data concerning the incidence and causes of morbidity among persons hospitalized for head and spinal cord injury are limited[3]. These data are based mainly on information collected in a few communities more than ten years ago.

There are other major limitations of existing epidemiologic and surveillance data. Case definitions of head injury and spinal cord injury differ from one study to another. Furthermore, the kinds of information collected vary among surveillance systems. This lack of consistency makes it difficult to compare data and makes it hard to generalize from individual studies.

To improve the surveillance and prevention of central nervous system injury, we now need to develop standard case definitions and a standard minimum set of data elements for use by public health and prevention agencies across the country.

The Development of Central Nervous System Injury Surveillance Guidelines

Work on surveillance guidelines for central nervous system injury began with a 1988 resolution of the Council of State and Territorial Epidemiologists (CSTE) advocating spinal cord injury surveillance[4]. Working with CSTE, CDC developed a case definition spinal cord injury in 1988 and drafted a case definition for head injury in 1989. The CDC also proposed sets of surveillance data elements for each type of central nervous system injury. These early guidelines--definitions as well as data sets--were distributed to state health departments and agencies developing injury surveillance systems. CDC provided technical support, grants and cooperative agreements for the surveillance of central nervous sytem injury in several states. As of 1993, 27 states have implemented surveillance of spinal cord injury, and 16 states have implemented head injury surveillance.

Surveillance is largely dependent upon existing sources of data that are usually collected for other purposes. Many agencies collect data that may be useful for the surveillance of central nervous system injuries. These include federal, state, and local agencies for public health and public safety; medical care and rehabilitation facilities; and academic institutions. Therefore, these agencies need to be involved in the development of standards for the surveillance of central nervous system injury.

Responding to the recommendations of the Interagency Head Injury Task Force and recognizing the interest that many agencies may have in data collection, CDC's Advisory Committee for Injury Prevention and Control appointed a Drafting Group to help develop standard reporting guidelines for head injuries. This Drafting Group, which met in November 1992, included delegates from multiple agencies having interests or responsibilities concerning the surveillance of central nervous system' injury. CDC submitted an amended draft of the guidelines to the group for its review. The changes were based on experience in state health departments as well as the advice of a wide range of public health and health care professionals. The Drafting Group received oral and written comments concerning the amended draft guidelines from many organizations. These organizations represented multiple perspectives concerning the surveillance of central nervous system injury'.

'These organizations were represented on the head injury surveillance guidelines Drafting Group:

CDC Advisory Committee for Injury Prevention and Control, American Academy of Physical Medicine and Rehabilitation,

On the basis of the decisions of the Drafting Group, CDC further revised its proposed guidelines for central nervous system surveillance. The new draft guidelines contain case definitions and a common data set addressing both head injuries and spinal cord injuries. These guidelines were distributed in March 1993 to state health departments conducting surveillance of central nervous system injury. Through CDC cooperative agreements, three jurisdictions, New York State, Rhode Island, and Maricopa County, Arizona, began field testing these guidelines.

In early 1994, the Drafting Group will review the guidelines and recommend final revisions based on the results of the field tests. Subsequently, CDC will publish and distribute the guidelines to all federal, state, and local agencies conducting surveillance and investigations on the occurrence of head injuries.

Description of Guidelines for surveillance of Central Nervous System Injury

We can paraphrase the current draft definitions of head injury and spinal cord injury as follows:

A case of **head injury** (traumatic brain injury) is defined as the occurrence of at least one of the following conditions, which has been attributed to head trauma through medical evaluation:

- observed or self-reported loss of consciousness,
- observed or self-reported retrograde amnesia or post-traumatic amnesia,
- skull fracture,
- objective neurologic abnormalities determined from neurologic examination,
- diagnosed intracranial lesions determined from radiological examination or other neurodiagnostic procedures, or
- observed seizure acutely following head trauma in the' absence of a nontraumatic seizure cause.

A case of head injury is also defined as the occurrence of death from trauma for which the death certificate, autopsy report, or medical examiner report lists head injury in the sequence of causes.

A case of spinal cord **injury** is defined as the occurrence of an acute traumatic lesion of neural elements in the spinal canal resulting in a temporary or permanent sensory deficit, motor deficit, or bowel or bladder dysfunction.

Most jurisdictions identify cases of central nervous system injury from death certificate data and hospital discharge data. Additional sources of data may also be useful for case identification. These sources include hospital trauma registries and medical examiner records. Hospital discharge data are usually computerized data sets with diagnoses coded according to the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM)[2]. The draft head injury clinical case definition corresponds to ICD-9-CM nature-of-injury codes 800.0-801.9 (fracture of the vault or base of skull), 803.0-804.9 (other and multiple fractures of skull), and 850.0-854.1 (intracranial injury). The draft spinal cord injury clinical case definition corresponds to the ICD-9-CM nature-of-injury codes 806 (fracture of vertebral column with spinal cord lesion) and 952 (spinal cord lesion without evidence of spinal bone injury).

For each case of central nervous system injury identified, the guidelines recommend collecting the following types of 'data:

Demographic Variables

- Age
- Sex
- Race or ethnicity

Council of State and Territorial Epidemiologists, National Head Injury Foundation, National Institute of Neurological Disorders and Stroke (National Institutes for Health), Oklahoma State Department of Health, University of Texas at Houston School of Public Health, and National Center for Injury Prevention and Control (CDC).

These organizations also provided comments or were consulted by the Drafting Group:

U.S. *Department of Health and Human Services*--Health Care Financing Administration, Health Resources and Services Administration, Social Security Administration, National Center for Health Statistics (CDC), and National Institute for Occupational Safety and Health (CDC),

Other U.S. government agencies--Department of Defense, Department of Veterans Affairs, National Highway Traffic Safety Administration (Department of Transportation), National Institute on Disabilities and Rehabilitation Research (Department of Education), and 'National Council on Disability,

State government agencies--Colorado Department of Health, Oklahoma.State Department of Health, and Missouri Head Injury Advisory Council,

Academic, professional and other organizations--American Associationof Neurological Surgeons, American College of Emergency Physicians, American Medical Association, American Public Health Association, University of Alabama at Birmingham Injury Control Research Center, University of California at Los Angeles Injury Control Research Center, University of Texas at Galveston Medical Branch, and JMA Foundation.

- State and county of residence

Injury Circumstance Variables

- State and county of injury
- Date of injury
- Work-relatedness
- Intentionality
- E-code (ICD-9-CM external-cause-of-injury code)
- Etiology
- Position in vehicle (if applicable)
- Personal protective equipment used (if applicable)
- Alcohol involvement

Nature of Injury **Variables**

- Type of injury (head, spinal cord or both)
- N-codes (ICD-9-CM nature-of-injury codes)
- Clinical severity indicators (existence of brain lesions, altered consciousness, neurologic and mental status abnormalities)
- Abbreviated Injury Score
- Glasgow Outcome Score
- Discharge disposition

Other Variables

- Case identification number
- Hospital admission date
- Hospital discharge date
- Payment source

Some of these variables are easily obtained from hospital discharge data or death certificates. Other variables, especially those that address severity and circumstances of injury in detail, can only be obtained from additional sources of data. Such additional information may be found in medical records, medical examiner records, police accident reports, and other records. Collecting more detailed data from noncomputerized sources. may be time consuming and thus feasible for only a sample of reported cases. Some jurisdictions with limited resources may only be able to collect basic information found in hospital discharge data and death certificates.

CDC proposes a standard computer database format for collecting and reporting these data. The format is compatible with the National Electronic Telecommunications Surveillance System (NETSS) currently used for reporting cases of communicable diseases to CDC. Reports of central nervous system injury thus can be transmitted to CDC, where data from multiple jurisdictions can be analyzed and compared.

Conclusions

Using standardized case definitions and data elements, CDC is promoting the development of surveillance systems for central nervous system injury in multiple jurisdictions in *the United States. The collection and analysis of national data will improve our understanding of the incidence, causes, risk factors, severity, and outcome of central nervous system injury. This understanding will provide a rational basis for the design and focus of more effective injury prevention interventions. Such interventions may include educational and legal measures to promote safer behaviors, as well as engineering and environmental changes to reduce possibilities of injury. By monitoring continued trends in the occurrence of head injury, the effectiveness of these programs in reducing the incidence and severity of injury can be evaluated.

References

[1] DeVivo MJ. The cost of spinal cord injury: a growing national dilemma. In: Apple DF Jr. and Hudson LM, editors. Spinal Cord Injury: The Model. Atlanta, Georgia: Shepherd Center for Treatment of Spinal Injuries, Inc., 1990; 109-113.

[2] International Classification of Diseases, 9th Revision, Clinical Modification, Third Edition. Washington DC: U.S. Department of Health and Human Services, 1989.

[3] Kraus JF. Epidemiologic features of injuries to the central nervous system. In: Anderson DW, ed. Neuroepidemiology: A Tribute to Bruce Schoenberg. Boca Raton: CRC Press, 1991; 334-353.

[4] National Committee for Injury Prevention and Control. Injury Prevention: Meeting the Challenge. New York: Oxford University Press, 1989: 290.

[5] Sosin DM, Sachs JJ, Smith SM.' Head injury-associated deaths in the United States from 1979-1986. JAMA 1989; 262: 2251-2255.

[6] Sosin DM, Nelson DE, Sacks JJ. Head injury deaths: the enormity of firearms. JAMA 1992; 268:791

[7] U.S. Department of Health and Human Services, Interagency Head Injury Task Force Report, 1989.

Session H

Measuring the Health and Disability of Older Americans

Toward the year

2000

Refining the measures

HEALTH EXPECTANCY: THE CANADIAN EXPERIENCE

Russell Wilkins, Statistics Canada

In Canada, we have been publishing results of health expectancy calculations since the early 1980s,[1-3] and participating in the work of the International Network on Health Expectancy (REVES) since its inception in the late 1980s.[4-6] There is currently a fairly widespread recognition of the potential utility of this kind of index, as well as a growing awareness of health expectancy terminology, concepts and measures.[7-9]

We have produced calculations at a variety of geographic levels,[10-16] disaggregated the results by important socioeconomic dimensions,[2,11] and paid some attention to the causes and origins of activity limitations,[3,16-18] each of which has added to the policy utility of the work. Explicit weighting of health states[19-21] and modelling of health determinants[16,22-23] have been pursued.

Nevertheless, much of what was originally expected of health expectancy indices has not yet been accomplished. In Canada, the most serious disappointment is the absence of health expectancy time series based on comparable data,[13,24-25] followed by a lack of international comparability of the calculations[26] which have been produced--each of which has reduced the policy utility of the work. Also, because of a lack of suitable longitudinal data sets, no calculations for Canada have been based *on* observed transition probabilities. However, we do the best we can to exploit whatever cross-sectional data are available, using a modified version of the Sullivan method.[1-3,10-15,27-30]

We calculate the expected number of years of life in each of several discrete health states. Sometimes the expected years in each health state are explicitly weighted and then summed into an index of "health-adjusted life expectancy" (HALE), which can be useful for comparisons across socioeconomic, regional and disease groupings. However, we feel that it is important to show the underlying expected years in each state prior to weighting, since a single summary index provides too little information for many practical purposes (such as long-term care services planning), and of course, not everyone will agree on the weights, regardless of how they were chosen.

Generally speaking, we have found that health expectancy calculations are useful for four different purposes:

- *Helping to shift the focus of outcome indicators from exclusive reliance on mortality-based measures to also include impairment, disability and handicap.* This enlarges the scope of interest from survival to independence and autonomy.
- *Revealing the true extent of health inequalities by sex and among different regions and socioeconomic groups.* When health status is considered as well as mortality, the differences can either increase or diminish.
- *Redefining priorities among disease entities.* Rankings by cause of death or loss of life are not the same as rankings by cause of disability or loss of health expectancy.
- *Shedding light on additional options for improving health, including non-medical options.* In terms of health expectancy, progress can be made not only by preventing or treating the occurrence of disease, but also by reducing its consequences. Interventions can include stopping or slowing the progression from disease or injury to impairment, from impairment to functional limitation or disability, and from disability to handicap.

Health expectancy of the elderly

Most recently, Owen Adams and I have been using the results of Canada's Health and Activity Limitation Survey (HALS) of 1986-87 to examine the health expectancy of the elderly. The survey included residents of health-related institutions as well as ordinary households.

We classified the population into six exclusive groups, which were defined so as to form a scale representing increasing degree of autonomy and independence. From most to least dependent, the categories were as follows:

- Institutionalized. Resident in a nursing home or other long-term care facility.
- *Heavily dependent.* Not institutionalized but dependent on others for personal care (such as eating, bathing, dressing, toileting) or for moving about within the home.
- *Moderately dependent.* Dependent on others for going out, light housework, or meals.
- *Somewhat dependent.* Dependent on others for heavy housework, groceries, or partially dependent in any of the above activities.

■ *Disabled but not dependent.* Neither dependent nor partially dependent in any of the above activities, but at least one **"yes"** to any of the HALS disability screening questions.

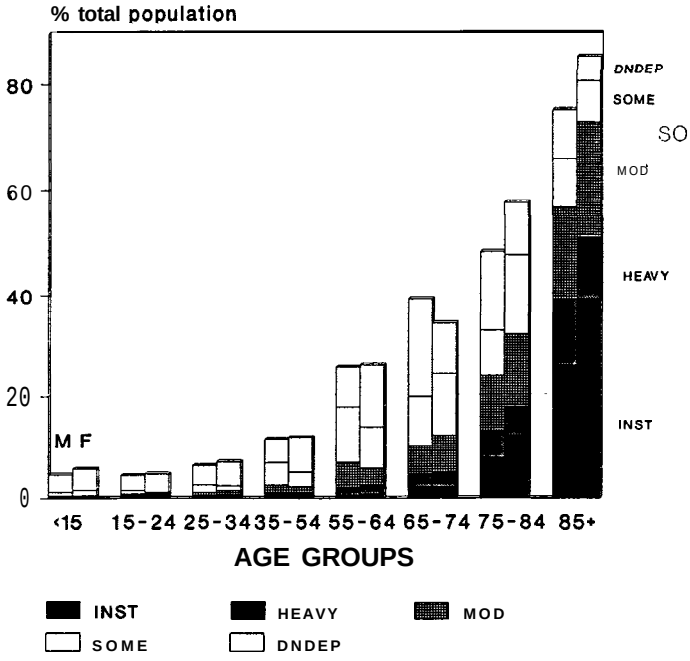
■ *Not disabled.* **"No"** to all of the HALS disability screening questions (and not dependent or partially dependent in any of the above activities).

Note that the categories are hierarchical. For example, a person who was dependent for personal care might also be dependent for shopping or housework, but would be classified only to the highest applicable degree of dependency category (heavily dependent). Persons were considered to be dependent if, for *reasons of health,* someone else helped them with any of the specified activities.

The cross-sectional prevalence rates of disability at each age and by sex are shown in Figure 1. Note the high rates of institutionalization at advanced ages, particularly for women. Without the institutional population, the rates of disability would falsely have appeared to increase much more slowly, since persons with the most severe disabilities would have been selectively excluded.

Figure 1

DISABILITY AND DEPENDENCE
CANADA, 1988



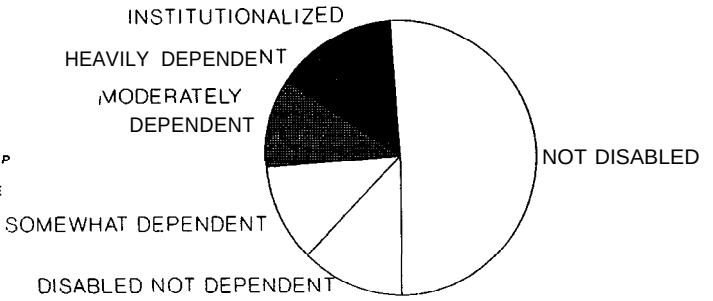
We then multiplied the expected number of life years lived within each age group (from the LL column of an abridged life table) by the corresponding rates of each category of disability, yielding the expected life years in each category. To obtain the expected years of life remaining in each health state at age 65, we summed the expected life years in each state at ages 65 and over, and then divided those sums by the number of survivors at age 65.

The results of these calculations are shown in Figure 2. At age 65 for both sexes together, Canadians could look forward to living another 17 years in all states of health, including almost 9 years disability-free, plus 2 years disabled but not dependent, about 4 years somewhat or moderately dependent, and 2.5 years heavily dependent or institutionalized.

Figure 2

HEALTH & INDEPENDENCE EXPECTANCY AT AGE 65, BOTH SEXES
CANADA, 1988

TOTAL OF 17.2 YEARS OF EXPECTANCY REMAINING



The six discrete states defined in terms of degree of dependence were then aggregated into four overlapping states of **"autonomy"** and **"independence",** which cumulated from least dependent to most dependent as follows:

■ *Disability-free.* **"No"** to all of the HALS disability screening questions.

■ *Independent.* Disability-free or disabled but not dependent.

■ *Autonomous.* Independent, somewhat dependent or moderately dependent.

■ *Total.* Autonomous, heavily dependent or institutionalized (all states of health).

Figure 3 shows the percentage of men and women surviving to age 65 and over in each of those four overlapping states of autonomy and independence. The topmost curve is for all states of health, the second highest curve is for autonomous, the third is for independent, and the Lowest curve is for disability-free. The area under each curve represents the expected life years in each state.
ing any of the three lower curves closer to the survival curve would represent progress against disability and/or handicap, and would result in a longer period

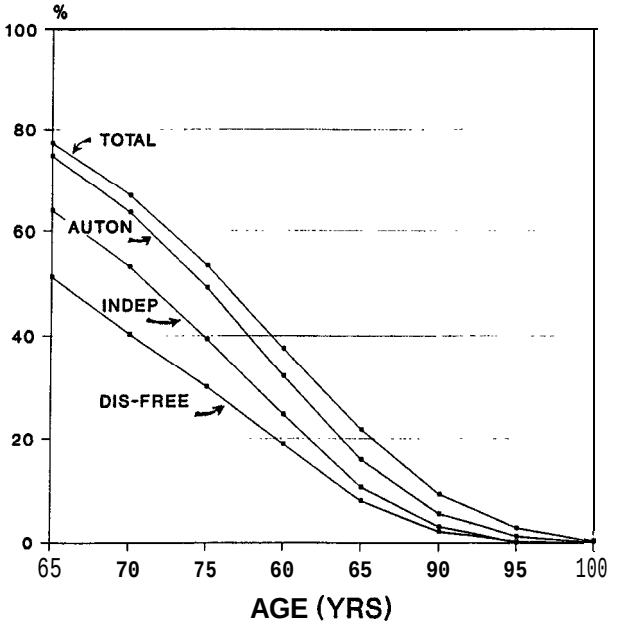
omous life (depending on which curve was

which one-quarter were years of independent life, and one-half were years of autonomous life.

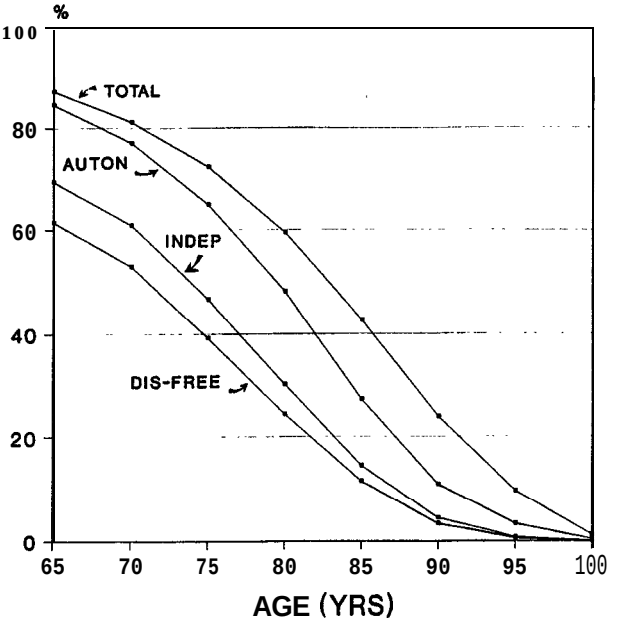
The disparity between men and women was much greater in terms of total life expectancy than in terms of autonomous, independent or disability-free life expectancy. While senior women in Canada could expect to live nearly 4 1/2 years longer than men, they could expect only 2 1/2 more years of autonomous life, and only 1/2 year more of independent life.

Health expectancy calculations also helped to reveal the true extent of socioeconomic and regional inequalities in health outcomes. Autonomous, independent, and total life expectancy was high-

SURVIVAL, AUTONOMY AND INDEPENDENCE
OF SENIOR MALES AGED 65+, CANADA, 1986



SURVIVAL, AUTONOMY AND INDEPENDENCE
OF SENIOR FEMALES AGED 65+, CANADA, 1986



Health expectancy calculations provide a useful summary of the health conditions of older persons. At age 65, the average remaining life expectancy of Canadian seniors (men and women together) was 17 years. Only half of those years were expected to be completely free of disability, but nearly two-thirds were years of independent life, and 85% were years of autonomous life. At age 75, remaining life expectancy was nearly 11 years, half of which were years of independent life, and three-quarters were years of autonomous life. At age 85, remaininq life expectancy was 6 years, of

est in the richest income quintile group, and lowest in the poorest quintile group. For both sexes together at age 65, the differences between the quintiles were 1 year for autonomous life, 1.7 years for independent life, and 2.3 years for disability-free life.

The differences between Quebec (predominately French-speaking) and the rest of Canada (predominately English-speaking) ***diminished*** when disability-free or independent life expectancy was considered rather than total life expectancy. However, the differences between the Atlantic provinces and the rest of Canada

increased when disability-free or independent life expectancy was considered rather than total life expectancy.

The major causes of loss of health expectancy among the elderly are not necessarily the major causes of death. Limb and joint disorders were the most frequently reported disabling conditions among dependent seniors (accounting for 56% of the total), followed by circulatory system diseases (20%). All cancers were of relatively little importance as disabling conditions (less than 2% of the total), despite their importance as causes of death.

From answers to supplementary questions relating to these main conditions, the origin of each problem was classified as "congenital" (existing since birth: 1%), "accidental" (including work-related accidents as well as motor vehicle accidents, other accidents and violence: 7%), or "other" (92%).

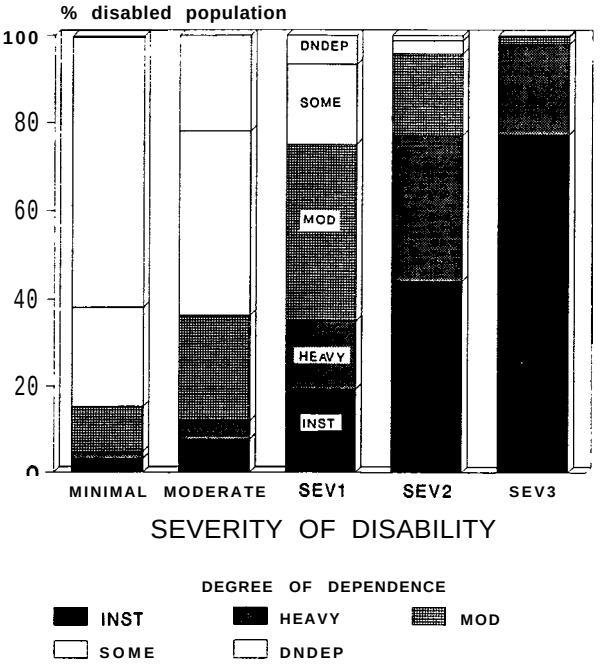
To examine the mental or emotional nature of disabling health problems, several questions were scanned. These included questions relating to learning disability, mental handicap, memory loss, general activity limitations due to mental or emotional causes, and main conditions of a mental or emotional nature. Disability of a mental or emotional nature was present in almost 40% of dependent seniors.

Another series of questions concerned technical aids or assistive devices needed but lacking. These included seeing and hearing devices, housing adaptations, mobility and agility aids, and special vans. We analysed these'data for disabled and dependent persons living in the community. Nearly a fifth of dependent seniors lacked needed technical aids or assistive devices which might have helped them overcome, at least to some extent, the handicapping effects of their disability. Moreover, we found that the greater the loss of autonomy, the more likely it was that needed technical aids would be lacking (27% of heavily dependent seniors lacked needed aids compared to 12% of somewhat dependent seniors).

Figure 4 shows the population aged 65 and over by degree of dependence and severity of disability. In other words, it cross tabulates handicap by functional limitation. Note that many elderly persons managed to maintain considerable autonomy despite fairly severe disability. Severity of disability was classified according to scales devised by McDowell and Brodeur.[31-32]

Figure 4

DEGREE OF DEPENDENCE
AGE 65+, BOTH SEXES, CANADA, 1986



Conclusions

Because most of our current indicators of health status are based on measures of ill-health, it is easy to create an excessively negative picture which puts too much stress on the burden of ill health, and not enough emphasis on the positive prospects for longer, healthier lives for most elderly Canadians. Health expectancy calculations can be used to help show both the positive and negative aspects of aging.

We have come to the conclusion that measures of dependency are needed as well as measures of functional limitation, since health objectives for the elderly tend to be stated in terms of independence and autonomy, rather than in terms of the absence of disease or impairment. On the other hand, very broad measures of limitations in major activities have proved too vague and ill-defined when applied to retired persons.

In respect to data requirements, we have learned the importance of starting with rates which describe the health or disability status of the entire population, including persons living in health-related institutions. Household surveys alone are not adequate for health expectancy calculations.

When longitudinal data are to be used, and the results are intended to represent the health expectancy of the

entire population, then it is necessary to include institutional residents in the initial cohort. Otherwise, the results will continue to reflect the relatively better health of the household population, even though later waves of the survey may include persons who have become institutional residents.

If institutional residents are not included in health and disability surveys, then compatible data from other sources are required for the institutional population. The institutional population should include persons in all kinds of residential care facilities; however classified for administrative purposes. Many elderly persons live in residential care facilities other than nursing homes, and these must be included as part of the health-related institutional population. Ideally, the same set of socioeconomic variables available for the household population should also be available for the institutional population.

For evaluating the extent of socioeconomic and regional differentials in health expectancy, in the foreseeable future it will probably continue to be necessary to have recourse to cross-sectional data in order to obtain adequate sample size for each group.

Survey data typically capture all of the main socioeconomic variables of interest, but corresponding life tables are also required by variables such as income, education, occupation, community size and urban and rural regions.

Selective sampling based on disability screening questions has proved to be efficient and cost effective. Our post censal surveys now include the 1986 and 1991 HALS, and the 1991 Aboriginal Peoples Survey.

Acknowledgement

The research on which this paper is based was supported by Health and Welfare Canada and Statistics Canada.

References

1. Wilkins R. La distribution de l'espérance de vie parmi les differents états de santé: composantes, méthodes de calcul, et résultats pour le Quebec, 1978. Cahiers québécois de dimographie 1982;11(2):253-274.

2. Wilkins R, Adams 0. Health Expectancy in Canada, Late 1970s: Demographic, Regional and Social Dimensions. American Journal of Public Health 1983; 73(9): 1073-1080.

3. Dillard S. Durée ou qualité de vie? Conseil des affaires sociales et de la famille: Collection "La santé des Québécois." Quebec: Editeur officiel du Québec;1983.

4. Robine JM, Blanchet M, Dowd JE, eds. Health Expectancy. First Workshop of the International Healthy Life Expectancy Network (REVES). Office of Population Censuses and Surveys (OPCS), Studies on Medical and Population Subjects, No. 54. London: HMSO;1992.

5. Adams 0, Wilkins R. Development of Health Expectancy Indicators: Meeting of the International Network on Heath Expectancy (REVES). Health Reports 1992: 4(1):67-72.

6. Robine JM, Mathers CD, Bone MR, Romieu I, eds. Calculation of health expectancies: harmonization, consensus achieved and future perspectives. Paris/Montrouge: Colloques INSERM/John Libbey Eurotext;1993.

7. Colvez A, Blanchet M, Lamarche P. Quebec planner's choice of health promotion indicators. In: Abelin T, Brzezinski BJ, Carstairs VD, eds.,, Measurement in health promotion and protection. WHO Regional Publications, European Series, No. 22. Copenhagen: World Health Organization Regional Office for Europe, 1987:480-494.

8. Brunelle Y, Rochon M. Limites, avantages et utilisation des EVSI dans le contexte actuel des systèmes de soins. Cahiers québécois de démographie 1991; 20(2):405-437.

9. Brunelle Y, Rochon M, Saucier A, Robine JM. Understanding changes in health status. In: Robine, Mathers, et al., [Ref. 6];1993:287-308.

10. Robichaud JB. Objectif 2000: vivre en santé en franqais au Nouveau-Brunswick. Moncton: Editions d'Acadie;1985.

11. Wilkins R. Health Expectancy by Local Area in Montreal. Canadian Journal of Public Health 1986;77:216-220.

12. Wilkins R, Sauvageau Y. Incapacité. Chapter 8 in Emond A, Guyon L, et al., eds., Et la santé, ça va? Rapport de l'enquête Sante Quebec 1987. Quebec: Les Publications du Québec;1988:175-192.

13. Wilkins, R. Health Expectancy in Quebec, 1987. In: Robine, Blanchet, Dowd, [Ref. 4];1991:41-52.

14. Lafontaine P, Pampalon R, Rochon M. L'espérance de vie sans incapacité en region au Quebec en 1987. *Cahiers québécois de demographie* 1991;20(2):383-403.

15. Wilkins R. L'Espérance de vie en santé au Quebec et au Canada en 1986. *Cahiers québécois de démographie 1991;* 20(2):367-382.

16. Berthelot JM, Roberge R, Wolfson MC. The calculation of health-adjusted life expectancy for a Canadian province using a multi-attribute utility function: a first attempt. *In:* Robine, Mathers et al., [Ref. 6];1993:161-172.

17. Reynolds DL, Torrance GW, Badley E, Bennett KJ, Chambers LW, Goldsmith CH, Tugwell P, *Modelling the reduction in quality of life from musculoskeletal diseases: arthritis.* Paper presented to the 5th Meeting of the International Network on Health Expectancy, Ottawa, February 1992. (REVES Paper No. 84)

18. Walker J, McColl MA, Wilkins R, Corey P, Stirling P. *Life expectancy in a disabled population.* In: Robine, Mathers et al., [Ref. 6];1993:245-253.

19. Torrance GW. Health status index models: a unified mathematical view. *Management Science* 1976;22:990-1001.

20. Torrance GW. *Years of healthy life: concept, aliases, weights and applications.* Paper presented to the NCHS Workshop on Measuring Years of Healthy Life, Washington, D.C.; 3-5 February 1993.

21. Wilkins R, Adams OB. Quality-adjusted life expectancy: weighting of expected years in each state of health. *In:* Robine, Blanchet, Dowd, [Ref. 4]; 1992:71-74.

22. Wolfson MC. *PQHEM - A framework for understanding and modelling the health of human populations.* Analytical Studies Branch, Research Paper No. 34, Statistics Canada, Ottawa;1991.

23. Wolfson MC, Manton KG. *A review of models of population health expectancy: A microsimulation perspective.* Analytical Studies Branch, Research Paper No. 45, Statistics Canada, Ottawa: 1992.

24. Wilkins R, Adams OB. Changes in the healthfulness of life of the elderly population: an empirical approach. *Revue d'épidémiologie et de santé publique* 1987;35(3-4):225-235.

25. Wilkins R, Adams OB. Health expectancy trends in Canada, 1951-1986. *In:* Robine, Blanchet, Dowd, [Ref. 4], 1992:109-112.

26. Qiao ZK, Wilkins R, Yang M, Lan Y, Chen X, Xu Y, Ng E. *Health expectancy of adults in Xichang, China, 1990: autonomy in various activities of daily living.* In: Robine, Mathers et al., [Ref. 63; 1993:359-370.

27. Newman SC. A Markov process interpretation of Sullivan's index of morbidity and mortality. *Stat Med 1988;* 7:787-794.

28. Newman SC. The analysis of hospital morbidity: data using life table methods. *Canadian Journal of Public Health* 1988;79:45-48.

29. Wilkins R, Adams OB. Health Expectancy in Canada, 1986. *In:* Robine, Blanchet, Dowd, [Ref. 4];1992:57-60.

30. Lamb VL. *Measuring health expectancy in Canada: an empirical assessment using the 1985 Canadian General Social Survey.* Paper presented at the Annual Meeting of the American Sociological Association, Washington, D.C.; August 1990.

31. McDowell, Ian. A disability score for the Health and Activity Limitation Survey [for adults in the household population]. Ottawa: Disability Datatbase Program, Statistics Canada: July 8, 1988.

32. Brodeur, Marie. Severity score for adult residents of institutions. Health and Activity Limitation Survey. Ottawa: Disability Database Program, Statistics Canada: October 28, 1988.

COMPRESSION OF DISABILITY: EVIDENCE FROM THE NATIONAL LONG TERM CARE SURVEYS

Larry S. Corder, Duke University

This paper briefly reports on the characteristics of the National Long Term Care Surveys (1982, 1984, and 1989), the policy uses to which the studies may be put, and their application to the study of compressionof disability, morbidity, andmortality. Additional materials are available through the Center for Demographic Studiesat Duke University and through the National Technical Information Service where public use data files for 1982 and 1984 currently reside and the University of Michigan where the 1989 study resides.

I willnowbriefly review the background, design, content, and policy uses to which these studies may be put. In the recent past, few multi-year longitudinalstudiesoftheagedpopulationhavebeen conducted to represent the United States. The only other studies currently available that fit this description are National Nursing Home Survey (NNHS) and Longitudinal Supplement on Aging (LSOA). Recognizingthatchanges inlongevityandthepattern of chronic disease'directly impacts the type and volume of expenditures for health care, several federal agencies sponsored the Long Term Care Survey in 1982 and 1984. The 1989 survey is funded by an NIA grant. Additional support was provided by ASPE for an informalcaregiver survey as well as a follow-on to the institutional sample component. Also, HCFA's support was instrumental in the development of the survey sample frame. Figure 1 illustrates the changesinhealthstateanduseof servicesovertime. Figure 2 illustrates the distribution of functional states in a survival curve partitioned by morbid state and Figure 3 illustrates level of care and functional level in a survival curve on an idealized basis where partitioning under the survival curve is represented. However, no data sourceavailable.prior tothedevelopmentofthe Long Term Care Surveyallows the analyst to actually estimate the previous two basic charts of the aged population's medical care use, disability and disease with multi-year longitudinal data for a cohort of older persons. Rather, what we have had in the past have beenmodels (i.e., Figure 4) based on cross-sectional surveys of noninstitutionalized persons' use of acute care services. Such models clearly require expansion to include institutionalized persons' long term care services, improved methods of measuring health status and noninstitutionalized persons use of home health services and associated expenditure. We, in fact, believethatthe Long Term Care Surveys provide aresourcethatmeetsa substantialpartofthatneed. First, we follow a national sample cohort of aged persons over many years (7) with plans to continue to follow them in 1994. Second, we develop cross-sectional estimates for each survey year. Third, we cover the entire aged population, e.g., those in the community and those in nursing homes. Fourth, we routinely assess these persons' health status and collect detailed social and economic information from them. Fifth, we link our study findings to the Medicare files for use of covered medical services, and for those who die, we will collect their death certificate. We have also linked our sample to Medicaid records in states where that is possible (MEDSTAT).

Medicare files are linked continuously to the NLTCS from 1982 to the present. They contain basic information on utilization of covered services, particularly inpatient services. Theydonot include institutional or pharmacy services as a rule.

Medicaid files containinfonnationon enrollment, inpatient services, outpatient services, pharmacy, and long term care organized into different files. NLTCS is currently linked to Medicaid data for 26 state in1989, the survey year. Thus, several states which contain NLTCS primary sampling units are not represented. Coverage of the states is improving over time. We expect coverage in excess of 75% of dollar volume in 1994, the next survey year.

These central survey design elements (a list sample, oversampling the disabled) provide exactly the right framework for conducting policy analysis of aged population issues for both research and legislative purposes. Specifically, we at Duke continuetocarryoutpolicyanalyses inasubstantial numberofpolicyareasusingthe 1982, 1984, and1989 NLTCS. These areas include out-of-pocket payments, AAPCC, home health services, agedpopulationmortality, quality of care, Medicaid spenddown and spousal impoverishment, estimation of financial liability for Medicare services vis forecasting, the risk of institutionalization, and Active Life Expectation (ALE). We are using the LTC Surveys because 1) we have a survey that follows persons over time when benefits have changed, e.g., 1982 marked the introduction of DRGs, 2) the survey collects data that is not available from Medicare files thus complementing that resource. Specifically, for covered services, you don't have to expend resources on data collection activities, 3) the Medicare enrolled list sample approach provides excellent population coverage. Further, it is less expensive than the alternative-area probability sampling approach and differentiates the Medicare population by disability status in the first sample stage. This design featureisparticularlyrelevanttolegislation which focuses on disability as an entitlement criterion, thus expanding the utility of HCFA administrative data as a policy tool, and 4) the surveyprovidesanexcellentsurveydesign, including record linkage, to evaluate the impact of changes in the Medicare program at low cost.

The design of the Long Term Care Surveys has been fairly stable overtime. It is difficult to describe this complex survey in a brief fashion. A detailed presentation is available in Manton, Corder, and Stallard (1993a). Its important to remember that in each year 1982, 1984, and 1989 we are simultaneously following a longitudinal cohort of aged persons replenished at younger ages, but also are developing estimates of important characteristics of the aged population for the survey year. Thus, each sample has longitudinalandcross-sectional components. By examining Figure 5 it is possibletotracethe status of each sample person from the beginning of the study until the most recent survey. you should note that this figure is reproduced and updated from Corder, Woodbury, and Manton (1993). Briefly, therespondent universe is the elderly (over 65 years of age). This universe comes from the Medicare files for the aged beneficiaries included in the health insurance master file, e.g., a list sample. Prior to any

subsampling, the sample is drawn in such a way to ensure proper representation of the agedpersons who were previously disabled. Stratification of the sample by these two factors --age group andoriginal reason for entitlementprovidesthis representation. Note at this point, that the Bureau of the Census carries out the samplinganddata collection for each of the Long Term Care Surveys. The basic components of the sample are illustrated in the columns of boxes in Figure 5. you should note that by focusing on the aged in generaland in the disabled aged, bothinthe community and in institutions, the long term care surveysare abletoachievelarge sample sizes of high risk populations at relatively low cost. The next table lays out how our periods of survey observation overlap with data available form Medicare Part A and death certificates which may be overlayed with PPS installation (discussion-itsexcitingamultifaceted natural experiment). (Figure 6) The next figure (Figure 7). characteristics of the United States cross-sectional Surveys, provides some summary detail on the other available studies that either in whole or in part, represent the aged population in the United States in recent years (Corder and Manton, 1991). Among them, only NMES has conducted Medicaid to a survey data file linkage. A brief review of the table illustrates that the medial utilization studies are multipurpose studies which cannot heavily represent the aged and cover the entire population while only the most recent study covers the nursing home population. Only LTC studies concentrate on the aged in the community and in institutions. In short, by narrowly addressing the needs of aging research and policy formulation LTC has become a very useful data base particularly for aged persons in the community at risk of intense service use/institutionalization and the oldest old. However, note that the NNHS covers the institutional populationwithrelatively large samples and the SOA-LSOA, which are based on the NBIS sample frame cover the healthy older population which the LTC doesn't attempt. In fact, the emphasesinthesethreestudies coordinate well.

We believe that the LTC survey is the appropriate vehicle for evaluating the impact of Medicare legislation in particular and health legislation in general while it remains a primary source for charting disability change among cohorts of aged persons. The effect of spending down to Medicaid, continuing shifts from inpatient to outpatient serviceuse, changingpatterns ofhomehealthuseand institutional stay patterns are available for examination because the cohort study and the instruments are in place to study these areas as we speak. Further, the LTC with its emphasis on functional assessment over time is an excellent vehicle for projecting service needs of the aged and associated expenditures under various financing and active life expectation scenarios.

Quite naturally, there has been substantial continuity in the content of the information collected over the three studies. Indeed, our predilection has been to err on the side of replication of questionnaire batteries over the survey years. Our view has been to address changing data needs by adding questionnaire batteries or supplementswhentheyhavebeenneeded. In the survey process, persons who are not disabled complete a screeningquestionnaireonly. Thisbasicinformation and linkable-file data are available for this group. Persons who are disabled or functionally impaired while living in the community receive a lengthy in person questionnaire which exceeds one hour on average. It is divided into sections. The longest concerns functional status (ADL, IADL), impairment and medical conditions. The second section asks a battery of questions concerning other functioning, informal caregivers, and social networks. The third sectionasksquestionsabouthousingandneighborhood characteristics while the fourth asks some general health insurance coverage questions. The fifth section asks aboutmedicalprovider andprescription medicines while Section 6 is a cognitive functioning battery. Thelastsectionelicitsavarietyofsocial andeconomicdataincludingdetailedincomeandasset information for 1989. Persons found to be in institutions receive a shorter questionnaire (reinterviewedata six-monthintervalin1989) which includes a section on cognitive functioning, ADL, IADL, admission to health facilities and source of payment and detailed income and assets. The questionnaireendswithasectiononthecharacteristics of the institution. Last, in 1982 and 1989 the communityquestionnairewasusedtoidentifyinformal caregivers and conduct an interviewwith a sample of them. In 1989, this activity is carried outwiththe main caregiver. We ask about the amount and kinds of help provided in great detail, help from others that the provider knows about, problems encountered in providing help, and the caregiver's living and work arrangements as well as caregiver personal characteristics. The interview lasts about 30 minutes. No questions are asked about deceased persons either from providers or relatives. We believe that we have a unique resource to report on problems with people becoming impoverished through institutionalization (6 months) and characteristics oftheinformalcaregiverandthe serviceswhichthey provide these data should directly address respite care legislation. In addition, this resource allows us to examine issues aboutolderpersons and Medicaid in detail.

Analyses

Next, webrieflyreviewtheevidence forcompression of disability in the NLTCS based on recently publisheddata (Manton, Corder, and Stallard, 1993a, b). Evidence from our longitudinal analyses which indicates compression of disability has occurred in the context of mortality decline (NCHS, 1991). However, in order to properly examine the behavior of interest, concepts of mortality, morbidity, disability and date of onset should be introduced. First, it is arguable that the NLTCS collects information in the form of ADLs and IADLs that represent the underlying etiology of disease in a population. As the disease progresses from date on onset, it expresses itselfthroughmorbidityandthen disability (functional limitation). (Figures 2 and 3) That is, we think morbidity and functional limitation reports for a population are closely tied (Manton, Corder, and Stallard, 1993a). Therefore, a change in the date of onset and/or a change in the lifeexpectationcouldaffectthedegreeofcompression of disability in a life table context (Fries, 1989). Further, date of onset as well as persistence in disability state are important concerns for the calculation of active life expectancy (ALE) (Manton, Corder, and Stallard, 1993b).

In order to fit our discussion more completely into the life expectancy context, mortality, disability, and morbidity compression trends are discussed separately. Mortality compression is

characterized by stalling of life expectation at a round a spectic age, say age 85, accompanied by reduction in the variance at age at death. This 'squaring' of the survival curve is no't being observed. Rather, current evidence from France, Sweden and the United States indicate increasing expectation of life at older ages. While we can speculate about the future,' overall, the survival curveappearstobeshiftingtotheright. Compression of disability can occur withmortality. In fact, we expect a strong relation between disability and mortality declines.

When disability compression is considered, questions cometomind. First, given increasing life expectation, how will those additional person years be spent? Second, can we observe changes in the date of disease onset as it is expressed in morbidity. Third, what is happening to the persistence of person's in health states, e.g., how long does a personhaveamild fonnofadiseasebeforeitbecomes more severe). Our recent paper shows an overall decline in disability across ages and increased persistence in health states (Manton, Corder and Stallard, 1993a) we believe this decline is linked to treatment (period) andeducation (cohort) effects as well as the strong relationbetween mortality and disability in the data. Next, morbidity compression should be examined in the context of functional limitation as the concrete expression of morbidity. However, intervention can alter, probably has alteredtherelationshipbetween functional limitation and morbidity as it is reported. Specifically, reportofadiseasemaynotmanifest itselfasquickly in functional limitation prior to development of an intervention.

With the above relationships in mind, our recent research clearly shows declines in disability among the aged in recent years. These declines are more or less consistent across functional limitations levels and age categories. Table 1 presents basic information on longitudinal change in disability. The table shows institutional counts are stable over time while accurately accounting for stock and flow issues. Clearly, if period data were used, ALE would increaseovertheperiodsrepresentedhere. Further, carefulworkwhich is presented in Manton, Corder and Stallard (1993a) clearly indicates that persistence instatehasincreasedovertime. Table2 illustrates the finding, which also directly influence certain ALE estimates. I should also note that it is clear that part of the decline in chronic disability incidence for nondisabled persons is -hidden- by increased survival, and prevalence of disabled persons.

Conclusion

Our 'evidence supports the notion of.disability compression in the context of increased life expectation. Further, persistence in health state has improvedandonsetofdisabilitymayhave shifted tolateryears. However, theseresultsdonotsuggest that recent projections of the increased need for services associated with the growth of the aged population may be ignored. Rather they suggest that years of disabled life are not increasing as rapidly as life expectation.

References

Corder, L.S., and Manton, K.G. (1991), "National surveys and the health and functioning of the elderly: The effects of design and content,* *Journal of the American Statistical Association,* 86, 513-525.

Corder, L.S., Woodbury, M.A., and Manton, K.G. (1993), "Health loss due to unobserved morbidity: A design based approach to minimize nonsampling error in active life expectancy estimates,. in *Calculation of Health Expectancies: Harmonization, Consensus Achieved and Future Perspectives,* eds. J.M. Robine, C.D. Mathers, M.R. Bone, I. Romieu I. Collogue INSERM/John Libbey Eurotext Ltd., Vol. 226, pp. 217-232.

Fries, J.F. (1989), 'The Compression of Morbidity: Near or Far?," *Milbank Quarterly, 67, 208-232.*

Manton, K.G., Corder, L.S., and Stallard, E. (1993a), 'Estimatesofchangeinchronicdisability and institutional incidence and prevalence rates in the U.S. elderly population from the 1982, 1984, and 1989 National Long Term Care Survey,' *Journal of Gerontology: Social Sciences, 48,* S153-S166.

Manton, K.G., Corder, L.S., and Stallard, E. (1993b), "Changesintheuseofpersonalassistance and special equipment 1982 to 1989: Results from the 1982 and 1989 NLTCS," *The Gerontologist,* forthcoming.

National Center for Health Statistics (1991b), 'Births, Marriages, Divorces, and Deaths for July 1991," *Monthly Vital Statistics Report,* Vol. *40, No. 7,* Hyattsville, MD: Public Health Services.

Figure 1

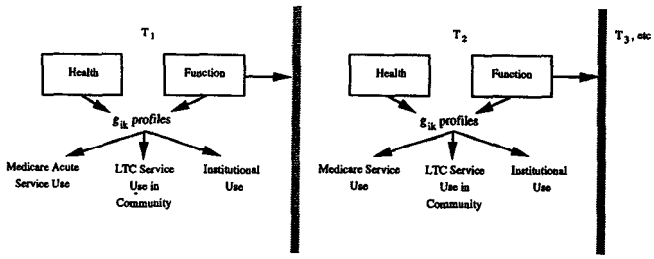


Figure 2

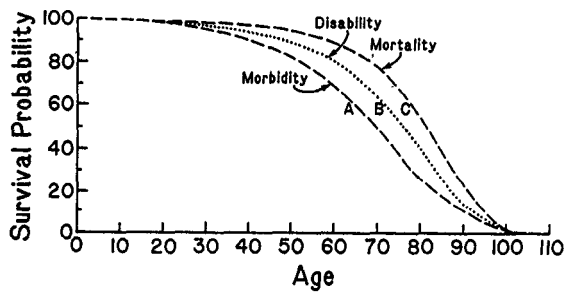


Figure 3

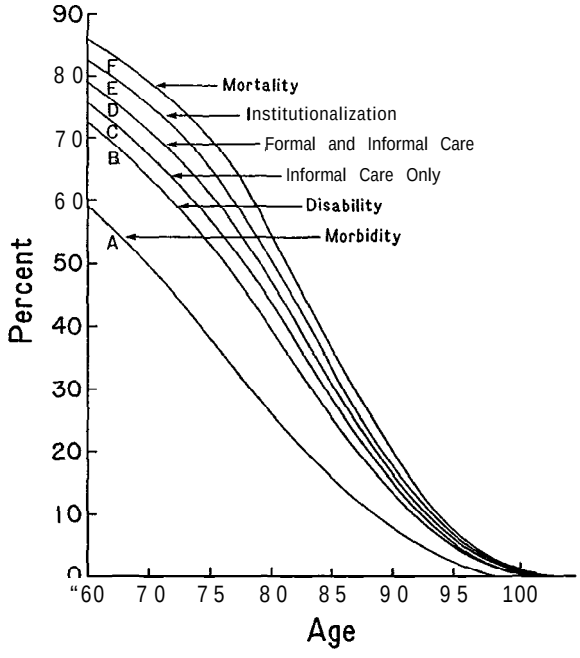
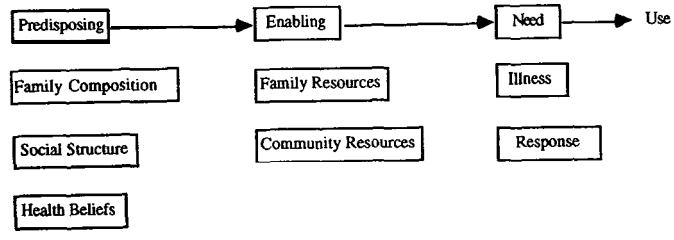


Figure 4: A behavioral model of health services utilization (Andersen, 1968)



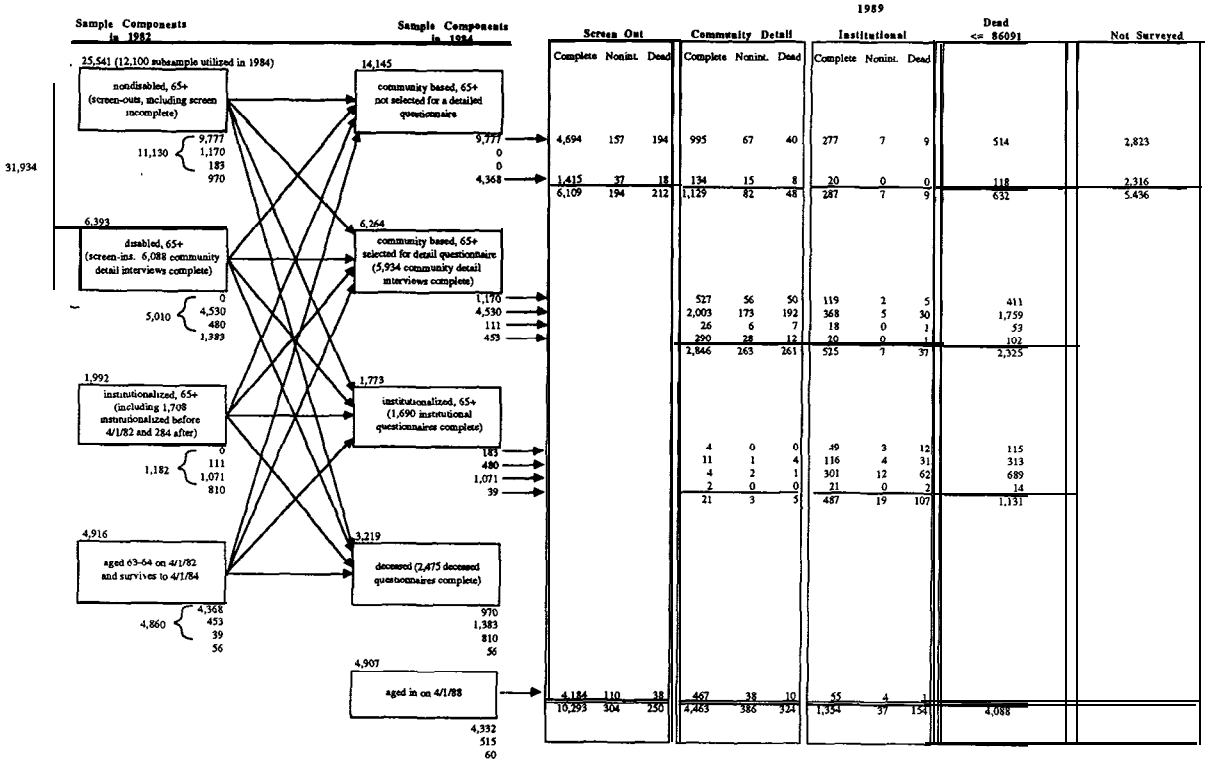
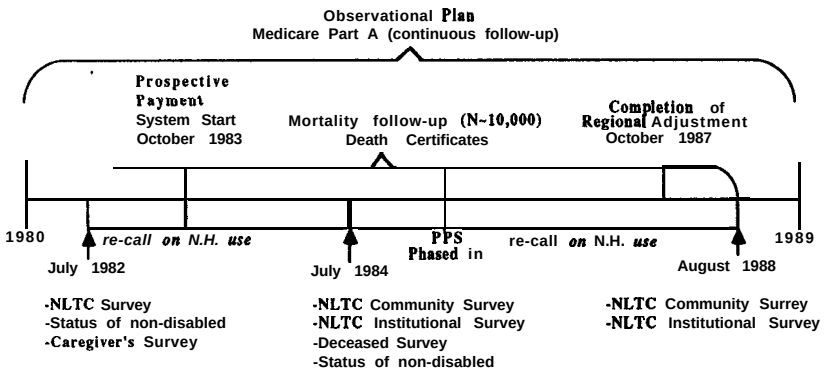


Figure 5: Component subpopulations of the 1982 and 1984 National Long Term Care Survey

Figure 6



1) Data on "state" of person at the three waves are multi-dimensional and involves both **multiple functional and medical conditions.**

2) **Data on service use and mortality are continuous with exact dates and with associated diagnostic data (e.g., hospital diagnoses -post October, 1983; multiple medical conditions and service use)**

3) **State can be expanded to include prior state characteristics (e.g., hospitalization or N.H. use in prior 12 months, Short Portable Mental Status Questionnaire)**

4) **There are special one time surveys (e.g., caregiver's help with cognitive assessments; deceased survey)**

Figure 7: Characteristics of the United States Repeated Cross Sectional Surveys

	Sample Size	Sample Type	Sample Size ≥65 Years of Age - Percent at Survey Date	Expectation of Life at Age 65 at Survey Dale Total	White Female	Expectation of Life at Age 85 at Survey Date Total	White Female
Medical Utilization							
CHAS 1963	7,803	self-weighting area probability	9% (719)[4]	14.3	16.0	4.5	4.6
CHAS 1970	11,619	multistage area probability	10.3% (1,197)[4]	15.2	17.1	5.8	6.0
NMCES 1977	40,320	stratified multistage probability		16.3	18.4	6.4	6.8
NMCUES 1980	17,123	stratified multistage probability	10.7% (1,830)[4]	16.4	18.5	5.9	6.3
NMES 1987	37,000[1]	stratified multistage	Not Known (~4,000)[4]	Not Known		Not Known	
Institutional							
RPS2 1963	3,143'	stratified two stage; list & cluster	100%	14.3	16.0	4.3	4.6
RPS3 1969	2,088[2]	stratified two stage; list & cluster	100%	14.8	16.6	5.0	4.8
NNHS 1973	21,180[3]	two stage: list & cluster	100%	15.3	17.3	5.4	5.7
NNHS 1977	7.0333	stratified two-stage; list & cluster	100%	16.3	18.4	6.4	6.8
NNHS 1985	5.000'	stratified two-stage; list&cluster	100%	16.8	18.7	Not Available	
NHIS							
NHIS 1985 (Other NHIS available 1963 lo 1985)	105,000	multistage area probability - continuous sampling	11.4% (11,983)[4]	16.8	18.7	Not Available	
LTC							
1982	6,000	stratified, list and cluster	100%	16.8	18.9	6.3	6.7
1984	10,000	stratified, list and cluster	100%	16.8	18.7	6.1	6.5
1988	10,000	stratified, list and cluster	100%	Not Available		Not Available	

[1]Approximate, civilian noninstitutional key persons: [2]Places; [3]Approximate; [4]Unweighted counts in parentheses

Table 1. The 1982, 1984, and 1989 Prevalence of Chronic Disability in the U.S. Elderly Population
Estimated from the 1982, 1984, and 1989 NLTCS

Disability Level	Population 1982	Population 1984	Population 1989	1984-1989 Change with S.E. and t test	1982-1989 Change (Age Standardized) with S.E and t-Test
Nondisabled	20,548,000	21,403,000	23,906,000	+1.1%	+1.7%
% (S.E.)	76.3 (±.30)	(±29)	(±.30)	(± 0.42; t =2.6)	(±0.42; t = 4.1)
N	(11,612)	(14,097)	(17,006)		
IADLs Only	1,434,000	1,590,000	1,360,000	-1.3%	-1.4%
%(S.E.)	5.3 (±.16)	5.7 (±.16)	(±1 .14)	(±.21; t = -6.2)	(±.21; t = -5.7))
			(975)		
1 - 2 ADLs	1,740,000	1,831,000	1,993,000	0%	-1.2%
% (S.E.)	6.5 (±.17)	6.5 (±.17)	6.5 (±.17)	(± .24; t = 0.0)	(±.24; t = -5.0)
N	(2,007)	(1,778)	(1,429)		
3 - 4 ADLs	732,000	797,000	1,079,000	0.7%	0.7%
% (S.E.)	2.7 (±.11)	2.8 (±.11)	3.5 (±.13)	(±.17; t = 4.1)	(±.17; t = 4.2)
N	(831)	(813)	(774)		
5 - 6 ADLs	937,000	882,000	848,000	-0.4%	-0.9%
% (S.E.)	3.5 (±.13)	3.1 4.12)	2.7 (±.11)	(±.16; t = -2.5)	(±.17; t = -5.3)
N	(1,027)	(926)	(608)		
Institutional	1,532,000	1.538,~	1,685,000	0%	-0.4%
% (S.E.)	5.7 (±.16)	5.5 (±.16)	5.5 (±.16)	(±.23; t = 0.0)	(±.23; t = -1.7)
N	(1,992)	(1,773)	(1,354)		
Total Population	26,924,000	28,042,000	30,871,000		
N	(19,142)	(20,474)	(22,146)		
Dead in Interval	—	2,705,000	7,274,000		
Age-in Interval	—	3,841,000	10,123,000		

Source: Manton, Corder and Stallard, 1993a

Table 2: Two-Year Changes in Functional Status in the U.S. Elderly Population (% Distribution):
Estimates from the 1982, 1984, and 1989 NLTCS

Disability Level	Nondisabled	IADLs only	J-2 ADLs	3-4 ADLs	5-6 ADLs	Institutional	Dead
Nondisabled							
1982-1984	85.3	3.4	3.2	1.1	0.9	1.2	4.9
Mortality Adjusted	89.7	3.6	3.4	1.1	1.0	1.2	- - -
1984-1989	86.5	2.4	2 5	.9	.8	1.3	5.7
Mortality Adjusted	91.7	2.5	2 6	.9	.9	1.4	- - -
IADLs only							
1982-1984	16.7	31.4	20.0	5.5	5.5	7.1	13.9
Mortality Adjusted	19.4	36.4	23.2	6.3	6.4	8.2	—
19841989	7.5	42.1	18.7	3.9	4.3	5.3	18.2
Mortality Adjusted	9.2	51.5	229	4.8	5.2	6.5	- - -
1 - 2 ADLs							
1982-1984	J0.6	12.5	30.5	11.0	6.8	a 1	20.5
Mortality Adjusted	13.3	15.7	38.3		8.6	10.2	- - -
1984-1989	3.4	5.5	46.9	12.8	4.1	8.0	19.5
Mortality Adjusted	4.2	6.8	58.2	IS.8	5.0	9.9	—
3 - 4 ADLs							
1982-1984	3.4	4.0	17.7	19.0	18.3	9.0	28.5
Mortality Adjusted	4.8	5.6	24.8	26.7	25.6	12.5	- - -
1984-1989	.2	1.2	10.9	44.7	11.1	8.9	229
Mortality Adjusted	.3	1.6	14.1	58.0	14.4	J 1.6	—
5 - 6 ADLs							
1982-1984	1.1	4.1	8.5	9.4	32.3	9.1	35.5
Mortality Adjusted	1.7	6.3	13.2	14.6	50.1	14.1	- - -
1984.1989	.4	1.3	3.9	6.0	41.5	7.6	39.2
Mortality Adjusted	.7	2.2	6.4	9.9	68.3	12.5	- - -
Institutional							
1982-1984	1.2	1.3	1.7	1.7	1.5	56.2	36.5
Mortality Adjusted	1.8	2.0	2 6	2.7	2.3	88.6	- - -
1984-1989	.2	.5	.9	A	.8	59.1	38.1
Mortality Adjusted	.3	.8	1.5	.6	1.2	95.5	—

Source: Manton, Cm-da and Stallard, 1993a

COMPRESSION OF DISABILITY? RESULTS FROM THE ESTABLISHED POPULATIONS FOR EPIDEMIOLOGIC STUDIES OF THE ELDERLY

Marcel E. Salive, National Institute on Aging, NIH
Jack M. Guralnik

This paper discusses findings from the Established Populations for Epidemiologic Studies of the Elderly (EPESE) that relate to the development of limitations in activities of daily living (ADL) in the years before death. In particular it addresses the question: How are the person-years of disability suffered by the older population distributed in relation to subsequent death? This may have significant implications for the compression of disability in the older population.

Compression of disability in this context is synonymous with compression of morbidity. Fries' defined compression of morbidity as the extension of active life expectancy and the compression of disability, which may imply that disability begins closer and closer to the time of death.

Active Life Exoectancv

Active life expectancy, a measure of health expectancy, is defined as the average number of years of remaining life which is disability-free that is expected by an individual of a specific age. Branch et al[2] expanded the analysis of active life expectancy through the use of an increment-decrement life table. The computations, which require longitudinal data, take into account the transitions to and from disability and to death. Studying the EPESE populations, Branch et al[2] demonstrated only slight differences between the three communities in health expectancy. Compared to men of the same ages, women have approximately equal percentages of remaining years of dependency, but given their longer life expectancy; this- is accompanied by a greater absolute number of years of dependency.

Recently, Guralnik and others[3] examined life expectancy and active life expectancy for men and women age 65 and older in the North Carolina EPESE population according to race and educational attainment. The authors concluded that higher education is associated with longer lives and longer active life expectancy. Although computations that employ increment-decrement life tables do take into account the transitions to and from disability and to death, the results are frequently shown schematically in a picture which suggests that all the years of disability occur in the last years of life. This is not true for many individuals.

Disability prior to death

Unpublished data from the National Institute on Aging's Survey of the Last Days of Life,[4] which is similar to the National Center for Health Statistics Mortality Followback Survey,[5] show that the prevalence of ADL disability increases dramatically as death approaches, with a prevalence of disability greater than 50% in the last month of life in some age and sex groups. This study was limited to decedents, however, and this does not permit comparison with persons of a comparable age who did not die. From a prevention perspective, the disability immediately prior to death may be "harder" and less amenable to prevention or rehabilitation.

Another EPESE study that focused on physical function in the years prior to death found that the prevalence of ADL disability increased with aging and with approaching death and that decedents had higher rates than survivors.[6,7] Disability rates prior to death increased with increasing age at death in this study and the Last Days of Life Survey.

Therefore, it has been demonstrated that the prevalence of disability increases in the years just prior to death, but there have been no studies that estimate the proportion of all person-years of disability in the older population that occur in each of the years prior to death.

Objectives

The objectives of the study were twofold. First, to examine the development of physical limitations in the years before death in a population of older adults, and second, to partition the total person-years of disability suffered by the older population according to the timing of the disability relative to subsequent death.

Study oooulation

The EPESE is a collaborative multi-center longitudinal study of older men and women, which was initiated and funded by the National Institute on Aging. Beginning in 1982, men and women age 65 and older were enrolled at three sites: East Boston, New Haven, and Iowa and Washington counties. Over 10,000 persons, between 80 and a4 percent of eligible community residents completed baseline interviews with trained lay interviewers.

Surviving participants were re-interviewed annually for up to six years. The followup response rates exceeded 95 percent in each community. The last

contact was in 1988-89 but to further increase the cohort duration, we used mortality followup through 1990 via the National Death Index. This resulted in over 54,000 person-years of observation for the cohort.

Activities of daily living

Every year the interviewer asked about the ability to perform six activities of daily living: walking, bathing, dressing, transferring from bed to chair, toileting, and eating.[8,9] Limitations in any ADL were defined as inability to perform the task or requiring help from another person.

Using the activities of daily living at each interview, the interview dates, and the relation of the interview with the date of death (obtained from the death certificate or a proxy) a matrix could be defined where the element was the disability prevalence, using single year of age and single year before death. Separate matrices were produced for men and women. Data for each element came from all times throughout the study. For example, any person at any time in the study who was age 70 years and destined to die within one to two years contributed disability data to the element, "age 70, l-2 years prior to death."

Statistical Methods

These data were employed in gender-specific life tables to estimate the distribution of person-years of disability suffered by the older population according to the timing of the disability relative to subsequent death. Starting with a hypothetical cohort of 100,000 adults age 65 years, the cohort was "aged" until extinction, using national death rates from 1979-81 according to single year of age.[10] Person-years for the hypothetical cohort were computed and divided according to the number of years before death. If l_x represents the number alive at the beginning of the interval, the number dying, $_nd_x$ is computed by multiplying 1, by q_x, the probability of death in a one-year period. The person-years of the cohort were partitioned according to the time until death (e.g. one year until death, two years, etc.) by bringing the person-years across and up a diagonal path. An assumption was made that people on average lived halfway through the final year. To determine person-years of disability according to both age and years until death we multiplied the disability rates by the number of person years for each cell of the life-table analysis. Then the person-years of disability were summed for the entire cohort according to years until death, and the proportion of disability in each year before death was computed.

Results

The proportion of participants at a given age who were disabled rose as death neared (see Figure for an example of data for women). The proportion with ADL limitations was higher in older persons and women.

Among women, the largest proportion of disability-years are expected in the period more than five years before death, which accounted for nearly half of the years of disability (Table). About ten percent of the disability-years occur in each of the five years prior to death. Among men, a higher percentage of disability-years are closer to death. Still, 40% of the disability-years do occur more than five years before death.

Stratifying these data according to age at death, it is clear that among younger deaths, an even larger fraction of the disability-years occur far from death (data not shown). Deaths at the oldest ages, however, have a greater percentage of disability years immediately proximate to death.

Table. Proportion of person-years of disability in relation to subsequent death according to gender.

Years before death	Percent of total disability-years: Men	Women
<1	11.5	9.1
l-2	16.6	12.9
2-3	12.4	11.2
3-4	10.3	9.6
4-5	8.4	8.5
>5	40.7	48.8

Discussion

Several conclusions can be drawn from the EPESE studies that apply to the distribution of disability in older populations. First, the majority of person-years of disability in older adults occurs several years before death. Second, preventive interventions that target ADLs will have their greatest impact several years prior to death. These years may be more amenable to prevention of disability or to rehabilitation of existing disability. Understanding the etiology of physical impairment and the pathway from disease to disability to death will be important in achieving national goals for an increased span of healthy life." If present trends of declining mortality persist, even greater reductions in disability will be necessary to produce a compression of morbidity as people die at older ages, where there is a greater disability prior to death.

Census bureau projections of life expectancy suggest a continued mild to moderate rise into the 21st century." This has significant implications, for the

age distribution of deaths among older women would change over time. The proportion of deaths over age 85 years is projected to climb from 45% in 1980 to 62% in the year 2040.

There is good news in our results. If our results had shown that most of the disability were those years immediately prior to death, this would be a grim picture, because this disability is the hardest to reduce. But since we found that the majority of years of disability occur more than five years prior to death, society has a real opportunity to intervene and decrease these years of disability.

Data needs

There are a number of data needs that will enable an accurate assessment of the compression of morbidity in the future. only consistent longitudinal studies will provide data that can be analyzed for evidence of a compression of morbidity. Other concerns fall into two general areas: measurement and methods.

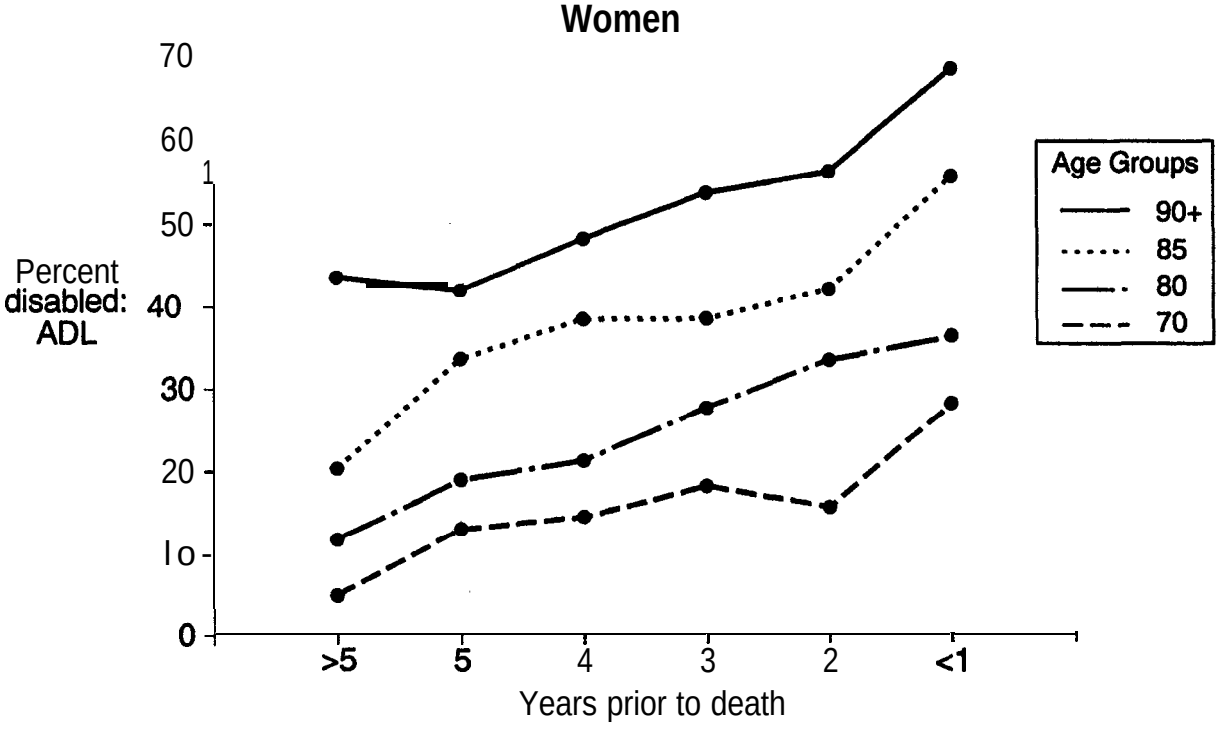
Measurement of Disability

- *Validity* of the disability measures is critical, so that we can know if a change in the prevalence of **ADLs** over time is real or an **artifact.**[13]
- *Stability over* time is a validity issue--if there is drift, this could compromise measurement of compression of morbidity.
- Is there a *bias* by educational **attainment that** might affect the reporting of ADL limitations? Secular changes are occurring in the educational levels of the older population that might then affect reported levels of disability.
- Use of *compensatory strategies* might **affect the** measurement of disability. For example, if new devices are developed that can compensate for activities that previously required human assistance, this might result in a change in the definition of what is a particular ADL disability and a consequent change in prevalence rates.
- A system is needed to measure severity *of disability,* to evaluate whether there might be a change in severity in conjunction with either change **or stability** in the prevalence of **disability.**

Statistical Methods

- A standard approach is needed for computation of active life expectancy. To date, all published papers have used different methods.
- Methods are needed to account for severity of disability in the computations of active life expectancy.

Disability Prevalence in the Years Prior to Death



Source: EPESE 1981-l 990

Acknowledgements
The authors thank their colleagues from the National Institute on Aging who initiated the EPESE populations, as well as the investigators and staff at Harvard Medical School, Yale University School of Medicine, Duke University and the University of Iowa, who have played a critical role in collecting the data that made this research possible.

References

1. Fries JF, Green LW, Levine S. Health promotion and the compression of morbidity. *Lancet* 1989;i:481-3.
2. Branch LG, Guralnik JM, Foley DJ, et al. Active life expectancy for 10,000 Caucasian men and women in three communities. *J Gerontol: Med Sci* 1991;46:M145-150.
3. Guralnik JM, Land KC, Blazer D, Fillenbaum GG, Branch LG. Educational status and active life expectancy among older blacks and whites. *N Engl J Med* 1993;329:110-6.
4. Brock DB, Foley DJ. Health status trends in the last year of life. *Proc Am Stat Assn. Soc Stat Sect.* 1988;418-423.
5. Seeman I, Poe GS, Powell-Griner E. Development, methods, and response characteristics of the 1986 Mortality Followback Survey. National Center for Health Statistics. *Vital Health Stat* 1(29). 1993.
6. Guralnik JM, LaCroix AZ, Branch LG, Kasl SV, Wallace RB. Morbidity and disability in older persons in the years prior to death. Am J Public *Health* 1991;81:443-7.
7. Guralnik JM. Prospects for the compression of morbidity. The challenge posed by increasing disability in the years prior to death. *J Aging Health* 1991;3:138-54.
8. Katz S, Ford AB, Moskovitz AW, Jackson BA, Jaffe MW. Studies of illness in the aged. The Index of ADL: a standardized measure of biologicalandpsychosocial function. *JAMA* 1963;185:914-9.
9. Branch LG, Katz S, Kniepmann K, Papsidero JA. A prospective study of functional status among community elders. *Am J Public Health* 1984;74:266-268.
10. National Center for Health Statistics: United States life tables. *U.S. Decennial Life Tables for 1979-81.* Vol. I, No. 1. DHHS Pub No. (PHS) 85-1150-l. Public Health Service. Washington, DC: U.S. Government Printing Office, Aug. 1985.
11. *Healthy People 2000: National Health Promotion and Disease Prevention Objectives.* Pub. No. (PHS) 91-50213. Washington, DC: U.S. Department of Health and Human Services, Public Health Service, 1991.
12. U.S. Bureau of the Census (Gregory Spencer): *Current Population Reports,* Series P-25, No. 1018, Projections of the Population of the United States, by Age, Sex, and Race: 1988 to 2080. Washington, DC: U.S. Government Printing Office, 1989.
13. Guralnik JM, LaCroix AZ. Assessing physical function in older populations. In: Wallace RB, Woolson RF. *The Epidemiologic Study of the Elderly.* New York: Oxford University Press, 1992:159-81.

MEASURING THE HEALTH AND DISABILITY OF OLDER AMERICANS

Discussants:

William Scanlon
Georgetown University

Lois M. Verbrugge
University of Michigan

(Not available for publication)

Second Plenary Session

Measuring Health Disparities

Toward the year

2000

Refining the measures

USING DATA ON HEALTH DISPARITIES IN THE FORMULATION OF PUBLIC HEALTH POLICY

Mohammad N. Akhter
D.C. Department of Health and Human Services

(Not available for publication)

CLOSING GAPS IN HEALTH & HEALTH CARE: IMPROVED DATA FOR A NEW HEALTH SYSTEM

David Mechanic, Rutgers University

Closing inequalities will be high on the Nation's future health care agenda. This requires good information on disparities and their underlying causes. At the most basic level we seek to monitor whether the health resources of the nation are equally accessible to varying categories of our population. Do persons who vary by age, sex, income, education, geography, color, or ethnicity have comparable access not only to basic preventive, physician and hospital services but also to specialized care of a variety of types? This question, of course, is only a rough proxy for the more meaningful question of whether the finite resources available are distributed in relationship to need as measured by existing patterns of morbidity, disability, and mortality. The latter question is more difficult than the descriptive question concerning utilization patterns because it embodies assumptions and values that are not fully shared.

The challenges go beyond such issues because our goal is the improvement of health and not simply access to services. In improving health, we seek to identify high risk groups who fail to achieve health outcomes comparable to the population as a whole and to intervene to achieve greater parity in health status.' This requires meaningful markers that identify risk categories that can serve as a focus for special health initiatives.

Certain markers have become conventional for data collection agencies. No reputable data collection effort would fail to acquire information on sex, age, or education and most make some effort to measure income, household arrangements, and marital status, color, and ethnicity. Age and sex, which are simultaneously important biological and social indicators, are of acceptable reliability although there are various distortions in reporting when people seek to present themselves in ways they regard as socially acceptable or personally advantageous, or heap their ages at the nearest 5 or 10. As one considers other sociodemographic characteristics such as color and ethnicity, the problems of measurement multiply and the meaning of the categories themselves is more vague and uncertain.

In the objectives for the year 2000, a range of descriptors is used to characterize disparities and health objectives: for example, the objectives use categories such as Blacks, Hispanics, single parents, education, blue collar, Southeast Asian, people without insurance, education, etc. Are these to be regarded as factors important in and of themselves, like age and sex, or are they more to be viewed as proxies for a variety of disadvantages that have a more proximal relationship to the health outcomes of concern? To the extent we can identify and measure the proximal factors, we are in a much stronger position to efficiently address the sources of health difficulties. Even large statistical differences among major population aggregates are difficult to address from the point of view of prevention and remediation because most people in each of the groups will not be at risk, and thus, interventions are very expensive. As we become clear about the nature of the underlying risk, we can target our efforts more specifically, and, thus, in a more cost-effective way.

Race and Ethnic Classification

Federalstatisticalagencies function under OMB Directive Number 15 which sets standards for race and ethnic classifications. The difficulty with these classifications is that they reflect political pressures and compromise as much as scientific criteria and are limited for subclassifying groups in a manner most informative for health or other social concerns. As counts of varying categories become relevant for financial and political reasons, individuals who ordinarily would not identify themselves in particular ways, and whose behavior and risks are little influenced by these identifications, may begin to do so.

One such problematic category is the Hispanic identifier which has attained increasing importance in reporting health statistics. Although Directive 15 allows, and the U.S. census collects, information that differentiates Mexican, Puerto Rican, Cuban and other Spanish/Hispanic groups, most health surveys and health records collect no more, or make possible no smaller disaggregations, than the category "Hispanic." Recent recommendations, for example, from the National Committee on Vital and Health Statistics and the Interagency Task Force on a Uniform Ambulatory Care Data Set[2] require no specifications beyond the designation of whether a person is of Hispanic origin. This is a designation of limited meaning that merges and confuses sociological differences likely to affect important social and health indicators. The term encompasses an extraordinary range of cultures, social conditions, and health practices. Average family income among Puerto Ricans, for example, is some 40% lower than among Cubans. While Mexicans receive less economic returns on their educations than the general population, Cubans do better than average.[3] As Alejandro Portes, one

of our nation's most distinguished students of ethnic influences, puts it:

> The immigration flows that created the "Hispanic" population originated overwhelmingly in Latin America, but this is a tenuous common denominator. For centuries, Colombians and Venezuelans, Peruvians and Ecuadoreans, Chileans and Argentineans have sharpened their national identity in repeated confrontations with each other. Upon arrival to the United States, they suddenly discover that they are all Hispanic. But the label does not dilute their strong national self-image; it simply creates a gulf between how people see themselves and their official designation. A parallel exists with the lumping together of such diverse groups as Vietnamese, Filipinos, Chinese and Indians under the label "Asian." These terms, so carelessly tossed about by government officials and the media, do symbolic violence to the groups so summarily aggregated.... (p. 39)[3]

Most of the data we have on disparities in health depend on classificatory information obtained by self-identification or occasionally by assessments made by treatment personnel or interviewers. Indirect analyses are also commonly used, as in the analysis of surnames for the purposes of inferring ethnicity. Observational and indirect methods are relatively crude and subject to large errors resulting from the difficulties in making valid observations, lack of reliability among observers, and, most importantly, because much of what has to be measured are not visible characteristics of the person.

Thus, most of our classifications are based on self-report and on how people choose to present themselves. To obtain valid information, it is essential that the questions be meaningful to respondents, that they know the answer to the question, and that they are motivated to accurately provide the requested information. Such terms as Hispanic, Chicano, Black, African-American, and American Indian mean varying things to different people and may not elicit the same response from one context or occasion to another. With each succeeding generation, Americans are less likely to know their ancestry, to be confused by backgrounds of parents and grandparents from varying countries, or to care little about the issue. Moreover, ancestry is an inherently confusing indicator. People from abroad may or may not share much, and origins from the same country may represent entirely different life experiences, habits and behavioral patterns.[4] It is difficult to see how unspecified ethnic identifiers can serve very usefully to help us move a health agenda forward.

The data become more useful as we focus on narrower subclassifications because the probability increases that we get closer to identifying persons more likely to share a common culture and perhaps more common behavioral patterns. But the United States is a highly heterogeneous nation, a product of immigration from all over the world, and except for a few major classifications we almost never have adequate samples to reliably describe the relevant subpopulations. In important instances, we can of course have special surveys as in the Hispanic HANES, or oversample, but this is expensive once we get beyond a few major groupings. Even then, immigrants often come to America from nations equally heterogeneous and complex, and the assumption that gross classifications can be useful is dubious. Knowing that a person is from India or China conveys little because of the complex social and cultural structures that characterize these large and diverse nations. Ethnicity measurement is an enormously complex area that serves political agendas more than health.' We need better proxy measures.

Socioeconomic Status Indicators

The single most important proxy, explaining much if not most of the effects we attribute to color and ethnicity, is social class. Social class, in the sociological sense, refers to the social location of a person in a community's stratification system and how it affects prestige, networks of association, and behavioral patterns as well as material welfare.[6] In mass society, the consciousness of class may be less developed or more shielded than in smaller communities where the social pecking order is more easily and clearly perceived, but it nevertheless affects almost every aspect of human endeavor including most important health outcomes. Social class predicts fetal wastage and perinatal death, infant mortality, developmental problems, overall mortality andmorbidity, well-being, and most indices of health care utilization. No study of health in general populations can be complete without careful consideration of the role of social class.

Americans, of course, like to believe there are no social classes, and in any case, we depend. almost completely on simple indicators as proxies in most surveys and studies. The major proxies have been education, income and occupation, although residence or ecological areas such as census tracts have been used

to represent the social homogeneity of particular neighborhoods that set the context for daily life. Identifying health districts in Harlem or neighborhoods in the District of Columbia by itself can convey a great deal about the likely risks for children and families and indices of health and social disorganization. It is always suspect to attribute to individuals the characteristics of their ecological location and, thus, whenever possible we seek individual measures.

Each of the key indicators of social class, or the more commonly used designation, socioeconomic status (SES), conveys somewhat different information. The main indicators-- education, income, and occupation--are reasonably highly correlated but each often has an independent predictive effect in multivariate studies. Researchers frequently combine them into a single index, but we clearly lose important information by doing so. If social class is a proxy for how we live, each SES indicator is a proxy for a somewhat different dimension of that life. Each conveys special information, but each also encompasses certain measurement difficulties.

Education is the indicator used most frequently, and no competent study of inequalities would exclude it as a covariate. It is consistently found to influence health outcomes. While most researchers view education as a proxy for other personal attributes or behaviors, the effects of education on health cannot usually be fully explained. Education is a proxy for such varied influences as knowledge, cognitive capacity, self-esteem, self-direction and sense of mastery. The acquisition of education selects to some extent for persons with greater innate abilities and persistence, but numerous studies also show that among youth with comparable abilities, those with better educated parents achieve more favorable outcomes at every point of the educational and occupational process.[7]

Education also provides familiarity and experience with bureaucracies and organizational routines. Better educated persons seem better able to manage the complex pathways of organizational systems and get what they need or want. This will be no small matter as the nation moves to more tightly organized health care arrangements, and where access may come to depend on how persons present symptoms and their persistence and ingenuity in managing bureaucratic gatekeepers.

Education is a relatively easy indicator to measure and most surveys simply seek to identify the number of years of schooling achieved. Common break points are high school graduation, some college, college graduation, and post-college professional education. Despite the success of education as an indicator, it's not clear that we fully capture its importance with the crude ways we represent it. American education is a highly diversified sector ranging greatly in quality and focus, and the character of education and its meaning within the culture change over time. Many Americans are at least partly educated abroad, and there may be little comparability or equivalence from one educational system to another. The meaning of a high school or college education was quite different fifty years ago than today, and every type of educational experience is confounded by the individual and social selective processes that bring people to particular levels and contexts.

My point is not that general surveys or record systems need to represent this complexity and subtlety, but that we need to ask thoughtful questions about the educational differences we find and the factors driving these relationships. In the Year 2000 Objectives, for example, objective 8.2 seeks to increase the high school graduation rate to at least 90 percent, "thereby reducing risks for multiple problem behaviors and poor mental and physical health."[8] The goal is a laudable one and probably worthy on its own terms, but it encompasses a causal expectation likely to be incorrect. We have no good basis for believing that school retention by itself would alleviate problems or improve health, and the causal pattern is more likely to go from problems and poor health to school dropout. More likely, both school drop out and poor health are part of a constellation of interrelated factors. A carefully thought out intervention strategy requires knowing a great deal about the processes underlying the relationships between level of schooling and these adverse events.

A second major SES indicator is income. Income is more difficult to measure than education because of its complexity as a concept, its many dimensions, and a reluctance in the population to report income. Do we measure income simply to represent the potential ability of an individual or family to access the necessities and amenities of daily life or do we also seek to estimate the person's standing in a hierarchy of prestige, power, influence and other resources. The concepts of need and poverty are relative, not absolute, and what people feel they need and lack depends in large part on what others around them have.

If income measurement is an attempt to represent access to necessities, it should take into account income in kind such as food stamps, Medicaid, housing subsidies, contributions of friends and relatives, and the like. If access to material resources is the major consideration, then family income is a more appropriate measure than personal income, but if income is also a proxy for control over resources, sense of self-worth, and over-

all well-being the individual measure will be more informative, and the distinction between earned and unearned income becomes relevant. Also, the use of family income as a measure assumes a reasonable sharing of income among attached individuals, but norms about sharing may vary by cultural group.

Many interesting ways of conceptualizing and measuring income are possible but measurement poses considerable challenge. Does one ask about weekly, monthly, or annual earnings and about before and after tax income? It is not clear that people really know the answers to the common types of income questions we ask. Do people in a household really know each other's incomes? Can people who think of their earnings in monthly or annual terms translate to weekly or vice-versa? Do people really know their income, before and after taxes, or do they simply know what their check or pay packet brings? What about people whose incomes fluctuate widely from week to week or who are unemployed for significant periods during the year. These, of course, are old questions and researchers have developed numerous solutions to them. In general, the quality of income data depends on the specificity with which they are collected, and accurate measurement requires a battery of carefully constructed questions. This is commonly not feasible in multipurpose surveys or in many administrative circumstances. Ideally, we would prefer that people consult their records and income tax returns to refresh their memories, but this is unlikely unless the study is especially designed with income as a major focus.

Because of public resistance, most surveys simply ask respondents to locate themselves or their families in broad income classes by identifying the income category. There is a belief that this procedure reduces resistance and leads to reasonably accurate ordinal categorization of persons and household units. Income has become such a routine measure that researchers give it little conceptual thought. With the growing literature on the importance of empowerment for a range of health outcomes distinctions between individual and family income and between earned and unearned income may serve different analytic needs. The income classification typically used by the NCHS includes all sources of household income among persons related to one another while unrelated individuals are classified according to their own incomes. This measure neglects important questions about how related and unrelated individuals within and outside of households share their resources and cultural differences in these patterns. It also has the disadvantage of not reflecting family size, which is rarely controlled in most data presentations. Since poor families tend to be larger than affluent ones, treating income this way underestimates the income effect.

Occupation is the third major SES indicator and provides somewhat different information than either income or education. The association between occupation and risk has important health implications and the indicator serves both as a general SES measure and a marker for specific types of risks. There is a long history of distinguished epidemiology and social analysis in Britain using the Registrar General's six categories of occupational classification varying from professionals at the top to unskilled manual labor at the bottom. Numerous studies demonstrated a robust association between this classification and disease and mortality outcomes. Moreover, these differences have persisted through time and have not been reduced with the National Health Service.[9] Studies of husbands and wives show that this classification is more than a measure of occupational risk. It captures to a large extent the circumstances of living that are shared by spouses."

Occupations, of course, change, and some become more important and others less important over time. Yet, occupational rankings over decades are relatively invariant" and provide an excellent basis for locating people's status positions in social structures. There are problems in coding new and. obsolescent occupations, and housewives and farmers pose special difficulties. In 1961, Duncan developed a socioeconomic index to rank occupations based on income and education data from the U.S. census,'* and this classification has been refined over time with new data.[13] In the United States, occupational coding is used infrequently in health in part because of the labor intensity of such classification. But as we become more concerned with issues in occupational health, it will be necessary to consider new classifications that capture risks associated with handling toxic materials, the stressfulness of work, the extent of autonomy and decision making, and other aspects of occupations associated with health outcomes. Epidemiologists have used occupational categories in creative ways that help us understand important disparities in health outcomes,[14] and such efforts can be instructive as we proceed.

Tracking Disparities

In identifying and characterizing disparities we should seek, to the extent possible, to identify generic factors that differentiate the population in terms of risk and outcomes, that can be reliably monitored over time and that are potentially modifiable. Establishing that persons of color or members of particular ethnic groups have less access to medical

care in the aggregate than others, tells us less than measures that identify the specific barriers that reduce help-seeking and adversely affect outcomes.

Geographic measures, similarly, do not tell us a great deal because of the diversity of conditions characterized by such designations as rural and urban. Population density is, of course, important because it represents the ability to concentrate resources people need and their accessibility to services they want. But areas so designated differ enormously in the availability of regional access to transportation, highways, and services. Moreover, density does not guarantee services (consider the case of urban ghettos), and car ownership, good highways, and personal resources give most rural residents accessibility to services. There are, of course, access barriers associated with low density populations. Small communities often have difficulty attracting physicians, and many rural hospitals have difficulty because of small markets and low occupancy. But the key issue in rural areas typically is not density per se, but the poverty of many inhabitants who then face compounded difficulties when they need and seek access to services.

In summary, our nation is now in the midst of major consideration of the future configuration of health care. Whatever the ultimate strategy and refinements, it is inevitable that improved data will be needed for planning and for monitoring access, use of services, and health outcomes. Disparities in care and outcomes will be used more commonly as indicators of deficiencies among health providers, as a way of focusing special efforts, or to identify impending problems. To do this well, we will need clear and relevant definitions of the types of concepts I have briefly reviewed, more uniform and reliable coding practices, and a higher level of validity. in general. We will need a capacity to monitor complex moving targets in a manner which is adaptable but also trustworthy. The challenge is large, and the course difficult, but the future of our health system depends upon 'it.

References

1. Bunker, J.P., Gomby, D.S. and Kehrer, B.H. (Eds.) Pathways to Health: The Role of Social Factors. Menlo Park: Henry J. Kaiser Foundation, 1989.

2. National Center for Health Statistics. The National Committee on Vital and Health Statistics, 1989. DHHS Publication No (PHS) 90-1205, Hyattsville, Maryland, June 1990, 47-73.

3. Portes, A. "The Longest Migration." New Republic 208:38-42, 1993.

4. Liberson, S. "The Enumeration of Ethnic and Racial Groups in the Census: Some Devilish Principles." Joint Canada-United States Conference on the Measurement of Ethnicity, April l-3, 1993, Ottawa, Canada.

5. Peterson, W. "Politics and the Measurement of Ethnicity." In Alonso, W. and Starr, P. (Eds.) The Politics of Numbers. New York: Russell Sage Foundation, 1987, 187-233.

6. Mechanic, D. "Socioeconomic Status and Health: An Examination of Underlying Processes." In Bunker et al. on. cit., 1989, 9-26.

7. Sewell, W.H. and Hauser, R.M. Education, Occupation and Earninss. New York: Academic Press, 1975.

8. U.S. Department of Health and Human Services. Healthy People 2000: National Health Promotion and Disease Prevention Objectives. DHHS Publication No. (PHS) 91-50212. Washington, D.C.: Government Printing Office, 1991, p. 619.

9. Townsend, P. and Davidson, N. (Eds.) Ineoualities in Health: The Black Report. New York: Penguin Books, 1982.

10. Benjamin, B. Social and Economic Factors Affectins Mortalitv. The Hague: Morton, 1965.

11. Hodge, R.W., Siegel, P.M. and Rossi, P.H. "Occupational Prestige in the United States, 1925-63." American Journal of Sociology 70:286-302, 1964.

12. Duncan, O.D. "A Socioeconomic Index for all Occupations." In Reiss, A.E. Jr. (Ed.) Occupations and Social Status. New York: Free Press, 1961.

13. Featherman, D.L. and Stevens, G. "A Revised Socioeconomic Index of Occupational Status: Application in Analysis of Sex Differences in Attainment." In Hauser, R., Mechanic, D., Halter, A.O. and Hauser, J.S. (Eds.) Social Structure and Behavior. New York: Academic Press, 1982, 141-181.

14. Link, B.G., Dohrenwend, B.P. and Skodal, A.E. "Socioeconomic Status and Schizophrenia: Noisome Occupational Characteristics as a Risk Factor." American Sociological Review 51:242-258, 1986.

Session I

What's News in Health Statistics?

Toward the year

2000

Refining the measures

TELLING THE TRUTH FROM THE TRASH

Victor Cohn, The Washington Post

Disraeli equated "damn lies" and statistics. This need not be. As a non-statistician, I set outtotrytounderstandwhatccmcepts journalists and others most need to test those who tell us something supposedly based on statistics or "studies" — or who say "most people do this or that," based on no knowledge.

I was prompted to do this by many examples of lack of understanding of statistics and statistical tests -- the tests of probable truth -- in daily reporting of the news of science, medicine, health, health risks, the environment and everything else.

I concluded that understanding statistics-- and science -- requires anunderstandingof six basic concepts; that understanding these rather simple concepts can teach us to ask "How do you know?" with a considerable degree of sophistication; and that the same attitude can help all of us examine our own assertions and decisions and help explain them to others.

The six concepts:

UNCERTAINTY. All science is almost always uncertain, or uncertain to a degree. Nature is complex and observation is inexact, so almost all anyone cancurrently say about any phenomenon or behavior is: There is a strong possibility that_ such-and-such is true, by all that is known at the moment. Better information may yield a better result tomorrow.

It is important to tell this to the public, including the media's readers, viewers and listeners, so they will understand why "they" say one thing today, another thing tomorrow.

PROBABILITY. Scientists live with uncertainty by measuring orobability. A "P value" is commonly used to help decide if a result could have seemed to occur just by chance, when there actually had been no effect. A P value of .05 or less -- .05 -- is most often regarded as low or desirable. It mans there are probably only five or fewer chances in 100 that this "statistically significant" result could have happened by chance. It also means that something like one actually negative study in 20 may yield a misleading false positive "result."

We should not trust a study that is not statisticallv significant. But remember that statistical significance alone may or may not mean practical -- or, in medicine, clinical -- significance.

POWER, the power of numbers. Power, in statistics, means the likelihood of finding something, of detecting a true effect, if it's there. Sample size -- numbers -- confers power. The greater the number of cases or subjects studied, the greater a conclusion's probable truth. We should be wary of studies with only a small number of subjects.

Also; we should ask about a study's "strength." The greater the odds against an association being a matter of chance, the greater its statistical strength.

BIAS. In science, bias means introducing spurious associations and reaching unreliable conclusions by failing to consider other influential factors or explanations -- "confounding variables." Among common biases to guard against: failing to take account of age, gender, occupation, nationality, race, income, health or influential behaviors like smoking. Such bias has been called "the most common cause of unreliable data'!

VARIABILITY. Almost everything that is measuredvaries from measurement to measurement. Every human experiment, repeated, has at least slightly and sometimes markedly different results. Among reasons: fluctuating physiologic response from moment to moment; measurement and observational limits or errors; biologic variations between persons arid populations.

This and all the above tell us that a single study rarely proves anything. The most believable studies and observations are those successfully repeated among different populations with much the same result.

THE HIERARCHY OF STUDIES. There is a hierarchy of studies from the least to the generally most believable, that is, from simple observation or "eyeballing" to controlled studies or experiments: those comparing one sample or population with another, under known or controlled conditions.

Is a study, or indeed any statement of supposed fact, worth reporting -- or believing? We can ask as a start: How do you know? What are your data, yournumbers? Hcwdidyougetthem? Have you, or anyone, done any studies or experiments? What is your degree of certainty or uncertainty? How valid --- in science, this means how accurate -- are your observations? And how reliable -- that is, how reproducible?

There is a lot more to know. As a start I must suggest reading NEWS & NUMBERS: a Guide to Reporting Statistical Claims and Controversies in Health and Other Fields (Iowa State university Press, 2121 S State Av., Ames IA 50010 / 515 292 0140). The ideas in it are those of many statisticians and defenders of health and truth. It's by --

-- Victor Cohn.

WHAT'S NEWS IN HEALTH STATISTICS?

Discussants:

Alan L. Otten
Wall Street Journal

Rhonda Roland
Cable Network News

(Not available for publication)

Session J

Race-Ethnicity Classification

Toward the year

2000

Refining the measures

AMERICAN INDIAN INFANT MORTALITY REVIEW IN WISCONSIN: BACKGROUND, PROBLEMS AND PROCESS

R.D. Nashold. Center for Health Statistics
R.A. Aronson, P.K. Wegehaupt, M.M. Shapiro

This paper provides background data, notes problems and outlines the process for American Indian (AmInd) infant mortality review in Wisconsin. It is hoped that the Wisconsin experience might be helpful elsewhere in the United States.

In America's Minorities -- The Demosranhics Of Diversitv, Wm. O'Hare noted "The United States is undergoing a transition from a predominately white population rooted in Western Culture, to a society composed of diverse racial and ethnic minorities. By the middle of the 21st century, today's minorities will comprise nearly one-half of all Americans."

Unless infant mortality is improved for minority populations, the United States will not attain the lower rates of other industrialized nations. The high African American (AfrAm) infant mortality rate in the U.S. is well known; however, data problems have at times obsured the comparably high AmInd rates.

The Amind infant mortality rate in Wisconsin is about double the white rate (Figure 1). This difference was not obvious, in part due to the practice of not publishing rates for fewer than 20 events annually. Three-year or five-year average rates were published only in special studies which were done about every five years.

Demographics. The subjectivity of race becomes obvious when considering the apparent five-fold increase in the U.S. AmInd population since 1950 (Table 1) and the three-fold increase in Wisconsin (Table 2). This increase cannot be explained by natural increase nor by migration; therefore it is necessary to look at self-identification and at changes in definition. Because of this rapid increase, it is now feasible to do AmInd infant mortality review independently in about 16 states (over 35,000 AmInd population, see Tables 3 and 4). Most other states could participate in joint efforts with neighboring states. New England and the East South Central states could undertake multiple state review (Table 5).

Review Process. The major participants in the review process in Wisconsin are as follows:

- Indian communities
- Great Lakes Inter-Tribal Council
- U.S. Indian Health Service
- Maternal and Child Health, Division of Health (DOH)
- Center for Health Statistics, DOH
- Local public health agencies
- Local social services agencies
- Coroners, medical examiners, pathologists
- Hospital medical records staff
- Clinic medical records staff
- Sudden Infant Death Center
- Infant Death Review Committee

Wisconsin is essentially following the recommendations in A Manual for Fetal and Infant Mortality Review by the American College of Obstetricians and Gynecologists (Paul H. Wise, editor: Louise M. Wulff, project director).

The ACOG Manual discusses the following subject areas:

- Transforming analysis into local action
- Developing a methodology that meets local needs
- The analysis of vital statistics
- Linking vital statistics to other data sets
- Performing an infant mortality medical record review
- Interview-based methods
- The use of expert panels in the infant mortality review process
- Legal considerations
- Critical definitions and criteria
- Tobacco
- Maternal drug use
- Assessing prenatal care
- Creating an integrated fetal and infant mortality review system
- Presenting findings and recommendations

Confidentiality policies are discussed under legal considerations and Wisconsin essentiallyv follows the recommendations set forth. Personal identifying information is removed and an identification number is assigned to each case. Similarly, identification numbers are assigned to physicians, hospitals and clinics. Reviewers who might identify particular cases have signed confidentiality agreements, not to reveal identity. Law enforcement agencies are not represented in the review process; however, the forensic viewpoint is represented by the Medical Examiner for Milwaukee County.

In Wisconsin, fetal mortality review is still at the discussion stage. By including fetal mortality (20 weeks or more gestation), a more complete picture of reproductive loss could be obtained. By including fetal mortality and thereby shifting to perinatal mortality review, the number of cases would increase by 5-10 annually. This would help overcome the small-number problem now encountered when analyzing AmInd infant mortality data. The negative aspect of including fetal deaths is the extra effort needed to assemble the data, and the dilution of the

summary data with less complete data. Under-reporting is greater for fetal deaths and the quality of the data reported is also lower as noted by the lack of complete cause of death information on many fetal death reports.

General **Findings.** When summarizing birth and death data by race it is appropriate to state that race and ethnicity are used as study variables because of the difficulty of obtaining more direct social or economic measures. State poverty data classified by several demographic variables are available in the decennial census. Figure 2 shows the high percentage of minority children living in poverty.

Table 6 lists the number of AmInd infant deaths for Wisconsin from 1987 through 1991. For this five-year period, when tabulating by race of mother, 61 infant deaths were reported for an average of 12 per year and a rate of 15.2 per 1,000 live births. When the race of either parent is used the total number increases to 70 and the annual average to 14; however, the rate falls to 13.4. This is because about 25 percent more live births were added to the base when the rate was calculated.

For American Indians, the definition of race is complicated by the high proportion of marriages with persons from other racial or ethnic groups. In Wisconsin, the last year for which there are data, 1988, shows this to be over 60 percent. The problem of defining race is only partially solved by considering either parent because for 63 percent of the AmInd live births, the father was not reported as being married to the mother. Therefore, race and other information are generally not available for the father.

Five-year moving averages (Table 7) provide a more accurate view of the infant mortality trend. From 1983 to 1991 the rate has fallen by nearly 27 percent (Figure 3); however, it remains at a level double the white rate.

Table 8 illustrates three major problems encountered in AmInd infant mortality review; interstate record exchange, definitions and accuracy. Any of these problems can seriously affect results when numbers are small.

High AmInd Mortality Rates. Table 9 lists causes of mortality that have called attention to the AmInd population and to maternal and infant risks. A study of the unusually high SIDS rates for Native Americans in Wisconsin was one of the reasons for undertaking AmInd infant mortality review. In a study using data from 1978-87, SIDS was highest for the AmInd population. For the five year period shown in Figure 4, AmInd and AfrAm SIDS rates are about equally high.

Figure 5 shows AmInd infant mortality below that of the AfrAm population. For the most recent years the AfrAm rate has declined so that it is similar to the AmInd rate. Appropriately, infant mortality review has now also- begun for the AfrAm population in Milwaukee.

AmInd infant deaths have the highest proportion (56 percent) occurring in the post-neonatal period (Figure 6.) When infant deaths occur in the post-neonatal period, there is a longer time for care factors and other environmental factors to affect survival.

The high diabetes death rate for American Indians reflects the high prevalance of this disease which in turn affects birth outcomes, including higher birth weights (Figure 7).

Age-adjusted heart disease mortality rates are highest for the AmInd population (Figure 8). Coronary heart disease morbidity, in the AmInd female population of childbearing age, and associated conditions may be risk factors for delivery.

Chronic liver disease includes cirrhosis due to alcoholism. Alcohol abuse during pregnancy and in the home following birth is a risk factor for infants. Chronic liver disease death rates for the AmInd population are four times higher than the white rate (Figure 9). Some of this difference may be due to variation in reporting.

AmInd unintentional injury mortality rates are more than double white rates (Figure 10). High rates of unintentional injury mortality and morbidity are linked to alcohol abuse and to poverty. During the 1987-1991 period four (6 percent) of the 70 AmInd infant deaths were reported as due to injury.

Chronic obstructivepulmonarydisease (COPD) death rates are higher for the AmInd population than for the AfrAm or white population (Figure 11). Smoking is a causal factor for COPD. Smoking is also a risk factor for SIDS and other causes of infant death including some deaths due to fires, burns and suffocation.

Birth Risk Factors. Table 10 is a list of traditional risk factors collected on the U.S. Standard Birth Certificate. When making comparisons for five race/ethnic groups the AmInd population generally ranks in the mid-range (second or third), Wisconsin AmInd women have the highest proportion of smokers but the lowest proportion of low-birthweight infants.

Medical risk factors include pregnancy-associated hypertension, gestational diabetes, previous large infants, pre-term infants, STD, urinary tract infections and a number of other medical conditions. Overall these risks are highest for the AfrAm population and second highest for AmInd women (Figure 12).

The AmInd birth rate in Wisconsin is similar to the Hispanic birth rate (Figure 13). Both are lower than the AfrAm birth rate and well below the Southeast Asian (largely Hmong) birth rate. The largely

Hmong crude birth rate is four times greater than that of the white population.

Fertility rates for women age 15-44 are in the same order as crude birth rates; however, differences are even greater (Figure 14). The AmInd fertility rate is 86 per 1,000 women in the child-bearing age range (15-44). In contrast, the Hmong rate is 368 and the white rate is 60. High fertility rates are similar to those of earlier generations of Wisconsin immigrants.

AmInd women are second in terms of high parity , (4th or higher order of birth). During 1991, 22 percent of the births to AmInd women were birth order four or above. This is in contrast to 63 percent for mothers from Southeast Asia and nine percent for white mothers.

The percent of AmInd mothers below age 18 is 8.4, or third among the five racial/ethnic categories (Figure 16).

The AmInd population is second highest for unmarried mothers at 63 percent. The percentage for the largely Hmong group is the lowest at 8.5 (Figure 17). It should be noted that there may be fundamental difference in marriage customs and traditions from one racial/ethnic group to another.

AmInd women rank third highest for lack of prenatal care during the first three months, of pregnancy. The largely Hmong group and AfrAm women rank higher (Figure 18).

Among the four minorities listed, AmInd mothers have the lowest percent not completing high school (Figure 19). Still, they are nearly four times less likely than white women giving birth to have completed high school.

AmInd mothers have a percentage of low-birthweight infants that is about the same as that of white mothers (Figure 20). There are likely to be intervening variable such as higher incidence of diabetes for AmInd women. Infants born to diabetic women are often fragile despite their high birth weights.

Figure 21 dramatically shows the high percent of AmInd women who reported smoking during pregnancy. At 48 percent it is more than double the white percent and more than 10 times the percentage reported by women from Southeast Asia who gave birth in 1991. Historically, tobacco has been part of AmInd tradition; however, cigarettes were not.

Linked Birth/Death Data. Data from the AmInd birth file can be compared with data from the linked birth/death file (Infant death file). Table 11 compares the same risk factors from the two files. As expected, most of these risk factors occur in higher proportion for infant deaths than for live births.

Figure 22 points to the association between smoking by AmInd women during pregnancy and infant survival. Of the infants who died, 58 percent of. their mothers smoked during pregnancy; whereas, for the infants who survived the first year of life, 48 percent of their mothers smoked during pregnancy. This simple association between reported maternal smoking and infant death yields an odds ratio of 1.5. For 85 percent of the AmInd SIDS deaths the mother smoked during pregnancy. For non-SIDS deaths, 56 percent of the mothers smoked during pregnancy (Figure 23).

General Discussion. The work completed to date provides summary data as background for more intensive case review by a medical review committee. The value of showing these associations again in a single state is primarily for health education and public health program purposes.

AmInd health workers in Wisconsin have commented that although they knew the AmInd infant death rate was higher than the State rate, they were unaware of the extent of the problem until the data were tabulated for this review. public health workers in Wisconsin are convinced that infant mortality review is a worthwhile endeavor.

References. Available upon request.

TABLES AND FIGURES

Table 1. American Indian Population United States, 1950-1990

Year	Number	Percent Increase
1950	357,499	
		46.5
1960	523,591	
		51.4
1970	792,730	
		72.1
1990	1,364,033	
		37.7
1990	1,878,285	

Table 2. American Indian Population Wisconsin, 1950-1990

Year	Number	Percent Increase
1950	,12,196	
		17.2
1960	14,297	
		16.2
1970	16,616	
		76.5
1990	29,320	
		34.3
1990	39,307	

Table 3. States with Over 40,000 AmInd Residents, 1990

Rank	State	Number
1	Oklahoma	252,420
2	California	242,164
3	Arizona	203,527
4	New Mexico	134,355
5	Alaska	95,699
6	Washington	91,699
7	North Carolina	80,155
8	Texas	65,077
9	New York	62,651
10	Michigan	55,639
11	South Dakota	50,575
12	Minnesota	49,909
13	Montana	47,679

Table 4. States with 20,000 to 40,000 AmInd Residents, 1990

Rank	State	Number
14	Wisconsin	39,387
15	Oregon	38.496
16	Florida	36,355
17	Colorado	27,776
18	North Dakota	25,917
19	Utah	24,283
20	Kansas	21,965
21	Illinois	21,836
22	Ohio	20.358

Table 5. **AmInd** Population, United States, Regions, and Divisions, 1990

	Total	Percent
United States	1,959,234	100.0
Northeast	**125,148**	6.4
New England	32,794	
Mid. Atlantic	92,354	
Midwest	337,899	17.3
B.N. Central	149,939	
W.N. Central	187,960	
South	**562,731**	28.7
S. Atlantic	172,281	
E.S. Central	40,839	
W.S. Central	349,611	
west	933,456	47.6
Mountain	480,516	
Pacific	452,940	

Table 6. American Indian **Infant Mortality** Wisconsin, 1987-1991

Year	Mother AmInd	Mother or Father American Indian
1987	11	12
1988	16	18
1989	16	18
1990	6	8
1991	12	14
Total	61	70
	5 Year Average	
Number	12.2	14.0
Rate/1000 L.B.	15.2	13.4

Table 7. **AmInd** Infant Mortality, WI Five Year Moving Average Rates, **1983-1991** By Race of Mother

Years	Total Infant Deaths	Average Rate*
1983-1987	76	20.7
1984-1988	70	18.6
1985-1989	67	17.4
1986-1990	61	15.5
1987-1991	61	15.2

*Number of resident deaths per 1000 resident live births.

Table 8. **AmInd** Infant Deaths Wisconsin, 1987-1991

	Number	Percent
Infant Deaths	70	100
Out of State	8	11
Father AmInd	9	13
Child listed as White on death certificate	13	19

Table 9. **AmInd Mortality, Rank Among** Other Racial/Ethnic Groups, WI, 1987-1991

Sudden Infant Death Syndrome	Highest
Infant Mortality	2nd Highest
Coronary Heart Disease	Highest
Unintentional Injury	Highest
Diabetes	Highest
Chronic Liver Disease	Highest
Chronic Obst. Pul. Dis.	Highest

Table 10. AmInd Birth Risk Factors, Rank Among **Five** Racial/Ethnic Groups, WI, 1991 **(AmInd, AfrAm,** Hisp, **HLCV*,** White)

Medical Risk at Delivery	2nd Highest
High Fertility	3rd Highest
Four or More Children	2nd Highest
Mother < age 18	3rd Highest
Single Mothers	2nd Highest
No First Trimester Prenatal Care	3rd Highest
Mothers Not Completing High School	2nd Lowest
Smoked During Pregnancy	Highest
Low Birth Weight	Lowest

*HLCV - Hmong, Lao., Cambodian, Vietnamese

Table 11. Risk Factors for **AmInd** Live Births and Infant Deaths in Wisconsin

Risk Factor	Percent of Live Births 1991	Percent of Infant Deaths 1987-1991
Single Mother	63.2	64.3
Mother <H.S. Ed.	42.6	44.3
No 1st Tri. Care	38.5	48.6
Birth Order 4+	21.8	**43.1***
Low Birth Wt.	4.9	35.8
Smoked During Pregnancy	48.2	**66.7****

* Excludes unknowns
** Available 1989-1991

Figure 1.
Infant Mortality Rates
American Indian and White Population
Wisconsin, Average 1987-1991

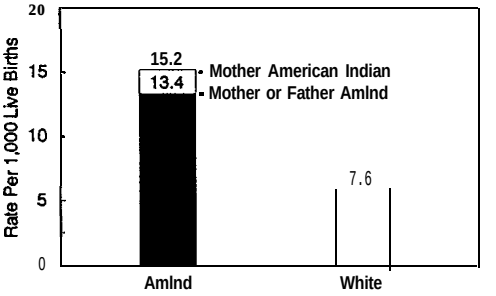


Figure 2.
Children Under Age 5 Living Below The Poverty Level
By Race/Ethnicity, Wisconsin 1989

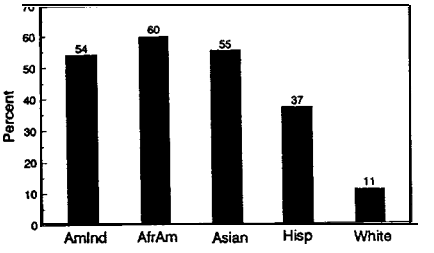
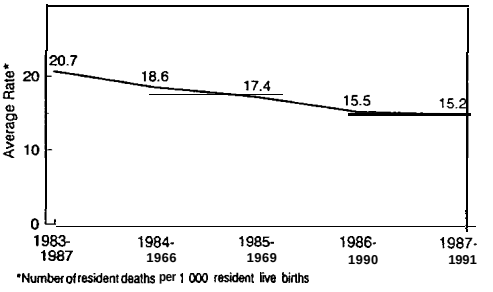


Figure 3.
American Indian Infant Mortality. Wisconsin
Five Year Moving Average Rates, 1983-1991
By Race of Mother



*Number of resident deaths per 1 000 resident live births

Figure 4.
SIDS Mortality Rates By Race/Ethnicity of Mother
Wisconsin 1987-1991

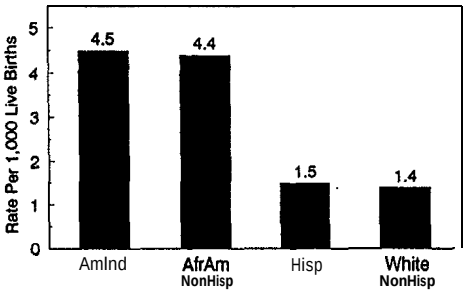


Figure 5.
Infant Mortality By Race/Ethnicity
Per 1,000 Live Births, WI 1989-1991

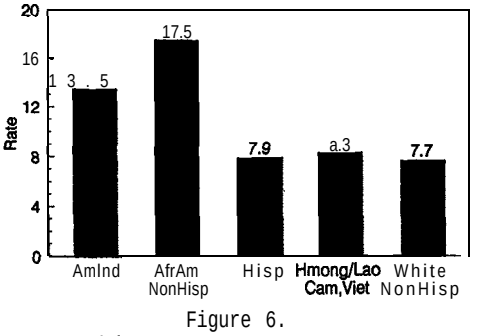


Figure 6.
Percent of Infant Deaths That Are Post-Neonatal Deaths
By Race/Ethnicity, Wisconsin, 1987-1991

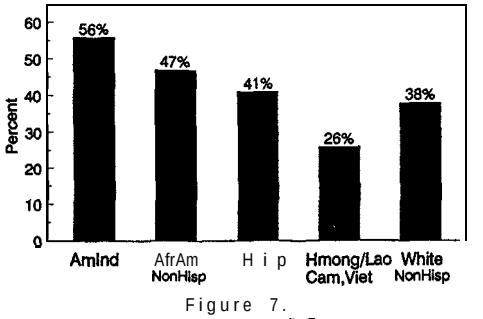
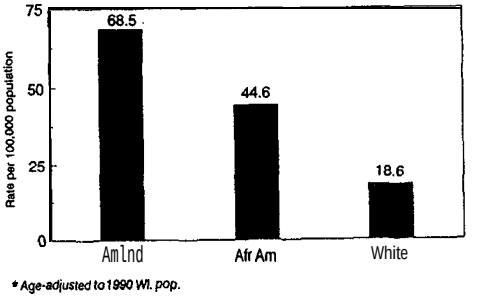
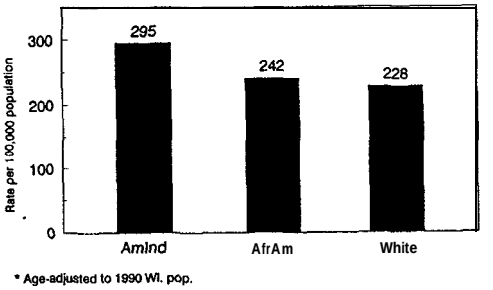


Figure 7.
Diabetes Deaths By Race
Age-Adjusted Death Rates*, WI 1987-1991



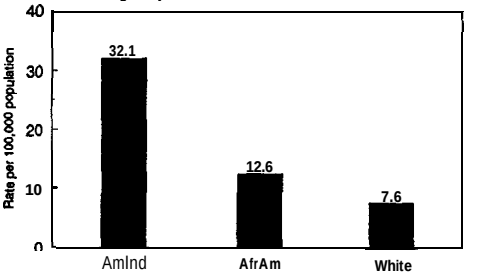
* Age-adjusted to 1990 WI. pop.

Figure a.
Coronary Head Disease (ICD 410-414) By Race
Age-Adjusted Death Rates'. WI 1987-1991



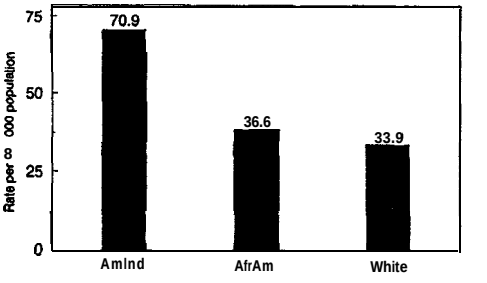
* Age-adjusted to 1990 WI. pop.

Figure 9.
Chronic Liver Disease Deaths By Race
Age-Adjusted Death Rates*. WI 1967-1991



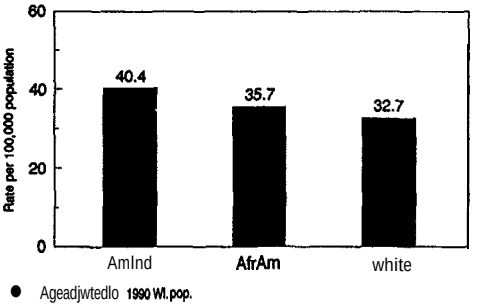
* Age-adjusted to 1990 WI. pop.

Figure lo.
Unintentional Injury Deaths By Race
Age-Adjusted Death Rates', WI 1967.1991



* Age-adjusted to 1990 WI. pop.

Figure 11.
Chronic Obstructive Pulmonary Disease Deaths By Race
Age-Adjusted Death Rates*, WI 1987-1991



● Ageadjwtedlo 1990 WI.pop.

Figure 12.
Medical Risk Factors indicated
By Race/Ethnicity of Mother, WI 1991

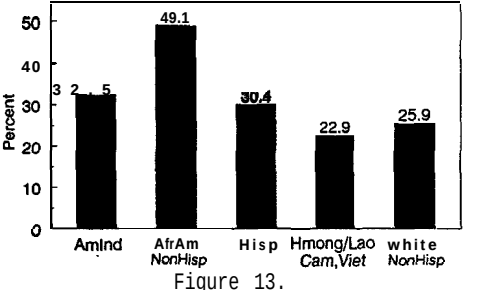


Figure 13.
Live Births Per 1,000 Population
By Race/Ethnicity, WI 1990

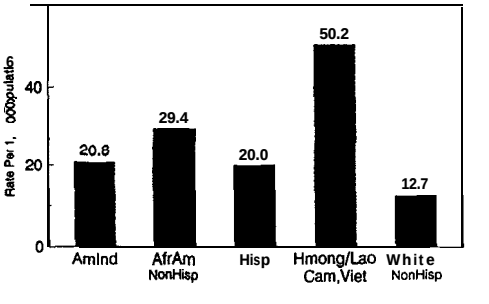


Figure 14.
Live Births Per 1,000 Females Age 15-44
By Race/Ethnicity, Wisconsin. 1990

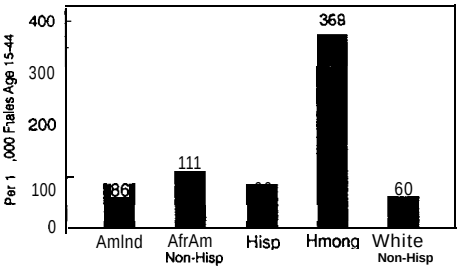


Figure 15.
Live Births That Are Birth Order Fourth or Above
By Race/Ethnicity, Wisconsin 1991

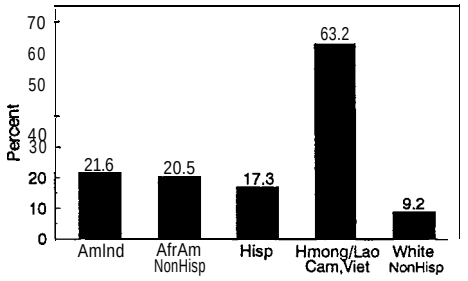


Figure 16.
Live Births To Women Below Age 18
By Race/Ethnicity, WI 1991

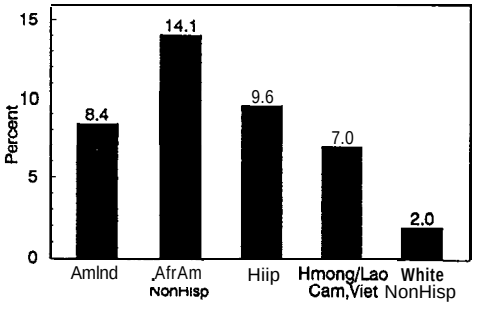


Figure 17.
Live Births to Unmarried Mothers
By Race/Ethnicity, WI 1991

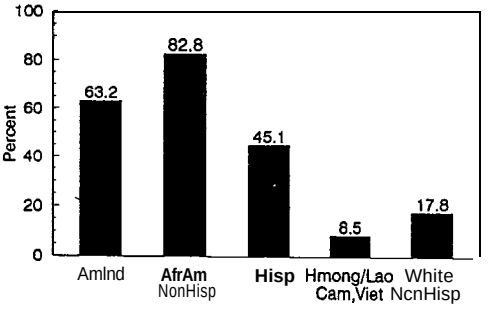


Figure 18.
No First Trimester Prenatal Care
By Race/Ethnicity, WI 1991

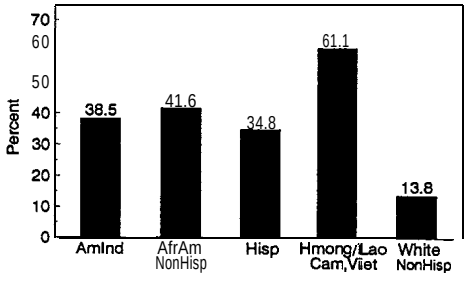


Figure 19.
Live Births to Mothers
Who Did Not Complete High School
By Race/Ethnicity, Wisconsin 1991

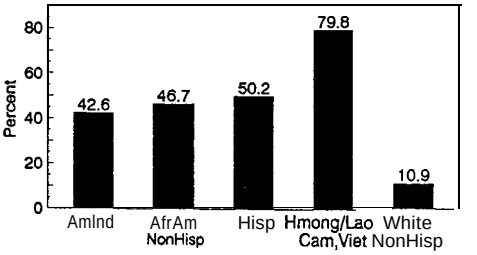


Figure 20.
Live Births with Low Birth Weight
by Race/Ethnicity, WI 1991

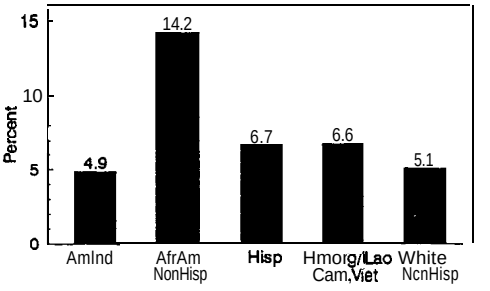


Figure 21.
Women Who Smoke During Pregnancy
Rates Per 1 ,000 By Race/Ethnicity, Wi 1991

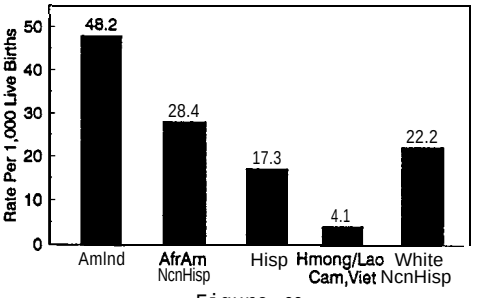


Figure 22.
AmInd Infant Deaths And Survivors
By Percent Smoking Mothers, WI 1989-1991

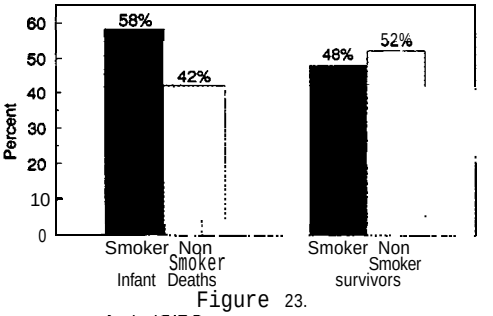
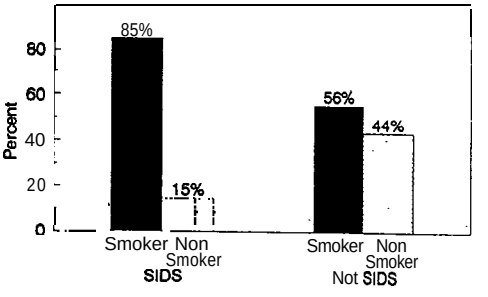


Figure 23.
AmIndSIDS And Other Infant Deaths
By Percent Smoking Mothers, WI 1989-1992



TRIBE-SPECIFIC CODING AND REPORTING IN NEW MEXICO

C.M. Halasan, NM Bureau of Vital Records & Health Statistics

Today I will discuss methodological issues on the collecting and reporting of Native American tribal vital events in NM. I will also try to share some context on why coding and reporting of births and deaths by tribe is needed and why it has been worthwhile.

RACE & ETHNICITY IN NEW MEXICO

In general, we use race and ethnic categories recommended by NCHS. These are our race categories for birth and death records: White, Black, Native American, Chinese, Japanese, Hawaiian, Filipino, Other Asian or Pacific Islander, and Other Non-White. (FIGURE 1) We also use NCHS' race algorithm to assign the race of the child.

Thirty-eight percent of NM residents are Hispanic. Thus, most of the ethnic categories we collect are Hispanic ethnicities: Puerto Rican, Cuban, Mexican, Spanish, European Spanish, Central or South American, Other Spanish. The only other category we collect and report is Non-Hispanic ethnicity. (FIGURE 2) Most persons of Hispanic origin within New Mexico identify themselves as Mexican or Spanish. Those who identify themselves as Spanish are generally those whose families have been in New Mexico for many generations and claim Spanish conquistador heritage. The majority of NM Hispanics are of the White race.

FIGURE 1:
NEW MEXICO RACE/ETHNICITY CODING
(Birth & Death Records)

RACE CATEGORIES:

100 White
200 Black
300 Indian, Aleut, Eskimo
400 Chinese
500 Japanese
600 Other Non-White
900 Filipino
000 Other Asian or Pacific Islander
900 Unknown, Not Stated

In NM, American Indian vital statistics by tribal groups have been collected since the 1950s; prior to this vital events were recorded by the single category of Indian. (FIGURE 3) In the last few decades all Native Americans (which I use interchangeably with American Indian) were coded into five tribal groups: Zuni, Pueblo, Navajo, Apache, and Other.

FIGURE 2:
NEW MEXICO RACE/ETHNICITY CODING
(Birth & Death Records)

ETHNICITY CATEGORIES:

01 Mexican
02 Puerto Rican
03 Cuban
04 Central or South American
05 European Spanish
06 Spanish
07 Non-Spanish White
08 Spanish
09 Unknown

Persons reported as Other generally are of non-NM tribes, such as Cree and Sioux. Since 1988, our Division has worked cooperatively with the Indian Health Service by providing IHS with vital statistics by tribal groups for their planning and evaluation needs. Out of this cooperative relationship arose the idea and the incentive of capturing vital events by specific tribe. Thus in 1990 NM began coding tribal affiliations by specific tribes. Our tribe-specific codes recognize all of NM's resident tribes and encompass one Navajo, 2 Apache, and 20 Pueblo tribes. (FIGURE 4) I emphasize that we developed this primarily because of requests from tribal leaders and service providers and because our Division Director in 1989 made it a priority to develop this capacity. Tribal leaders are interested in answering basic vital statistics and service delivery questions about their own tribes. Anthony Ortiz, Statistical Director of our Vital Records, and Anthony Yepa, of the Albuquerque

FIGURE 3:
NEW MEXICO TRIBAL CODING
(Birth & Death Records)

300 INDIAN, ALEUT, ESKIMO:

***PRE-1990 (TRIBAL GROUPS):**

30 Pueblo
50 Navajo
60 Zuni
70 Apache
80 All Other Indian

FIGURE 4:
NEW MEXICO TRIBAL CODING

300 INDIAN, ALEUT, ESKIMO:

*1990 TO PRESENT (TRIBE-SPECIFIC):

01 Apache, Jicarilla
02 Apache, Mescalero
03 Navajo
04 Pojoaque Pueblo
05 Acoma Pueblo
.
.
23 Ysleta 'Del Sur Pueblo
98 Other

Indian Health Service's Office of Planning and Evaluation, worked together to implement tribe-specific coding for NM. Mr. Yepa, with input from tribal leaders, provided the tribal names to use. Mr. Ortiz oversaw the incorporation of necessary modifications into our mainframe vital statistics registration systems. This included developing coding structures (that is, each tribe was assigned a 2-digit numeric value in addition to its race and ethnic codes), working with our Data Processing folks to modify key entry, modifying record layouts, and expanding mainframe edits to include the new data elements. These edits, which are run daily for births and monthly for deaths, allow us to review and correct errors before they are finalized in our systems. Our electronic birth certificate system also has a screen for collecting the same data elements.

(FIGURE 5) This map shows locations of American Indian reservations in NM. Pueblo lands are in blue, Navajo in purple! and Apache in brown. For most births, it is clear which tribe to assign a child to because the parents are usually from the same NM tribe. If both parents are of different NM tribes then we code the child according to the residence reported on the birth record. For example if one parent is Navajo and the other is Santo Domingo Pueblo, and the address given is Santo Domingo, then the child is coded Santo Domingo Pueblo. If one parent is from a New Mexican tribe and the other is of a non-NM tribe, then we code the child to the NM tribe. If both parents are of non-NM tribes, then we code the child's ethnicity as All Other Indian rather than a specific tribe. For deaths, it is straightforward: the tribe we code is whatever tribe the decedent is reported to be.

SIMILARITIES AND HETEROGENEITY OF TRIBES

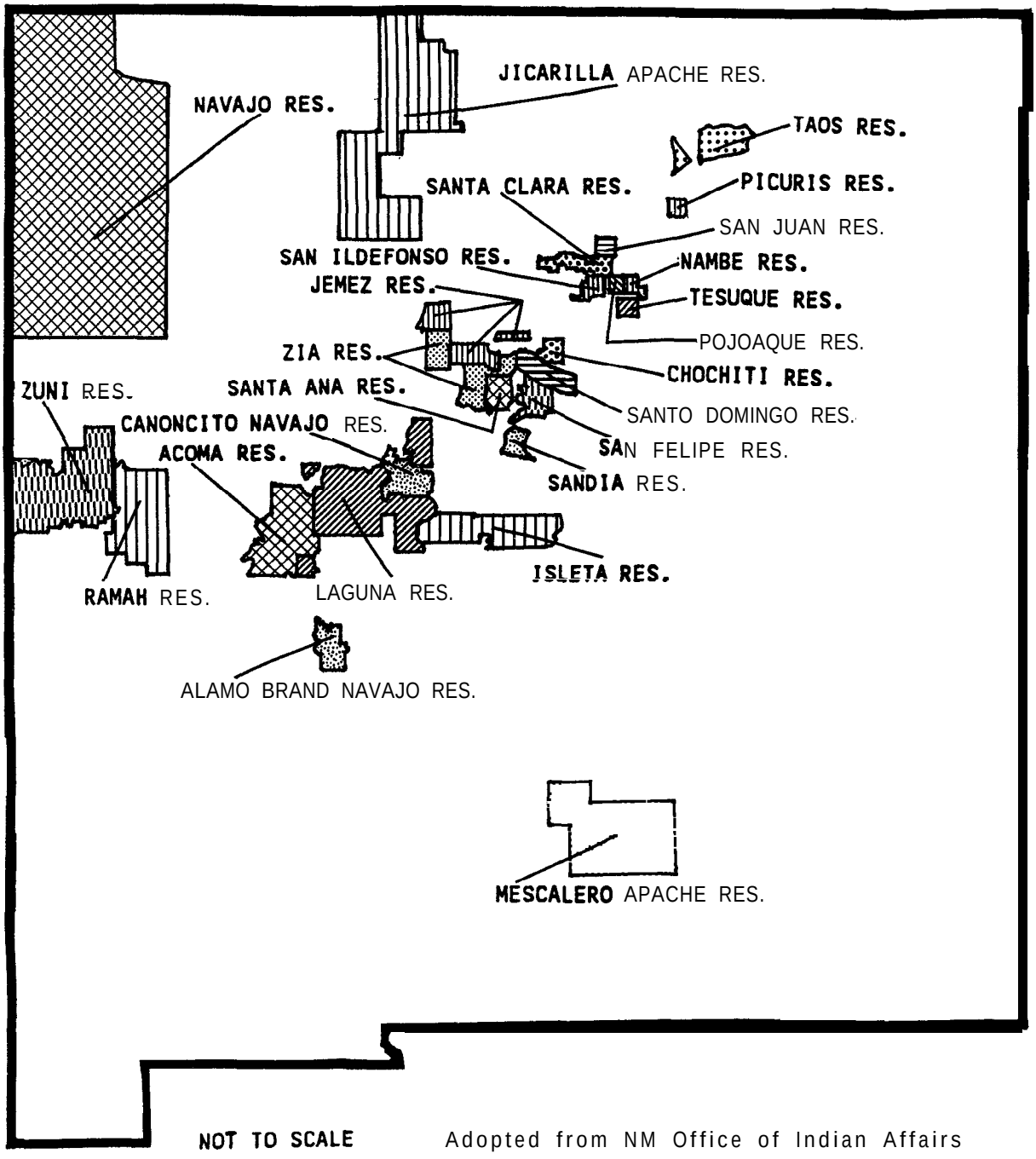
To provide some context I want to briefly discuss the differences and similarities among the tribes. I am not, to any extent, knowledgeable about Native Americans; so I restrict myself to general comments. The Navajo and Apache languages are of the Athapaskan language family. These tribes were originally small highly mobile bands with no central political structure and were not confined to reservations until the late 19th century. The Navajo Nation (which is in three states including NM) has the most people and as such has the most extensive system of governance of tribes within NM. Both the Navajo Nation and the Apache tribes have constitutional tribal governments. Unlike the Pueblos, Navajos live in highly dispersed residences throughout their huge reservation. However, a significant percentage (around 20%) of Navajos live in urban areas in and outside reservation land. Traditional livelihoods in this century include sheep raising and silver work and weaving. The Navajo people have also adopted some basic agricultural and technological practices from the Pueblo Indians. Though there have been intermarriage between Navajo and Pueblo peoples, each have maintained their own traditions. The two Apache tribes in New Mexico now also live in communities within their reservations.

NM's 20 Pueblo tribes subdivide into three unrelated linguistic groups and can trace their agriculturally-based village life back through at least the 1200s and 1300s. All Pueblo tribes share common traditions and beliefs but the expressions of these are unique to each tribe. Traditionally Pueblos have been farmers along the Rio Grande but since the turn of this century have produced distinctive craft products for income from tourism. Unlike the Apache and Navajo governments, some Pueblo tribes have written constitutions while others operate under unwritten rules. Just as with the other states of this country, NM tribes act in concert with and separately from other tribes.

ADVANTAGES OF TRIBAL BREAKDOWNS

Why break out Native Americans beyond the single race stratum? Why do we break out different Hispanic groups? The ability to look at each specific tribe confers the same advantages as the ability to look at any ethnicity/group separate from the majority population(s). We see how these groups may differ from the larger groups -- differences that may be obscured when a smaller group is placed within a larger group or when smaller groups are defined and characterized by the language of the larger group. With tribal breakouts we can describe some of these inter-group

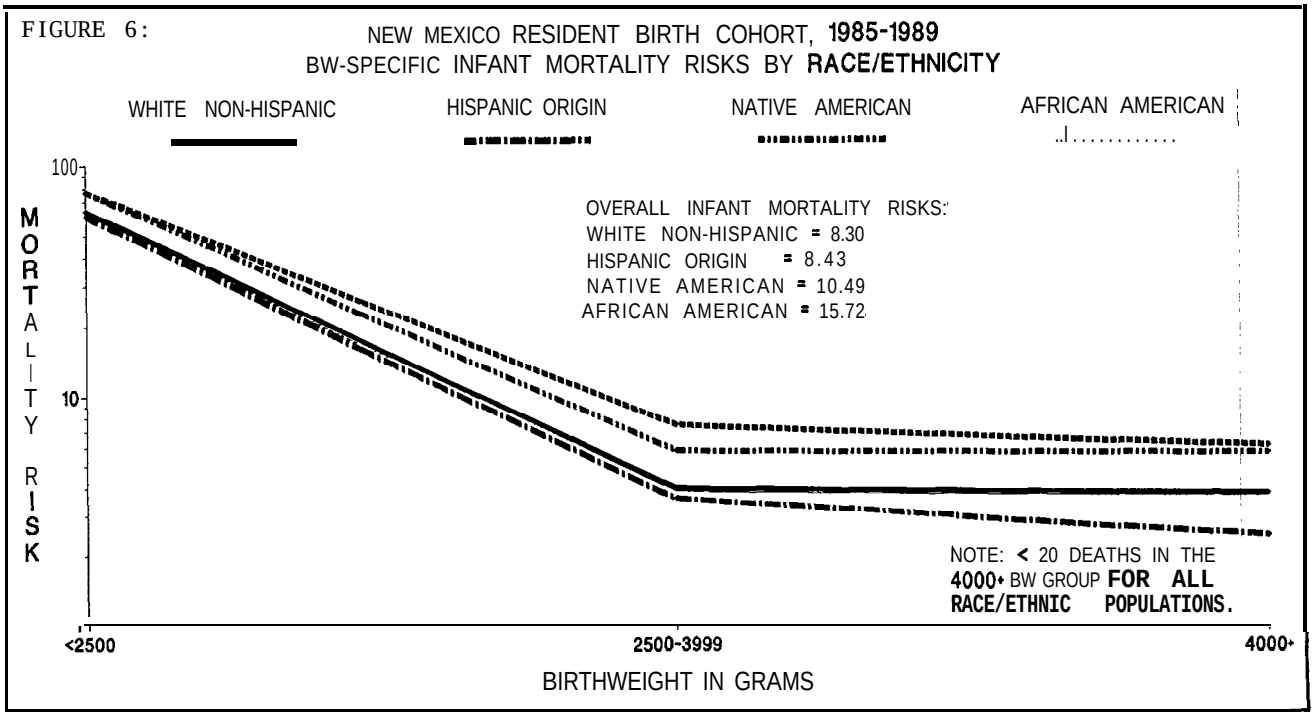
FIGURE 5: INDIAN LANDS in NEW MEXICO



Adopted from NM Office of Indian Affairs

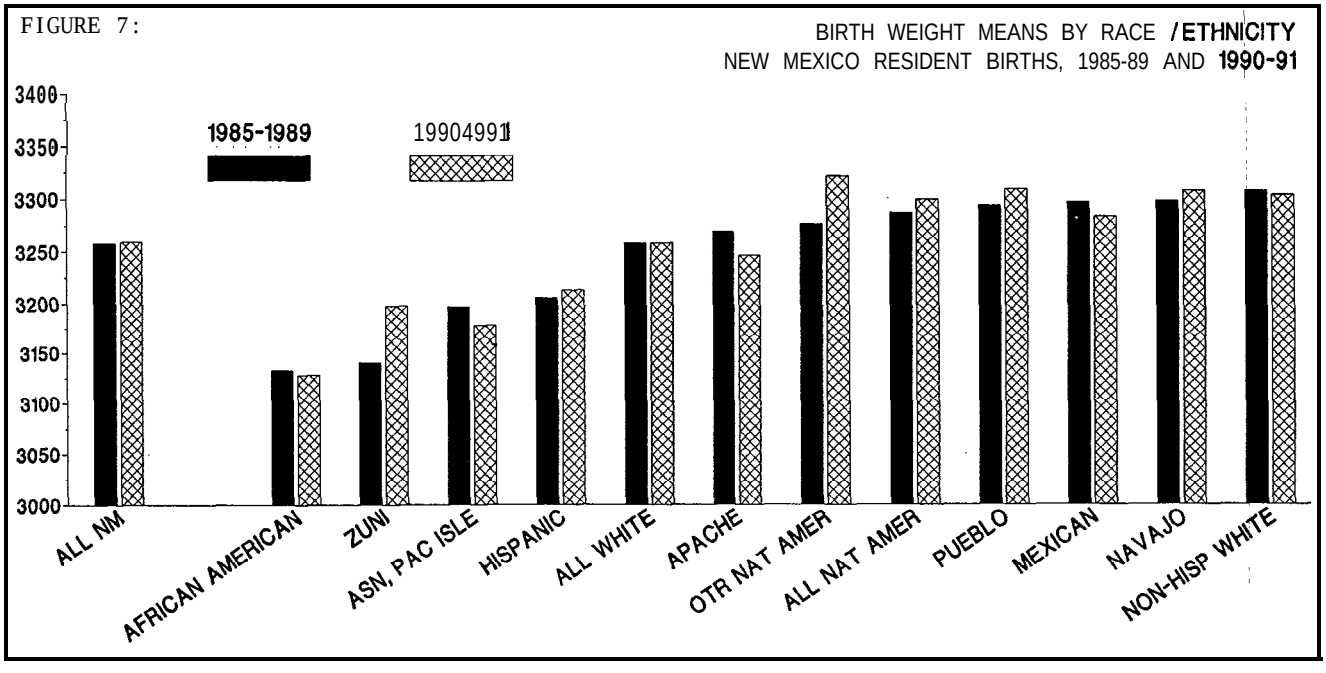
differences in our limited language of vital statistics.

What do we know about New Mexico Native Americans? Let's look first at the group as a whole. They comprise 9% of all NM residents, had the highest percentage of persons under 35 years old of any group *in* NM in 1990, and had the highest rate of natural increase (birth to death ratio) from 1980 to 1990. [FIGURE 6] Infant mortality rates among Native Americans for the past decade have generally been higher than the majority White population. Accidents were the leading cause of death for Native Americans in 1990, in contrast to being the 3rd or lower cause of death for other NM populations. This group had the highest percentage of low prenatal care in 1990-1991.



Let's see what we find by tribal groups. [FIGURE 7) The birthweight mean of tribal and other NM ethnic groups are shown in Figure 7 for 1985-89 and 1990-91. The birthweight mean of Native Americans is similar to the means of Non-Hispanic White and Mexican. However, the Zuni and Apache tribal groups have lower means than the Navajo and Pueblo tribal groups. We can address the issues expressed in this shorthand measure -- in this case, birthweight mean --more meaningfully because now we can ask more precise questions.

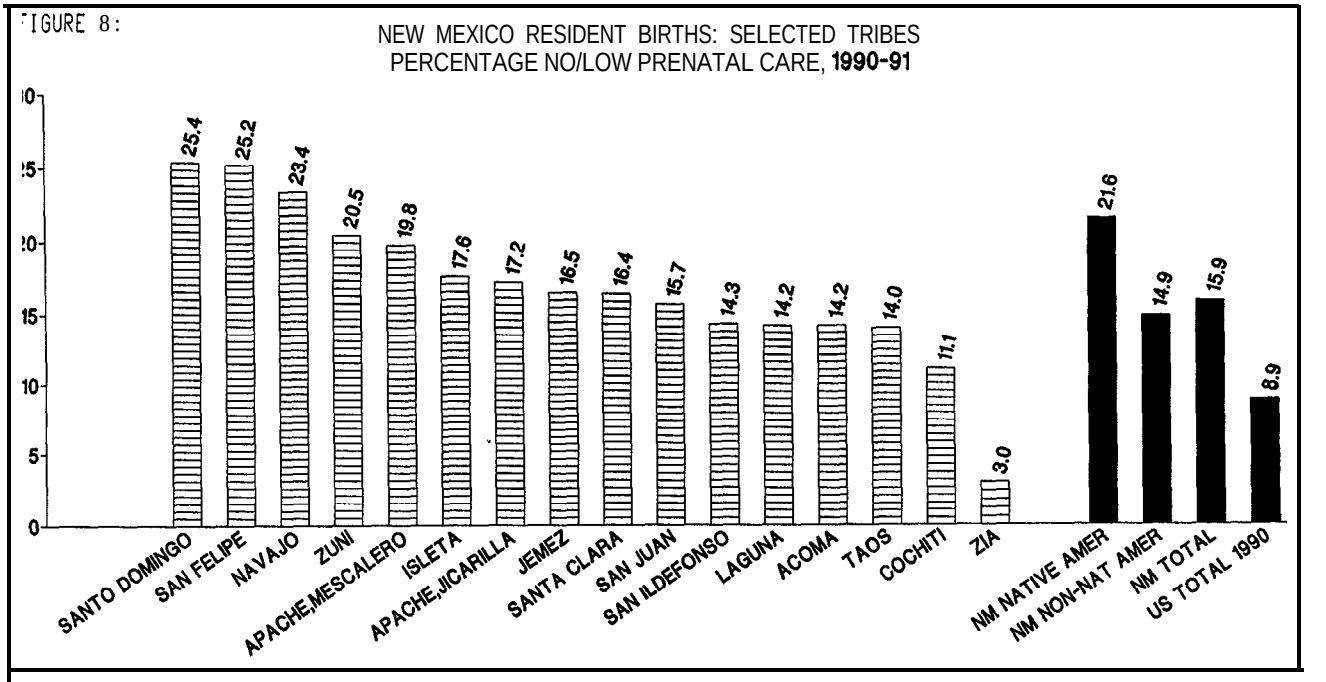
Thirdly, let's look at specific tribes. We have published two reports thus far on tribal vital events based on two years of data. Vital statistics residence data indicate that Pueblo mothers differ in their degree of urbanization. In 1990-91, 87% of Zuni mothers reported living on pueblo lands and only 9% reported living in urban areas. In contrast, 33% of Isleta Pueblo mothers reported living in Isleta Pueblo and 36% reported their residence in urban Albuquerque. obviously this reflects only the residence distribution of women of childbearing age and which may be

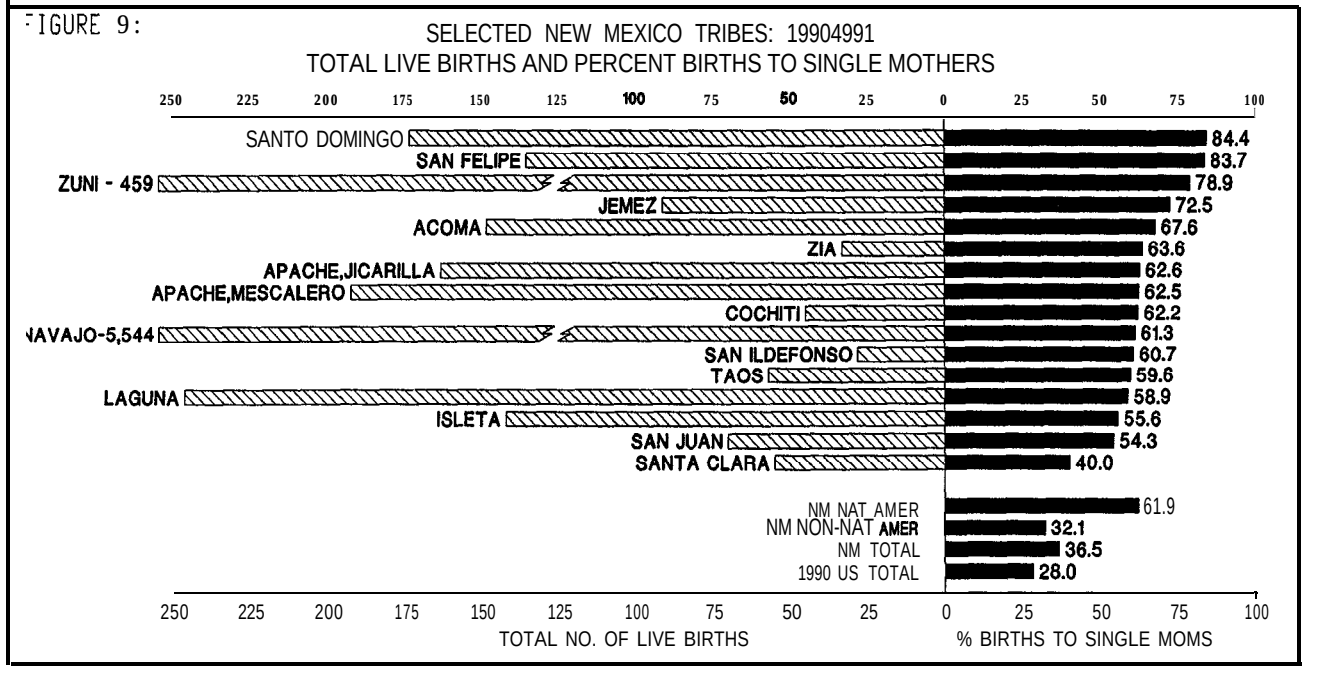


different from the residence patterns of their mothers, aunts, or grandmothers. A partial explanation may be that the Zuni reservation is located 40 miles from the nearest city while Isleta Pueblo is just outside Albuquerque, our biggest city. This residence distribution probably affects how mothers from these two pueblos utilize all kinds of services.

Earlier I mentioned that American Indian mothers as a whole had the highest rate of low prenatal care of all NW races. [FIGURE 8] We see some preliminary indications of differences among tribes when we look at no and low prenatal care rates. In 1990-1991 the range of no/low prenatal care rates was quite wide, from 3% for the Zia tribe to over 23% for the Navajo, San Felipe, and Santo Domingo tribes. Only tribes with 25 or more births are shown on this slide. 63% of the 16 tribes had no/low prenatal care percentages higher NM's non-Native American population while the other 37% had similar or lower percentages than the NM's non-Native American population. The mother's origin -- Pueblo, Navajo, or Apache -- didn't seem to make a difference. From this I ask what is different in many of the tribes that is being reflected in this measure?

[FIGURE 9] This slide shows two things about tribes that had at least 25 births





in 1990-1991. On this side are the total number of live births to each tribe and should give you an idea of the size of each tribe's population. However, the point I'd like you to focus on is the side showing percent births to single moms. All these tribes had higher rates of births to single moms than the state's non-American Indian population or than the US 1990 population. I am tempted to conjecture that perhaps something common throughout the American Indian community as a whole is being expressed. However, what we see in these three examples is that it doesn't always work to assume, for any particular measure, that what seems valid for the entire population remains appropriate for each component group.

We have known for some time now that for most if not all reservations there are many needed and wanted basic public health, environmental health and primary care services that are not now or are minimally offered where they are most accessible --- on the reservations. My point is that while race breakouts are instructive they still hide inter-tribal differences crucial in planning

something operative in the society/environment (whether it is the Native American or New Mexican society) or just within some of the tribes.

Tribe-specific measures increase each tribe's ability to compare itself to any other population(s) **for their** own purposes and to see how their individual vital statistics change over the longterm. They can decide on a tribal and/or **population-**wide basis, something they had little choice about before, which. or how vital statistical comparisons are important for them. For tribes --and I'm sure that's why they asked us to do this -- this can be a **tool to** better define common and disparate problems as well as actions and collaborations that could be undertaken or strengthened.

I have been asked if we would recommend using tribal breakdowns. Our answer is yes, unequivocally! NM will continue to publish tribe-specific reports as more years of data are available.

RACIAL DISPARITIES IN BIRTH OUTCOMES: RACE, ETHNICITY, ORNATIVITY

T. Joseph Sheehan, PhD, University of Connecticut School of Medicine
Judith R. Kunisch, R.N., M.B.A., Charles J. Ingardia, M.D.

At the outset, I would like to draw attention to an important aspect to any discussion of race, ethnicity, or nativity, and that is the emotional dimension. This is a scientific meeting and we are supposed to be factual and objective, so we should at least acknowledge that it exists. My hero, the personification of grace, articulateness, and cool athleticism, Arthur Ashe has written that the burden of being Black was harder for him than the burden of having Aids. It is hard to discuss race without raising emotions. And, ethnicity is no easier; just look at Northern Ireland, Lebanon, or Bosnia Herzegovina! Tonight in Washington, they are dedicating a tennis stadium to Arthur Ash. I think he would be even more pleased if we did something to ease his burden.

Racial disparities in birth outcomes are commonly described in terms of Black-White differences[1,2]. Former Secretary of Health and Human Services, Dr. Sullivan, outlined some of the major disparities between Blacks and Whites in the foreword of a 1992 book providing details of these discrepancies'. He describes a gap in the life expectancy of black and white Americans with the disparity "mirrored in disproportionately higher mortality from many causes. Cancer, heart disease and stroke, cirrhosis, diabetes, infant mortality, unintentional injuries and homicide were identified as the largest contributors...". While racial disparities are described simply by comparing Blacks to Whites, Ethnic disparities are approached more discreetly. The National Center for Health Statistics publishes a Catalogue of Presentations to "foster communication between NCHS and the university community.."[4] One is titled Issues to Consider When Analyzing NCHS Data on Hispanics. It says that when comparing disparities in morbidity and mortality for Hispanics and non-Hispanics, it is important to take into account "cultural differences" such as acculturation, ethnic group identification, birthplace, generation, and language usage." The description implies what is now well known, that there is considerable heterogeneity among Hispanics; this presentation will draw attention to heterogeneity among those of African ancestry.

In 1989, I was faced with the task of trying to uncover useful information from data, over 200 measures from some 1600 pregnant women involved in a preterm birth prevention project in Hartford's inner city. The women answered questions about race and ethnicity. To race, they could choose White, Black, Asian, American Indian, or Other. To ethnicity, they could choose Hispanic, West Indian, or Other. In search for mutually exclusive and exhaustive categories, I came up with four distinct groups: Whites, U.S. Blacks, West Indian Blacks, and Hispanics. There was a small number of Asians and others, who did not fit into one of these groups but, not enough for any meaningful analysis.

The way these groups were first constructed is shown in Table 1. Hispanics consisted of those who chose Hispanic as their ethnicity, irrespective of their choice of race. West Indians consisted of those who chose West Indian as their ethnicity. U.S. Blacks chose Black as their race and other as their ethnicity, while Whites chosen White as their race and other as their ethnicity. We labelled this classification Race/ethnicity. I have since learned that this classification is consistent with Directive 15[5]. According to Hahn's scholarly critique, the Directive is ambiguous about self-identification, nor is there any attempt to "define the concepts of 'race' and 'ethnicity' or consider the scientific principles on which classification schemes are based."[6] Hahn sketches some of the research required if we are to attain scientific validity; nevertheless, these groupings seemed useful and revealed interesting findings with respect to risk factors.

Table 1

Race/Ethnicity =	RACE +	ETHNICITY
U.S. Black	Black	Other
West Indian	Black	west Indian
Hispanic		Hispanic
white	White	Other

Within-race differences in birth outcomes began to emerge between our U.S. and British West Indian Blacks. We did a search and found other studies showing within-race birthweight differences[7,8].

Since the mothers in our sample also answered a question about birthplace with responses of 1) U.S.A,

2) Puerto Rico, 3) West Indies, 4) Italy, 5) Poland, and 6) Other, we looked more carefully at our Race/ethnicity groupings and found a number of individuals who's Race/ethnicity status was questionable. The inclusion of the birthplace question, along with careful checks on the original interviews, allowed us to resolve discrepancies and more clearly to delineate membership in the racial/ethnic groups. Moreover, place of birth also suggests another classification of racial/ethnic groups according to nativitv.

Table 2 shows how the addition of information about birthplace allows us to classify according to nativity, with four groups of particular interest: 1) U.S. born Blacks, 2) Non-U.S. born Blacks, 3) U.S. born Hispanics, and 4) Hispanics born in Puerto Rico. For our sample, there are a small number of U.S. born Blacks with West Indian ethnicity; and a somewhat larger number of non-U.S. born Blacks who were not born in the West Indies.

Table 2

NATIVITY =	RACE + ETHNICITY + BIRTHPLACE		
U.S. born Blacks	Black		U.S.
Non-U.S. born Blacks	Black		Outside U.S.
U.S. born Hispanics		Hispanic	U.S.
Non-US. born Hispanics		Hispanic	Outside U.S.

As I mentioned earlier, racial discrepancies are commonly described by comparing differences between Blacks and Whites. And, it is implicit in these comparisons that race is a major independent variable'. Other risk factors need to be considered, but if the disparity remains after adjusting for these risk factors, then race remains as a major independent variable. Does this mean that the disparity is due to race? Or, is the disparity associated with race or due to something else that is associated with race, as Marcia Angel1 speculates in her recent New England Journal Editorial?"

One way to address this question is to focus on within-race heterogeneity. If we can demonstrate disparities within race as large as disparities reported between races, we can establish that disparities are not due to race, but merely associated with race. Using the groupings shown in Table 2, we have found differences of that magnitude in the birthweight of Black infants born of mothers born in the U.S. as compared to mothers born outside the U.S. That difference is 233 grams and is larger than the often reported black/white difference of 217 grams. We have found similar differences, though not so dramatic, between Hispanics born in the U.S. compared to those born in Puerto Rico.

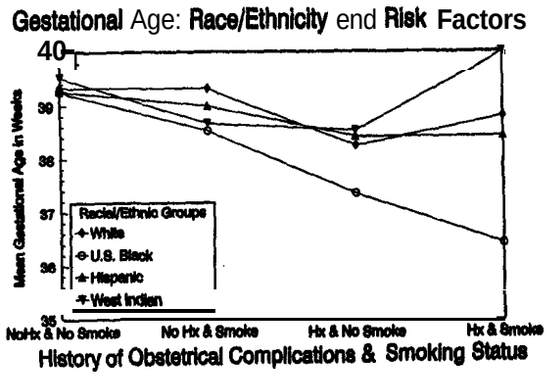
Other studies can be seen from this perspective. Ahmed[11] tried to demonstrate the "usefulness of studying intra-black differences rather than relying solely on race comparative research," by comparing urban-suburban differences in the incidence of low birthweight in a metropolitan black population. He did show great variability, especially in prevalence of risk factors, but his logit model showed that differences in low birthweight rates were almost eliminated when he controlled for risk factor differences.

Attempts to study within-race variability need not preclude between race studies, but even between-race studies might be conducted from a slightly different perspective. That perspective might be described as looking for synergistic effects rather than main effects only. For example, Collins and David" studied the differential effect of traditional risk factors on infant birthweight between Blacks and Whites in Chicago and found that among high-risk mothers living in the poorest areas, the proportion of low birthweight infants for Blacks and Whites was less diveraent that in higher income areas. This might be interpreted as a synergistic effect betweeen race and income, where affluence was not as advantageous for Blacks as for Whites. Li and Daling[13] examined maternal smoking, low birth weight and ethnicity in relation to sudden infant death syndrome. After controlling for confounders, they found that the effects of maternal smoking were stronger for Blacks and Hispanics than for Whites or Asians, and negligible for American Indians, suggesting a synergistic effect between race/ethnicity and smoking. Another study of the risk of smoking on low birth weight reported a higher relative risk of low birth weight in Black smokers than in White smokers[14] or, a synergistic effect between race and smoking.

In our studies, we have shown differences among racial/ethnic groups for both birthweight and gestational age, as well as low birth weight and preterm birth, and a synergistic effect between smoking and having a history of obstetrical complications affecting U.S. Blacks most severely. Figure 1 plots the mean gestational ages associated with each risk factor for each of the four racial/ethnic groups. Analysis of variance revealed significant

differences in birthweight and gestational age between-those with a history of obstetrical risk and those with no history, and between smokers and non-smokers. There is also a significant synergistic effect due to the combination of race/ethnicity together with having a history of obstetrical complication. This synergistic or "interaction" effect implies that a history of obstetrical complication affects some worse than others. It is clear from the figure that U.S. born Blacks with a history of obstetrical complications who also smoke have the worst outcomes.

Figure 1

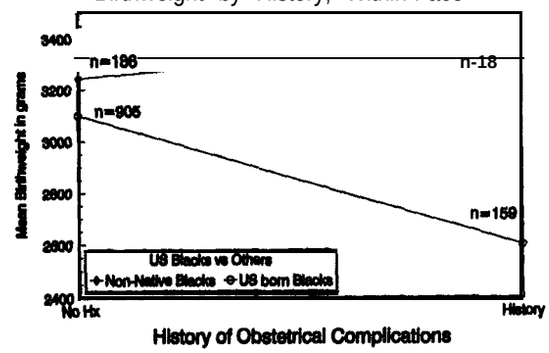


Since our charge is to concentrate on methodology rather than results, please regard these next preliminary findings as additional examples of how we might focus on synergism and within race differences. We are currently examining how a number of stressful life events might affect birthweight. Limiting our analysis to Blacks, we compare the effects of birthplace, having a history of obstetrical complications, having been abused, naving social support, and having gained no weight by the 20th week of pregnancy.

As mentioned above, there is a 233 gram difference in birthweight between infants of Black mothers born in the U.S. and Black mothers born outside the U.S., a considerable within-race difference. We also find a 208 gram difference associated with being abused, a 166 *gram* difference associated with social support, a 156 gram difference associated with having gained no weight, and a 437 gram difference associated with having a history of obstetrical complications. We did not include smoking in this analysis because we are focusing on the stress factors for the moment., All of these differences are statistically significant as well as clinically important. And, there are a number of interactions between pairs of stressors, but we are looking for interactions between each risk factor and birthplace. There is only one. The interaction of history and birthplace is shown in Figure 2. There is a within-*race* difference of 144 grams for those with no history, which grows by a factor greater than five to 798 grams for those with a history of obstetrical complications.

Figure 2

Birthweight by History, Within-Face



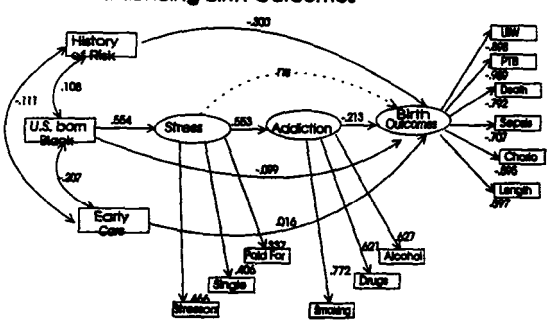
What we see, then, is that all of these factors adversely affect both groups, more or less proportionately, whereas the effects of history are differential and severe for U.S. Blacks. The non-synergistic effects exert their influence more or less uniformly on all racial/ethnic groups. The synergistic effects act differentially and provide the basis for understanding and perhaps eliminating racial/ethnic disparities. The non-synergistic factors affect everyone; the synergistic factors are the causes, or associated with the causes of the racial disparity.

Before concluding, I would like to back-up to those first 1600 cases and 200 measures. In approaching the analysis of these data, we worked out a conceptual model of how each measure related to a risk factor and how the risk factors related among themselves to influence birth outcomes. That work was presented at a CDC sponsored symposium on the statistical evaluation of intervention strategies", and titled *Structural Equation Modeling as a Heuristic in Analyzing Large Data Sets." This heuristic guided our systematic analysis of groups of risk factors. Within a couple of months, using the conceptual model as a guide, we had constructed an actual structural equation model showing how risk factors affected birth outcomes in our sample. That model is shown in Figure 3 and described fully in another paper now in press[16]. Briefly, the circles represent theoretical or latent variables and the rectangles represent measured variables or observations. The directed arrows depict influence quantified by standardized regression coefficients. Note the strong role of history, the role of addictive behavior - a theoretical variable affecting smoking,

alcohol, and drug use, the role of race/ethnicity, and the role of stress. model to the data statistically, and have cross-validated the model with another 1500 cases from the same population. Today's presentation on race/ethnicity, along with the preliminary findings on stressful events, represent an attempt to look more deeply at the effects of race/ethnicity and stress. Note that in the model stress does not have a direct effect on birth outcomes; its affect is indirect, through addictive behaviors. We have another 2100 cases which can be used for cross-validation.

Figure3

A Structural Equation Model of Factors Influencing Birth Outcomes



I show the structural equation model because we are seeking support now to build models reflecting our best understanding of the data and to compare the fit of the models to the data within-racial-ethnic groups. For instance, the above analysis might suggest that the history of obstetrical risk behaves differently for U.S. Blacks than for non-U.S. Blacks. What about addictive behaviors? We can compare the within-group fit of the model statistically to determine their similarity or difference. If they are similar or different for U.S. born Blacks and non-U.S. born Blacks, do they behave similarly for Hispanics according to their birthplace?

Finally, on the basis of what we've found so far it seems that the inclusion of nativity helps to demonstrate within-race heteroaeneitv and provides evidence that disparities are associated with, but not due to race. Further analyses in pursuit of synergistic factors may identify the underlying risk factors. If this line of work continues to be fruitful, we may some day see an entry in the Catalog of University Presentations, "Issues to Consider When Analyzing Data on Race."

The authors would like to thank the following colleagues for helpful suggestions on an earlier draft: Drs. Gary King, Hal Marks, Holger Hansen, and David Gregorio.

References

1. Schoendorf KC, Hogue CJR, Kleinman JC, Rowley D. Mortality Among Infants of Black as Compared to White College-Educated Parents. **The New England Journal of Medicine;** June 4, 1992, Vol 326, No 23, 1522-1526.

2. Kleinman JC, Kessel SS. Racial Differences in Low Birth Weight. The New **England Journal of Medicine,** September 17, 1987, Vol 317, No 12, 749-753.

3. Braithwaite RL, Taylor SE (eds.). Health Issues in the Black Community. San Francisco: Jossey-Bass, 1992.

4. Catalog of University Presentations 1991-92, U.S. Department of Health and Human Services, Public Health Service. Hyattsville, Marvland. 1991

5. Office of Management and Budget. Directive NO. 15: Race and ethnic standards for federal statistics and administrative reporting. In: Statistical Policy Handbook, Washington, DC: Office of Federal Statistical Policv and Standards, US Dept of Commerce; 1978:37-38.

6. Hahn RA. The State of Federal Health Statistics on Racial and Ethnic Groups. **Journal of the American Medical Association,** January 8, 1992, Vol 267, No. 2, 268-271.

7. Ethnic Variation and Maternal Risk Characteristics Among Blacks - Massachusetts, 1987, 1988. **MMWR**, Vol **40, No 24, 403-411.**

8. Cabral H, Fried LE, Levenson S, Amaro H, Zuckerman B. Foreign-Born and US-Born Black Women: Differences in Health Behaviors and Birth Outcomes. **American Journal of** Public Health, January, 1990, Vol 80, No 1, 70-72.

9. Behrman RE. Premature Births Among Black Women. The New England Journal **of** Medicine, Vol 317 No. 12, Sept. 17, 1987, 763-765.

10. Angel1 M. Privelege and Health - What Is The Connection? The New England **Journal of Medicine,** July **8, 1993, vol. 329, No. 2, 126-127.**

11. Ahmed F. Urban-Suburban Differences in the Incidence of Low Birthweight in a Metropolitan Black Population. Journal of the National Medical Association, 81(8), 1989, 849-855.

12. Collins JW, David RJ. The Differential Effect of Traditional Risk Factors on Infant Birthweight Among Blacks and Whites in Chicago. **American Journal of Public Health,** 80(6), **1990, 679-81.**

13. Li Dk, Daling JR. Maternal Smoking, Low Birth Weight, and Ethnicity in Relation to Sudden Infant Death Syndrome. American Journal of Epidemiology, 134(9), 1991, 958-64.

14. Anonymous. Cigarette Smoking and the Risk of Low Birth Weight: A Comparison in black and White Women. Epidemiology, 1(3), 1990, 201-5.

15. Sheehan TJ, Ingardia CJ, Kunisch JR, Hansen H. Structural Equation Modeling as a Heuristic in Analyzing Large Data Sets. Poster Session presented at the Symposium

on Statistical Methods for Evaluation of Intervention and Prevention Strategies, Sponsored by the Centers for Disease Control and Agency for Toxic Substances and Disease Registry, December 5-6, 1990, Atlanta, Georgia.
16. Sheehan TJ, Ingardia CJ, Kunisch JR, Hansen H, Gershman S, Peterson M. A Structural Equation Model of Factors Affecting Neonatal Outcomes. **Quaderni Di Statietica E Matematica Applicata Alle Scienze Economico-Sociali**, in press.

UNDERESTIMATION OF RACIAL MINORITIES IN U.S. ARMY RECORDS

Claire C. Gordon
Natick Research, Development, and Engineering Center

(Paper was not presented)

Session K

People with Disabilities

Toward the year

2000

Refining the measure?

CONTRIBUTION OF DEVELOPMENTAL DISABILITIES TO CHILDHOOD MORTALITY IN THE UNITED STATES: A MULTIPLE CAUSE-OF-DEATH ANALYSIS

Coleen A. Boyle, Centers for Disease Control and Prevention
Pierre Decoufle, Patricia Holmgreen

Developmental disabilities, a heterogeneous group of childhood conditions involving neurologic, cognitive, neuromuscular, and sensory impairments, are relatively common.' According to a 1988 National Health Interview Survey, 17% of children less than 18 years of age were reported to have at some time had a developmental disability.[2] Data from the U.S. Department of Education for the same year indicated that 10% of U.S. children received special education services for one or more developmental disabilities.[3]

Many, if not most, developmental disabilities are chronic and can have a major impact on a child's quality of life. Recent analyses of National Health Interview Survey data indicated that compared with other children, children with developmental disabilities have 50% more doctor visits, 3.5 times more hospital-days, twice the number of school-days lost and a 2.5 fold increase in grade retention.[2] Further, about 30% of children with developmental disabilities have multiple disabilities that have an even greater impact on their school functioning and use of medical care.[2]

Because of the high prevalence of developmental disabilities, the chronic nature of these conditions, and the heavy burden on the child's functioning, one might anticipate that developmental disabilities are important contributors to childhood mortality. We used the National Center for Health Statistics (NCHS) multiple cause-of-death data to explore this issue.

Our choice of a multiple cause-of-death strategy for the analysis was based on findings for other chronic diseases. Because of the long duration between disease onset and death for many of these conditions, such conditions are often underrepresented as the underlying causes of death in mortality statistics.[4,5] The one relevant study of a developmental disability, a British cohort study of cause of death among children with cerebral palsy, found similar results: cerebral palsy was underrepresented as the underlying cause of death, and other conditions representing immediate causes or "terminal events" were overrepresented, especially among older children.[6]

METHODS

We used the NCHS multiple-cause-of-death data for children ages 1 to 19 years for the calendar years 1980 and 1983 through 1989. We did not include 1981 and 1982 because multiple-cause-of-death coding was incomplete for these 2 years.

A death was considered to be related to a developmental disability if one or more developmental disabilities were coded as either the underlying or a contributing cause. NCHS multiple-cause-of-death coding includes up to 20 causes of death, one of which is assigned as the underlying cause.[7]

The conditions we included in the analysis and the corresponding ICD-9 codes are listed in Table 1. These included the major developmental disabilities, such as mental retardation and epilepsy, as well as developmental disabilities less commonly associated with increased risk of mortality, such as learning disorders and attention deficit disorder. We did not include childhood psychiatric disorders other than infantile autism or pervasive developmental disorder. We examined the rank order for the leading nontraumatic underlying causes of death using the classification scheme of Hockelman and Pless.[8] This scheme collapses the 72 standard cause of death categories used by NCHS into 16 categories more appropriate for childhood mortality. In addition, we included only nontraumatic causes of death in the ranking of underlying cause so as to view developmental disabilities relative to other nontraumatic causes of death. We did, however, include traumatic causes of death in the multiple-cause analysis. For clarity, only the five leading nontraumatic causes were displayed in the Figure.

We did not examine calendar year trends, since only a 10 year period was represented. However, we did examine separately the contribution of developmental disabilities to childhood mortality for children within four 5-year age groups; for whites, blacks, and other races; and for males and females.

RESULTS

The Figure shows the average annual death rates by 5-year age group for the 5 leading nontraumatic or disease-related causes of death. The age-specific curve for developmental disabilities is J-shaped which is similar to that for the other leading causes except congenital malformations. On the basis of the underlying cause of death alone, developmental disabilities rank fifth as a cause of death for

children less than 15 years of age and third for children 15-19 years of age.

Table 2 shows the results of our multiple-cause-of-death analysis. When we included the contributing causes of death the number of deaths related to developmental disabilities increased 74%, from 6,049 to 10,514. In a similar analysis for congenital malformations and neoplaems, two major causes of childhood mortality, and the proportion of deaths coded only to contributing causes was considerably smaller: 27% for congenital malformations and 7% for neoplasme.

The largest percent increases for the individual developmental disabilities were found for the least frequent causes of death, namely autism, attention deficit disorder, learning disorders, blindness, and deafness (Table 3). These developmental disabilities were listed as the underlying cause of only 32 deaths, but they were associated with 475 deaths based on a multiple-cause-of-death approach, with blindness and learning disorders coded as contributing causes for the vast majority (83%) of these deaths. Cerebral palsy was the most common cause of death with either the underlying or multiple-cause-of-death approach. Mental retardation was the fourth most common underlying cause but moved up to the second most common when we included underlying and contributing cause information.

Of the 10,514 deaths related to developmental disabilities that we examined, 14% had more than one developmental disability listed as a cause of death. The most frequently occurring combination was mental retardation and cerebral palsy which accounted for 65% of the deaths involving multiple developmental disabilities. The joint occurrence of mental retardation and epilepsy and of cerebral palsy and epilepsy each accounted for about 12% of deaths involving multiple developmental disabilities.

Table 4 shows the major underlying causes for those deaths where a developmental disability was coded as a contributing cause. Infectious diseases were the leading underlying cause for such deaths, with 22% coded to an infectious disease. other major causes included congenital malformations, respiratory diseases, cardiovascular diseases, and injuries. As a comparison, a similar type of analysis was done for congenital malformations and neoplasms. Again, infectious diseases was the leading underlying cause for both, with cardiovascular diseases, respiratory diseases, and injuries also ranking as major causes.

Table 5 shows the rates of developmental disability-related deaths by age, race, and sex. The age-specific curve for all developmental disability-related deaths combined was U-shaped, with the mortality rates highest in the youngest and oldest age groups. This is in contrast to a more J-shaped curve (Figure) when only underlying cause of death is used, reflecting a decrease with age in the proportion of developmental disability deaths coded to contributing causes. Rates were highest among blacks and lowest among race groups other than whites or blacks, and rates were higher among males than among females.

The age-specific mortality rates for children with each of the four leading developmental disability causes of death, cerebral palsy, mental retardation, muscular dystrophy, and epilepsy were somewhat different from the overall patterns. The rates of mortality associated with mental retardation and cerebral palsy were highest in the youngest age group, whereas the mortality associated with muscular dystrophy .increaeed with age. Except for white children having higher rates of muscular dystrophy-related deaths, the race-specific patterns did not vary considerably from the overall developmental disability patterns. Finally, males had higher rates for deaths associated with cerebral palsy and muscular dystrophy but not for deaths associated with the other developmental disabilities.

DISCUSSION

We showed that on the basis of underlying cause of death alone, developmental disabilities ranked as the fifth most common nontraumatic cause of death overall for children 1 to 19 years of age and the third most common cause for children 15-19 years of age. We did not attempt to use the multiple-cause-of-death approach to rank the major nontraumatic causes of death. Obviously, the usefulness of this approach for causes of death other than developmental disabilities would vary considerably.

We also showed that using a multiple-cause-of-death strategy increased the absolute number of deaths classified as associated with developmental disabilities by 74%. Thus, in close to 40% of all deaths involving developmental disabilities, the developmental disability was coded as a contributing cause. Developmental disabilities were more likely than other major non-traumatic causes of childhood mortality (e.g., neoplaems and congenital malformations) to be coded as contributing causes of death. This increased likelihood, however, may be related to the interval between the

onset of a developmental disability and death being longer than that between the onset of neoplasms and congenital malformations and death.

Two issues should be addressed regarding the reporting of developmental disabilities on death certificates. The first is the possibility that developmental disabilities were underreported even with the multiple-cause-of-death approach. According to the World Health Organization rules of death certification, if the developmental disability "unfavourably influenced the course of the morbid process," even if it was not directly related to the disease or condition that caused the death, it should be mentioned as a contributing cause on the certificate.[10] However, evidence from a British cohort study of mortality in children with cerebral palsy, found that in 38% of deaths there was no mention of cerebral palsy on the death certificate, a proportion which increased with the age of the child.[6] Perhaps, our finding of a decrease with age in the proportion of developmental disability-related deaths coded only as a contributing cause is further evidence that developmental disabilities are less likely to be listed as a contributing cause as the interval between disease onset and death increases.

The second issue is whether developmental disabilities were underreported as the underlying cause of death. Obviously, the long interval between disease onset and death which occurs for many of the developmental disabilities, increases the likelihood of another disease developing which is then recoded as the underlying cause.[11] However, methodologic studies of other chronic debilitating conditions, such as, diabetes and rheumatoid arthritis, suggest that they are underreported as the underlying cause from death certificate information.[4,5] A medical record review of deaths among children with cerebral palsy has suggested that respiratory diseases were overrepresented and cerebral palsy underrepresented as the underlying causes of death.[6] We examined the types of underlying causes for developmental disability deaths coded only as a contributing cause. The major causes represented included infectious, respiratory, and cardiovascular diseases. A study comparing autopsy and medical examiner's reports with death certificate information found that causes, such as respiratory diseases, are often terminal events rather than underlying causes.[12]

cerebral palsy was the developmental disability most often mentioned as a cause of death. This may not be surprising given the neurologic and neuromuscular involvement and secondary conditions associated with cerebral palsy. Even though the prevalence of mental retardation is many times higher than the prevalence of cerebral palsy, mental retardation ranked only second as a cause of death among the developmental disabilities examined here -- and even then only when rankings were based on results of the multiple-cause-of-death analysis.' This may reflect a real difference (i.e., that the mortality rate associated with cerebral palsy is considerably higher than that associated with mental retardation), or, it may reflect the tendency of the certifier to attribute death to the more obvious physical disability associated with cerebral palsy. Similarly, the fourfold increase that we found in the number of deaths related to mental retardation when we included the contributing causes of death may reflect a reluctance on the part of the certifier to attribute the death to mental retardation because of the limited physical disability associated with this condition.

The age-specific mortality curve for developmental disability-related deaths is similar to the pattern of childhood mortality in general-- a pattern influenced by an increasing incidence of disease with age. Since, the majority of developmental disabilities have their onset in the first few years of life, the high mortality in the l-4 year olds probably represents the most severely affected children. The high mortality rate in the 15-19 year olds may represent the increased vulnerability of the older disabled child.

Regarding the higher disability-related mortality in blacks and in males, the higher prevalence rates of some developmental disabilities in these subgroups could explain the mortality patterns.[13-15] Differential rates of survival, however, may also play a role.

In summary, we found that developmental disabilities are among the leading nontraumatic causes of childhood mortality and that a multiple-cause-of-death approach is necessary to appreciate the full magnitude of this impact. Although we believe that multiple-cause-of-death information more accurately reflects the.burden of mortality attributed to developmental disabilities, specific patterns in our results suggest that even this approach may underestimate the degree to which mortality is associated with developmental disabilities.

REFERENCES

1. Yeargin-Allsopp M, Murphy CC Oakley GP, et Al. A multiple source method for studying the prevalence of developmental

disabilities in children: The Metropolitan Atlanta Developmental Disabilities Study. Pediatrics 1992;89:624-30.
2. Boyle CA, Decoufle' P, Yeargin-Allsopp M: Prevalence and health impact of developmental disabilities in U.S. children. Pediatrics, in press.
3. Tenth Annual Report to Congress on the Implementation of the Education of the Handicapped Act. U.S. Department of Education, Washington DC, 1988.
4. Tokuhata G, Miller W, Digon E, Hartman T. Diabetes mellitus: An underestimated public health problem. J Chronic Dis 1975;28:23-35.
5. Ingemar B, Lindahl B, Allander E. Problems in the classification of death diagnoses affecting the reliability of mortality statistics for rheumatoid arthritis. J. Chronic Dis 1985;38:409-18.
6. Evans PM, Alberman E. Certified cause of death in children and young adults with cerebral palsy. Arch Dis Child 1991;66:325-9.
7. Public Use Data Tape Documentation: Multiple Cause of Death for ICD-9, 1985 Data. National Center for Health Statistics, Hyattsville, MD, 1987.
8. Hoekelman RA, Pless IB. Decline in mortality among young Americans during the 20th century Prospects for reaching national mortality reduction goals for 1990. Pediatrics 1988;82:582-95.
9. Intercensal Population Estimates, Bureau of Census, 1980-1989, Unpublished data, 1992.
10. International Classification of Diseases. Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death. Volume 1, World Health Organization, Geneva, 1977.
11. Kelsey JL, Thompson WD, Evans AS. Methods in Observational Epidemiology. Oxford University Press, New York, 1986:50.
12. Kircher T, Nelson J, Burdo H. The autopsy as a measure of accuracy of the death certificate. NEJM 1985;313:1263-69.
13. Kiely M. The prevalence of mental retardation. Epidemil Reviews 1987;9:194-218.
14. Hauser WA, Hesdorffer DC. Epidemiology of Epilepsy. in: Neuroepidemiology: A Tribute to Bruce Schoenberg. CRC Press, Boca Rotan, 1991.
15. Murphy CC, Yeargin-Allsopp M, Decoufle P, Drews CD. The prevalence of cerebral palsy among lo-year-old children in metropolitan Atlanta, 1985 through 1987. J Pediatrics, in press.

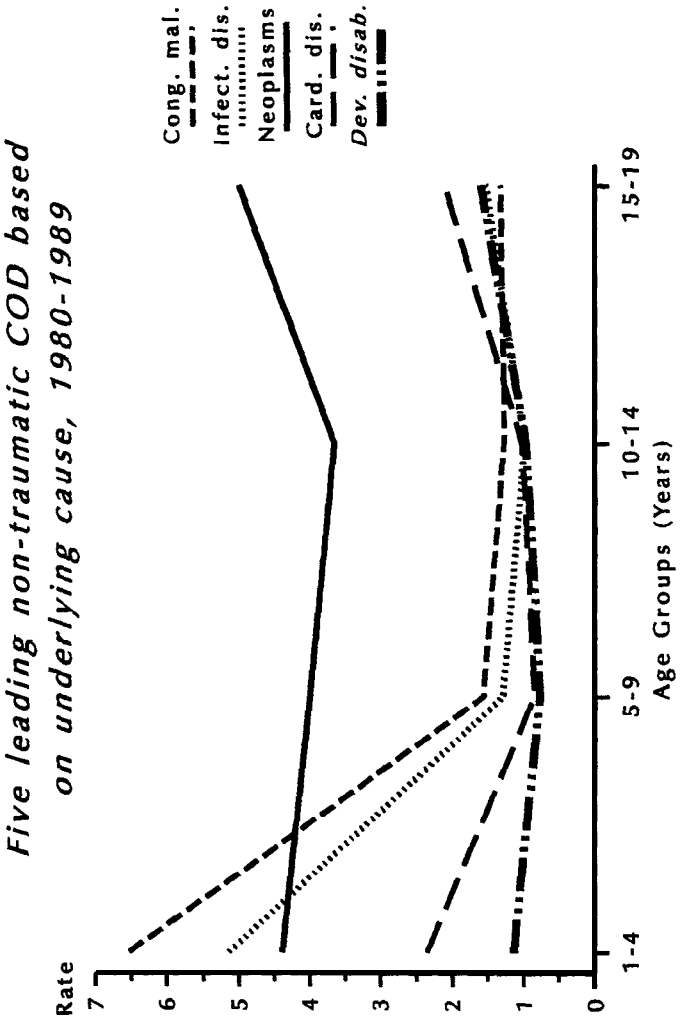


Table 1: Specific conditions and corresponding ICD-9 codes used in the developmental disabilities analysis.

Condition	ICD-9 Codes *
Autism	299
Attention deficit disorder	314
Learning disorder	315
Mental retardation	317-319
Cerebral palsy	333.2, 333.7, 342-344
Epilepsy	345
Muscular dystrophy	359
Blindness	369
Deafness	369

* Includes all 4th-digit codes unless otherwise indicated.

Table 2. Number and rate/100,000 for deaths involving a developmental disability congenital malformation, and neoplasm among children l-19 years of age based on underlying and multiple cause of death, United States, 1980, 198389.

	Underlying cause of death		Underlying or contributing cause of death	
Cause (ICD-9)	No.	Rate	No.	Rate
Developmental disabilities	6,049	1.13	10,514	1.96
Congenital malformations	13,242	2.47	16,671	3.14
Neoplasms	22,936	4.27	24,527	4.57

Table 3. Number and rate/100,000 for deaths due to developmental disabilities (DD) among children 1-19 years of age based on underlying and multiple cause of death by specific type of DD death, United States. 1980. 1983-89.

	Underlying cause of death		Multiple cause of death	
DD cause* (ICD-9)	No.	Rate	No.	Rate
Autism (299)	2	+	21	+
Attention deficit disorder (314)	0	-	6	+
Learning disorder (315)	29	0.01	227	0.04
Mental retardation (317 - 319)	741	0.14	3166	0.59
Cerebral Palsy (342-344)	2665	0.50	5049	0.94
Epilepsy (345)	1039	0.19	1654	0.31
Muscular dystrophy (359)	1552	0.29	1789	0.33
Blindness (369)	1	+	168	0.03
Deafness (389)	0	-	55	0.01

Table 4. Underlying cause in developmental disability, congenital malformation, and neoplasm deaths coded as a contributing **cause for children l-19 years of age, United States, 1980, 1983-89.**

Underlying cause	Developmental disabilities		Congential malformations		Neoplasms	
	No.	%	No.	%	No.	%
Infectious diseases	995	22	843	23	550	31
Congenital malformations	906	20	-	-	72	5
Cardiovascular diseases	466	9	607	17	205	13
Injuries	380	9	314	9	116	7
Nervous system diseases	250	6	180	5	68	4
Gastrointestinal and other digestive system diseases	231	5	162	4	55	3
Respiratory system diseases	821	18	279	8	230	14
Neoplasms	142	3	220	6	-	-
Developmental disabilities	-	-	203	6	13	1

Table 5. Number and rate per 100,000 for deaths involving developmental disabilities* among children ages **1-19** years of age by age race, and sex, United states, 1980, 1984-1989.

	No.	Rate
Age :		
1 - 4 years	2,524	1.96
5- 9 years	2,010	2.23
10 - 14 years	2,289	1.47
15 - 19 years	3,691	2.46
Race:		
White	8,092	1.86
Black	2,181	2.67
Other	241	1.25
Sex:		
Male	6,409	2.34
Female	4,105	1.56

*Based on multiple causes of death

DEFINITIONS OF DISABILITY AND THEIR MEASUREMENT AND OPERATIONALIZATION IN SURVEY DATA

Barbara M. Altman
Agency for Health Care Policy and Research

Capturing the essential medical, physical and social aspects of disability by means of survey data is a difficult task. An extensive literature has examined and elaborated the definition of disability into various conceptual models in order to facilitate the operationalization and eventual measurement of that phenomenon (Nagi, 1965; Wood, 1975; Wan, 1974; Williams, 1979). The purpose of these models is to reduce confusion of the terms impairment, disability and handicap by explicating the exact relationships among these terms.

The commonly accepted model in use originated with Nagi (1965). It identifies a pathology or disease state which can lead to an impairment which becomes the basis which can lead to some functional limitation. In the original model, these elements are associated with a disability which represents a disjunction between the individual and their environment and is recognized by a modification in social role behavior. An extension of the model by Wood (1975; 1980) also incorporates the concept of handicap which reflects the disadvantaged status a person with a disability can experience in their social role environment. Pathology and/or impairment then, represent the biophysical manifestation of a problem and functional limitation is the resulting physiological outcome. Disability is the individual behavior associated with the pathology/impairment and/or functional limitation and handicap is the resulting experience for the individual with a disability as the result of behaviors of persons and organizations toward the individual with a disability.

Theoretical models of disability are useful because as abstractions of the area of focus in the real world they identify the key concepts and their relationships. However, in order to develop empirical representations, the process of defining the concepts that make up the theory is the first step in narrowing the conceptual field. In a social situation, multiple definitions of a concept may exist depending on the vantage point of the observer. At the point that one definition is chosen, it creates a sizable gap between the theoretical concept and the empirical representation by limiting the form of questions or indicators to ideas relevant to the unidimensional definition.

The more complex the social phenomenon being defined, the greater the possible variety in aperationalization. Blalock (1968; 1979) discussed the concept of prejudice and showed that one definition will not suffice for all circumstances. In the same way, disability is a complex social phenomenon that involves the interaction of individuals with specific limitations, possibly due to a physical impairment, with their identity, their physical and social surroundings, and the societal culture. Disability, then, is undefinable empirically unless one reduces the focus of the definition to a specific aspect of experience. Blalock (1968;1979) offered the possibility that a theoretical concept could have a number of different operationalizations and associated indices, "provided it 'does not lead to endless bickering about which of the procedures is really measuring prejudice (the essence of which is presumed understood). "[1] Blalock's point is that the research process can be used to clarify theoretical concepts.

This paper follows from Blalock's discussion. Though much effort has been spent on refining the definition of disability and the development of a classification system to improve its operationalization (Nagi, 1965; 1979; Wood, 1975; 1980), less attention has been paid to what can be learned from measures already available and in use on a regular basis. Confusion exists also because of the origin of measures of disability already in use which have been developed to address specific policy issues. The case of administrative decision making requires an exact demarcation between persons with and without disabilities for the distribution of scarce resources. This explains the emphasis on the measure of work limitation that has been used extensively (Krause, 1976; Duckworth, 1982). Demographic estimates and epidemiological research require a much different kind of measure and indicators of types of conditions and types of functional limitations or task limitations have been collected (Census and NCHS).

Slater et al (1974) identified six different types of definitions in their operationalization of the concept of disability for a study designed to measure prevalence of persons with disabilities. These definitions ranged from subjective definitions of the individual to community definitions. Krause (1976) more concisely identified three basic groupings of definitions of disability in terms of biophysiological, administrativeorlegal, or social role aspects of the concept. It

is my objective to use five indicators representing the three types of definitions proposed by Krause (1976) and to examine their convergence or divergence in identifying a population which can be called disabled. The importance of this convergence/divergence to theoretical development, future data collection and policy analysis will be discussed.

DATA AND METHODS

A. Discussion of NMES

The data used for this study are from the Household Survey Component of the 1987 National Medical Expenditure Survey which was based on a representative sample of the civilian noninstitutionalized U.S. population. The entire household survey was fielded in four rounds of interviews at approximately four month intervals with a fifth short telephone interview at the end. All four rounds of data were used here to provide estimates of some of the measures, particularly indicators of chronic conditions or impairments and disability days. The last round of data provides the information on receipt of disability benefits.

The survey also included a self-administered questionnaire on health status which was completed during the third round of data collection. Variables measuring disability via role limitations and functional limitations were drawn from this supplement. Other details and information on the sampling and survey procedures are available in Edwards and Berlin (1989).

B. Description of Five Measures

Five indicators were used to represent the three types of disability definitions. They were created as follows:

Bionhvsiolosical Measures

1. Chronic Conditions and Impairment - Data on conditions associated with physician visits, hospitalizations, emergency room use, home health care, medication use and disability days were identified by survey respondents and coded according to the ICD9 specifications, using procedures developed for the National Health Interview Survey. This indicator was created starting with all these ICD9 codes. In conjunction with a physician, Dr. James Panagis, this author identified physical or biological conditions which are recognized medically as chronic but controllable or degenerative (such as diabetes, arthritis) or life threatening at onset with residual damage or threat after control (such as stroke) or impairments which resulted from disease, accident or birth defects but are no longer associated with an active pathology (such as amputations, blindness). Short-term and infectious diseases were not included. The exact ICD9 codes used are listed in the appendix. The limitation of this measure is that health conditions were self-reported and associated with some form of medical treatment or disability days. Persons with conditions who reported no disability days during the year or did not have need for medical care will be missing from this measure.

2. Functional Limitations - This measure focuses on physical functioning and uses three questions from the self-administered questionnaire which ask about trouble walking, climbing stairs, bending, lifting and stooping. The precise questions and methods of constructing the variable are detailed in the appendix.

Social Role/Behavior Measures

1. Role Limitations - This measure is a classic one used in the National Health Interview Survey. It is based on four questions from the self-administered questionnaire that ascertain if the individual is kept from working, is limited in the kind or amounts of work, housework or schoolwork they can do because of their health or is limited in activities apart from the work role because of health. The questions and the methods of construction of this variable are detailed in the appendix.

2. Bed Disability Days - This measure is limited to any days for which the respondent was bedridden or hospitalized because of a health problem. The total number of days was accumulated for the year. For use as a categorical variable, those with 30 or fewer bed days were considered not disabled, while an indication of 31 or more bed days was considered to reflect a disability. Disability days have usually been considered a measure of illness and used by economists in that way. Bed days can be considered a measure of severity or functional limitation if the individual does not have the ability to leave bed. However they also reflect behavior if an individual chooses to remain in bed or if circumstances are such that the individual does not have the assistance necessary to get out of bed or get dressed. In this latter case the concept of handicap rather than disability may be an interpretation (Chamie, 1993, personal communication).[2]

Administrative/Lesal Measure

1. Disability Benefits - This measure was based on information that the individual received, SSI, Social Security or a Pension because of a disability.

RESULTS

One of the first uses of a measure of disability is to estimate its prevalence in the society. Since each measure being used here is associated with one aspect or definition of disability, it is logical that the prevalence estimates generated by each measure should differ. The results are exhibited in Figure 1. The measure which reflects the biophysiological definition associated with medical care, chronic conditions/impairments, defines the largest group, or about 30 percent of the working age population. At the other extreme, disability bed days define disability in terms of behavior associated with severe limitation in all activity or actual number of days spent in bed during the year. By this measure only 1.4 percent of the working age population are so limited that they spend 31 days or more in bed or in the hospital.

In Figure 1, we also see that persons who have applied for and met administrative or legal definitions of disability and receive disability benefits account for 4 percent of the working age population. Persons with role limitations which include other activities besides work, are more prevalent than those who indicate functional limitations; the reverse might have been anticipated from the accepted theoretical model of disability which indicates the hierarchical sequencing of pathology/impairment - functional limitation - disability.

If one were to examine the congruence of these measures based on the accepted model of disability discussed in the introduction, the expected relationship would be a hierarchical one as represented in Figure 2. A nested effect would result from the continued narrowing of the definitions. Persons with chronic conditions or impairments would not all be expected to experience functional limitations, while persons with chronic conditions and functional limitations would not necessarily be expected to have role limitations and so forth.

Combining those measures using actual data, however, results in a somewhat different picture shown in Figure 3.[3] The actual prevalence distribution differs from that expected in two very important ways: (1) those with role limitations are a larger group than those with functional limitations, so that their position in the model is reversed; (2) many of the respondents indicated a disability on only one measure, including but not limited to chronic conditions, so that a good proportion of the population identified by the five measures do not overlap at all.

Examining. the pattern of congruence in Figure 4, we find that the proportions of persons that have convergence of from two through five of the measures follows the hierarchical pattern expected, at least numerically, based on the proportions in each group of combinations. As would have been predicted from the expected nested model, those convergent on all of the five measures would be the smallest group. In this instance they make up 1.2 percent of the disabled population. However, among all persons identified as disabled by any of the five variables, the largest group, 53.5 percent, are identified by only one indicator.

When we estimate self-reported prevalence of disability based on measures that theoretically should be congruent, we find that all five measures taken together identify as much as 43.7 percent of the working age population as having some indication of disability.[4] In Figure 5, the left hand column shows that 46.5 percent of the total group identified as disabled are categorized as having a disability on two or more measures. The remaining 53.5 percent are identified as disabled on only one measure. In the same graph, those two groups are used to estimate prevalence separately. A prevalence rate of 20.3 percent of the working age population is associated with those who are defined as disabled on two or more measures, while an additional 23.4 percent of the working age population are identified by only one measure.

Since the total prevalence rate, 43.7 percent, far exceeds what would have been considered the maximum rate of prevalence associated with the chronic condition measure, 29.7 percent, it suggests either the nesting effect predicted by theory is not necessarily accurate or the possibility of inadequate operationalisation of the definitions or other methodological problems. This combined with the indications of prevalence of limitations beyond those associated with chronic conditions/impairments describedby Figure 3 leads us to examine the relationship of the chronic condition/impairment measure with the other measures used here.

Some patterns of congruence have been noted in the literature but not fully examined. McNeil (1989) found that many with functional limitations do not have work limitations and many with work limitations report no functional limitations. A report for the U. S. Department of Health and Human Resources (1989) demonstrated the lack of total convergence of measures of work limitation, functional limitation and receipt of disability benefits for the working age population. An underlying assumption in both instances is that such indications of work or functional limitations were directly related to health status. The value of the current analysis lies in its ability to use data on chronic conditions/impairments which

can be incorporated into the modeling.

Figure 6 displays the results when the congruence between the chronic condition/impairment measure and the four other measures of disability were examined. In keeping with the theoretical model pictured in Figure 2, we find that 35 percent of the individuals have a chronic condition or impairment only. A second group shows some congruence between the chronic conditions/impainnentsmeasure and at least one other measure that the theory and model predict. Thirty-four percent of the individuals have the combination giving us convergence between the major biophysiological measure and either the other biophysiological measure or the social role or legal measures. These individuals may be considered to represent, in actuality, the closest approximation to the definition of disability as identified by the current theoretical model.

However, there is a third group identified in Figure 6, those who indicate a disability on one of the role, functional or administrative measures but who do not indicate a chronic illness or impairment within the bounds the operationalization of that variable in the NMES survey. They account for 32 percent of the persons identified by these measures or 13.8 percent of the working age population. Taken at face value, this indicates some limitations with our definitions or problems with measurement, or both.

DISCUSSION

The results presented confirm that measurement of disability in survey data is difficult and very dependent on the conceptual definitions employed and the operationalization of those definitions in terms of questions and methodology. This finding from the National Medical expenditure Survey supplements a long series of methodological work such as that by Haber (1990) examining the factors affecting variations in prevalence estimates from various surveys and by McNeil (1993) on the continuing questions of validity and reliability as reflected in the different estimates of prevalence associated with different measures. The results take on renewed importance in light of the current policy issues. Health care reform and the implementation of the Americans with Disabilities Act require increased research immediately. The results are also relevant for the collection of an improved data set with a much broader elaboration of the concept of disability and the social context in which it is produced. The following discussion explores theoretical, methodological and analysis issues in the context of some of the results.

One interesting result of this analysis is the pattern of congruence of the chronic condition/impairment measure with a role limitation functional limitation or administrative measure of disability. In one third of the respondents it is apparent that chronic conditions and impairments are not associated with any disability. Since illness need not produce disability this result is not unexpected. A second group demonstrates at least some convergence among the variables in accordance to the theoretical model. However, the final group indicates a disability without a corresponding condition or impairment. While some of this discrepancy may be explained by mental and emotional conditions which are not included in the chronic condition/impairment measure, it is not known if that is the explanation. In fact, preliminary examination of that group indicate this is not the case. Further examination of this group is necessary to ascertain what characteristics of the population or methodological problems contribute to the discrepancies.

This raises questions about the underlying theory and the assumption that health problems are the only causes of functional and role limitations and thereby disability. Since disability is actually a relational concept which goes beyond individual limitations and attributes and includes the actual physical and social environment as well as the reactions of others (Nagi, 1989), the causative agents may be outside the individual. Therefore measurement of disability can not take place only at the individual attribute level although it could be measured as an outcome, i.e. the inability to work. Theoretically, is this disjunction of the individual from the social context central to the concept of disability or is the causal link with ill health an essential ingredient? If health is an essential ingredient, how are we to understand the group with limitations but without reported associated health conditions? Or is it our understanding of health that is in need of expansion to include more than clinically defined manifestations?

From a methodological perspective, it is important that all the questions associated with the role limitation and functional limitation measures in this analysis did include the concept of health as the causative agent; "Do you have any trouble walking one block because of your health?". What then causes the contradiction that limitations are health related but no evidence of a condition is provided? Are respondents ignoring the health aspects of the limitation measures or is health interpreted differently? In NMES the role and function questions directly followed a generalized question about health status. Would it be more

appropriate to use a question asking about limitation without reference to cause following it by one that asked the cause? The effects of location and wording of questions are known to influence the responses, the combinations of concepts in a question, such as limitations and health, may also be a factor that needs investigation.

These theoretical and methodological issues are important to the improvement of data that is collected in the future. However, our immediate problem is the influence of definitional and thereby measurement inconsistencies on analysis related to the policy decision process taking place now. One area of concern that follows from this analysis is how do the various measures, which represent different facets of the concept of disability and identify different individuals, differ in their explanatory capacity in research, particularly policy research. Are the results as varied as the prevalence rates associated with each measure? In light of the lack of convergence of the measures, what is the generalizability of findings from the use of a single definition to the population with disabilities or the target population of the particular policy in question? We have taken a multidimensional experience and reduced it into a few of its major components in order to accomplish the measurement process. Although single indicators are useful for comparing different populations and examining what may be happening in an aggregate population, it is important to recognize that they are not the most sensitive measures of disability in a population.

A true understanding of the limitations of the measure being used goes a long way toward interpreting the findings appropriately. Investigation of the relationship of the measure of choice to other measures of disability may help improve this interpretation process. The consideration and investigation of composite measures, using available indicators, may be worthwhile. Some feel that disability, like "quality of life," contains so many facets of objective and subjective experience that it can not be measured at all. However, understanding the strengths and limitations of the measures currently available and choosing the appropriate single or combination of indicators should be useful in informing the policy issues that face us today.

APPENDIX and BIBLIOGRAPHY are available from the author on request.

ENDNOTES

1. Blalock was discussing the conceptualization and operationalization of prejudice, however, the discuddion is easily applied to the concept of disability as well.

2. Disability days are included in this analysis becasue it is a measure commonly used in some areas. Its accuracy as a measure of disability is not the question here, rather I am concerned with examining what currrently used measures can contribute to furthering our understanding of the concept of disability and the measurement process.

3. Neither this figure or the next one is drawn to actual scale.

4. The size of the population identified as disabled by the combination of all five measures demonstrates that the various definitions of disability used to guide the empirical measures may lack specificity, or the questions or methods used to elicit the data from respondents may not adequately reflect the definitions. If the measures were truly convergent, the prevalence rate for all five measures should not be greater thatn for the single measure which estimated the highest prevalence rate.

Self-reported Disability Among Population 18 to 64 - Five Measures

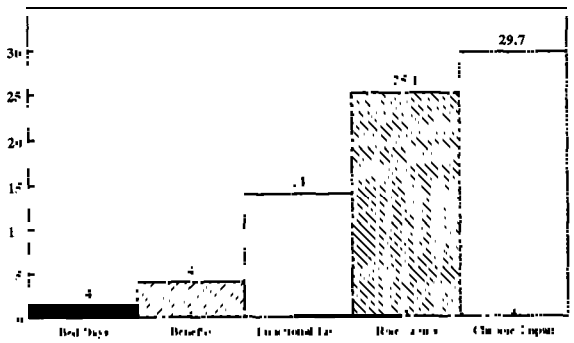


Figure 1: Percentage with a Disability

Ideal Convergence of Five Measures

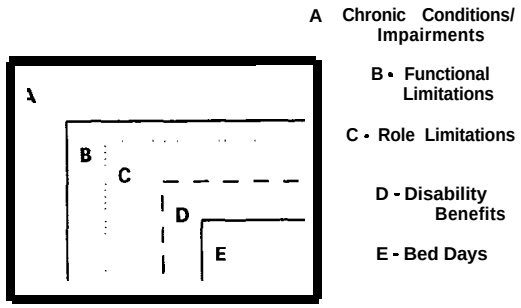


Figure 2

Actual Convergence of Five Measures

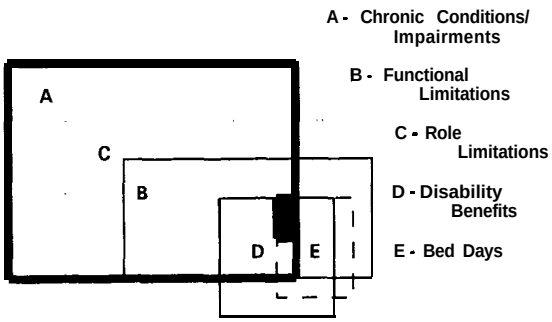
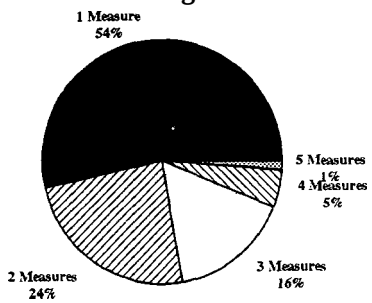


Figure 3

Percentage of Disabled Population who have Convergent Measures



All Persons Identified by Measures as Disabled

Figure 4

Self-reported Disability Among Pop. 18 to 64 - Five Measures Combined

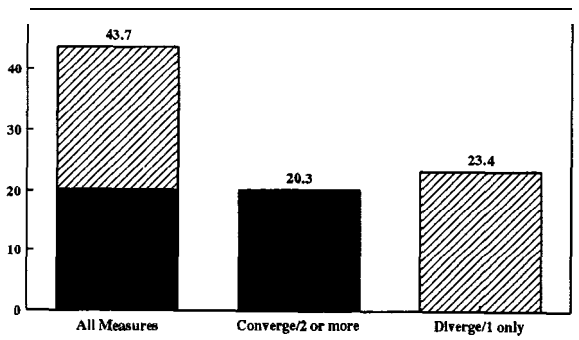
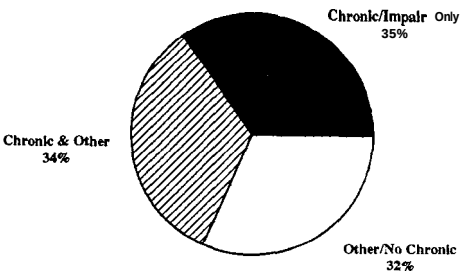


Figure 5: Percentage with a Disability

Convergence of Chronic Conditions/ Impairments with other Measures



All Persons Identified by Measures as Disabled

Figure 6

APPLICATIONS OF THE CHART: CRAIG HANDICAP ASSESSMENT AND REPORTING TECHNIQUE

Gale G. Whiteneck, Craig Hospital

Introduction

Within the World Health Organization's (WHO) model of impairment, disability, and handicap, handicap has received the least attention. Similarly, within most rehabilitation facilities there hae been far more attention paid to documenting organ system functioning (or impairment) and independence in activities of daily living (or disability) than in documenting the fulfillment of social roles (or handicap). It has been my effort to focus greater attention on handicap as a primary outcome of rehabilitation for people with disabilities and to develop a measure to quantify handicap. To that end, this presentation endeavors to: explain why it is important to measure handicap; develop a conceptual basis for handicap; describe various attempt5 to measure handicap with particular attention to our strategies for designing the Craig Handicap Assessment and Reporting Technique (CHART); describe a few applications of CHART in the research; and suggest the next steps needed in the field of handicap assessment.

Why Measure Handicap

Over the years, the great disparity between what we say we do in rehabilitation and what outcome5 we spend our time measuring has been troubling. The rhetoric of rehabilitation is that we reintegrate people with disabilities back into their communities as active, independent, productive members of society with a high quality of life. But what is really measured in the process, as evidenced by patient records, is motor and sensory function (is. impairment), and a variety of ADL and functional skills such as eating, bathing, and transfers, (ie. disability). But little time is spent documenting the outcome5 that are broader in scope. If community integration, independent living, social role fulfillment, productivity, and quality of life are the ultimate reasons for the existence of rehabilitation, why don't we measure these outcome5 more often? Why do we have so many measure5 of impairment and disability and so few measures of handicap?

The Conceptual Basis For Handicap

The World Health Organization model of disablement has provided considerable clarification differentiating between the key concept5 of impairment, disability, and handicap. Obviously, there is a lot of dissatisfaction with the term handicap, because of its pejorative meaning in every day use. Although theorists and agencies may choose other word5 to describe outcomes in the community and the fulfillment of normal societal roles, the World Health Organization's model, conceptualization, and terminology will be retained for this presentation.

The disablement process from impairment to disability to handicap provides an understanding of the outcomes of two integrated but distinct components of rehabilitation, which Dr. Mark Fuhrer has labeled medical rehabilitation and psychosocial rehabilitation. These two rehabilitation processes have two distinct goals: the goal of medical rehabilitation is to minimize disability for given levels of impairment; psychosocial rehabilitation goals focus around minimizing handicap for a given degree of disability.

These two components of rehabilitation occur simultaneously, but have different interventions. Medical rehabilitation follow5 the medical model with medical management, treatment and procedures, rehabilitation nursing, physical and occupational therapy, adaptive equipment, and substantial skills training. Psychosocial rehabilitation follows the educational model, utilizing patient education, counseling, vocational service, recreational therapy, public education, and advocacy.

Additionally, medical and psychosocial rehabilitation components are different in their constraints. Medical rehabilitation is constrained by associated injuries, medical complications, pre-existing conditions, inadequate funding, and poor access to services. Psychosocial rehabilitation , on the other hand, is limited by such thing as poor support networks, environmental barriers, societal discrimination, financial disincentives, and a host of pre-injury factors.

From all of this, it can be concluded that handicap should be the primary outcome measure of psychosocial rehabilitation just like disability is the primary outcome measure of the medical rehabilitation component of the rehabilitation process.

Quality of Life and Handicap

Quality of life is the catch-phrase to describe the broadest goal5 of rehabilitation. Because, as a term it may be too expansive, two separate components of quality of life might be differentiated: Subjective and objective quality of life. In the first component, objective observable criteria exist which society uses to evaluate the quality of anyone's existence. These objective criteria include where and with whom people live, what friends they have, what their jobs are, how much money they have, and how they spend their time. From these observable factors, conclusions are drawn about the quality of an individual'5 life. The second component of quality of life contain5 the subjective perceptions individuals have about the quality of their own existence. This component has more to do with personal satisfaction than with objective criteria others might observe.

Quality of life, therefore, can be conceptualized as both the degree to which an individual's life meets the expectations of others and the degree to which an individual's life meets his or her own expectations. The objective component of quality of life, it seems, may be synonymous with the concept of handicap in the World Health Organization model of disablement.

Just like handicap needs to be differentiated from the broader concepts of quality of life and subjective well-being, handicap also must be differentiated from its sources. There is no denying that impairment and disability are major causes of handicap, but they are not synonymous with handicap. Everyone can think of people who have significant impairments and disabilities, but are not handicapped because they are active, productive members of society, fulfilling many social roles just as well as non-disabled persons. Likewise, handicap is influenced by many internal factors - such as demographic factors, pre-injury characteristics, personality and motivation - even though handicap is also distinct from these factors. Many external factors also contribute to handicap including environmental barriera, societal discrimination, financial disincentives, and limited resourcea, but these, too are distinct from the resultant handicap. Handicap is the failure to fulfill normal non-disabled role expectations for whatever reason.

The Measurement of Handicap

So, how do we go about meaauring handicap? Three measurement atrategiea are suggested. It is important to document what people do in the community, decide what value society places on those behaviors and, finally, compare the value of what is observed with the expectations of non-disabled members of society.

Efforts to assess handicap have included a variety of individual gueationa. Almost any follow-up survey asks gueationa about marital atatus, employment, and living situation, all of which constitutes a portion of the handicap domain. There are a variety of instruments that look at activities that are performed by individuals in the community including the Activities Pattern Indicator, and the Self-Observation and Report Technique. The Independent Living Outcomes Measure not only assesses alternative outcomes related to independent living, but values them to create a score. Finally, more recently, the CHART and the Community Integration Questionnaire have been developed to examine at the concept of handicap directly. The Community Integration Questionnaire has focused on the traumatic brain injury area, looking at the dimension of family functioning, social integration, and productivity.

CHART, has specifically attempted to follow the conceptual framework of the World Health Organization and has developed scales for five of the six dimensions identified by the WI-IO, including physical independence, mobility, occupation, social integration, and economic self-sufficiency. The sixth dimension of handicap - orientation - is currently under development.

In developing CHART, efforts were made both to identify the correct content for the scale, and to utilize the conceptual model of WHO to accomplish that. Objective questions, that are easily answerable and are of a behavioral nature, were designed to guarantee that CHART would be an actual measure, there was a commitment from the start to developing scales, or values of O-100, so that there was a number attached to it. Too many of the individual items in the past simply asked questions without scoring of any kind, let alone interval scales. And, CHART was designed to be norm-referenced. Conceptually, it was important that the expectations of persons with disabilities be the same expectations that would be applied if they were not disabled. Thus, the norm-referencing that was chosen was the nondisabled community. Finally, CHART was intended to be a practical instrument.

Dimension of Handicap and CHART

Briefly, WHO defines physical independence as "an individual's ability to sustain a customary effective independent existence". Within CHART, physical independence is measured by looking at the amount of time that assistance is provided, particularly through attendant care.

Mobility is defined by the WHO as "the individual's ability to move about effectively in his surroundings." CHART, therefore measures mobility by documenting the time an individual spends out of bed, the accessibility of his/her living environment, the days that he/she gets out of the houae, the independent use of adequate transportation, and the time spent traveling outside of the community.

The WHO defines occupation as "the individual's ability to occupy his time in a manner customary to his sex, age, and culture." CHART, assesses the hours per week that are spent in employment, school, homemaking, volunteer work, recreational activities, and other self-improvement activities. The intent was to select the items which the general public views as productive; therefore, time spent watching TV is not assessed. Furthermore, employment, school, and homemaking generally are valued more by our society than volunteer work, recreation, and self-improvement, so in the scoring procedure for CHART, these weigh more heavily.

Social integration is defined by the WHO as "an individual's ability to participate in and maintain customary social relationships." Within CHART, social integration is assessed by whether individuals are living with a spouse, significant other, or a room-mate; whether they are involved in a romantic relationship; and by a series of questions about whether they have regular contact with friends, relatives, business associated, and acquaintances.

Finally, economic self-sufficiency is defined as "the individual's ability to sustain customary socioeconomic activity and independence." CHART assesses this area through the combined annual income of all family members in the household, considering income from all sources after discounting the unreimbursed medical care expenses paid out for the disability.

Application of CHART

CHART was designed to be a community-based instrument, thus it is not intended to be

instrument, thus it is not intended to be administered until after the individual has returned to the community. CHART is currently being used throughout the lifetime5 of disabled clients, beginning with the first year post-injury. It was designed to be simple and objective and its reliability and validity have been assessed. Initially, norming studies were conducted with CHART to refine the scoring procedures, ensure that most persons without disability demonstrate no handicaps, and to establish expected values for various group5 with spinal cord injury.

Next, a multi-center collaborative study investigated the psychometric properties of CHART. The test-retest reliability is 0.93. Correlation between self-report5 of the client and information provided by family member5 is 0.83. It differentiates clients that rehabilitation professionals globally identify as either of high handicap or low handicap, it differentiates persons with disability from those without disability, and it differentiates among persons with varying degree5 of impairment. Rasch analysis supported the item structure.

The interrelationships of handicap and other concepts were then investigated in a study of spinal cord injury. Impairment was assessed by the ASIA motor score, disability was assessed using the Functional Independence Measure, and handicap was assessed using the CHART. Information about a variety of pre-injury factors and a variety of post-rehabilitation factors also was collected. It was learned that impairment was a very good predictor of disability (68% of the variance in disability is explained by impairment), but neither impairment nor disability were very good predictors of handicap. The two of them combined explained a little over 1/3 of the variance in the handicap score. The further inclusion of the pre-injury factors and the post-rehabilitation factors brought the total explained variance in handicap up into a higher range. The lesson to be learned from this is that handicap is really quite distinct from impairment and disability, and that a variety of other factors need to be considered in attempting to build a predictive model of handicap.

The CHART has been used in studies of aging with spinal cord injury. Clear evidence of increasing handicap has been seen with both increasing age and increasing years post injury. CHART may be clinically useful in identifying deterioration in social role fulfillment that warrant5 intervention. A growing number of longitudinal studies of spinal cord injury are utilizing CHART a5 an outcome measure. In particular, the Rehabilitation Research and Training Center in Community Integration has examined the positive relationships between handicap and subjective well-being. CHART and the Community Integration Questionnaire are also being utilized in studies of traumatic brain injury outcome including Model System's and Regional Center's research, and studies to validate the appropriateness of these instruments in this population are underway.

Future Research

The next step5 in handicap assessment are to improve the comprehensiveness of the tool5 that currently are available. Within the area of CHART itself, we are still not assessing one of the key dimensions of handicap - orientation or a cognitive dimension. CHART and the Community Integration Questionnaire seen to be the two measures that are most directly attempting to measure the concept5 of handicap. These instruments need to be validated on additional impairment groups. Third, because of the interplay of the many variables, we need to understand the influencing factor5 that relate to handicap. We also need measures of handicap included in program evaluation and in batteries of rehabilitation outcome measures.

Finally, handicap assessment is needed in the evaluation of ADA goals. The Americans with Dieabilities Act expresses its goals in terms of assuring equality of opportunity, full participation, independent living, and economic self-sufficiency. It has been suggested by Scott Brown that the six dimensions of handicap identified in the WHO model closely parallel the four stated goals of the ADA. Indeed, orientation, physical independence, and mobility roughly correspond to the concept of independent living. Occupation corresponds very closely to equality of opportunity. Social integration correspond5 very closely to full participation, and economic self-sufficiency is a term used in both documents. CHART or other improved measures of handicap are clearly needed to assess both the success of specific rehabilitation program5 and the overall success of society in reducing handicap through such legislative efforts as the ADA.

APPLICATIONS OF THE INTERNATIONAL CLASSIFICATION OF IMPAIRMENTS, DISABILITIES AND HANDICAPS: FROM CONCEPTS TO TAXONOMY TO DATA

Scott Campbell Brown, United States Department of Education

Historically, health statisticians have been strongly interested in tracking mortality and its causes. As life expectancy has increased and attention has shifted somewhat from acute to chronic diseases, the monitoring of chronic morbidity has attracted greater attention.

Only recently has disability been recognized as a distinct concept from chronic morbidity. There have been at least three reasons for this recognition and for the increasing attention that health statisticians are paying to disability-related phenomena.

First, passage of the Americans with Disabilities Act implies that persons with disabilities will have some unique protections under the law. As a protected group, it is thus important for persons with disabilities to be monitored for health and other disparities as, African Americans, women and persons in particular age groups are.

Second, disability has been recognized as being at least partially a consequence of health-related phenomena. Related to this is the growing recognition that although a majority of persons with impairments or disabilities are in good health, such persons are more likely than the general population to characterized themselves in fair to poor health (Ries & Brown 1991).

Finally, greater attention is being paid to the World Health Organization's 1948 definition of health as "...not merely the absence of disease, but total physical, social and psychological well being." Under this view.of health, even Civil Rights legislation, such as the Americans with Disabilities Act, can be viewed as legislation for health promotion, because it seeks to ensure well being for persons with disabilities (WHO 1964).

The purpose of this paper is to demonstrate how the World Health Organization's (1980) International Classification of Impairments, Disabilities and Handicaps can be employed to evaluate progress towards the goals of the ADA. First, the article shows how both the conceptual framework and the taxonomy are relevant to the ADA as written. Then, it is shown how the taxonomy can be used to reconcile different disability data sources for children in order to monitor progress towards the ADA goals for children.

The ADA states that "...the Nation's proper goals regarding individuals with disabilities are to assure equality of opportunity, full participation, independent living and economic self-sufficiency for such individuals (Public Law 101-336, July 26, 1990, 104 STAT. 327, Section 2, 42 USC 12101)." It is clear that the ADA is aimed at achieving outcomes for persons with disabilities, as opposed to preventing disability. But, who are the people with disabilities for whom these goals are to be achieved?

The ADA defines disability as ". . ..with respect to an individual

(A) a physical or mental impairment that substantially limits one or more of the major life activities of such individual;

(B) a record of such an impairment; or

(C) being regarded as having such an impairment."

Note that under this definition, impairment or the history or recognition thereof is the entry point for being classified as disabled under the ADA. However, the impairment must be linked to a limitation in one or more major life activities.

The Equal Employment Opportunity Commission has determined that major life activities means functions such as caring for oneself, performing manual tasks, walking, seeing, hearing, speaking, breathing, learning and working (EEOC 29 CFR Chapter XIV Part 1630, Section 1630.2(i)). The EEOC notes that this term adopts the definition of the term "major life activities" found in the regulation implementing Section 504 of the Rehabilitation Act at 34 CFR part 104.

It further elaborates that major life activities are those basic activities that the average person in the general population can perform with little or no difficulty. Noting that the above list of major life activities is not exhaustive, the EEOC has stated that other major life activities include, but are not limited to, sitting, standing, lifting and reaching (EEOC 29 CFR Chapter XIV Part 1630.2(i)).

It is clear from this definition that the EEOC has defined a major life activity limitation as a person-level function. It is also based on what the average person can do.

These statutory and regulatory definitions correspond well to the ICIDH framework. The ICIDH uses the following three definitions:

Impairment - any loss or abnormality of psychological, physiological or anatomical structure or function.

Disability - any restriction or lack (resulting from an impairment) of ability to perform an activity in the manner or within the range considered normal for a

human being.

Handicap - a disadvantage that limits or prevents fulfillment of a role that is normal (depending on age, sex and social and cultural factors).

From these definitions it can be first observed that the presence, record or recognition of an impairment refers to an organ level function and serves as the point of entry in both the ADA and ICIDH. Then, the ADA definition of major life activity reduction corresponds to the ICIDH definition of disability, as both refer to person-level functioning. At this point, a person is established as disabled. Finally, the ADA goals seek to eliminate disadvantages which are Handicaps in the ICIDH.

As there is good correspondence between the ADA and ICIDH concepts, so, too, the ICIDH are applicable to the ADA examples of major life activities. Disability codes and classifications can be used for each of the major life activities noted by the EEOC, as shown in Table 1. Out of thirteen activities, eight have some word overlap. Moreover, there are Disability codes and classifications for activities not listed. This allows for a more expansive definition of Disability.

Likewise, there appears to be some agreement between the ADA goals and the ICIDH survival roles which serve to capture the dimensions of Handicap. As discerned in Table 2, the linkage to the six survival roles is an attempt to separately define the goals so that they are not merely interchangeable. These linkages have been further elaborated elsewhere (Brown 1993).

Given the good correspondence between the ICIDH taxonomy and the ADA, the ICIDH would appear to be a useful mechanism for monitoring progress towards the four goals outlined in the ADA. The Impairment and Disability classifications could be used to help determine who is covered under the Act and the Handicap dimensions can be used to determine whether the Act and other pieces of legislation providing services are improving the quality of life for persons so covered.

In order to provide an example of how such monitoring can be established, children have been selected for analysis for at least two reasons. One reason for selecting children is to emphasize that they, too, are covered under the ADA. The second is to note that, because it is not clear as of this writing what all the ramifications of the ADA are for children, their situation provides an interesting laboratory for testing the applicability of the ICIDH to the ADA.

For this data analysis, six data sets with data on children with Disabilities have been selected. As demonstrated by Table 3, each of the survey's Disability categories have been retroactively coded

Table 1: ADA MAJOR LIFE ACTIVITIES AND CORRESPONDING DISABILITY CODES AND CLASSIFICATIONS FROM THE INTERNATIONAL. CLASSIFICATION OF IMPAIRMENTS, DISABILITIES AND HANDICAPS (ICIDH)

ADA Major Life Activity	Disability Code and Classification	
Caring for Oneself	3:	Personal Care
Performing Manual Tasks	6:	Dexterity
Walking	40:	Walking
Seeing	25-27:	Seeing
Hearing	23-24:	Listening
Speaking	20-22:	Speaking
Breathing	59:	Other Body Disposition
Learning	16:	Other Educational
Working	18:	Occupational Role
Sitting	46.1:	Transfer from Sitting
Standing	46.2:	Standing Transfer
Lifting	48:	Lifting
Reaching	53:	Reaching
Other	10-15:	Awareness
	17:	Family Role
	19:	Other Behavior
	41-45:	Ambulation
	47:	Transport
	49:	Other Locomotor
	50-51:	Domestic
	52:	Retrieval
	54-57:	Body Movement
	58:	Postural
	7:	Situational
	8:	Particular Skill
	9:	Other Activity Restrictions

Table 2: ADA GOALS AND CORRESPONDING HANDICAP CODES AND CLASSIFICATIONS FROM THE INTERNATIONAL C-CATION OF IMPAIRMENTS, DISABILITIES AND HANDICAPS (ICIDH)

ADA Goal	Survival Role and Definitions
Equality of Opportunity	4: Occupation Extent to which time-use is sustained, active and normative
Full Participation	5: Social Integration Participation in/maintenance of customary social relationships
Independent Living	1: Orientation Signal receptivity, assimilation and expression of response 2: Physical Independence Sustained and customary independent existence 3: Mobility Degree of effective movement within surroundings
Economic Self-Sufficiency	6: Economic Self-Sufficiency Sustained/customary socio-economic activity or independence

Table 3: DATA SETS ON DISABILITY FOR CHILDREN IN THE UNITED STATES CODED TO THE INTERNATIONAL CLASSIFICATION OF IMPAIRMENTS, DISABILITIES AND HANDICAPS (ICIDH DISARILITY CODE AND CLASSIFICATION AND SURVEY CATEGORY)

Disability Code and Classification	National Educational Longitudinal Study, 1988	National Assessment of Educational Progress, 1988	Current Population Survey, 1992	Survey of Income and Program Participation, 1990-91	National Health Interview Survey/Child Supplement, 1988*	Office of Special Education Programs, 1987-88
15: Knowledge acquisition	Mental retardation	Mentally retarded	Mental retardation	Mental retardation	Delay in Growth/development	Mentally retarded
16: Other educational	Specific learning problem	Learning disabled	Learning disability	Learning disability; Autism	Learning disability	Learning disabled
19: Other behavior	Emotional problem	Emotionally disturbed	Serious emotional disturbance	Mental or emotional problem or disorder	Emotional! behavioral problem	Emotionally disturbed
21: Talking	Speech problem	Speech impaired	Speech impairment	Speech problems	*Speech impairment*	Speech impaired
23: Listening to speech	Hearing problem/ Deafness	Hard of hearing/ Deaf	Other hearing impairment/ Deafness	Serious trouble hearing or deafness	*Other trouble hearing or deafness in one or both ears*	Hard of hearing or deaf
26: Detailed visual tasks	Visual handicap (not correctable with glasses)	Visually handicapped or blind	Other visual impairment/ Blindness	Vision problems or blindness	*Other trouble seeing or blindness in one or both eyes*	Visually handicapped
23 and 26		Deaf/blind				Deaf-blind
4: Locomotor				Limited ability to walk. run. or use stairs		
57: Other body movement	Orthopedic problem	Orthopedically impaired	Orthopedic impairment	Paralysis of any kind; Cerebral Palsy; Impairment or deformity of back, side, foot or leg; Impairment or deformity of finger, hand, or arm; missing legs, toes, arms, hands, or fingers	*Paralysis of extremities (complete or partial)/ Orthopedic impairment*	Orthopedically impaired
9: Other activity restrictions	Other physical disability/ Other health problem	Multidisabled/ other	Other health impairment	Other		Multi-handicapped/ Other health impaired

**For the Child Supplement, data items in italics (Codes 21, 23, 26 and 57) might be more properly classified as Broad Impairment Codes 3, 4,s and 7, respectively.*

to the ICIDH Disability classifications. An attempt was made to group the studies as closely to children thirteen years old in 1988 as possible.

The first survey is the National Educational Longitudinal Study, 1988, conducted by the Department of Education (1990) National Center for Education Statistics. Although this survey excluded, 5.4% of children with disabilities, parents of eight graders were asked if their child had "any of the following problems." This screening for problems allows for an assessment of the prevalence of particular Disabilities for children included in the NELS.

Another study included that was conducted by NCES is a sample of thirteen year old students from the National Assessment of Educational Progress, 1988. Here, too, 5% of students were excluded due to the presence of a Disability. Unlike the NELS, the information available about students with Disabilities are for those who were excluded, not included.

The third study is for youth three to fourteen years old in the Current Population Survey, 1992. Disability categories used in this study correspond well to those used by the Department of Education's Office of Special Education Programs. However, it is possible that, in some cases, the parents may have been referring to organ-level, as opposed to person-level, function. This may imply that the data would be more Impairment than Disability based. Because parents were focused on their child's education, it is more likely that the frame work of reference was the functioning of the child.

The fourth study is the Survey of Income/ProgramParticipation, 1990-91. As with the NELS, this data used the concept of problem in collecting data. As with the CPS, the study might have had an Impairment-based focus. Little data are currently available, except to note that 1.6% of all children surveyed had a Locomotor Disability (Code 4).

The National Health Interview Survey Child Supplement, 1988 is the fifth survey and it covered children under eighteen years. Of all of the surveys, it is the most likely to have an impairment-based focus. It has been included for comparative purposes. It is also important to note that some of the disability categories were asked in terms of whether the child ever had the condition, often referred to as lifetime prevalence. This can result in non-comparable rates that are relatively large as compared to current prevalence.

The last set of data is from the Department of Education (1989) Off ice of Special Education Programs, 1987-88. To be included, students with disabilities must have an Individual Education Program. This implies that the disability is severe enough to warrant special education services.

Observation of the six data sets reveals a great deal of similarity in the categories surveyed. Terms for learning, emotion, speech and hearing are captured in all six, while mental and vision are captured in five. These similarities made assignment to ICIDH disability categories rather easy.

While the five surveys with data were similar in what the categories were called, they were quite different in the resulting prevalence rates. Rates for the NAEP were generally low, as demonstrated in Table 4. Except for the mentally retarded and speech categories, the OSEP rates were generally the second lowest and relatively close to those in the NAEP. The NHIS Child Supplement's rates were very high in ever disability categories of developmental delay, learning disability and emotional or behavioral problem (Codes 15, 16 and 19, respectively).

These differences in rates could be due to differences in a variety of factors definitions for terms, the operationalization of similar definitions, age coverage, year coverage and sample selection. It also appears that multiple choices within the Disability classifications raise the overall prevalence for classification totals.

Some similarities between surveys should be noted, however. Talking disabilities (Code 21) had a range of 1.8%-2.7% in four surveys - NELS, CPS, NHIS/Child Health and OSEP. Knowledge acquisition disabilities (15) had a range of 0.7%-1.4% in three surveys - NAEP, CPS and OSEP. The visual categories (21) had a range of 0.9%-1.9% in three surveys - NELS, CPS and NCHS/CH. Likewise, other body movement (57) range of 0.1%-0.9% in three surveys - NELS, CPS and OSEP. Note that there was some agreement between CPS and NELS in four categories (deafness, speech, vision and orthopedic) and between CPS and OSEP in four categories (mental retardation, emotional disturbance, speech and orthopedic).

Turning from the prevalence of Disabilities to ADA goals, two data sets have been selected for analysis - OSEP and the National Longitudinal Transition Study of Special Education Students, conducted by SRI. The NLTS uses the OSEP categories but follows a cohort of students over time. Thus, OSEP rates represent a period analysis while NLTS rates represent a longitudinal study.

Two ADA goals have been selected. To represent equality of opportunity by promoting occupational roles (Handicap 4), graduation rates are examined. To represent full participation by promoting social integration (Handicap 5), rates of participation in regular classes are analyzed.

The data reveal overall differences

in ADA Goal achievement between the two data sets. OSEP data graduation rates and participation in regular classes higher than for the National Longitudinal Transition Study, as given in Table 5. These differences are large, even though survey categories are the same in both studies. This difference is due to the longitudinal nature of the NLTS data versus the cross-sectional OSEP data.

This study has revealed certain problems in reconciling disability data sets for children. First, there clearly is a range of disability - from simple action to complex activities. This problem is inherent in both the ADA and the ICIDH. Second, the research has demonstrated some problems in the retrospective application of ICIDH to disability statistics. A prospective application would be better. Finally, an important question remains: Can similar definitions be operationalized in a reliable fashion between surveys? Two important considerations here are the severity of the disability and the nature of the data set (i.e., cross-sectional versus longitudinal data).

There are, however, several potential solutions to the problem of reconciling differences in disability data. First, the use of ICIDH has demonstrated that a) there is a good conceptual fit with the ADA, b) the ICIDH taxonomy fits well with some ADA groupings and c) there is a solid potential for producing reliable data analyses. Second, positive measures of function can be developed to explore what individuals can do, as opposed to what they cannot. Finally, a set of indicators for progress towards ADA goals can be developed to assess the appropriate of mechanisms to reduce Handicaps for persons with Disabilities.

REFERENCES

Brown, Scott Campbell. 1993. Revitalizing "Handicap" for disability research: Developing tools to assess progress in quality of life for persons with disabilities, Journal of Disability Policy Studies, Volume 4, Number 2.

Ries, Peter and Scott Brown. 1991. Disability and health: Characteristics of persons by limitation of activity and perceived health status, United States, 1984-88, Advance Data from Vital and Health Statistics, Number 191, Hyattsville: National Center for Health Statistics.

United States Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics. 1990. National Education Longitudinal Study of 1988. Base Year: Parent Component Data File User's Manual, NCES 90446, Data Series DR-NELS:88-88-2.2. Washington: Government Printing Office.

United States Department of Education, Office of Special Education and Rehabilitative Services, Office of Special Education Programs. 1989. Eleventh Annual Report to Congress on the Implementation of the Individuals with Disabilities Education Act. Washington: Government Printing Office.

World Health Organization. 1964. Constitution of the World Health Organization. Geneva: World Health Organization Basic Documents, Fifteenth Edition.

World Health Organization. 1980. International Classification of Impairments, Disabilities and Handicaps: a manual of classification relating to tbe consequences of disease. Geneva: World Health Organization.

SOURCES FOR TABLES

Table 1: Equal Employment Opportunity Commission 29 CFR Chapter XIV Part 1630, Regulations to Implement the Equal Employment Provisions of the Americans with Disabilities Act, Section 1630.2(i) Major Life Activities and World Health Organization, The International Clessitication of Impairments. Disabilities and Handicaps, Geneva, 1980, pages 144-180.

Table 2: Public Law 101-336, The Americans with Disabilities Act of 1990 (104 STAT. 321 42 USC 12101, Section 2(b) and World Health Organization, The International Classification of Impairments, Disabilities and Handicaps, Geneva, 1980, pp. 183-205.

Table 4: See Table 3.

Table 5: Public Law 101-336, The Americana with Disabilities Act of 1990 (104 STAT. 321 42 USC 12101, Section 2(b) and World Health Organization, The International Classification of Impairments, Disabilities and Handicaps, Geneva, 1980, pp. 144-205; United States Department of Education, Office of Special Education and Rehabilitative Services, Office of Special Education Programs, Fifteenth Annual Report to Congress on the Implementation of tbe Implementation of The Individuals with Disabilities Education Act, Washington: Government Printing Office, 1993; Values, Kathryn A., Cynthia L. Williamson and Mary Wagner, The National Longitudinal Study of Special Education Students. Statistical almanac, Volume 1: Overview, Menlo Park: SRI International, 1990, Page 89, Table 15A; Page 5-l, Figure 5-2, Wagner, Mary, Sticking it out: Secondary school completion, in Mary Wagner, Lynn Newman, Ronald D'Amico, E. Deborah Jay, Paul Butler-Nalin, Camille Marder and Robert Cox, Youth with disabilities: How are thev doing? The first comprehensive report from the National Longitudinal Study of Special Education Students, Menlo Park: SRI International, 1991, Page 5-4, Figure 5-2.

FOOTNOTE INFORMATION FOR DATA SOURCES FOP. TABLES 4 AND 5

1. Disability rates for the National Educational Longitudinal Study are for included students only. A total of 5.4 percent of students were excluded due to a physical disability, mental disability, or language problem. Approximately 64 percent of those excluded had a disability, the remainder having only a language problem. A "*" means the data were suppressed for confidentiality reasons and m-coded under "Any other health problem."

2. Disability prevalence among the approximately 5% of students excluded from the National Assessment of Educational Progress because of a physical disability, mental disability, or language problem. Approximately7996 of excluded 13 year old students had a disability, the remainder having only a language problem. This may provide a basis for approximations of the disabilities of students excluded fmm the NELS. NELS collected no specific disability data on excluded students. Tbe percentage is the number of excluded students with a disability as a per cent of total students. This bracketed figure when added to the NELS percentages, approximates the total population, included and excluded, with a disability. Bath NELS AND NAEP excluded approximately the same percentage of students for similar reasons.

3. Data are from the Current Population Survey.

4. *For the National Health Interview Survey Child Supplement. data items* in italics *(Codes 21, 23, 26* and *57) might be more* properly *classified as Broad Impairment Codes 3, 4, 5* and 7, *respectively.* Data items not in italics (Codes 15, 16 and 19) include children who ever bad the condition (lifetime prevalence). Code 19 for emotional or behavioral problem was only asked for those aged 3 to 17 years but it includes children who a) aver had, b) ever been treated for or c) have been suggested by anyone for help for any emotional, mental or behavioral problem (three separate questions, with a positive response to any one counted). All rates are conditions per 100 persons. A person may have more than one condition pet category, especially for orthopedic impairments group which consists of many subgroups.

5. Figures from the Office of Special Education Programs are for public school children who have an Individualized Education Program (IEP). This excludes children and disabilities unrelated to an IEP.

6. Data are from the National Longitudinal Transition Study of Special Education Students.

ACKNOWLEDGEMENTS

The author wishes to thank Robert C. Ficke for compiling the data, Scott Crosse and Rob Ficke for ICIDH backcoding, Marsha Brauen for serving as Project Director and William Frey for consulting. The opinions expressed in this paper, however, are solely those of the author and do not express those of the United States Departments of Education or Health and Human Services.

Table 4: PREVALENCE RATES FOR DISABILITIES F'ROM DATA SETS ON DISABILITY FOR CIIILDREN IN THE UNITED STATES CODED TO THE INTERNATIONAL CLASSIFICATION OF IMPAIRMENTS, DISABILITIES AND HANDICAPS (ICIDH DISABILITY CODE AND CLASSIFICATION, SURVEY CATEGORY AND RATE)

Disability Cede and Classification	1988 NRLS (8th grade)	1988 NAEP (Age 13)[a]	1992 CPS (Age 3-14)[a]	1988 NHIS/CH (Under age 18)	1988 OSEP (Age 6-17)[a]
15: Knowledge acquisition	.	1.0	0.7	4.0	1.4
16: Other educational	6.6	2.1	3.6	6.5	5.0
19: Other behavior	3.4	0.3	0.9	13.4	1.0
21: Talking	1.8	0.1	2.1	1.8	2.6
23: Listening to speech	2.9 2.5 0.4	0.0 0.0 0.0	1.6 1.2 0.4	1.7	0.1
26: Detailed visual tasks	1.8	0.0	1.9 1.6 0.3	0.9	0.1
23 and 26		0.0			0.0
57: Other body movement	0.9	0.0	0.9	3.1 0.2 2.9	0.1
9: Other activity restrictions	5.5 1.2 4.3	0.5 0.2 0.3	1.8		0.3 0.2 0.1

Table 5: PERCENTAGES OF EXITING STUDENTS WITH INDIVIDUALIZED EDUCATIONPROGRAMS WHO GRADUATR AND PERCENTAGES OF STUDENTS WITH INDIVIDUALIZED EDUCATION PROGRAMS SERVED IN REGULAR EDUCATION CLASSES IN THE UNITED STATES CODED TO THE INTERNATIONAL. CLASSIFICATION OF IMPAIRMENTS, DISABILITIES AND HANDICAPS (ICIDH DISABILITY CODE AND CLASSIFICATION, SURVEY CATRGORY AND GOAL PERCENTAGE)

Disability Cede and Classification	Survey Category	Equality of Opportunity 4: Occupation *Graduated*		Full Participation 5: Social Integration *Regular Class*	
		OSEP[a]	NLTS[b]	OSEP[a]	NLTS[b]
15: Knowledge acquisition	Mental retardation	63.3	49.8	6.6	1.5
16: Other educational	Learning disability	62.5	60.9	21.8	7.9
19: Other behavior	Emotional disturbance	38.6	41.7	15.6	1.8
21: Talking	Speech impairment	50.4	62.7	70.6	21.8
23: Listening to speech	Hearing impairment Hard of hearing Deaf	73.1	 72.3 71.6	27.0	 7.3 1.9
26: Detailed visual tasks	Visual impairment	74.8	69.5	40.0	16.1
23 and 26	Deaf-blind	70.4	43.1	14.6	2.7
57: Other body movement	Orthopedic impairment	68.3	75.1	28.3	10.5
9: Other activity restrictions	Multiple disabilities Other health impairment	64.9 65.0	32.2 65.4	6.6 29.6	1.6 18.5

SUMMARY OF DISCUSSION

Mary Chamie, United Nations statistical Division'

This session has demonstrated that a discipline is evolving for the study of human function and performance in daily life and in society. The field of disablement covers a wide range of academic territory and programmatic concerns. Its diversity as a field has been amply demonstrated in the four presentations of this session.

In the past most medical and health-related research focused on the disease process itself. The outcomes of disease were investigated through study of the disease or injury process and through assessment of survival probabilities. This is very much exemplified by the first presentation in this session. In contrast, in sociological and economic spheres, attention is focused on the study of differential marginality, poverty, and opportunity loss among population groups. This area of research has studied the effects of stigma, prejudice and ignorance upon the relative distribution of power. Attributes studied and classified for their predictive power in sociological and economic research have focused on such issues as differential status according to race, ethnicity, gender, educational attainment and income of individuals and of population groups. Little attention has been previously given to the analysis of disability, marginality, and the relative distribution of power.

How then, could these two large and energetic approaches of medical and socioeconomic research be combined and usefully applied to the study of disablement, or the assessment of human function, performance and participation?

Barbara Altman remarked in her paper, that quality of life research dwelt in the domain of subjectivity through the use of weighted estimates of the value of life, and remained largely incomparable across research studies. She feared that disability might inadvertently fall into this domain of subjectivity - perhaps never to return as an objectively measurable event - if attempts were not made to evaluate and validate existing disability measures.

She proposed that these widely divergent and existing disability measures be distilled and that they be more accurately focused to meet the needs of policy issues facing us today. She did not, however, concretely propose how the divergent and existing measures were to be distilled and selected.

Scott Brown and Gale Whiteneck both chose a common framework for their selection of issues, while using very different assessment tools. How were they able to do this? They both recognized the usefulness of working within a common conceptual and classificatory framework, even when using alternative theories, research methods and statistical measures. The common framework that they chose was the ICIDH, or the International Classification of Impairments, Disabilities and Handicaps. The ICIDH allows researchers interested in the disease and injury process to study the long-term consequences of disease and injury, through assessment of remaining organ function (impairment), personal function and activity (disability), and through the study of the relationship between society, culture and people with impairments and/or disabilities (handicap).

The framework of the ICIDH, provides a logical and measurable link between medical and public health research on the disease and injury process, and the socioeconomic and psychological research on quality of life outcomes from long-term reductions in function or from changes in appearance through experiences with scarring, limping, deformity or modifications of human behaviour.

In the policy arena, programmes of *prevention* have focused mainly on the study of impairment, with some attention to disability. Programmes of *rehabilitation* have focused largely on the issue of disability, with some attention given to the impairment through attention to access of clients to appropriate technical aids such as prostheses, wheelchairs, hearing aids and spectacles. Programmes aimed at equal opportunity and human rights focus on the issue of the elimination of discrimination, disadvantage and marginality through the study of *handicap* (relationship of person and society in life-roles).

This is a very large field of study. It cannot, in these early stages, be limited by researchers to one or two agreed-upon indicators, only. Freedom to experiment with emerging issues, to argue the importance of one major arena over another, is essential. Policies focused on the impairment, disability and/or handicap must be allowed and encouraged.

This experimentation does not, however, have to proceed in a vacuum of language and concepts. Very little science has taken place without benefit of systematized language and common concepts for the discussion of diversified theories and studies.

Common concepts, definitions and assessment tools are emerging in the field of disablement. You saw much evidence of this in today's session. The ICIDH allows researchers to indicate which facet of disablement they have chosen to study. It allows policy makers to objectively describe the aims of their policies, be they

prevention of disease or injury; reduction of disability; or elimination of handicap through human rights measures and through elimination of stigma and barriers. These activities would be instituted in addition to those programmes dedicated to modification of the behaviour of impaired and/or disabled persons themselves.

Having made these general comments, I would now like to offer several specific observations about the papers presented in this session.

I am concerned with the use of bed_days as a disability measure, as was presented in Altman's paper, because it implies sickness. Disability, as a concept, does not measure illness, nor does it consider death or mortality as a disability outcome. How then, could bed-days be considered in the field of disablement, when disability is a measure of long-term reduction in human function and/or activity among persons who survived illness, and/or injury, rather than an assessment of illness per se?

In my opinion, and similar to the way in which Altman embedded the concept of "bed-days" in social roles, long-term stays in bed would be best studied through the assessment of handicap, or reduced life-roles. Being in bed, such as an elderly person in a nursing home, might be the result of severely reduced function; it might also be due to loss of opportunity to get up from lack of access to personal assistance; it might be due to lack of access to mobility devices such as wheelchairs, or perhaps because of negative attitudes of nursing staff toward abandoned elderly persons who have impairments. In any case the end result is a constricted concentric circle of movement and socialization, or reduced life roles, expressed in bed-days. Policy interventions to reduce handicap measured as long-term bed days would then take the person and his/her environment into account when attempting to reduce bed-days.

My question to Dr. Altman, is also, to know whether she would consider placing these diverse measures of "disability" presented in her paper into the common ICIDH framework so that they could be sorted out conceptually? Which of these measures are impairment specific; which are disability or handicap-specific; and which appear to be a blend of concepts? Knowing the purpose of the measure might help to more objectively measure it, even when using alternative assessment tools. For example, social security policy for the provision of work disability benefits may actually be determined through medical assessment of impairment states, rather than through any direct assessment of work disability. Knowing the purpose of programme indicators would sharpen the description of assessment tools when using ICIDH indicators. It would also assist to establish the underlying policy or programmatic goal attached to the measures selected.

Scott Brown's application of the ICIDH to ADA policy is an excellent example showing how common concepts assist in linking policy and research goals to programme monitoring and assessment. Regarding the codification of programme activities, however, I have specific questions concerning Table 3. In Table 3 - Dr. Brown placed impairment concepts, i.e. mental retardation, into the disability code of knowledge acquisition; knowledge acquisition is a disability code which may or may not be associated with mental retardation. Visual impairments or hearing impairments may also result in limited knowledge acquisition. It might be better for Dr. Brown to indicate that programmes of special education aimed at mental retardation, are impairment-specific.

Similarly, hard of hearing and deafness may or may not be the only impairments associated with difficulties in listening to speech (disability) as laid out in Table 3. Other psychological problems could also be associated with this disability. Again, in this case, programmes addressing deafness are impairment specific. We must be careful to not impute disability from impairment states.

One must caution coders that there is no such thing as a mental or physical disability in the ICIDH. Disability is the description of human performance at the level of personal behaviour and activity, both simple and complex. Impairment describes physical and mental functioning of organ or other systems of the body. Thus, difficulties in knowledge acquisition (disability) may be attributable to physical impairment, psychological impairment, or both. The underlying impairment of any disability may not necessarily be readily imputed from description of the disability itself.

The CHART of Gale Whiteneck's presentation is a very interesting and creative attempt to have impairment and disability research assess long-term outcomes of programme intervention through the study of handicap. I would appreciate learning from Dr. Whiteneck, how the normative aspects of the CHART measure would be controlled across social class? For example, how might one control in the analysis, for predicting SES differences in individual life-roles? Has stratification of the results by social class been tried in the study of comparison of scores between people with and/or without disability?

I conclude by applauding these research papers for delving into the issue of disablement in systematic and conceptually creative ways. Disablement is rapidly becoming firmly established as a

field of research, no longer viewed solely as an appendage to the field of medicine; no longer adrift in a sea of quality of life measures, ungrounded in the assessment of human performance, and participation.

---

1. The opinions expressed by the Discussant are those of the author and do not necessarily reflect those of the United Nations.

Session L

Health Statistics from Record Systems

Toward the year

2000

Refining the measures

Access to Prenatal Care Following Major Medicaid Eligibility Expansions

Paula Braveman, MD, MPH; Trude Bennett, DrPH; Charlotte Lewis, MNS, RD; Susan Egerter, PhD; Jonathan Showstack, MPH

Objective.-To determine whether lack of financial access was a significant barrier to prenatal care following major expansions of Medicaid eligibility in California.

Design.—Retrospective analysis of birth certificates, assessing risks of inadequate prenatal care by insurance, controlling for maternal race/ethnicity, birthplace, age, parity, education, and marital status.

Sample.—Singleton live births to California residents occurring in-state in 1990 (N=593 510).

Outcome Measures.—Untimely initiation of care, too few visits, and no prenatal care.

Results.—Despite major Medicaid expansions, nearly 11% of live births were uninsured for prenatal care. Being uninsured and having Medi-Cal were both risk factors of sizable magnitude, controlling for maternal characteristics Compared with women who had private fee-for-service coverage, uninsured women were at elevated risk of untimely initiation (odds ratio [OR], 2.54; 95% confidence interval [CI], 2.47 to 2.69) and too few visits (OR, 2.49; 95% CI, 2.44 to 2.55). Women with Medi-Cal had a high risk of untimely care (OR, 3.33; 95% CI, 3.26 to 3.40); their risk of too few visits was also elevated (OR, 1.33; 95% CI, 1 .60 to 1.66) but less than for the uninsured. Lack of private insurance was a strong risk factor for no care (OR, 6.70; 95% CI, 6.00 to 7.47).

Conclusions.-In spite of major Medicaid expansions, access to prenatal care was limited for women without private insurance. Medicaid was associated with untimely entry but with improved continuity. The findings suggest that financial barriers were salient even when controlling for many factors related to care-seeking behavior. Policy initiatives need to address continuing financial barriers along with other obstacles.

(*JAMA.* 1993;269:1285-1289)

A SUBSTANTIAL and increasing proportion of women fail to receive adequate prenatal care in the United States, contributing to this country's poor international ranking on infant mortality.[2,3] Adequate prenatal care has been shown to reduce the risk of poor birth outcomes,[4] especially for women in the highest risk groups," end financial barriers have been a key determinant of insufficient care.'"

Concern about financial access led to federal mandates in the late 1989s to expand eligibility criteria for Medicaid coverage for maternity care. California has utilized federal options and state tobacco tax funds-to nearly double the Medi-Cal (California's Medicaid program) income eligibility level for pregnant women. In July 1989, eligibility for coverage of maternity care under Medi-Cal was raised from near the federal poverty level to 188% of the poverty level; on January 1, 1990, eligibility was boosted to 299% of the poverty level.

Medicaid coverage may not confer financial access, however, if a limited number of providers are willing to accept Medicaid obstetric patients. Obstacles to enrollment may delay lnltiation of care for Medicaid-eligible women regardless of provider availability. Institutional and "personal" barriers, as distinguished from strictly financial issues, have been shown to be powerful predictors of inadequate prenatal care for low-income women.[7] Several studies have shown that Medicaid coverage may not be sufficient to overcome these obstacles.[8-11]

It is not known to what extent California's Medicaid policies have improved access and thus enhanced prenatal care utilization. This study was conducted to determine whether lack of financing remains an important risk factor for inadequate prenatal care in Caliiornla and whether coverage by Medi-Cal is associated with timely entry into care and an adequate number of visits once care is initiated. If financing strategies have achieved their maximum benefit, future reforms should focus primarily on nonfinancial access barriers related to the organization and quality of services, logistical problems with transportation and child care, and health knowledge, attitudes, and beliefs.[3,4,7,8] However, the persistence of financial barriers suggests the need for further state and national policy initiatives to expand financial access while addressing other barriers as well.

Findings from birth certificates have indicated that maternal racial or ethnic group (African American, Latina, or Native American), age less than 20 years, low educational attainment, high parity, and being unmarried are prominent risk factors for inadequate prenatal care? Although all of these characteristics are correlated with low income and lack of private insurance, birth certificates traditionally have lacked information on fi-

From the Department of Family and Community Medicine (Drs Braveman, Bennett, and Egerter) and the Institute for Health Policy Studies (Drs Braveman end Bennett and Mr Showstack), School of Medicine (Ms Lewis). University of California. San Francisco.

The conclusions expressed in this report are those of the authors and do not reflect the views of the University of California or the W. K. Kellogg Foundation, Battle Creek. Mich. The Health Data and Statistics Branch of the California Department of Health Services was the original source of the vita statistics data. All analyses, interpretations, and conclusions are the responsibility of the authors. not the California Department of Health Services.

Reprint requests to Department of Family and Community Medicine, University of California, San Francisco, Box 0900. San Francisco. CA 94143 (Dr Braveman).

Reprinted from JAMA ® The Journal of the American Medical Association
March 10, 1993 Volume 269
Copyright 1993, American Medical Association

nancial access. Income is not recorded in US vital statistics, and, prior to 1989, only New York City and the state of Massachusetts included information on insurance in their birth certificate files. Therefore, it generally has not been possible to ascertain the independent role of financial access using vital records data; the only exception in the literature is a study by Cooney[12] using 1981 New York City birth certificates. Studying the utilization of prenatal care by women on Medi-Cal has depended on expensive linkages of vital statistics with Medicaid administrative data. There is limited literature describing the use of prenatal care by women with Medicaid coverage subsequent to recent major eligibility expansions up to 185% of the federal poverty level in many states.

In January 1989, California began to collect health insurance data in its birth files. We examined birth certificate data on all singleton live births occurring in-state to California residents during 1990 to study whether insurance coverage was a significant predictor of prenatal care utilization after controlling for multiple maternal sociodemographic risk factors. The availability of these data in the most populous state in the United States provided a unique opportunity to explore the role of financial access in predicting the adequacy of prenatal care among a demographically diverse population.

METHODS

Data Source and Sample

The source of data for this study was the public-use computer tape provided by the California Department of Health Services containing information from certificates of live biih for 1990. The sample consisted of all singleton live births to California residents occurring in California during calendar year 1990 (N=593 510).

Variables

Measures of Adequacy of Prenatal Care.-The dependent variables were quantitative measures of the adequacy of prenatal care. We studied the timeliness of initiation of prenatal care and the adequacy of the number of prenatal visits among women who received some prenatal care. Care was considered timely if it was initiated during the first trimester of pregnancy. Adequacy of the number of visits was determined using Kotelchuck's Index,[13] which controls for the timing of entry into care and the gestational age at delivery, giving a distinct measure of utilization independent Of timeliness of initiation. We categorized the number of visits as adequate if it met Kotelchuck's criteria for being "adequate" or "adequate plus"; we called the number of visits less than adequate if it met Kotelehuck's criteria for being "intermediate" or "inadequate." We separately studied the likelihood of receiving no prenatal care.

Insurance Coverage.-As of January 1, 1989, California birth certificates include confidential information on the principal source of payment for prenatal care and the expected principal source of payment for delivery. Based on information from the biih records, the following categories of insurance coverage were defined: uninsured (self-pay or medically indigent); coverage by Medi-Cal; private prepaid coverage, including the Northern California Kaiser system (Kaiser-North), the Southern California Kaiser system (Kaiser-South), and other prepaid plans as subcategories; and private fee-for-service insurance (by any private carrier). Other types of coverage (Medicare, worker's compensation, Title V, other government, other non-government, and no charge) were grouped together with unknown and missing information on coverage in descriptive analyses; this combined group was not studied in multivariate analyses. For some analyses, the uninsured and Medi-Cal groups were combined to form a category characterized as lacking private insurance. For women who received any prenatal care, prenatal care insurance coverage was determined by the principal payer for prenatal care. Women who had no prenatal care were&&lied by expected payer for delivery.

Maternal Characteristics.—Other independent variables obtained from the data source were maternal demographic characteristics (racial or ethnic group, birthplace, age, parity, education, and marital status). Based on the information in the biih certificate files, we defined six mutually exclusive groups according to racial or ethnic origin: (1) African American (used for all women whose race was recorded as black in birth tiles, regardless of Hispanic origin); (2) Asian American/Pacific Islander (race recorded as national origin in any Asian/Pacific Island nation, regardless of Hispanic origin); (3) European American (race recorded as white, and Hispanic origin recorded as "no"); (4) Latina (race recorded as white or other/missing/refused/unknown and Hispanic origin recorded as "yes"); (5) Native American (race recorded as American Indian, Eskimo, or Aleut, regardless of Hispanic origin); and (6) other/unknown (Hispanic origin not recorded as "yes," and race missing, recorded as refused or unknown, or not classifiable to the above categories). Mother's birthplace was included as an independent variable, classifying women as US-born (born in the 50 states or Washington, DC) or foreign-born (born elsewhere, including Guam, Puerto Rico, or the Virgin Islands).

Other maternal characteristics were classified according to established risks for perinatal outcomes. Four age groups were distinguished: 17 years and under, 18 and 19 years, 20 through 34 years, and 35 years and over. Categories defined for parity were no previous live births (primiparous), one to three previous live births, or four or more previous live births. Five levels of educational attainment were considered, measured by years of completed schooling. 0 to 9 years, 10 or 11 years, 12 years, 13 to 15 years, and 16 or more years. Women's marital status (unmarried or married) was imputed in California vital statistics from information on parents' names.

Statistical Analyses

Multiple logistic regression models (SAS program LOGIST[14]) were used to assess the likelihood of untimely care, a less-than-adequate number of visits, or no prenatal care, associated with insurance coverage and controlling for other maternal characteristics. For each variable, the group found to be at lowest risk for untimely initiation of care in descriptive analyses was selected as the reference group for constructing odds ratios. Adequacy of the number of visits was studied in the entire sample. Analyses of timeliness and no prenatal care excluded women with prepaid coverage who delivered at Kaiser-South hospitals because the unique Kaiser-South information system makes information on these variables not comparable to that from other hospitals.

When studying untimely initiation and less-than-adequate number of visits, births that were uninsured and those covered by Medi-Cal for prenatal care were examined as two separate insurance groups. Among Medi-Cal-funded deliveries of women who had no prenatal care, it was not possible to determine which women were uninsured while pregnant and which women had Medi-Cal coverage but did not receive care during pregnancy. For this reason, when studying "no prenatal care," the uninsured and Medi-Cal groups were combined and considered as a single group lacking private insurance.

Limitations

This study could not analyze content or quality of care, which, when combined with access, are known to be crucial determinants of birth outcomes.

Birth certificates do not contain adequate information on numerous factors that could influence the utilization of care, such as preexisting medical risks, perceived barriers to utilization, health behaviors, or knowledge, attitudes, or beliefs of individuals.

The recorded birth-weight data are believed to be more accurate than data on gestational age,[15,16] which are needed to calculate adequacy of the number of visits. For this reason, we imputed gestational age from birth weights for improbable gestational age values using methods developed by Williams et al.[17] We also recoded improbable maternal ages and implausible age/education and age/parity combinations as missing values. (A detailed description of the methods is available on request from the authors.)

The quality of data for new elements added or restored to California birth records in January 1989, including sources of payment for prenatal care and delivery, number of prenatal visits, and maternal education, has not been determined. Medi-Cal-covered births cannot be distinguished according to eligibility categories, and limitations of private insurance coverage (eg, copayments, deductibles, and exclusions) cannot be determined.

Changes in insurance coverage during pregnancy cannot be assessed, because only the principal payer for prenatal care is given and the duration of coverage by the principal payer is not specified. The most frequent change in coverage during pregnancy, especially given California's recent expansions of Medicaid maternity eligibility, is likely to be from uninsured status to Medi-Cal coverage. The distinction between Medi-Cal and uninsured is not always precise, since the group classified as Medi-Cal-covered for prenatal care contains many women who were uninsured during part of their pregnancies. In addition, births with the payer for prenatal care recorded as private prepaid include an unknown number of births to Medi-Cal-covered women enrolled in health maintenance organizations.

RESULTS

Table 1 displays maternal characteristics of the study population. Among women who gave birth who received some prenatal care, 10.7% had no insurance coverage for prenatal care, 32.8% were covered by Medi-Cal, and 50.5% had private, third-party coverage. Among women who had some prenatal care, 26.2% of women had untimely initiation of prenatal care and 27.9% had a less-than-adequate number of visits. Among the total sample, 2.1% of all women had no prenatal care (Table 2).

Uninsured status and Medi-Cal coverage were both significant risk factors for inadequate care that were of sizable magnitude even when controlling for maternal racial or ethnic group, birthplace, age, parity, education, and marital status (Table 3). In findings from the adjusted analyses in the entire sample, uninsured women were 2.54 times more likely (96% confidence interval [CI], 2.47 to 2.60) and women with Medi-Cal coverage were 3.33 times more likely (95% CI, 3.26 to 3.40) to have untimely care than women with private fee-for-service coverage.

Controlling for all other risk factors in the model, the risk of a less-than-adequate number of visits for the uninsured (odds ratio, 2.49; 95% CI, 2.44 to 2.55) was essentially the same as that group's risk of untimely care. For women covered by Medi-Cal, the risk of a less-than-adequate number of visits (odds ratio, 1.63; 95% CI, 1.60 to 1.66) was sign&ant, but its magnitude was about half that of the risk of untimely initiation in the same group.

Lack of private insurance (combined uninsured and Medi-Cal-covered) was a dramatically significant risk factor for receiving no prenatal care (odds ratio, 6.70; 95% CI, 6.00 to 7.47). Women in the Raiser-North system bad a far lower risk of no care (odds ratio, 0.49; 96% CI, 0.35 to 0.68) than those with private fee-for-service coverage statewide. However, risks of untimely care and a less-than-adequate number of visits were slightly elevated for women in each of the private prepaid groups compared with private fee-for-service coverage.

Racial or ethnic group, while generally statistically significant, did not appear to be one of the more important risk factors for inadequate care when insurance coverage and other characteristics were taken into account. Odds ratios for these groups relative to European Americans ranged from 1.11 to 1.42 for untimely care, from 0.91 to 1.21 for a less-than-adequate number of visits, and from 0.57 to 1.32 for no prenatal care; these odds ratios were of relatively low magnitude compared with those for several other risk factors in the model (Table 3).

Analyses of timeliness repeated for subgroups stratified by racial or ethnic group confirmed the role of insurance coverage., For each racial or ethnic group considered separately, both the uninsured and the Medi-Cal groups *were* at least twice as likely to receive untimely care as those with fee-for-service coverage. Results of multivariate analyses in the sample stratified by separate age groups and levels of education also confirmed that lack of private insurance was highly associated with untimely initiation. Insurance status was more important for some groups than for others; for example, the risk of untimely care associated with being uninsured or covered by Medi-Cal was significant but of small magnitude for women under 18 years of age. No changes were seen in the findings on timeliness from supplementary analyses of the entire sample with untimely care redefined as initiation after the fourth month of pregnancy. Results of subgroup analyses witb adequacy of number of visits as the outcome, stratified by race/ethnicity, age, and education, confirmed the role of lack of private insurance (both uninsured and Medi-Cal) as an independent risk factor

Table 1.—Maternal Characteristics for California Resident Singleton Live Births, 1990

Maternal Characteristic	No. (%)
Insurance status	
uninsured	63 708 (10.7)
Medi-Cal	194 928 (32.8)
Kaiser-North	30 834 6.21
Kaiser-South	32 356 (5.5)
Other private prepaid	83 098 j14.b)
Private fee-for-service	152888 (25.8)
Other/unknown	23 554 (4.0)
No prenatal care	12 166 (2.0)
Racial or ethnic group	
African American	46 818 (7.9)
Asian American/Pacific Islander	54 326 (9.2)
European American	247 702 (41.7)
Latina	239 292 (40.3)
Native American	2897 (0.5)
Other/unknown	2475 (0.4)
Birthplace	
US-born	350113 (59.0)
Foreign-born	243124 (41.0)
Unknown	263 (0.0)
Age group, y	
≤17	25 619 (4.3)
18-19	43 690 (7.4)
20-34	451677 (77.5)
≥35	62 309 (10.5)
Unknown	215 (0.0)
Previous live births	
0	239 697 (40.4)
1-3	320 089 (53.9)
a4	33251 (5.6)
Unknown	443 (0.1)
Education (completed schooling), y	
0-9	125597 (21.7)
10-11	71301 (12.0)
12	185 342 (31.2)
1515	114 602 (19.3)
≥16	89 489 (15.1)
Unknown	4181 (0.7)
Marital status	
Unmarried	186 819 (31.5)
Married	406 691 (55.5)
Total	593510 (100.0)

Table 2.—Prenatal Care for California Resident Singleton Live Births. 1990

Prenatal Cam	No. (%)
Initiation after third month*†	143 642 (26.2)
Less-than-adequate number cd visits*‡	161 941 (27.9)
No caret	11 957 (2.1)

*Excludes births with no prenatal care or missing information on prenatal care.
†Excludes Kaiser-South; their data were not comparable to those of the rest of the sample.
‡According to Kotelchuck's Index.

Table 3.—Adjusted Likelihood of Untimely Initiation, Less Than Adequate Number of Visits, or No Prenatal Care, by Maternal Characteristics, California Resident Singleton Live Births, 1990

Maternal Characteristics	Adjusted* Odds Ratio (95% Confidence Interval)		
	Untimely Initiation of Prenatal Care†‡	Less Than Adequate No. of Visits‡§	No Prenatal Care
Insurance status‖ (compared with private fee-for-service)			
Uninsured	2.54 (2.47-2.60)	2.49 (2.44-2.55)	6.70 (6.00-7.47)¶
Medi-Cal	3.33 (3.26-3.40)	1.63 (1.60-1.66)	
Kaiser-North	1.05 (1.01-1.09)	1.09 (1.06-1.13)	0.49 (0.35-0.68)
Kaiser-South#		1.18 (1.15-1.22)	.
Other private prepaid	1.14 (1.11-1.18)	1.07 (1.05-1.10)	0.91 (0.77-1.08)
Race/ethnicity (compared with European American)			
Africa" American	1.20 (1.17-1.23)	1.06 (1.03-1.09)	1.21 (1.12-1.31)
Asian American/Pacific Islander	1.11 (1.08-1.15)	0.91 (0.89-0.94)	0.57 (0.51-0.64)
Latina	1.32 (1.29-1.35)	1.19 (1.17-1.21)	0.98 (0.92-1.05)
Native American	1.42 (1.30-1.55)	1.21 (1.11-1.32)	1.32 (1.05-1.65)
Birthplace (compared with US-born)			
Foreign-born	1.29 (1.27-1.32)	1.33 (1.31-1.36)	1.37 (1.29-1.46)
Age group, y (compared with ≥35 y)			
≤17	2.14 (2.05-2.23)	1.32 (1.27-1.38)	2.03 (1.81-2.27)
is-19	1.37 (1.80-1.94)	1.49 (1.44-1.54)	2.05 (1.85-2.27)
20-34	1.29 (1.26-1.33)	1.34 (1.31-1.38)	1.47 (1.35-1.59)
Previous live births (compared with none)			
I-3	1.26 (1.24-1.28)	1.18 (1.16-1.20)	1.74 (1.66-1.83)
54	1.94 (1.M-2.00)	1.46 (1.42-1.50)	3.25 (3.06-3.52)
Education, y (compared with ≥16 y)	2.77 (2.68-2.87)	1.93 (1.88-1.98)	3.81 (3.21-4.52)
0-9			
10-11	2.57 (2.48-2.66)	1.59 (1.54-1.63)	2.74 (2.31-3.25)
12	2.23 (2.16230)	1.36 (1.33-1.39)	2.31 (1.95-2.74)
13-15	1.70 (1.64-1.75)	1.14 (1.12-1.17)	1.54 (1.29-1.85)
Marital status (compared with married)			
Unmarried	1.44 (1.42-1.46)	1.20 (1.18-1.21)	2.05 (1.97-2.14)

*Adjusted for maternal insurance status, race/ethnicity, birthplace, age, parity, education, and marital status.
†Care initiated after the third month of pregnancy.
‡Excludes births with no prenatal care or missing information on prenatal cars.
§According to Kotelchuck's Index.
‖Prenatal payer for those with some prenatal cars; delivery payer for those with no prenatal cars.
¶Many women with Medi-Cal at delivery were uninsured during pregnancy; thus, the uninsured and Medi-Cal subgroups were analyzed as a single group when assessing the outcome "no prenatal care."
#Data on untimely initiation of prenatal cars and no prenatal cars for Kaiser-South were not comparable to moss of the rest of the sample.

of relatively high magnitude compared with other risks in the model. Analyses of all three outcome measures using the entire sample of data from 1989 also produced consistent results.

COMMENT

Findings from this study demonstrate that financial barriers were an important determinant of whether or not a woman received adequate prenatal care in 1999 despite major Medicaid reforms in California. Nearly 11% of pregnant women were uninsured. The uninsured were at very high risk both of starting care late and of having a less-than-adequate number ofvisits, even when controlling for the mother's education, age, parity, marital status, birthplace, and racial or ethnic group.

It is not surprising that uninsured women enter care late and lack continuous care, given their multiple financial barriers. Most have limited financial means[18] and thus would be further burdened by out-of-pocket expenses for care. Working women without health insurance are also likely to lack "sick leave" benefits[19]; loss of pay for prenatal visits would thus be an additional disincentive to obtaining care, especially for the low-wage worker. Before the Medi-Cal eligibiity expansions, financial barriers for uninsured pregnant women were offset by the state-funded Community Based Perinatal Services program. Assuming that all those in need would be covered by Medi-Cal after the expansions, the state phased out the Community Based Perinatal Services program and in fact may have reduced access for women who remaineduninsured.

The findings also indicate that Medi-Cal coverage under conditions prevailing in 1990 did not necessarily confer financial access to prenatal care, especially early in pregnancy. In the entire sample, women with Medi-Cal coverage were the group most likely to initiate care after the first trimester, controlling for the maternal characteristics in the model. Obstacles known to delay care for women with Medi-Cal coverage Include bureaucratic problems in the application and certification process and a shortage of obstetric providers willing to accept Medi-Cal clients. These systemic barriers may have been exacerbated by the major eligibility expansions, which were not accompanied by commensurate expansions in system capacity. The federal option for presumptive eligibility was adopted in California in 1992 (Presumptive eligibility guarantees reimbursement to providers for prenatal care received by Medi-Cal-eligible women during the Medi-Cal certification process.); implementation should facilitate earlier initiation of care but is unlikely to resolve the problem of limited provider availability. It is important to note that, while women with Medi-Cal coverage had a significant risk of a less-than-adequate number of visits, that risk was generally higher among uninsured women. This suggests that Medi-Cal coverage does enhance the continuity of prenatal care once care is initiated and that women will take advantage of care if financial barriers are removed.

Haas and colleagues[20] studied changesin care and outcomes before and after the initiation of a program called Healthy Start, which gave prenatal coverage to uninsured women in Massachusetts up to 188% of the federal poverty level. Observing no improvement in disparities between all uninsured women and women with coverage, the authors concluded that the program had not demonstrated beneficial effects. However, they were unable to identify among the uninsured in 1984 the subgroup comparable to the Healthy Start population. Changes in risks of the uninsured overall along with deterioration in access for uninsured women not enrolled in Healthy Start could account for the apparent lack of effect. It is very likely that such changes (worsened poverty, unemployment, and substance abuse along with increased foreign immigration and increasing strains on providers to the uninsured) occurred during the study period and were not taken into account by the methods used. Moreover, prenatal care was called "satisfactory" if care began before the third trimester and there were five or more visits-not acceptable according to the standards of any industrialized country. This misclassification biased the study strongly against finding a difference.

Despite these biases, Haas et al[20] dismissed observed improvements in all parameters when comparing the 1987 Healthy Start and 1984 uninsured groups, because they might "merely reflect . . . enrollment [into Healthy Start] of healthier women who were more motivated to seek prenatal care." However, Healthy Start disproportionately enrolled higher-risk women. Our study, which was not hampered by some of the limitations of the project of Haas and colleagues, indicates that the improvements suggested by the Massachusetts data were real and that they support the role of financial access as a significant barrier. We agree with Haas et al that providing coverage alone may not be sufficient to overcome the multiple barriers faced by low-income women and that it is insufficient to provide access to women only during pregnancy.

It is unlikely that our findings of elevated risk associated with lack of private insurance coverage can be explained by differences in socioeconomic status independent of financial access barriers, because the other variables included in our multiple logistic regression models should have controlled for those differences to a large extent. Education alone has been widely used as a proxy for socioeconomic status; we also controlled for mother's racial or ethnic group, birthplace, age, parity, and marital status. Language, literacy, and important behavioral, knowledge, and attitudinal factors would be expected to be highly correlated with these control variables. Thus, the pattern of findings suggests that financial barriers were salient even when controlling for many individual factors related to care-seeking behavior.

Racial or ethnic group, often considered a key determinant of utilization of care, was relatively unimportant as a predictor compared with insurance coverage and other maternal characteristics. In addition to insurance, teenage motherhood, low educational attainment, and high parity were important risk factors. It is not surprising that insurance coverage was a significant but relatively weak risk factor for untimely care amongmothers 17 years of age or younger, a group at extremely high risk for psychosocial, educational, and economic reasons. Many factors in addition to financial access are known to be powerful predictors of utilization of preventive health services. Common sense suggests that economic barriers are likely to interact with personal and institutional obstacles to deter optimal utilization of prenatal services.

The importance of financial access might be seen by some as self-evident. However, given the current emphasis on the role of "life-style" factors in health and severe constraints on public spending, a widespread and convenient belief has emerged that places the blame for poor prenatal care on irresponsible behavior by pregnant women. This explanation downplays the importance of financial barriers that can be modified through public policy, while emphasizing the role of motivation and attitudes for which individuals are held responsible. Results of this study support the need for universal financial coverage to achieve adequate utilization of prenatal care for women in very diverse demographic groups. Blaming the women themselves diverts attention and resources away from effective solutions that would simultaneously address financial, systems, and personal obstacles to receiving high-quality care.

This research was an outgrowth of activities supported by a grant from the W. K. Kellogg Foundation, Battle Creek, Mich.

The authors acknowledge the contribution of Ronald Williams, PhD, who provided consultation on methods of analysis, and Susan Scheer, MPH, for assistance with research. We thank Linda Remy, PhD, for assistance with data management and Robert Lynch for assistance in preparing the manuscript.

References

1. Brown SS. Drawing women into prenatal care. *Fan Plann Perspect.* 1989;21:73-80.
2. Hughes D, Johnson K, Rosenbaum S, Liu J. The *Health of America's Children: Maternal and* Child *Health Data Book.* Washington, DC: Children's Defense Fund; 1988.
3. *Healthy People 2000: National Health Promotion and Disease Prevention Objectives.* Washington, DC: US Public Health Service; 1991. US Dept of Health and Human Services publication PHS 91-50212.
4. Committee to Study the Prevention of Low Birthweight. *Preventing Low Birthweight.* Washington, DC: National Academy Press; 1985.
5. Greenberg RS. The impact of prenatal care in different social groups. *Am J Obstet Gynecol.* 1983; 145:797-801.
5. Murray JL, Bernfeld M. The differential effect of prenatal care on the incidence of low birth weight among blacks and whites in a prepaid health care plan. N *Engl J* Med. 1988;319:1385-1391.
7. Committee to Study Outreach for Prenatal Care; Brown SS, ed. *Prenatal Care: Reaching Mothers, Reaching Infants.* Washington, DC: National Academy Press; 1988.
8. *Prenatal Care:* Medicaid *Recipients and Uninsured Women Obtain Insufficient Care. Washington, DC: Human Resources* Division, US General Accounting Office; 1997.
9. Schwethelm B, Margolis LH, Miller C, Smith S. Risk status and pregnancy outcome among Medicaid recipients. *Am J Prev* Med. 1989;5:157-163.
10. Piper JM, Ray WA, Griffin MR. Effects of Medicaid eligibility expansion on prenatal care and pregnancy outcome in Tennessee. *JAMA.* 1990;264:2219-2223.
11. Oberg CN, Lia-Hoagberg B, Skovholt C, Hodkinson E, Vanman R. Prenatal care use and health insurance status. *J Health Care Poor Underserved.* 1991;2:270-292.
12. Cooney JP. What determines the start of prenatal care? prenatal care, insurance, and education. *Med* care. 1985;23:986-997.
19. Kotelchuck M. The mis-measurement of prenatal care adequacy in the US and a proposed al-&native two-part index. Presented at the 116th Annual Meeting of the American Public Health Association's Annual Meeting; October 23, 1987; New Orleans, La.
14. *SAS User's Guide: Statistics, Version 5.* Cary, NC: SAS Institute Inc; 1985.
15. Jamison H, Birth *Certificate Data Quality: An Exploratory Study. Sacramento:* Health Demographics Unit, Center for Health Statistics, California Department of Health Services; unpublished.
19. David RJ. The quality and completeness of birthweight and gestational age data in computerized data files. Am *J Public Health.* 1980;70:964-973.
17. Williams RL, Creasy RK, Cunningham GC, Hawes WE, Norris FD, Tashiro M. Fetal growth and perinatal viability in California. *Obstet Gynecol.* 1982;59:624-632.
19. Brown ER, Valdez RB, Morgenstern H, Nourjah P, Hafner C. *Changes* in *Health Insurance Coverage of Californians,* 19794986. Los Angeles, Calif: UCLA School of Public Health; 1988.
19. Braveman P, Oliva G, Miller MG, Schaaf VM, Reiter R. Women without health insurance: links between access, poverty, ethnicity, and health. *West J Med.* 1988;149:708-711.
20. Haas JS, Udvarhelyi IS, Morris CN, Epstein AM. The effect of providing health coverage to poor uninsured pregnant women in Massachusetts. *JAMA.* 1993;269:87-91.

Printed and Published in the United States of America

COUNTY-WIDE SMALL AREA ANALYSES: ESTIMATION OF DENOMINATORS FOR USERS OF A COUNTY HEALTH CARE AGENCY AND THEIR APPLICATION IN HEALTH STATUS MEASURES

Arthur J. Davidson, Denver Health and Hospitals,
Mark P. Doescher, Gary Knapp, Cliff T. Foster, Richard A. Wright.

INTRODUCTION

Limited public and community health care resources result in implicit or explicit decision0 about which health intervention0 are available at a local level. While health promotion and disease prevention goals and objectives are often guided by national consensus, the process of setting priorities should be tempered by demonstrated need0 a0 defined at a national level as well as for each particular community. Decisions should be baaed on knowledge of the health status of the community served.

Health care planner0 often establish policy without the benefit of health status measures due to the complexity of ever expanding data resources. Synthesis and integration of public information such as vital statistics, disease registries and morbidity measures into health indicators for particular communities would assist in rational health care policy decision-making.

Definition of the service community or denominator permits estimation of the total population at risk and thereby assist in calculation of health statue measures or rates. Comparison of these measures with other or larger communities such as city, state or national rates may be useful in identification of health care problems among inhabitants of defined geographic commmunities.

As many other public health facilities or community health centers, we have noted a majority of our user population come0 from specific geographic areas. We have also noted that a one year count of unduplicated users is an underestimate of the actual number of people who would use the health facility. Over the course of multiple year0 many more people from those geographic areas may actually be counted as clients of our system.

This study was undertaken to test the association between residence in a census tract and being a registered user of Denver Health and Hospitals (DHH). This estimation of a denominator population is then applied to a computer aided geographic analysis for Denver County. Small area analyses are performed to compare health indicator0 between those areas associated with higher versus lower than expected Denver County health care facility utilization.

NRTRODS

Users are defined as registered Denver Health and Hospital0 (DHH) clients who may access care at an inpatient/emergency facility or any of nine satellite community health clinics. A large number of both Medicaid and medically indigent patient0 are serviced through DHH. Due to legislative mandate essentially all registered patients are Denver residents.

DATASOURCES

The primary data source for this study was an electronic database which included all outpatient clients who had a billable encounter during 1991. Information available included an address and demographic factors such as age, race/ethnicity and gender. A geographic information system (GIS; MAPINFO, Troy, NY) was used to geolocate clients to a census tract.

Secondary data sources included natality and mortality tapes from the Colorado Department of Health for 1990 and 1991 and a modified age-race/ethnicity-gender 1990 United States Census Bureau report for Denver County. Each of the vital statistios tapes contain0 information about the census tract of residence, age, race, ethnicity and gender. For the natality tape0 information is available on maternal age and level of education, month of onset of prenatal care, birth weight and clinical gestation. Mortality tape0 included information about underlying cause of death.

Census tract-specific penetration rates were calculated for each of the 181 census tract0 or 78 planning neighborhoods in the county of Denver. The penetration rate was defined as the number of 1991 DHH clients resident in the particular census tract divided by the number of census tract inhabitants as enumerated during the 1990 census. A chi-square teat was then performed comparing those rate0 versuO the null hypothesis ($p < 0.05$). Those census tracts with statistically higher rates are aggregated and then considered the primary zone of influence (PEOI) and those with less than expected values will be considered the secondary zone of influence (SZOI). The SZOI are those areas within the county where our health care agency is responsible for care yet has a lesser patient load when compared to an expected value.

Natality health status measures included fertility rates and onset of prenatal care rates while mortality measures included an analysis by diagnostic groupings using the ICD-9 coded underlying cause of death. Based on these analyses we calculated age-adjusted and age-specific natality and age/gender-adjusted mortality rates and compare the two zones of influence. Age-adjustment is performed by the direct method using the total 1990 Denver population as the standard. Stratification by race/ethnicity has been included for comparison. Relative risks and 95% confidence interval were calculated as suggested by Fleiss(1).

Age adjusted census tract or planning neighborhood-specific rates were then imported into a GIS to generate cartographic representation0 of the data.

RESULTS

In 1991 there were 105,000 ambulatory users of whom 92,818 were successfully geolocated. These clients made nearly 400,000 visit0 to Denver Health and Hospitals outpatient facilities. After penetration rates were calculated for each census tract there were 56 census tract0 or 32 planning neighborhoods found to have a significantly higher user penetration. Over 50,000 or 54% of the DHH client0 resided in these 32 neighborhoods. These clients constituted nearly 25% of the total population in

those neighborhoods. The remainder of our clients resided in the secondary service area or zone of influence. These 42,370 clients constituted 16% of the total population in the remaining 46 neighborhoods.

Figure 1 is a map of Denver with darker areas demonstrating higher user penetration and significant association between residency and DHH utilization. The primary zone of influence were the areas to the west, northwest and north east of the city and county.

NATALITY MEASURES

Table 1 demonstrates the results of the fertility analysis for women 15 to 44 years stratified by zone of influence. While the PZOI contains 42% of the fertile population, these women deliver 58% of the births for the entire county. Note that the crude and age adjusted rates for the primary ZOI are significantly higher than the secondary ZOI. Our service area has a large share of minority and indigent populations who have traditionally had higher fertility rates than the more affluent.and non-minority populations. Both fertility rate (per 1000 women) and late prenatal care rates (onset of care after the 1st trimester, per 100 births) were higher in the PZOI even after adjusting for age.

Among women 15 to 19 years of age, there were relatively high rates of fertility among the primary ZOI (134 vs 47) compared with the secondary ZOI. However, the difference in onset of prenatal care is essentially the same with much higher rates regardless of site Of residence.

Figure 2 shows fertility measures for women 15 to 44 year of age stratified by race/ethnicity. The bars represent the race/ethnicity comparison while the relative risk and 95% confidence intervals are tabulated in the box below. Primary ZOI is represented on the left of each bar pair. Minority populations continue to demonstrate higher fertility rates in the primary service area.

Figure 3 illustrates the rate of appropriate onset of prenatal care, (during the first trimester) where there were no signifcant differences for Afro-Americans or Hispanic women based on zone of influence. Caucasian women seem to have the best compliance with this health status measure. There was a significant difference for white women between the two zones with the secondary zone having the highest rate of all groups. The bar at 900 out of a 1000 births indicates the Healthy People 2000 objective.

Figure 4 demonstrates rates of no prenatal care with less discrepancy between the two service areas for minority women while rates *were* much higher for white women in the PZOI compared to white women living outside that area.

There were significant differences in age specific fertility for adolescent8 when stratified by race. Afro-American teens had high rates regardless of site of residence while the likelihood of pregnancy was nearly 5 times greater for a white teen living within the PZOI

Figure 5 demonstrates the neighborhood specific rates for fertility with much higher rates of fertility in the PZOI. *Figure 6* demonstrates onset of prenatal care during the first trimester with an inverse relationship where areas outside of the PZOI are darker indicating earlier onset of prenatal care.

MORTALITY MEASURES

Table 2 summarizes the mortality analysis. The marked disparity (PZOI being higher) for both crude and age-adjusted overall mortality measures based on ZOI is maintained throughout much of the disease specific or race stratified measures.

Figure 7 demonstrates the overall mortality rates for Denver County. Infant mortality rates (data not shown) by race showed no statistically significant difference. This is probably a reflection of low numbers of events for each race/ethnicity category during this 2 year time frame.

Disease specific underlying cause of death analyses demonstrated findings consistent with previously reported differences based on SES and race/ethnicity. Unintentional injury deaths (Figure 8) were nearly 2 and one half times higher in the PZOI where Afro-americans had a rate nearly 10 times higher than blacks living in the SZOI.

Rates of homocide (Figure 9) as a cause of death were markedly elevated in the PZOI with nearly a 3 fold increased relative risk for Afro-Americans. Among Hispanics the rates were elevated and relatively equal regardless of area of residence.

Rates of death due to diabetes, (Figure 10) were elevated in general for the PZOI with race/ethnicity differences. Rates for both Afro-Americans and Hispanics were much higher than whites within the PZOI. Although no inter-racial comparison8 are presented here the relative risk was signficant. On the other hand, rates of death due to diabetes were relatively equal across the race/ethnicity stratification in the secondary zone of influence.

LIMITATIONS

Worthy of commentary was the fact that about 14% of our users had addresses that were not geocoded - some were homeless, others had incorrect addresses in our computer file. The effect of these missing data is difficult to determine. *We* presume the estimate would remain essentially unchanged however a systematic error might result in the erroneous exclusion of an area from the PZOI.

Minority populations have asserted that US Census counts are less complete in lower SES when compared with higher SES areas. This would result in underestimates of denominators and thereby increase penetration rates and health status measures.

After the progression of a disease, persons who have severe disabilities may be forced economically to reside in lower SES areas consequently increasing mortality measures in the PZOI.

There may be a greater likelihood of persons who live in particular areas to have certain diagnoses placed as the cause of death or the omission of diagnoses in other areas. It seemed surprising that rates of pneumonia/influenza deaths (data not presented) were significantly lower for blacks and Hispanics in the PZOI yet higher for whites. Might there be a bias in recording underlying cause of death?

Future analyses should include the ACME coded cause of death for more complete case finding.

Some of these estimates are based on very small numbers of events. As reflected in some of the 95% confidence intervals, there may be great variation. Larger series or more years in subsequent analyses would be helpful. While we present age-adjusted rates in the maps, calculated confidence intervals for these rates were not presented. We have yet to develop a suitable method of demonstrating this variability.

CONCLUSIONS

In summary, we have combined a user based file with census information to calculate geographic regions associated with greater DHH client residence. The combination of data sets may provide a more accurate estimate of denominators for users of a county health care agency. These estimates permit us to suggest a method of representing health status measures using a geographic information system. Further research is indicated to validate this estimation process.

Complex and voluminous quantities of data may be presented in a more palatable format for politicians and decision and policy makers through the use of a geographic information syetem. This graphic representation may assist local health care providers begin to refine the measures useful in planning intervention and evaluation strategies for the Healthy People 2000 objectives.

This small area analysis demonstrated areas of increased need due to increased fertility, poor prenatal care rates, increased mortality and significant differentials based on race/ethnicity. Further efforts should develop analyses for cancer registries, morbidity registries and health care utilization datasets.

Limited resources may be more effectively used to target specific populations (based on race/ethnicity or SES) with proposed interventions around early prenatal care and teen pregnancy initiatives; alcohol abuse, violence and unintentional injury prevention as well as efforts to reduce mortality due to diabetes mellitus in the primary zone of influence.

1. Fleiss JL, (1981) *Statistical Methods for Rates and Proportions,* 2nd Ed. New York: John Wiley and Sons.

FIGURE 1. DENVER COUNTY - ZONES OF INFLUENCE

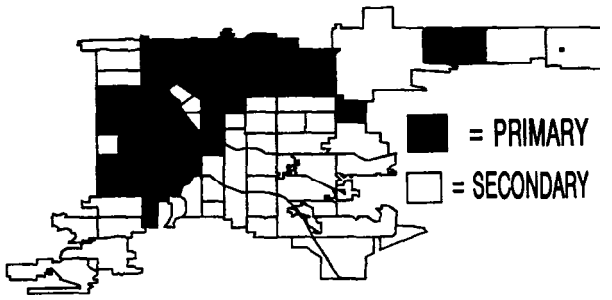
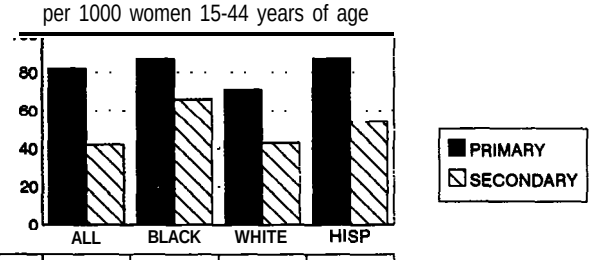


TABLE 1. NATALITY MEASURES, WOMEN 15-44 YEARS OLD 1990-91: Denver County by Zone of Influence (ZOI)

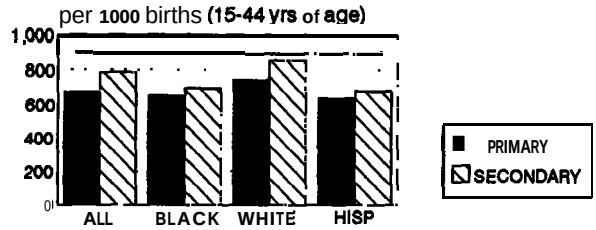
	DENVER	15 - 44 YRS PZOI	15 - 44 YRS SZOI	15 - 19 YRS PZOI	15 - 19 YRS SZOI
POPULATION	114,159	46,637	65,522	6,697	5,632
%	100	42	59	6	5
BIRTHS	7,254	4,197	3,057	925	272
%	100	58	42	13	4
FERTILITY RATE					
CRUDE	64	86	47	134	47
AGE-ADJUSTED		82	48		
LATE PRENATAL CARE					
CRUDE	26	34	17	42	43
AGE ADJUSTED		33	21		

FIGURE 2. AGE-ADJUSTED FERTILITY RATES 1990 - 91: Denver County by Zone of Influence



	ALL	BLACK	WHITE	HISP
RR	1.73	1.31	1.65	1.59
LOW	1.66	1.16	1.54	1.43
HIGH	1.82	1.49	1.76	1.78

FIGURE 3. AGE-ADJUSTED 1st TRIMESTER PRENATAL CARE RATES
1990-91: Denver County by Zone of Influence



	ALL	BLACK	WHITE	HISP
RR	0.85	0.94	0.86	0.94
LOW	0.84	0.88	0.84	0.9
HIGH	0.87	1.02	0.88	1

FIGURE 6. AGE-ADJUSTED 1st TRIMESTER PRENATAL CARE RATE
1990 - 91: Denver County by Neighborhooc

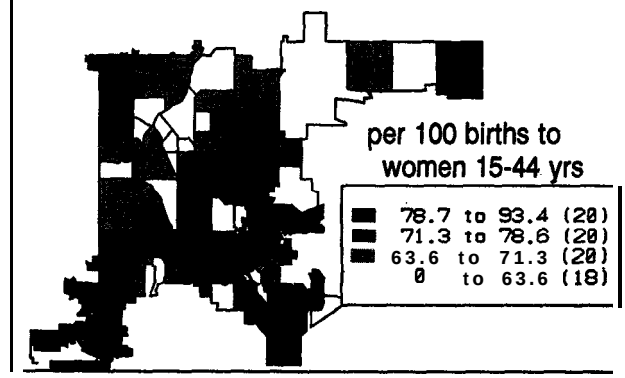


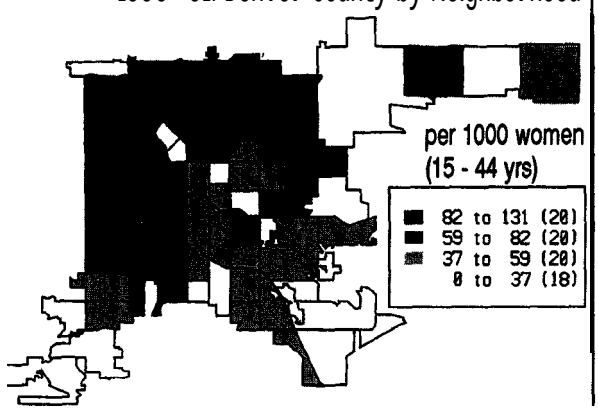
FIGURE 4. AGE-ADJUSTED NOPRENATAL CARE RATES
1990-91: Denver County by Zone of Influence

	ALL	BLACK	WHITE	HISP
RR	2.17	1.03	3.04	1.79
LOW	1.62	0.57	2.24	0.92
HIGH	2.91	1.91	5.97	3.63

TABLE 2. MORTALITY MEASURES
1990-91: Denver County by Zone of Influence

AREA	DENVER	PZOI	SZOI
POPULATION	467,610	206,322	261,288
%	100	44	56
DEATHS	3,426	1,777	1,649
%	100	51	49
AU CAUSE MORTALTIY RATE			
CRUDE	733	861	631
AGE-ADJUSTED		892	516

'IGURE 5. AGE-ADJUSTED FERTILITY RATES
1990 - 91: Denver County by Neighborhood



'FIGURE 7. AGE-ADJUSTED MORTALITY RATES
1990 - 91: Denver County by Neighborhooc

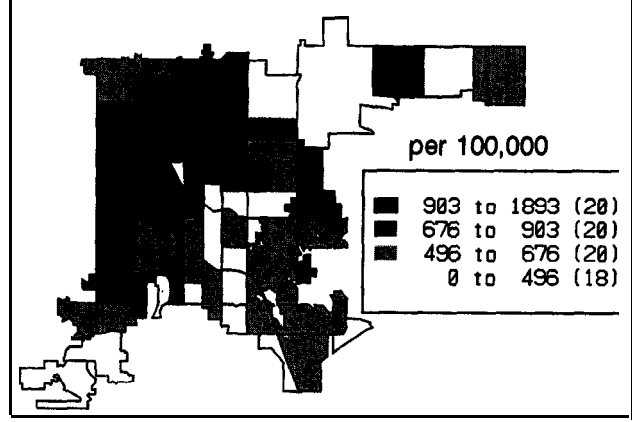
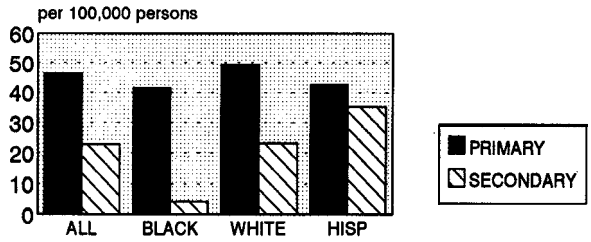
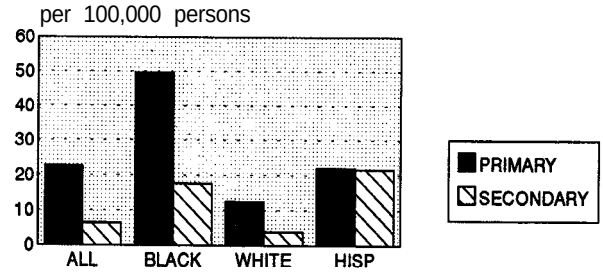


FIGURE 8. AGE-ADJUSTED UNINTENTIONAL INJURY MORTALITY RATES
1990-91: Denver County by Zone of Influence



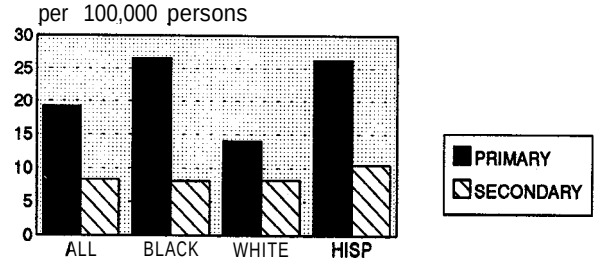
	ALL	BLACK	WHITE	HISP
RR	2.03	9.65	2.12	1.21
LOW	1.97	7.7	2.04	1.13
HIGH	2.09	12.1	2.22	1.3

FIGURE 9. AGE-ADJUSTED HOMOCIDAL MORTALITY RATES
1990-91: Denver County by Zone of Influence



	ALL	BLACK	WHITE	HISP
RR	3.59	2.81	3.17	1.02
LOW	3.39	2.51	2.89	0.93
HIGH	3.79	3.16	3.49	1.12

FIGURE 10. AGE-ADJUSTED DIABETES MORTALITY RATES
1990-91: Denver County bv Zone of Influence



	ALL	BLACK	WHITE	HISP
RR	2.31	3.27	1.71	2.5
LOW	2.19	2.76	1.58	2.21
HIGH	2.44	3.89	1.84	2.84

CI=0.4-1.4), and machinists (OR=1.0, 95% CI=0.5-1.7), occupations with high risk of suicide for white men, did not have elevated odds ratios. The other occupational groups with high odds for white men had fewer than lo deaths and were not analyzed.

In the analyses of women, also with a criterion of 10 deaths by suicide, there were four occupations at high risk for white and one for black women. In the health care industry, elevated risks were seen for white physicians (OR=5.69, 95% CI=1.2-27.1) and registered nurses of both races (white, OR=1.57, 95% CI=1.3-2.0; black, OR=2.7, 95% CI=l.l-6.5). In addition, elevated risks were seen for white women employed in real estate sales occupations (OR=2.45, 95% CI=1.4-4.2) and in the military (OR=infinite, 95% CI=7.1-infinite). Because of small sample sizes, odds ratios were not calculated for black female physicians, or those in real estate sales occupations or the military.

Breast cancer

A second NIOSH study looked at breast cancer among women in *23* states, 1979-1987. The purpose of this study was to assess whether occupational groups with excess breast cancer mortality could be identified and targeted for prevention efforts (13). The PMRs for breast cancer in white women were significantly elevated within three broad occupational groups: executives, including administrators and managers; professionals; and administrative support workers; including clerical personnel. Significant PMR elevations for black women were found in these same groups and also in sales and precision production occupations.

Breast cancer mortality for selected occupations within the broader occupational groupings are shown in Figure 1. Black women teachers had a statistically significant doubling of the expected proportion of deaths due to breast cancer. PMRs for black religious workers and librarians were non-significantly elevated more than 50%. For white women the PMRs for teachers, librarians, and religious workers were similar to those for black women. other occupations showed different risks for the two racial groups.

Cancer and other causes of death among farmers

NCI investigators examined cancer risks among farmers from 23 states for 1984-1988 (14). There were 119,648 deaths among white men and 11,446 among nonwhite men. Specific minority racial or ethnic groups were not examined. The results for the white and nonwhite men were fairly similar (Table 1). Excess risks for both groups were seen for cancers of the pancreas and prostate, vascular lesions of the central nervous system, asthma, and accidents. White men also had excess deaths from cancers of the lip, kidney, and brain. The major difference between the two racial groups was in the results for lymphatic and hematopoietic tissue cancers. Multiple myeloma, non-Hodgkin's lymphoma, and leukemia were significantly elevated for white men but were less than expected for the nonwhites.

Leukemia

In a fourth study, non-governmental researchers (Loomis and Savitz) used the NOMS 1985-1987 data from 16 states in a case-control study to assess occupational associations with leukemia mortality (15). The cases included 4739 white men and 390 black men. The age and race-adjusted odds ratios were computed and supported previously reported associations with employment in petroleum refining (OR=1.3, 95% CI=0.6-2.6, 8 cases) and rubber manufacturing industries (OR=1.3, 95% CI=1.0-1.8, 39 cases). Race-specific analyses of broad occupational groups found the risk for white professional men was elevated (OR=1.3, 95% CI=l.l-1.4, 403 cases), as was the risk for executives, administrators and managers (OR=1.2, 95% CI=l.l-1.4, 586 cases). Elevated risks were also found for black professional men (OR=1.8, 95% CI=1.2-2.7, 26 cases) and black mechanics and repairers (OR=1.8, 95% CI=l.l-2.9, 20 cases).

Discussion

Health disparities in minorities may be caused mainly by differences in social class, reflecting such factors such as income, education, and access to medical care (5). However, studies which have controlled for many of the factors for which race may serve as a surrogate still find unexplained racial differences in health outcomes (16). As mentioned earlier, data from the National Health Interview Survey and from the New Jersey Department of Health indicate that minority workers may experience greater exposure to hazardous work conditions than do white workers. In the studies described here, for some causes of death, the occupations exhibiting high risks for blacks were different from 'those for whites. For these reasons, we believe that it is important that results from occupational studies be reported as completely as possible for gender, race, and ethnicity.

Researchers attempting to study occupational disease in minorities can encounter several problems. Adequate sample sizes are difficult to obtain, since minorities represent a small proportion of the population: 12% of the U.S. population is black, 2.9% Asian, 0.8% American Indian, and 9% is of Hispanic origin. The quality and comparability of racial and ethnic information available for health research is uneven, with some data systems relying on self-reported data, which is sometimes inaccurate (17). For example, Hispanics sometimes choose "other" for race, rather than black or

white. Other systems use external assessment, such as interviewers' judgment or use of Hispanic surname. For death certificates the information is collected by funeral directors from the next of kin. And finally, the quality and completeness of the health outcome data and data on risk factors, including occupation, may vary among racial and ethnic groups.

The NOMS system has the potential to meet many criteria for an acceptable data source for studies of minority workers. For the most effective use, the NOMS system needs several enhancements and improvements. Currently, the NOMS system is acquiring data from only 21 states. Figure 2 shows the NOMS states and the states with the largest numbers of deaths for Blacks, Hispanics, American Indians, and Asians. Many of the states with large minority populations, such as California, Florida, Illinois, New York, and Texas, are not included in the NOMS system. To improve the ability of the NOMS system to identify occupationally-related health problems of minority workers, these states should be added.

Researchers have questioned the accuracy and completeness of the death certificate information for race and hispanic origin (17). NCHS has assessed the ethnicity data for Hispanics and in 1990 46 states and the District of Columbia submitted data that met its quality standard (18). The Hispanic data were at least 90% complete and sufficiently comparable for analysis.

NIOSH researchers made a rough assessment of the occupational data, which are also questioned for accuracy and completeness. We examined the percentage of deaths having an occupation code of not reported or retired. The death certificate requests the "usual" lifetime occupation, even if the person is retired at the time of death, so that retired is an invalid occupational response. Using this yardstick, we found the quality of the occupational data to vary somewhat by race and hispanic origin, although the percent was not unacceptably high (less than 5%) for any group (Table 2). In the 21 states that provided data in 1990, this ranged for males from 0.2% not reported and 0% retired in one state to 6.9% not reported and 0.3% retired in another. Some of the state differences might be due to special programs to educate funeral directors in the importance of collecting detailed and accurate industry and occupation information.

A study currently underway may help in better understanding the data and its usefulness for occupational health surveillance. NIOSH is collaborating in a Centers for Disease Control and Prevention survey of vital registrars and funeral home directors to determine the ways in which race, ethnicity, occupation, and other demographic information is collected on death certificates. We are examining the factors that influence quality and completeness. Recommendations for improved and standardized collection of these data will be developed.

Measuring disparities in health outcomes among racial and ethnic minority groups is a difficult task. It becomes more difficult when attempting to do so for occupational or industrial groups. No one source of data can provide sufficient information. The NOMS data can give a part of the picture. Analyses of the data may be used as a screening tool to identify associations of cause of death and occupation (or industry) that warrant further study, to evaluate the results of other studies, to target occupational groups for health promotion activities, or for descriptive purposes. If we gain a better understanding of the data, improve the quality of the data, and increase the number of contributing states with large minority populations, we will improve the utility of the NOMS system for occupational health surveillance in minorities.

References

1. Public Health Service. Healthy People 2000: National Health Promotion and Disease Prevention Objectives. DHHS Pub. No.(PHS) 91-50212. Washington, DC, 1991.
2. Friedman-Jimenez G. Occupational disease among minority workers. AAOHN Journal 37:64-70, 1989.
3. Park CH, Wagener DK, Winn DM, Pierce JP. Health conditions and other health-related characteristics among the currently employed, United States, 1988. National Center for Health Statistics. Vital Health Stat 10(186). 1993.
4. New Jersey Department of Health. Occupational Health Surveillance Update. Trenton, NJ. Winter 1993.
5. Williams DR. Race in the Health of America: Problems, issues, and directions. Presented at the Workshop on the Use of Race and Ethnicity in Public Health Surveillance, Centers for Disease Control and Prevention. Atlanta, GA, 1993.
6. NIOSH. NIOSH Occupational Disease, Injury and Hazard Surveillance Activities FY91. DHHS. Cincinnati, OH, 1993.
7. Guralnick L. Mortality by occupation and cause of death among men 20 to 64 years of age: United States, 1950. Vital Statistics-Special Reports, Vol 53, No.3. Washington: U.S. Government Printing Office. 1963a.
8. Reidmiller K, Doebbert G, Lashuay N, Rudolph L, and Glazer E. California Occupational Mortality 1979-81. Sacramento, CA: California Department of Health Services. 1987.

9. MacCubbin P, Herzfeld P, Therriault G. Mortality in New York State., 1980-1982: A Report by Occupation and Industry. New York State Department of Health Monograph No. 21. Albany, NY: New York State Department of Health. 1986.
10. Pennsylvania State Health Data Center. Mortality Experience of Pennsylvania Workers 1983-1985. Harrisburg, PA: Pennsylvania Department of Health. 1987.
11. Milham S. Occupational Mortality in Washington State 1950-1979. DHHS (NIOSH) Pub. No. 83-116. Cincinnati, OH: National Institute for Occupational Safety and Health. 1983.
12. Burnett CA, Boxer PA, Swanson NG. Suicide and occupation: Is there a relationship? Presented at the Annual Meeting of the Society for Epidemiologic Research. Keystone, Colorado. 1993.
13. Rubin CH, Burnett CA, Halperin WE, Seligman PJ. Occupation as a risk identifier for breast cancer. Am J Public Health. In press.
14. Blair A, Dosemeci M, Heineman EF. Cancer and other causes of death among male and female farmers from twenty-three states. Am J Ind Med 23:729-742, 1993.
15. Loomis DP, Savitz DA. Occupation and leukemia mortality among men in 16 states: 1985-1987. Am J Ind Med 19:509-521, 1991.
16. Polednak AP. Racial and ethnic differences in disease. New York: Oxford University Press, 1989.
17. Hahn RA. The state of Federal health statistics on racial and ethnic groups. JAMA 276:268-271, 1992.
18. National Center for Health Statistics. Advance report of the final mortality statistics, 1990. Monthly vital statistics report; vol 41 no 7, suppl. Hyattsville, Maryland: Public Health Service. 1993.

Figure 1. Breast Cancer Proportionate Mortality Ratios for Selected Occupations. Black and White Women: 23 States, 1979-1987

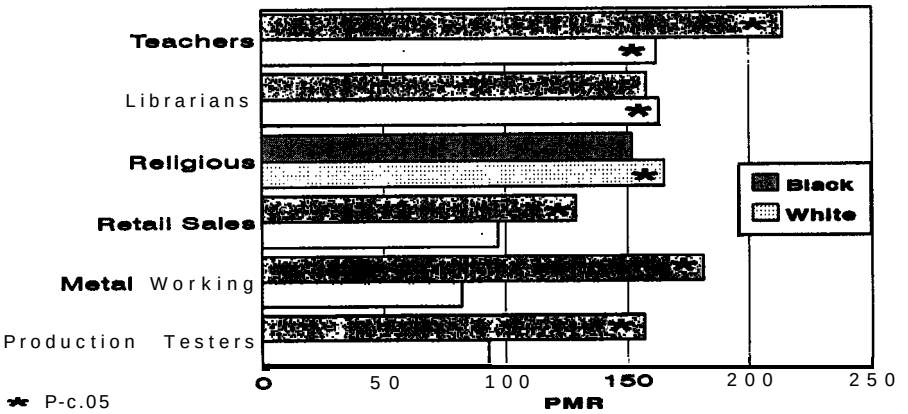


Figure 2. States with Large Numbeq ofBlack,Hispanic, American Indian, &Asian Deaths

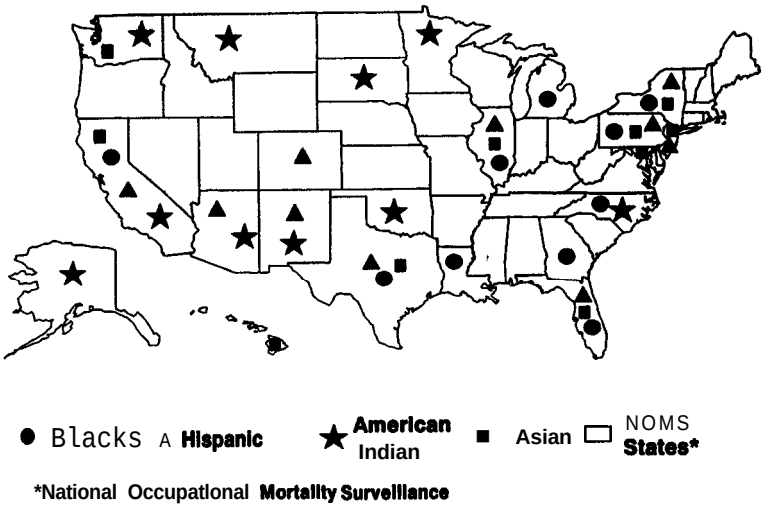


Table 1. Proportionate Mortality Among Male Farmers From 23 States, 1984-1988, by Race and Sex (PMRs for Non-Neoplastic Diseases and PCMRs for Specific Cancers)

CAUSE	WHITE MEN	NONWHITE MEN
Pancreas	1.13[a] (1133)[b] 1.07-1.20[c]	1.20 (125) 1.00-1.43
Lip	2.31 (21) 1.43-3.53	-- (0) 0 expected
Prostate	1.18 (3765) 1.14-1.22	1.14 (564) 1.05-1.24
Kidney	1.10 (552) 1.01-1.19	0.77 (30) 0.52-1.10
Brain	1.15 (447) 1.05-1.26	0.97 (16) 0.56-1.58
Non-Hodgkin's Lymphoma	1.21 (843) 1.13-1.30	0.71 (24) 0.46-1.06
Multiple Myeloma	1.15 (413) 1.04-1.27	0.92 (51) 0.69-1.22
Leukemia	1.27 (1072) 1.20-l-35	0.90 (55) 0.73-1.27
Vascular Lesions of Central Nervous Sys	1.15[d] (9874) 1.13-l-17	1.38 (1275) 1.30-1.46
Asthma	1.39 (171) 1.19-1.61	1.66 (36) 1.16-2.29
All Accidents	1.29 (3972) 1.25-1.33	1.34 (493) 1.23-1.47

[a] PCMRs for cancers
[b] Number of deaths
[c] 95% confidence interval
[d] PMR

Table 2. Percentage of "Not Reported" and "Retired" Occupations by Race or Ethnicity for Males, 1990 NOMS data

Race/ Ethnicity	Not Reported	Retired
American Indian	2.31	0.83
Asian	3.21	0.52
Black	3.76	0.48
White	1.83	0.39
Hispanic	3.21	0.82
Non-Hispanic	1.97	0.21

COGNITIVE ISSUES IN IMPROVING THE QUALITY OF MORTALITY STATISTICS

Jared B. Jobe, National Center for Health Statistics
Albert F. Smith, David J. Mingay, James A. Weed and Richard M. Roberts

Cause-of-death statistics in the United States are produced from death certificates by state and local vital registration offices (50 states, New York City, District of Columbia, Puerto Rico, Virgin Islands, and Guam) and by the National Center for Health Statistics. They are perhaps the most widely used national, state, and local data for identifying and dealing with the nation's health problems. Cause-of-death data are central and fundamental to measuring health status and evaluating the effectiveness of intervention and prevention efforts (Israel, Rosenberg, & Curtin, 1986; Sirken, Rosenberg, Chevarley, & Curtin, 1987).

Although each vital registration area in the United States has its own death certificate, the medical certification sections of the various death certificates are essentially identical. The certification section consists of two parts (see Figure 1). Part I contains either three or four lines, and the physician is instructed to "enter the diseases, injuries, or complications that caused the death." On the top line of the section, the physician is to enter the immediate cause, defined on the U.S. Standard Certificate of Death as the "final disease or condition resulting in death." On the remaining lines of Part I, the physician is to "sequentially list conditions, if any, leading to the immediate cause," entering the underlying cause on the *last used line. The* internationally accepted technical definition of the underlying cause is "the disease or injury which initiated the train of morbid events leading directly to death, or the circumstances of the accident or violence which produced the fatal injury" (World Health Organization, 1977, p. 700). In Part II of the certificate, the certifier is to list "other significant conditions contributing to death but not resulting in the underlying cause given in Part I." All of the information entered by the certifier on the death certificate is potentially valuable. However, by far the most important use of the death certificate for public health purposes is to extract the underlying cause of death from the information provided by the certifier. Although selection and modification rules applied by nosologists during coding of death certificates may override what physicians report on the last used line of Part I (World Health Organization, 1977, pp. 701-712), valid data on the underlying cause of death dependents on the ability and tendency of certifiers to report accurately the underlying cause of death.

The current format of the cause-of-death section of **the** death certificate is based on a particular set of assumptions about how a physician does (or should) make judgments about the cause of death so as to provide information that will result in selection of the "most appropriate" disease as the underlying cause of death. Specifically, the certificate was designed to encourage the physician to reason from the immediate cause back to the underlying cause. It instructs the physician to specify in Part II significant conditions that unfavorably influenced the course of the morbid process and thus contributed to the fatal outcome, but that were not related to the underlying cause given in Part I. The intent of Part II is to elicit from the physician statements about other conditions present in the patient, and to induce the physician to assert that these were not the underlying cause of death.

Although it has been assumed that physicians will identify most effectively the underlying cause of death by engaging in such a backward reasoning process and by explicitly excluding other alternatives, no empirical evidence supports these assumptions. For this reason, among others, there has been considerable concern about the quality and completeness of reported information on cause of death. Numerous methodological studies have raised serious questions about the accuracy of reports of underlying cause of death on death certificates (e.g., Percy, Stanek, & Gloeckler, 1981; Corwin, Wolf, Kannel, & McNamara, 1982).

This paper describes preliminary results from two components of a program of research designed to explore physicians' understanding of causes of death and of the death certification process. One component of the research involves interviews with physicians. These are intended to collect information concerning: 1) physicians' understanding of the concept of the cause or causes of death; 2) physicians' understanding of the validity of sequences of causes of death; 3) what training physicians receive, if any, in certifying the cause of death; 4) physicians' practices and attitudes about completing medical certification of the cause of death; and 5) underreporting of sensitive diseases or conditions, such as alcoholism, drug abuse, and AIDS. The second component of the research program consists of experiments intended to: 1) further elucidate our understanding of the cognitive processes of physicians as they complete death certificates; and 2) evaluate the effects of changes in the format of death certificates.

Interviews on Physicians' Knowledge of Death Certification

Method

One set of interviews (pilot interviews) was conducted with physicians in the mid-Atlantic region. The physicians were sampled by the vital registrars of Maryland, Virginia, and the District of Columbia from completed death certificates. A majority of the interviews was conducted face-to-face in physicians'

offices. A questionnaire was administered, asking physicians about their medical backgrounds, their experiences in certifying death, and their usual practices in death certification. Physicians were asked also about their knowledge of certification requirements and about their opinions regarding possible changes to the certificate. In addition, the interview protocol was designed to assess physicians' understanding of terminology used in death certification.

The physicians' comprehension of key terminology used in death certification is extremely important. If they do not understand the terminology, they may be more likely to make errors in completing certificates. Physicians were asked to define four terms: ***immediate cause of death*** (the final disease or condition resulting in death); ***underlying cause of death*** (the disease or injury that initiated the events resulting in death): ***other significant conditions*** (those conditions that contributed to death, but did not result in the underlying cause of death, that is, were NOT part of the causal sequence): and ***mode of dying*** (the mechanism by which death occurred).

Results

We report data for 16 interviews for which audible recordings were made. (Four other interviews were conducted, but not analyzed due to inaudible audiotapes.) The interviews averaged around 45 minutes in length. Most physicians were very cooperative and forthright in discussing death certification, although scheduling times for interviews was often difficult.

The mean age of the 16 respondents was 39 (range 28-65). Three were residents and two were fellows; of the remaining eleven, seven described their practices as primarily office-based, and four as primarily hospital-based. Internal medicine (one with geriatrics) was indicated by 6 respondents as the specialty in which they had spent the major portion of their working time in the past 12 months: family practice was indicated by 2; and each of the following specialties was indicated by one respondent: otolaryngology, hematology, hematology and oncology, oncology, general medicine, and urology (two respondents were not asked). Twelve respondents reported being certified by at least one specialty board. Excluding one respondent who was a deputy medical examiner as well as a primary care physician and who reported having certified about 250 deaths in the past 12 months, mean number of deaths certified in the past 12 months was reported as 16.2 (range from 0 to 40). Many respondents indicated considerable uncertainty about the actual numbers.

Several findings emerged from these interviews: First, physicians tended to answer questions using examples from their personal experience. Second, many described causes of death using terminology other than that used on the death certificate, such as primary or proximate cause of death instead of immediate cause, and secondary or fundamental cause of death instead of underlying cause. Physicians often described one condition as being secondary to another condition, instead of caused by or as a consequence of another condition.

Third, standard certification procedures vary from hospital to hospital. For example, at many hospitals, resident physicians and interns complete virtually all of the death certificates, whereas at other hospitals, attending physicians complete all certificates. At some hospitals, a hospital report of death is completed by the physician, and the hospital records office enters the information on the death certificate which the physician signs.

Fourth, physicians had trouble understanding the terminology used on the death certificate. These results are described below.

Mode of Dying and Immediate Cause of Death. All of the physicians demonstrated an understanding of at least one of the two terms, ***immediate cause of death*** and ***mode of dying.*** However, only 4 of the 16 respondents were able both to define the two terms and to demonstrate an understanding of the distinctions between them. For example, respondent S1 stated that, "The immediate cause is the actual diagnosis that describes why somebody expired, and it's not the mechanism by which they expired... I guess the term mode of dying would be the actual mechanical--or physical--event that happened.. . something like respiratory arrest or cardiogenic shock."

Six respondents demonstrated an understanding of the distinction between immediate cause and mode of dying, but did not define adequately one or both terms. For example, respondent M4 stated that the immediate cause was, "What actually caused the death.. . . the event that caused the patient's death; disregards any of the events that may have led up to the death.. . ;" and stated regarding the mode of dying, "I have no idea what this means. Could mean many things; what the actual cause was, whether the patient died by accident, or by natural causes, or by trauma."

Six respondents could not differentiate between the terms, but appeared to have some understanding of the meaning of immediate cause. For example, respondent S4 defined the immediate cause as, "The precipitating factor that took the patient's life at that particular time;" and for the mode of dying stated that, "I'm not sure what they mean by that--mode of dying--I have no idea."

Underlying Cause and Other Significant Conditions. Only 3 of 16 physicians demonstrated an understanding of the definitions of and distinctions between ***underlying cause of death*** and ***other significant conditions.*** For example, respondent M3 defined underlying cause as, "The disease process underlying the terminal event, for instance, in this man it is the arteriosclerotic heart disease... Underlying cause is a cause that does not require another to explain it;" and defined other significant conditions as, "Like a disease which may aggravate or worsen the underlying cause,

may make the patient more susceptible to underlying conditions, but which are not pathophysiologically a cause of or result of the underlying cause... ."

Five physicians demonstrated an understanding of the distinction between underlying cause and other significant conditions, but did not define adequately one or both terms. For example, respondent W1 stated that the underlying cause was, "The cause of the cause of death. Longitudinal diagnosis associated with the acute event., .;" and defined other significant conditions as, "...the diseases associated with patients' medical history that are contributory: they can alter the course of the underlying cause."

Four physicians did not differentiate between underlying cause and other significant conditions, but appeared to have some understanding of the meaning of one of the terms. For example, respondent S3 stated that the underlying cause was, "The condition or disease that led to or predisposed the patient to other comorbid conditions, including the one that may have caused their death;" and defined other significant conditions as, "Other comorbid conditions that were debilitating."

Two physicians did not demonstrate an understanding of the distinction between the terms. For example, respondent R1, when asked to define the underlying cause, stated that, "Underlying medical conditions are those that may have contributed to the patient's overall state of health or to have led up to contributing causes of death, but not necessarily to the principal illness or diagnosis." When asked to define other significant conditions, the respondent stated, "Nebulous. It depends on a multitude of diagnoses that the patient has. Could have many things, including chest pain and then expires.. ."

Three respondents gave definitions of underlying cause of death that were, clearly incorrect, and four gave incorrect definitions of other significant conditions.

Due to the small numbers of respondents and the use of responses to open-ended questions to infer physician knowledge, these findings should be regarded as preliminary. These interviews elicited considerable information about the attitudes and practices of physicians regarding death certification, as well as additional information about their knowledge of the task.

Using a revised version of the questionnaire, we plan to conduct a larger number of additional interviews that will permit a more definitive portrait of physicians as certifiers of death.

Evaluation of Certificate Format

As we indicated earlier, a primary goal of the experimental component of this project is to evaluate the effect of certificate format on statements about cause of death. The impact of certificate format on some aspects of certification practice may be revealed through natural experiments. For example, any relationship between the number of causes of death entered on the death certificate and number of lines in Part I may be examined by comparing recorded number of causes in states that use a 3-line cause-of-death section to that in states that use a 4-line section. However, other aspects of the comprehension and practice of certification may be evaluated only through controlled studies.

In an experiment that is currently underway, we are comparing the effectiveness with which three different certificate formats elicit reports of the correct underlying cause of death. Specifically, we are comparing reports of cause of death on the standard certificate to those on two experimental certificates. We call one of the experimental cause-of-death sections the reverse-order certificate, because the causal sequence requires the physician to enter the underlying cause of death on the top line and the immediate cause of death on the last-used line. Requiring the underlying cause of death on the top line may improve physicians' reporting of the underlying cause. Our interviews have indicated that many physicians think of the underlying cause first, especially for patients that they have treated for a chronic health condition. Moreover, with this format, physicians are less likely to be distracted by the conditions that immediately surround the death. We call the second experimental format the ***underlying-cause-only*** certificate because it has but one line that is to be used by the physician to enter the underlying cause of death. This certificate is intended as a more extreme way of focusing the certifier's attention on the underlying cause--the cause of paramount importance from a public health perspective.

In another experiment, we will compare reports of underlying cause on the standard certificate to those on a form that we call the ***checkbox*** certificate. On this form, the physician will be asked to check one of several modes of dying (e.g., cardiac arrest, respiratory arrest, exsanguination) prior to writing in the causes of death. Physicians often enter a mode of dying on the top line of Part I of the cause-of-death section. This study will test the hypothesis that better reports of underlying cause of death can be obtained if physicians state formally the mode of dying at the outset of the certification process.

In this section, we describe the methodology and preliminary results of the comparison of standard-order, reverse-order, and underlying-cause-only certificates. We anticipate that the general methodology for subsequent experiments will be similar.

Method

Materials were distributed at a board review course in family practice in Chicago. Each physician received an envelope containing a case vignette, one of the three cause-of-death sections, and a brief demographic questionnaire. The assignment of physicians to certificate conditions was random. Instructions were printed on the envelope. All physicians received the same vignette: the vignette was modified from one of the examples in the ***Physicians' Handbook on Medical Certification of Death*** (NCHS, 1987). Physicians were asked to read the vignette and to complete the enclosed cause-of-death section.

Results

We describe interim results from 76 subjects. The 76 physicians had been out of school for a median of 10.5 years. They reported completing a median of 2.5 death certificates in the past 12 months. Not surprisingly, 62 listed family practice as their specialty.

According to an expert pathologist (Lloyd Reich, personal communication, June 15, 1993), the correctly completed death certificate for this case should list non-insulin dependent diabetes mellitus as the underlying cause of death. In the standard-order condition, 14 of 25 physicians wrote diabetes mellitus as the underlying cause of death. In the reverse-order condition, 15 of 25 wrote diabetes mellitus. In the underlying-cause-only condition, only 6 of 26 wrote diabetes mellitus, whereas 13 wrote renal failure. These conditions dlffered significantly in the frequency with which dlabetes mellitus was reported as the underlying cause (X? (2) = 8.48, $p < .025$). The standard-order and reverse-order conditions did not differ in eliciting reports of diabetes mellitus as the underlying cause, but the rate of reporting of diabetes mellitus was significantly lower for the underlying-cause-only condition.

Because it would be imprudent to draw conclusions about certificate format on the basis of a single vignette, this experiment will be replicated using additional case histories.

Discussion

The research described in this paper investigated: 1) whether physicians understand the meanings of important terms on the death certificate; and 2) whether physicians' specifications of the underlying cause of death depend on certificate format. In this section, we summarize the most important interim findings of this study, and briefly discuss their implications.

Mode of Dying and Immediate Cause of Death

Perhaps the most striking finding is that, although the instructions on the cause-of-death section of the death certificate instruct the certifier not to enter the mode of dying as a cause of death, and provide examples of modes of dying (see Figure 1), only six respondents were able to correctly define mode of dying.

It is noteworthy, however, that comments made by three respondents who did not understand the term indicated a belief that they were not allowed to enter a mode of dying on the death certificate. This may reflect either memory for the certificate form or efforts by some state health departments and hospital records departments to correct physicians who make such entries as the sole condition in Part I of the cause-of-death section.

The term immediate cause of death was better understood than mode of dying. All sixteen respondents offered definitions of immediate cause of death, and none were clearly wrong. Nevertheless, the vagueness and ambiguity of many of the respondents' definitions suggest that most respondents had an incomplete understanding of the meaning of the term. Of course, even individuals as verbally skilled as physicians may have difficulty expressing their knowledge of a concept in words; thus many respondents may have a better functional understanding of ***immediate cause of death*** than their definitions would suggest.

Given the anecdotal evidence that physicians often enter a mode of dying on the top line of Part I of the cause-of-death section, we examined respondents' ability to discriminate between the terms mode ***of dying*** and ***immediate cause of death.*** Six of the 16 respondents failed to demonstrate any ability to discriminate between the terms. Should these results generalize to all physicians who certify deaths, numerous physicians may indeed enter erroneously a mode of dying rather than the immediate cause on the death certificate. This would suggest the need for educational efforts on how to complete death certificates.

Underlying Cause and Other Significant Conditions

As discussed earlier, the primary purpose of the death certificate is to obtain the physician's opinion about the decedent's underlying cause of death. Thus, it is troubling that three respondents defined incorrectly ***underlying cause of death,*** and only three respondents gave definitions that were fully adequate. It should be noted, however, that other respondents may well have been able to give correct definitions, including many of those who gave examples rather than definitions in the interview.

Examining respondents' definitions of ***underlying cause*** and ***other significant conditions*** revealed that only two respondents failed to demonstrate any ability to distinguish between these terms. This suggests that respondents were somewhat better at discriminating between underlying cause and other significant conditions than between ***immediate cause of death*** and ***mode of dying.***

In summary, the physicians interviewed appeared to vary greatly in their understanding of the cause-of-death concepts used on the death certificate. Many exhibited significant error or uncertainty in their understandings of the terms ***immediate*** and ***underlying cause of death, other significant conditions,*** and ***mode of dying.***

In addition, although no systematic attempt was made to measure within-subject reliability in defining the cause-of-death concept, there appeared to be substantial within-person inconsistency. An appreciable number of respondents contradicted themselves while defining one or more of the terms. This adds to the evidence that many respondents lacked well-articulated definitions of these terms.

Other Terminology

Another important finding is that the terms used by many physicians to refer to causes of death differ from those used on the death certificate. Physicians typically referred to primary or proximate cause instead of immediate cause of death and to secondary or fundamental cause instead of underlying causes of death

(e.g., acute respiratory failure secondary toemphysema). This may account in part for physicians' difficulty in defining the terms immediate cause of death and underlying cause of death.

Certification Procedures

Standard certification procedures vary from hospital to hospital, with residents and interns completing certificates at many teaching hospitals. In some cases, "cross covering" residents complete the death certificates for patients who die on evenings, nights, or weekends, even though they have not been involved centrally in the patient's care. This may increase the likelihood of erroneous reporting of the underlying cause of death, especially when these residents have had no training and little experience in death certification.

Format Experiment

The interim findings of the format experiment provide no support for the hypothesis that listing the underlying cause of death first improves the accuracy of reporting the underlying cause. The rate of reporting diabetes mellitus as the underlying cause on the reverse-order certificate did not differ from that on the standard-order certificate. It is possible, as our interviews indicated, that many physicians think first of the underlying cause of death for patients that they know had chronic health conditions. However, for many other physicians, especially those who are not the patient's primary physician, the immediate cause of death may be most salient. Thus, when multiple lines are available in the cause-of-death section, neither the standard- nor the reverse-order certificate may have a clear advantage.

Moreover, the underlying-cause-only certificate resulted in significantly less frequent reporting of diabetes mellitus. This may indicate that the lack of a causal sequence leaves physicians with less structure within which to decide on the underlying cause, increasing the chances that they will list the immediate cause or an intermediate cause on the certificate. This result may also reveal that physicians lack comprehension of the term ***underlying cause***. Given the less than clear understanding by 'physicians of the terminology demonstrated by our interviews, this interpretation appears to have considerable merit.

Liitatious

Three cautions should be stated about these interim results. First, pertaining to the interviews, the data reported in this paper were obtained from a small opportunity sample of physicians. Regarding the task of defining the cause-of-death terminology, we know of no reason why respondents would tend to perform more poorly on this task than would a random sample of certifiers, and indeed, can think of a number of reasons why these respondents should perform better. Nevertheless, we cannot know the extent to which these preliminary findings would generalize to all physicians who certify deaths. This could be ascertained only by means of a larger scale replication of this study using a more formally assembled sample of physicians.

Second, the classification of responses from the interviews involved an element of subjectivity. This is true of all analyses of responses to open-ended questions. We suspect, however, that the reported findings are robust against reasonable variations in the coding scheme.

Third, the results of the format experiment show that certificate formats differ in their effectiveness at eliciting diabetes mellitus as the reported underlying cause for the vignette we used. However, to evaluate the generality of this format effect, the study should be replicated with other vignettes. This project is currently underway.

Conclusions

The results of this project should indicate whether certain changes to the cause-of-death section of the standard death certificate would yield more accurate mortality statistics. In addition, evaluating physicians' knowledge of certification practices may suggest areas that should be emphasized in training physicians to complete death certificates.

References

Corwin, L.I., Wolf, P. A., Kannel, W.B., & McNamara, A.B. (1982). Accuracy of death certification of stroke: The Framingham study. ***Stroke,*** 13, 818-821.

Israel, R., Rosenberg, H., & Curtin, L. (1986). Analytic potential for multiple cause-of-death data. ***American Journal of Epidemiology, 124, 161-179.***

National Center for Health Statistics. (1987). ***Physicians' Handbook on Medical Certification of Death.*** Hyattsville, MD: U.S. Department of Health and Human Services.

Percy, C., Stanek, E., & Gloeckler, L. (1981). Accuracy of cancer death certificates and its effect on cancer mortality statistics. ***American Journal of Public Health, 71, 242-250.***

Sirken, M.G., Rosenberg, H.R., Chevarley, F.M., & Curtin, L.R. (1987). The quality of cause-of-death statistics. ***American Journal of Public Health, 77, 137-***139.

Smith, A. F., .Mingay, D. J., Jobe, J. B., Weed, J. A., & Clark, G. (1992). A cognitive approach to improving mortality statistics. ***Proceedings of the Section on Survey Methods Research, American Statistical Association,*** pp. 812-817.

World Health Organization. (1977). ***Manual of the International Statistical Classification of diseases, injuries and causes of death, Volume 1.*** Geneva, Switzerland.

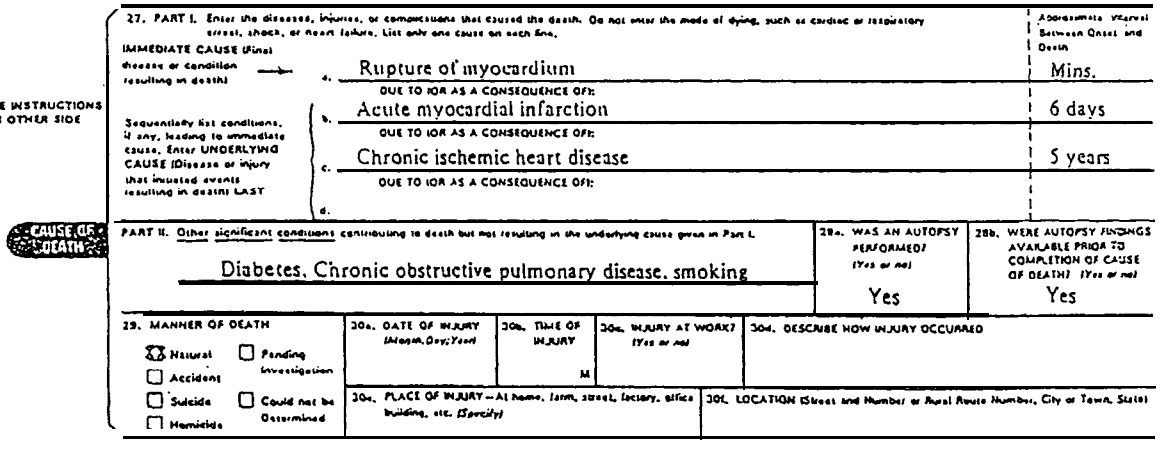
Author Notes

The research reported in this paper was supported by DHHS Contract 292-91-7031 between the National Center for Health Statistics, Centers for Disease Control and Prevention, and the Research Foundation of the State University of New York, and by a subcontract to the National Opinion Research Center, University of Chicago. Some of these results are described Smith, Mingay, **Jobe**, Weed, and Clark (1992).

Jared B. **Jobe** is a psychologist in the Office of Research and Methodology, National Center for Health Statistics (NCHS). Albert F. Smith is a psychologist at the State University of New York at Binghamton. David J. Mingay is a survey methodologist with the National Opinion Research Center, University of Chicago. James A. Weed is Deputy Director, Division of Vital Statistics, NCHS. Richard M. Roberts is a service fellow in the Office of Research and Methodology, NCHS.

We appreciate the valuable assistance of Harry Rosenberg and George Gay in conceptualizing this project. We thank Geneva Clark for assistance in collecting and analyzing data, and Mark Kloberdanz for assistance in collecting data. We thank Sheldon Waldstein, National Center for Advanced Medical Education, Chicago, and Findley W. Brown, Logan Square Medical Center, Chicago, for permission to conduct the experiment. We thank Monroe G. **Sirken** for constructive comments on an earlier version of this manuscript. We thank the staff of the vital statistics offices, Daniel Hughes (Maryland), Russell Booker and Kay Brooks (Virginia), and Carl Wilson (District of Columbia) for providing the names of physicians who certified deaths. Finally, we are deeply indebted to the physicians who gave freely of their valuable time.

Figure 1. Medical Certification Section of the U.S. Standard Certificate of Death


27. PART I. Enter the diseases, injuries, or complications that caused the death. Do not enter the mode of dying, such as cardiac or respiratory arrest, shock, or heart failure. List only one cause on each line.

Approximate Interval Between Onset and Death

IMMEDIATE CAUSE (Final disease or condition resulting in death) → a. Rupture of myocardium — Mins.

DUE TO (OR AS A CONSEQUENCE OF):

Sequentially list conditions, if any, leading to immediate cause. Enter UNDERLYING CAUSE (Disease or injury that initiated events resulting in death) LAST

b. Acute myocardial infarction — 6 days

DUE TO (OR AS A CONSEQUENCE OF):

c. Chronic ischemic heart disease — 5 years

DUE TO (OR AS A CONSEQUENCE OF):

d.

SEE INSTRUCTIONS ON OTHER SIDE

CAUSE OF DEATH

PART II. Other significant conditions contributing to death but not resulting in the underlying cause given in Part I.

Diabetes, Chronic obstructive pulmonary disease, smoking

28a. WAS AN AUTOPSY PERFORMED? (Yes or no): Yes

28b. WERE AUTOPSY FINDINGS AVAILABLE PRIOR TO COMPLETION OF CAUSE OF DEATH? (Yes or no): Yes

29. MANNER OF DEATH: ☒ Natural ☐ Accident ☐ Suicide ☐ Homicide ☐ Pending Investigation ☐ Could not be Determined

30a. DATE OF INJURY (Month, Day, Year)

30b. TIME OF INJURY M

30c. INJURY AT WORK? (Yes or no)

30d. DESCRIBE HOW INJURY OCCURRED

30e. PLACE OF INJURY – At home, farm, street, factory, office building, etc. (Specify)

30f. LOCATION (Street and Number or Rural Route Number, City or Town, State)

Session **M**

Numerator-Denominator Linkage

Toward the year

2000

Refining the measures

THE RELATIONSHIP BETWEEN MATERNAL HIV SEROPREVALENCE AND THE HOSPITALIZATION OF HIV INFECTED CHILDREN IN THEIR FIRST YEAR OF LIFE

Dr. Gerald Kaufman, New York State Health Department
Vito Logrillo, John Whitaker

INTRODUCTION

The care of HIV infected children is an area of public health concern in New York State (NYS). As of March 1993, 1,256 children under the age of 13 have been identified as infected with AIDS in NYS (1). Maternal, or vertical, transmission is the major route of HIV infection in the under 13 population, reported in 93% of the AIDS cases (1). For public health planning, it would be useful to know how many children annually will become HIV infected, when they will become sufficiently ill to require hospitalization, and the hospital resources these children utilize.

Information is available from'1988 to date regarding the number of HIV infected mothers who delivered liveborn children each year in NYS. This is determined by testing for maternal HIV antibodies in each NYS newborn. Approximately 7,330 HIV seropositive newborns were born in NYS from 1988 through 1991 as reported from NYS Seroprevalence data (2). Seropositivity rates have remained essentially constant at 0.6% for NYS, 1.2% in New York City (NYC) and 0.2% in the Rest of NYS from November 1987 through December 31, 1991. If the mother-to-child HIV vertical transmission rate was known, an estimate of HIV infected newborns via maternal transmission could be obtained. In studies performed in Rwanda (3), Zaire (4), Zambia (5) and the Congo (6), vertical transmission rates have been reported ranging from 25% to 49%. In studies performed in Europe (7), France (8) and the British Isles (9), lower vertical transmission rates have been reported, ranging from 14% to 19%. In the study of Glebatis, et al (10) of NYS children, it was concluded for births from 12/87 through 6/88, that 15% of the seropositive newborns were hospitalized with an AIDS related diagnosis during their first year of life. However, the calculation of a NYS specific vertical transmission rate is still lacking.

In this study we have attempted to estimate a minimum vertical transmission rate and the hospital resources utilized by HIV infected children during their first year of life. This was based on available information regarding children who were found to be HIV infected in the NYS hospital system during their first year of life. The data files utilized include seroprevalence data files, hospitalization abstract data files from the Statewide Planning and Research Cooperative System (SPARCS), and Vital Statistics matched birth-infant death files. The limitations of these data sources will also be discussed. Described below is the data preparation, methodology, and the results from the analysis.

DATA PREPARATION

Creation of a Longitudinal File Describing NYS Hospital Usage by HIV Infected Newborns During their First Year of Life.

All hospitalization records of children born during the period 1987-1992 with an indication of HIV infection in the discharge records were selected from the Statewide Planning and Research Cooperative System (SPARCS) and linked to form case specific hospitalization histories. The SPARCS collects an abstract describing each hospital stay in NYS except for stays in psychiatric and Federal Government hospitals. The SPARCS data file was considered over 98% complete for the time period studied and contained information regarding the patients date of birth, gender, race/ethnicity, hospital utilization, admit date, discharge date, medical record number, diagnoses, and discharge disposition. Also available for some admissions was information from the first matched bill describing the patients insurance codes and 'address. Previously longitudinalized SPARCS files which linked these records with case numbers for the years 1987 through 1990 were also utilized for linkage purposes. The SPARCS files do not contain any explicit patient identifiers such as name or social security number so linkages between records were indirectly determined.

The goal of the longitudinalization effort was to estimate the number of children who were born in a particular month and year and hospitalized within the first year of life with a diagnosis of HIV infection. To identify HIV infection, it was required that a principal or secondary diagnosis ICD-9-CM code of 042, 043 or 044 be present in any hospital admission within the first year of life. Children with stays where only seropositivity (ICD-9-CM code of 795.8) was present were not counted as HIV infected since this may be due only to maternal

anti-bodies during the first 15 months of life.

Cases were considered to be either from a distinct person (unique) or possibly from another person already represented by another case (uncertain). Records were categorized by unique dates of birth and gender and then formed into cases within each category. Records with different dates of birth or different genders were considered to be from different persons. Date of birth and gender are two variables which have been shown to be reported with high accuracy and reproducibility so they form the basis for determining cases. Records with the same date of birth, gender, medical record number and hospital of stay were considered to be from the same person (intra-hospital case formation), as well as records with the same case number and/or insurance company code (inter and intra- hospital linkage). If more than one case was determined to exist in the same category, uniqueness was established among the cases by the presence of multiple deaths, overlapping stays, or admissions in one case after death in a different case. These were the only criteria used to determine uniqueness of cases. An estimate of the number of HIV infected newborns hospitalized in the first year of life per month of birth was calculated by the number of unique cases plus one half of the uncertain cases found for the children born in each month. This is midway between our minimum and maximum estimated number of cases.

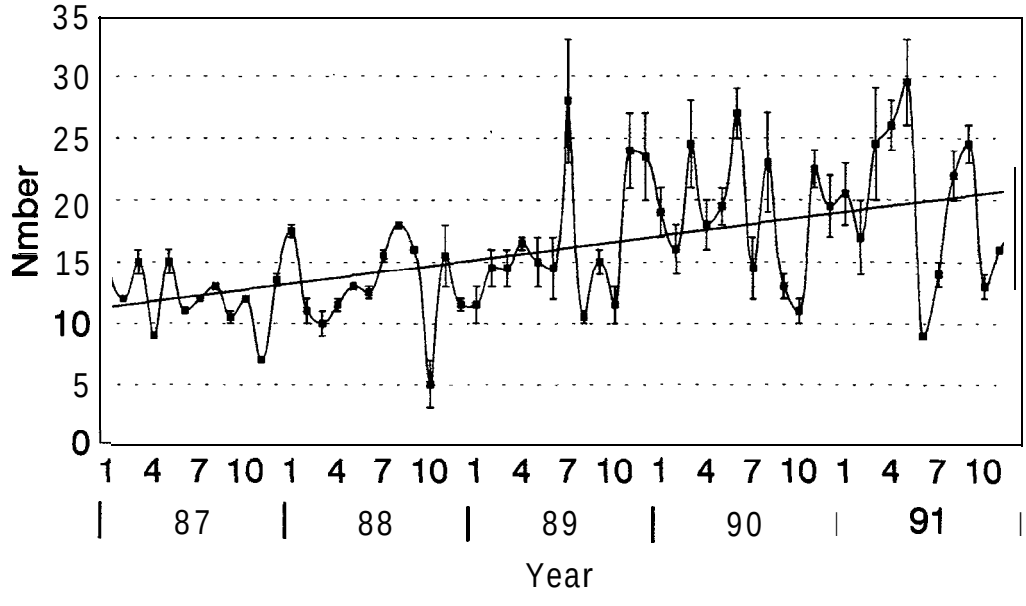
Figure 1 summarizes the results of the longitudinal file creation process; it contains the estimate of the number of HIV infected newborns hospitalized by birth month and birth year from 1987 through 1991. Table 1 displays the demographics of the cases.

METHODOLOGY

The Minimum Percentage of HIV Seropositive Newborns That Become HIV Infected During the First Year of Life.

To estimate the minimum percentage of HIV infected children that become HIV infected during the first year of life, a comparison of the number of children identified with HIV infection in the NYS

Figure 1
Number of HIV Infected* Children Hospitalized Within the First Year of Life By Month of Birth



***HIV Infected = the presence of a 042.X-044.X ICD-O-CM code as a principal or secondary diagnosis in any of the hospitalizations during the first year of life.**

Table 1
Demographics Of All Pediatric HIV Cases Hospitalized Within Their First Year Of Life
1987-1991 (1,048 Cases)

Race	Percent Of Cases
Black	46.9
Hispanic	34.8
White	11.2
Other	6.6
Asian	0.2
American Indian	0

Gender	Percent Of Cases
Male	53.1
Female	46.9

Region	Percent Of Cases
New York City	89.0
Upstate	11.0

hospital system to those born seropositive was performed for matched birth month cohorts from 1987 through 1991. Seropositive case counts were obtained from the NYS seroprevalence file and HIV infected hospitalized case counts were obtained from longitudinalized SPARCS file data (See Data Preparation Section). This ratio of the case counts results in a minimum estimate of the vertical transmission rate as HIV infected children are not necessarily all hospitalized during their first year of life.

For this analysis to be valid the HIV infected children identified in the NYS hospitals would have to been born in New York State and therefore be tested in the seroprevalence study. To validate the accuracy of this assumption, the birth location of HIV infected children born in 1988 who died in the hospital within the first year of life was obtained by matching these deaths to the NYS birth-infant death file maintained by the Vital Statistics Unit of the NYS DOH. This file contains information from the birth and death certificates of children who died within the first year of life in NYS. Of the 45 children who were born in 1988, identified from the SPARCS as HIV infected and died under the age of 1, 91% (41) were found in the NYS birth-infant death file. All of these 41 children were born in NYS. This estimated migration rate is conservative as there was insufficient information to link some children who died to their birth certificate records and obtain a place of birth. We consider this result sufficient justification of the assumption of little migration of HIV infected children under the age of 1 to NYS. If there is significant migration of HIV infected newborns from NYS, this would result in the estimate of the vertical transmission rate being more conservative.

Hospitalization Utilization During the First Year of Life

For each year, 1987 through 1991, the number of stays, number of cases, number of stays per case, days spent in the hospital per case, total days spent in the hospital per year and the average length of stay per admission were calculated for all HIV infection cases.

RESULTS

Minimum Vertical Transmission Rate

Figure 2 describes our findings of the relation of HIV Infected newborns hospitalized within the first year of life to Maternal HIV seroprevalence by matched month of birth. The data indicate an increase in the percent of seropositive newborns hospitalized from 1/88 to 12/91. Table 2 presents these data on an annual basis. The minimum vertical transmission rate appears to have changed from 9% in 1988 to 12.9% by 1991.

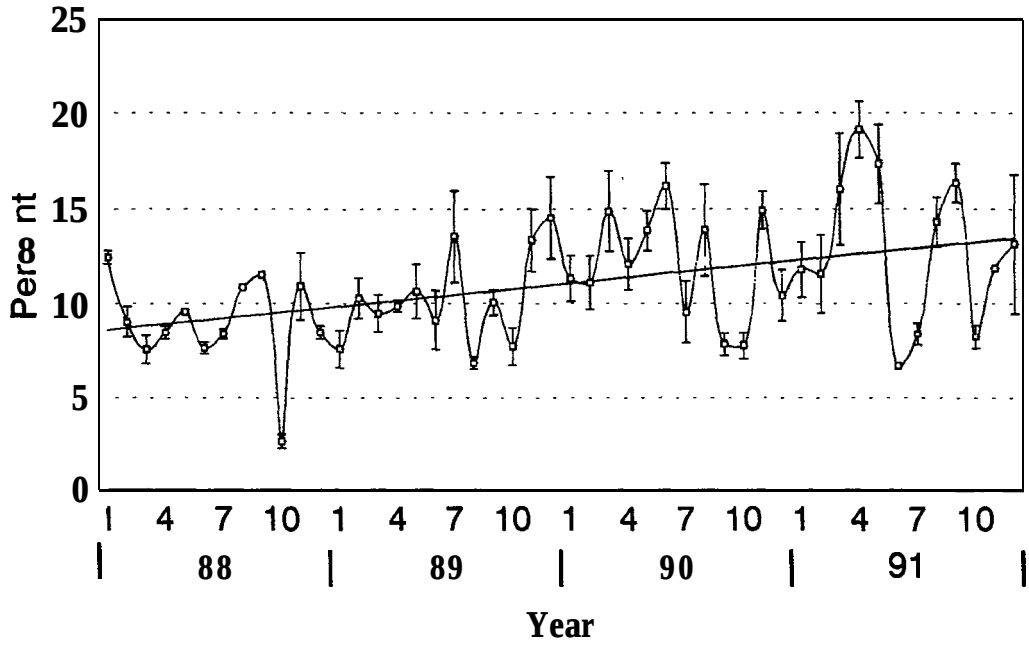
Hospital Utilization During the First Year of Life

The hospital utilization characteristics over time are shown in Table 3. The median length of stay per admission has dropped from 13 days in 1987 to 7 days in 1991. The number of children hospitalized has increased each year during this period. The net result is that the total days spent in the hospital does not show any significant trend.

DISCUSSION

The results from the analysis indicate that the vertical transmission rate has increased. The most likely reason for this phenomenon is that the identification of HIV infected children has become more complete over time. A previous study by the NYS AIDS Epidemiology Program (11) found that of the 042-044, 795.8 codes found in SPARCS, 94% were on the medical face sheet and of the primary discharge codes 98% were so listed on the face sheet. Of the AIDS cases reported to the NYSDOH, 81% were identified in SPARCS (559 of 687 in 1987). However, children under the age of 1 have different problems in terms of accurate identification of HIV infection and a study of the completeness and accuracy of HIV coding in SPARCS of this group is needed. However, it is also possible that the vertical transmission rate has changed due to other factors e.g. changes in breast feeding patterns.

Figure 2
Ratio of HIV Infected* Children Hospitalized Within the First Year of Life to Maternal HIV Seroprevalence By Matched Month of Birth



*HIV Infected = the presence of a 042.X-044.X ICD-9-CM code as a principal or secondary diagnosis in any of the hospitalizations during the first year of life.

Table 2
HIV Infected Children Under The Aae Of One In New York State

Year Of Birth	Number Of Live Births	Sero-positive Newborns	Expected' Number of HIV Infected Children	HIV Infected Children Hospitalized Under The Age Of One	Ratio Of Observed To Expected
1987	271,355	•	•	146±4	•
1988	279,976	1732	272	156±8	57.35%
1989	290,528	1884	295	199±24	67.46%
1990	297,468	1898	298	228±26	76.51%
1991	292,523	1816	285	234±26	82.11%

1- Estimated number of HIV infected cases based on 15.7% vertical transmission rate.
• -Not Collected.

Table 3
New York State Hospital Utilization By HIV Infected Children During Their First Year Of Life

Year Of Birth	HIV Cases	Stays Total	Avg/ Case	Median Time To First Stay	Total Days	ALOS Per Record	Median LOS Per Record	Average Days Per Case
1987	146±4	220	1.47	146.0	6877	31.26	13.0	47.10
1988	156±8	243	1.56	181.0	5518	22.70	10.0	35.37
1989	199±24	350	1.76	143.0	8175	23.36	11.0	41.08
1990	228±26	411	1.80	135.0	7389	17.98	9.0	32.40
1991	234±26	387	1.66	150.0	6499	16.79	7.0	27.77

The proportion of the HIV infected children who were hospitalized during the first year of life is still unknown and our present estimate of a minimum vertical transmission rate of 12.9% is a conservative estimate. It is interesting to note that it is not inconsistent with the initial estimate of 15.7% from the other Studies conducted in industrialized countries. Current plans are to continue the study to include hospitalizations beyond the first year, if the necessary migration and coding study results support this extension. Other data related issues which precluded more recent results include lags in reporting to the SPARCS and lags in updating the birth-infant death file.

The hospitalization utilization studies should prove useful for estimating the burden on the hospital system due to the HIV infected children. The utilization trends will be monitored to determine any changes due to improving care modalities or differing opportunistic infection rates.

It was also observed that there was a lack of information regarding HIV status of the mother and the child in the birth certificate files, and inconsistent reporting of in-hospital mortalities between the SPARCS and the death certificate files. These issues have been raised with the Vital statistics Unit of the NYS Department of Health and are being investigated.

ACKNOWLEDGEMENTS

we would like to acknowledge the assistance of the Bureau of Program Development and the AIDS Epidemiology Program for helpful discussions and the Information Systems and Health Statistics Unit for assistance in the migration analysis.

REFERENCES

1. New York State Department of Health: Bureau of Communicable Disease Control. AIDS Surveillance Quarterly Update. March 1993.

2. New York State Department of Health: AIDS in New York State. 1991

3. Ryder RW. Nsa W. Hassig SE et al. Perinatal transmission of the human immunodeficiency virus type 1 to infants of seropositive women in Zaire. N Engl J Med 1989:320:1637-42.

4. AIDS Task Force (European Economic Community) Workshop on mother-to-child transmission of HIV. Ghent. Belgium. February 17-20 1992. Final report. Brussels: European Economic Community-AIDS Task Force. 1992.

5. Hira SK. Kamanga J. Bhat GJ et al. Perinatal transmission of the HIV-l in Zambia. Br Med J 1989:299:1250-2.

6. Lallemant M. Lallemant-Lecoeur S. Cheynier D. et al. Mother-to-child transmission of HIV-l and infant survival in Brazzaville, Congo. AIDS 1989:3:643-6.

7. Lepage P. Van De Perre P. Mselllati P. et al. Mother-to-child transmission of human immunodeficiency virus Type 1 (HIV-l) and its determinants: a cohort study in Kigali, Rwanda. Am J Epi 1993:137:589-99.

8. European Collaborative Study. Risk factors for mother-to-child transmission of HIV-l. Lancet 1989:1:941-3.

9. Ades AE. Davison CF. Holland FJ. et al. Vertically transmitted HIV infection in the British Isles. Br Med J 1993:306:1296-99.

10. Glebatis DM. Novick LF. Stacy A. Hospitalization of HIV seropositive newborns with AIDS-related disease within the first year of life. Am J Public Health 1991:81(Suppl):46-9.

11. Muse A. Smith PF. Mikl J. The quality of computerized hospital discharge data. Am J Public Health 1992:82:1044.

SURGICAL PROCEDURE RATES WITH NUMERATORS FROM HOSPITAL DATA AND DENOMINATORS FROM SEVERAL DIFFERENT "AT RISK" POPULATIONS

Jane F. Gentleman, Statistics Canada
Greg F. Parsons, Michael N. Walsh

Ongoing at the Canadian Centre for Health Information at Statistics Canada is a study of surgical procedure rates across Canada, by Census Division. A Census Division (CD) is a geographical region similar to a county. There were 266 CDs in all of Canada in the years covered by this study. Procedure counts are derived from national hospital discharge data at Statistics Canada. In all, about 70 surgical procedures are being analyzed. For further information about the data and methodology, see Gentleman, Wilkins, Nair, and Beaulieu (1991a and 1991b).

The rate for a particular surgical procedure is calculated as the number of surgical procedures divided by the population at risk of undergoing that surgical procedure. As examples, this paper discusses two procedures: cesarean section and lung surgery. For each of these, two or more different denominators, representing different definitions of the population at risk, are used to calculate procedure rates. Not only is each denominator from a different data source than the numerator (which was the subject of the conference session at which this paper was presented), but the sources of the denominators are different from each other.

Procedures were coded using the Canadian Classification of Procedures (CCP). Procedure rates were calculated based on the residence of the patient, not on the location of the hospital. The rates presented here are age-sex-standardized, using as the standard population the combined area for all CDs used in the analysis. Different rates are based on different numbers of CDs; CDs were excluded from the analysis if their populations were smaller than 5,000, and/or for other reasons given below. Nineteen age groups were used: <1, 1-4, 5-9, 10-14,..., 85+.

In further analyses not discussed in detail here, age-sex&standardized rates are *normalized* to have mean zero and variance one under the hypothesis that all CDs have the same procedure rates, to adjust for differences in CD population sizes.

Characteristics of Numerators and Denominators

Table 1 describes the numerator and three denominators used to calculate three versions of rates of cesarean section. Six characteristics of the numerator and of each denominator are summarized in the table: the event or persons counted, the source of the data, the date(s) covered by the source, the date(s) used in the rate calculation, the population covered by the source, and the population used in the rate calculation.

For cesarean section rates, the numerator was the count of cesarean sections (CCP = 86.0-86.2 and 86.8-86.9) obtained from hospital discharge data. Denominator 1 was the number of females of child-bearing age (15-49 years), obtained from post-censal estimates. (Canada conducts its Census of Population every five years, in years ending in 1 and 6.) Denominator 2 (obtained from national vital statistics birth files) narrows the definition of population at risk to those females aged 15-49 who actually gave birth to a live child. Denominator 3 slightly widens the population at risk to include not just live births. but also stillbirths (from vital statistics stillbirth files).

Statistics Canada's hospital data are collected on a fiscal year basis. To smooth the procedure rates somewhat, two fiscal years (1988-89 and 1989-90) were used for the numerator. (Rates were appropriately calculated so that they could be interpreted as annual rates.) Denominator 1 for cesarean sections was an estimate at a single point in time, close to the midpoint of those two fiscal years. Denominators 2 and 3 were compiled from appropriate portions of birth and stillbirth data (which are compiled on a calendar year basis) for 1988-1990. With Denominator 1, the rate was calculated as the number of procedures over two fiscal years divided by twice the population count on June 1, 1989; with Denominators 2 and 3, the rate was the total number of procedures over two fiscal years divided by the respective population count over two fiscal years. With Denominator 1, procedure rates were calculated for CDs in nine provinces; Prince Edward Island, which is Canada's smallest province, was not included because of its small CD sizes and low procedure rates. With Denominator 2, Newfoundland was also excluded because (until recently), the mother's age was not captured in vital statistics data received from Newfoundland.

Note that hospital discharge data are case-based rather than person-based; that is, the numerator was the number of cesarean sections rather than the number of different females receiving cesarean sections during the reference period. This tends to inflate the numerator

slightly relative to a population of individual women at risk.

Table 2 describes the numerator and two denominators used to calculate two versions of rates for lung surgery. ("Lung surgery" includes partial and complete lung lobectomy and pneumonectomy: CCP = 44.3-44.5.) In this example, lung surgery rates were calculated for males only. (Lung cancer rates are appreciably higher for males than females.) Denominator 1 was the number of males, obtained from post-censal estimates. Denominator 2 (obtained from national cancer registry files) narrows the definition of population at risk to males who were diagnosed with a new primary case of lung cancer. This is actually an approximation insofar as not all lung surgery is performed for treatment of lung cancer, and because cancer registry data (before 1992) are case-based.

Continuing the concept of narrowing the population at risk, an additional denominator will be used in a future analysis: it will count new cases of *non-small cell* lung cancer (which, as opposed to small cell lung cancer, is treatable by surgery). These data are readily available from cancer registry files, which record the morphology (cell type) of the lung cancer. It would also be desirable to use as a denominator counts of a subset of non-small cell lung cancers, i.e., those diagnosed at earlier stages of the disease (stages I or II), which are usually treated by surgery, as opposed to those diagnosed at later stages (III or IV), which are usually too far advanced to be treated by surgery. However, the national cancer registry does not (yet) capture staging information, so this fourth denominator is not available at this time.

With Denominator 2, lung cancer surgery rates were calculated as the number of procedures over two fiscal years divided by twice the total number of lung cancer cases in calendar year 1989. More recent cancer registry data were not yet available for use in Denominator 2 at the time the analysis was conducted.

As was the case for cesarean section rates, lung surgery rates using Denominator 1 were calculated for CDs in 9 provinces (omitting Prince Edward Island). With Denominator 2, lung surgery rates were only calculated for 6 provinces: New Brunswick, Ontario, and British Columbia were omitted because the required data were not yet available at the time of the analysis.

Selected Results

Figure 1 is a stem-and-leaf plot of cesarean section rates in 255 CDs, calculated using Denominator 1 as described in Table 1.

A stem-and-leaf plot is a special kind of histogram on which the individual values of the observations are displayed. The final digit of a procedure rate is used as the graphing symbol. The prior digits of a rate are given on the vertical axis. In Figure 1, the decimal point is assumed to be two places to the right of the vertical line so, for example, the smallest value shown on the plot (at the top of the graph) is 260 cesarean sections per 100,000 females of child-bearing age (15-49 years), and the largest rate is 2,010.

The rates in Figure 1 represent a total of 145,740 cesarean sections in the 255 CDs over two years.

Observations which were outside the interval [median ± 2(interquartile range)] were identified as (possible) outliers. In Figure 1, the median rate was 1,123, and, using the above criterion, four CDs were identified as being outliers: the lowest rate (260) was in Annapolis County, Nova Scotia, and the three highest rates (1,960, 1,980, and 2,010), were in Division #19, Manitoba: Okanagan-Similkameen Regional District, British Columbia: and Dufferin County, Ontario, respectively.

Annapolis County is in the southern portion of Nova Scotia, on the Bay of Fundy. Division #19 in Manitoba is a large area straddling Lake Winnipeg. The Okanagan-Similkameen Regional District in British Columbia borders on the United States. Dufferin County, Ontario is northwest of Metropolitan Toronto.

Figure 2 is a stem-and-leaf plot of cesarean section rates in 245 CDs, calculated using Denominator 3 as described in Table 1. These rates represent a total of 142,404 cesarean sections in the 245 CDs over two years. As expected, the procedure rates are considerably larger than in Figure 1 because the population at risk is smaller. The median rate is 18,248 cesarean sections per 100,000 females of child-bearing age who gave birth to at least one live or stillborn child. Only one outlier was detected using the above criterion; this was the highest rate of 36,000, occurring in Cumberland County, Nova Scotia, which borders on the province of New Brunswick.

Changing from Denominator 1 to Denominator 3 resulted in different CDs being designated as outliers and, more generally, in a different ranking of procedure rates among the CDs. The CD that had the highest rate and was selected as the sole outlier in Figure 2 ranked 4th highest and was not an outlier in Figure 1. The CD designated as having an unusually small rate in Figure 1 did not have an unusually small rate in Figure 2. The CDs that had the highest, second highest, and third highest rates and were selected as outliers in Figure 1 ranked 6th, 2nd, and 17th highest,

respectively, in Figure 2 and were not selected as outliers.

Figure 3 is a stem-and-leaf plot of lung surgery rates for males in 163 CDs, calculated using Denominator 2 as described in Table 2. These rates represent a total of 2,123 lung surgeries in the 163 CDs over two years. The median lung surgery rate was 12,920 per 100,000 new lung cancers. One CD (Division #7, Newfoundland, on Canada's east coast at Bonavista Hay), was identified as having an unusally high lung surgery rate of 61,000.

It should be noted that the observations plotted in Figures 1-3 are not identically distributed, because the CDs have different population sizes and thus their procedure rates have different variances. For that reason, these outlier-identification techniques are also used to analyze the normalized procedure rates mentioned earlier. Just as the use of different populations at risk results in different CDs being designated as outliers, so does the use of the normalization procedure. A typical effect of normalization is to move rates for the smaller CDs closer to the median, so that larger CDs are more likely to emerge as designated outliers. For examples, see the references listed below.

Once outliers are detected using these statistical techniques, our medical colleagues assist us in interpreting the results. Various reasons that surgical procedure rates may vary across geographical areas are given in the references listed below.

Concluding Remarks

For two surgical procedures (cesarean section and lung surgery), data from five different data sources were used to calculate five types of procedure rates: the different denominators represented different definitions of the population at risk of undergoing the particular surgical procedure. The data sources used were (1) hospital discharge data; (2) Census data; (3) vital statistics birth files; (4) vital statistics still-birth files: and (5) cancer registry data.

When the numerator and denominator are taken from different sources, the counts must be adjusted to fit each other, and in the process, compromises and approximations must be made. As the above examples show, the two files may not cover the same time intervals and/or populations. All of the needed variables may not be available in both files. A file may be case-based when a person-based count is preferred. One file may be ready for use before the other is available.

Lack of comparability of data element definitions can also be a problem, either among data sources or across time. For example, geocode definitions undergo periodic revisions, especially in Census years, so, e.g., a given CD may be coded differently in data years before and after 1991.

Obtaining denominator data to represent the population at risk can be a complex procedure. The population at risk may be difficult to conceptualize. If it can be identified, it is often non-unique, and it is specific to the numerator (here, to the particular type of surgery). Appropriate data for the denominator may be difficult or impossible to obtain.

The problems of using rate numerators and denominators (which are aggregate data) from different sources are also encountered more generally when multiple data sources are used for analysis via record linkage (of microdata). Analysis using multiple data sources is, in a sense, a form of meta analysis which will become more common with the increased availability of captured data and continuing advancements in statistical methodology and information technology, driven by demands for multivariate and longitudinal data.

Referenaes

Gentleman, Jane F.; Wilkins, Russell; Nair, Cyril: and Beaulieu, Sylvana (1991a). An Analysis of Frequencies of Surgical Procedures in Canada. Proc. 1991 I.S.I. Meetings. Bull. of the I.S.I., Vol. LIV, Book 1, 2/2: 1-17.

Gentleman, Jane F.: Wilkins, Russell: Nair, Cyril; and Beaulieu, Sylvana (1991b). An Analysis of Frequencies of Surgical Procedures in Canada. Health Reports, 3(4), 291-309.

TABLE 1. Data Used to Calculate Rates of Cesarean Section

	Event or Persons Counted	Source	Date(s) Covered by Source	Date(s) Used	Population Covered by Source	Population Used
Numerator	Cesarean sections	Hospital discharge data	Fiscal Years (Apr-Mar) 1988-89 + 1989-90	← same	10 Provinces + visitors	8 or 9 Provinces
Denominator 1	Females of child-bearing age (15-49)	Post-Censal estimates	June I, 1989	← same	10 Provinces + 2 Territories	9 Provinces (not PEI)
Denominator 2	Females 15-49 who gave birth to at least one live child	Vital Statistic5 live birth data	Calendar Years 1988 + 1989 + 1990	1988: Apr-Dee +1989: Jan-Dec +1990: Jan-Mar	10 Provinces + 2 Territories + visitors	8 Provinces (not PEI or Nfld)
Denominator 3	Females 15-49 who gave birth to a least one live or stillborn child	Vital Statistics live birth and stillbirth data	↑ same	↑ same	↑ same	↑ same

TABLE 2. Data Used to Calculate Rates of Lung Surgery

	Event or Persons Counted	Source	Date(s) Covered by Source	Date(s) Used	Population Covered by Source	Population Used
Numerator	Lung surgeries (males)	Hospital discharge data	Fiscal Years (Apr-Mar) 1988-89 + 1989-90	← same	10 Provinces + visitors	6 or 9 Provinces
Denominator 1	All males	Post-Censal estimates	June 1, 1989	← same	10 Provinces + 2 Territories	9 Provinces (not PEI)
Denominator 2	Lung cancer cases (males)	Cancer Registry	Calendar Year 1989	← same	10 Provinces + 2 Territories	6 Provinces (not PEI, ND, Ont, or BC)

FIGURE 1. Number of Cesarean Sections per 100,000 Females of Child-bearing Age (15-49 years) (Denominator #1)

```
 2| 6
 3|
 4| 9
 5| 89
 6| 02668
 7| 1124456789
 8| 00122233455555666667999 9
 9| 01112222234456666778888999
10| 00001111122222223333444445556666777888888888999 9
11| 00001122222223344455556666777788999999
12| 00111112333333444445556667778888899
13| 0022222333333445667777 88
14| 1111222233445567999
15| 12244
16  000234566678
17|
18| 6
19| 68
20| 1
```

Decimal point is 2 places to the right of the line.
Number of CDs = 255.

FIGURE 2. Number of Cesarean Sections per 100,000 Females of Child-bearing Age who Gave Birth to at Least One Live or Stillborn Child (Denominator #3)

```
0| 5
0| 7
0| 89
1| 0111
1| 2222223333333333333
1| 44444444444445555555555555555555555
1| 66666666666666666666666677777777777777777
1| 88888888888888888888888888888888999999999999999999999
2| 00000000000000000011111111111111111111111
2| 22222222222222222233333333333
2| 4444444444555555555
2| 666
2| 89
3| 1
3|
3|
3| 6
```

Decimal point is 4 places to the right of the line.
Number of CDs = 245.

FIGURE 3. Number of Lung Surgeries per 100,000 New Lung Cancers (Males) (Denominator #2)

```
0| 00000000000011222233333444
0| 555555555666666777888888888889999999
1| 000000111111112233333333344
1| 5556667777788888899999
2| 00001111122233444
2| 555677778999
3| 00122234
3| 56778
4| 001144
4| 57
5|
5|
6| 1
```

Decimal point is 4 places to the right of the line.
Number of CDs = 163.

GENERALIZED LINEAR MODELS AND SIMULTANEOUS SURVEYS

Wilfred L.Rosenbaum, Simon Fraser University
Theodor D.Sterling, James J.Weinkam

The joining together of data obtained from separate samples is not a particularly recent idea, although it has not really seen a unified statistical treatment. In epidemiological research this is the key to case-control studies, in which separate samples of diseased and non-diseased persons are compared. In econometric analysis, the process is called choice-based sampling, and may be used for example, to study reasons for prefering one mode of public transportation over another. In this case it is more efficient to take separate samples of bus, train and car-riders at bus stops, train stations or parking lots, than to conduct a household survey to obtain the required information.

In epidemiological research, possibly the earliest study to combine information from independent surveys was a study conducted in the late 1950's by Haenszel [3]. Haenszel and his coworkers examined lung cancer deaths from the Current Mortality Sample, and used denominator data obtained from a supplement to the Current Population Survey in studying the effects of smoking and residence on lung cancer mortality in the US. In the early 1970's Godley and Krugel [2] used data from the 1966-68 National Mortality Follow-back Survey (NMFS) and the 1967 Current Population Survey to study differential mortality due to cigarette smoking.

Frenkel and Aronson [1] also used the 1966-68 NMFS in conjunction with data from the Current Population Survey in an analysis of the effect of income on mortality from a number of different diseases.

Our own analyses have used the NMFS with the National Health Interview Survey (NHIS) as the source of the denominator data in studies investigating mortality risks associated with cigarette use, with occupation, with smokeless tobacco use, with social class [6, 7].

For those who may not be aware of these surveys, the NMFS is a stratified probability sample of all deaths occurring in the US during a given time period. The most recent and most comprehensive NMFS was conducted in 1986, sampling 18,733 decedents. Death certificate informants, family members or friends supply various information on a number of demographic and health related variables for each sampled decedent.

The NHIS is a large ongoing household survey, which samples the non-institutionalized population of the US. Each year, in the order of one hundred thousand persons are sampled, and information on a number of demographic and health related variables is obtained.

When the NHIS can be used to estimate the distribution of one or more risk factors among the population at risk, and the NMFS can be used to estimate the distribution of the same risk factors among the decedents of that population, this information may be combined to yield estimates of annual mortality rates (probabilities of death).We assume that these rates, being based on nationally representative samples, should be good approximations to the to the true, usually unknown, national rates. In some cases, however, where the national rates are available, we can compare directly the survey estimated rates with the true national rates. For example, Vital Statistics publishes its own estimates of annual US Sex,Age,Race and Cause-Specific mortality rates. Figure 1 shows the all causes age specific mortality rates estimated from the 1986 NMFS and the 1986 NHIS for all US white males, and the corresponding rates published by Vital Statistics. We note there is very close agreement between the two sets of rates, although it would be reassuring to have some objective statistical evidence for this claim. To present the required statistical results we require some mathematical notation.

Suppose the population at risk and the decedents from that population are classified on the basis of one or more categorical variables, say sex, race, age and smoking status. This gives us two multiway tables of totals, one for the population and one for the decedents. We can assume the categories of the classifying variables are given once and for all in some specific order, so we write each multiway table as a vector with a single index. Thus we have two vectors of unknown totals, P which gives the distribution of the classifying variables among the population at risk, and D which gives the distribution of the same variables among the decedents.

We use the NHIS to estimate the vector P and the NMFS to estimate the vector D. We denote these estimates by $\hat{P}$ and $\hat{D}$ respectively. Finally for a vector v we let D_v denote the diagonal matrix with v on the diagonal and 0 elsewhere.

Theorem 1 *Let $\pi_0 = (\pi_{10}, \ldots \pi_{I0})$ be a fixed vector, $0 \leq \pi_{i0} \leq 1$ for $i = 1, \ldots, I$, and let*

$$X^2 := n \sum_{i=1}^{I} \frac{\hat{\mathbf{P}}_i(\hat{\mathbf{D}}_i/\hat{\mathbf{P}}_i - \pi_{i0})^2}{\pi_{i0}(1-\pi_{i0})}$$

Then under the null hypothesis

$$H_0 : \frac{\mathbf{D}}{\mathbf{P}} = \pi_0$$

$X^2 \sim \sum_{i=1}^{I} \delta_i \chi_i^2$ where the χ_i^2 are independent χ_1^2 variables, and the δ_i are eigenvalues of the matrix

$$D^{-1}_{P\pi_0(1-\pi_0)} \left(\hat{\Sigma}_{\mathbf{D}} + D_{\pi_0} \hat{\Sigma}_{\mathbf{P}} D_{\pi_0} \right)$$

In a more traditional statistical setting we would compute the usual Pearson χ^2 statistic to compare the observed survey-based rates with the expected vital statistics rates. If we performed independent binomial sampling at each level of the classifying variables (ie at each i) this statistic would have a χ^2 distribution with I degrees of freedom. However, the sampling structure is much different here, and we shouldn't be too surprised that the classical result no longer applies.

Theorem 1 above, based on a theorem in [5], shows that if we naively compute the Pearson statistic using weighted-up data from the surveys, the resulting statistic. is distributed like a weighted sum of independent χ^2 variables, each with one degree of freedom, where the weights are the eigenvalues of the indicated matrix, which is a kind of weighted sum of the covariance matrices of the individual surveys.

The relatively complicated distribution of the naively computed statistic makes it difficult to interpret, in general. However, we can apply the following correction to make the picture clearer. If we multiply the statistic X^2 by the sum of the eigenvalues divided by the sum of the squared eigenvalues, the resulting statistic has an approximate χ^2 distribution with $\nu = (\sum \delta_i)^2 / \sum \delta_i^2$ degrees of freedom.

This gives us a test for evaluating how well a set of survey-obtained rates fits a given fixed set of rates, say those supplied by Vital Stats.

- Compute the naive goodness of fit statistic
- compute the eigenvalues of the weighted covariance matrix
- apply the correction to the naive statistic
- refer the result to standard χ^2 tables

We can also use a modified version of the usual 6 method to obtain estimates of the variance and confidence intervals for functions of the population totals, whenever we can write down an explicit formula for the quantities we wish compute. Suppose $F = F(\mathbf{D}, \mathrm{P})$ denotes a (possibly vector-valued) function of the vectors of true population and decedent totals, and $\hat{F} = F(\hat{\mathbf{D}}, \hat{\mathbf{P}})$ is its obvious estimator. For example, $\hat{F}$ may represent the age-specific mortality rate, an age standardized morality ratio or an attributable risk. These quantities all have explicit expressions as functions of the number of persons at risk and the number of persons who died.

For large sample sizes in each survey (which is not a problem here) $\hat{F}$ will have an approximate normal distribution, with mean F and covariance matrix given

$$\Sigma_{\hat{F}} \approx \frac{\partial F}{\partial \mathbf{P}} \Sigma_{\mathbf{P}} \frac{\partial F}{\partial \mathbf{P}}^T + \frac{\partial F}{\partial \mathbf{D}} \Sigma_{\mathbf{D}} \frac{\partial F}{\partial \mathbf{D}}^T$$

This provides us with a means of computing approximate confidence intervals for risks based on data from the independent surveys.

The corrected Pearson statistic and the ability to compute confidence intervals for estimated mortality rates provide much information regarding how well a set of survey-estimated rates match vital statistics. For example, Figure 2 shows the 1986 age specific lung cancer mortality rates for white females, along with corresponding 95% confidence intervals. The corrected pearson goodness of fit statistic for this data is 7.96 on 7.5 degrees of freedom, which shows a good fit of the survey estimated data to the vital statistics rates. The naive uncorrected pearson statistic has a value over 10,000 on 10 degrees of freedom. Based on that evidence, of course, we would reject the claim that the survey data fits the vital statistics rates. However, the naive computation, using the weighted up population estimates does not properly reflect the true amount of data on which the computation was based. Thus, even small departures from the expected rates appear as significant.

We can't directly use the previous linearized variance estimate when the population parameter *or* quantity of interest can't be defined as an explicit function of population totals. For example a vector of model parameters may be defined implicitly as the solution to a given set of likelihood equations, as in a linear, loglinear or logistic regression problem. The particular class of models described here will be limited to a specific subset of generalized linear models with binomial errors and arbitrary link function. [4] considered Poisson errors with log link, and [5] considered binomial errors with logistic link. The techniques can be applied to other models as well.

Let $\pi_i = D_i / P_i$ *be* the true proportion of deaths among persons in class i. The model we assume in this case is that $\pi_i = h(x_i^T \mathbf{B})$ for some smooth monotone function h, where B is a vector of unknown model parameters, and x_i is a vector of known quantities for the ith class. x_i is just the ith row of the model matrix X and h is the inverse of the link function. Appropriate choice of h leads to logistic, to probit or to extreme-value models. If the deaths in each class i are regarded as indpendent Bernoulli trials with a fixed probability of success, we would determine an estimate for B by solving the resulting maximum likekihood equations, and plugging this estimate back into the equation $\pi_i = h(x_i^T \mathbf{B})$ would give us a model based estimate of the π_i. But we don't know the actual values of the D_i and P_i, so instead we solve the pseudo maximum likelihood equations

$$\sum_{i=1}^{I} \left(\hat{\mathbf{D}}_i - \pi_i \hat{\mathbf{P}}_i / \pi_i (1 - \text{xi}))) \; h_i' x_{ij} = 0 \right.$$

obtained by replacing the D_i and P_i in the maximum likelihood equations by their survey based estimators. The resulting estimator $\hat{\mathrm{B}}$ is an estimate of the implicitly defined finite population parameter B, and has variability due to the fact that it is computed from sample data only.

The usual questions in any model fitting problem are of interest:

1. How well does the model fit the data?
2. Is a particular model coefficient equal to O?
3. Could a simpler model also provide an adequate fit to the data?

Again, as in the case of the simple hypothesis test, the usual statistical assumptions under which such questions are answered are not applicable here.

Theorem 2 *Let* $\hat{\pi} = h(X\hat{\mathbf{B}})$ *be the model based estimator of a.*

$$X^2 := n \sum_{i=1}^{I} \frac{\hat{\mathbf{P}}_i (\hat{\mathbf{D}}_i / \hat{\mathbf{P}}_i - \hat{\pi})^2}{\hat{\pi}(1 - \hat{\pi})}$$

Then under the null hypothesis

$$H_0 : \frac{\mathbf{D}}{\hat{\mathbf{P}}} = \hat{\pi}$$

$X^2 \sim \sum_{i=1}^{J} \delta_i \chi_i^2$ *where the* χ_i^2 *are independent* χ_1^2 *variables,* and *the* δ_i *are* eigenvalves of *the matrix*

$$\left(\mathbf{C}^T \mathbf{A} \mathbf{C}\right)^{-1} \left(\mathbf{C}^T \mathbf{B} (\hat{\mathbf{\Sigma}}_{\mathbf{D}} + D_{\hat{\pi}} \hat{\mathbf{\Sigma}}_{\mathbf{P}} D_{\hat{\pi}}) \mathbf{B} \mathbf{C} \right)$$

where

$$\mathrm{A} = \mathrm{diag} \left\{ \hat{\pi}_i (1 - \hat{\pi}_i)) / h_i' \hat{\mathbf{P}}_i \right\}$$

$$\mathrm{B} = \mathrm{diag} \left\{ 1 / h_i' \hat{\mathbf{P}}_i \right\}$$

and C is the matrix complement to X, *relative to the saturated model matrix.*

Once we have obtained the estimate $\hat{\mathbf{B}}$ as the solution to the pseudo max likelihood equations, and get the model based estimator of the cell specific mortality rates by substituting $\hat{\mathbf{B}}$ Lack into the modelling equation, our next concern should be how well does the model fit. This is a question that is glossed over in much of the epidemiological literature. The tool we use is the usual pearson goodness of fit statistic. As in the case of a goodness of fit test for a simple hypothesis, the usual distributional results do not apply. Again the goodness of fit statistic based on the weighted-up data from the two surveys is distributed as the weighted sum of independent χ^2 variables each with one degree of freedom. But this time, the weights are the eigenvalues of the more complicated matrix shown here. The A and B matrices are diagonal matrices with entries determined by the fitted cell probabilites, the estimated population at risk totals and the derivative of the model function h evaluated at the estimated parameter vector $\hat{\mathbf{B}}$. The matrix C is the complement of the model matrix X relative to the saturated model matrix. What this *means* is that if we write down the saturated model matrix, then the first say J columns will be the matrix X and the last I $- J$ columns will be C.

So this now gives us a way to evaluate the goodness of fit of the model to the survey data estimates.

. Fit the model to the weighted up totals

- Compute the eigenvalue matrix given here
- Compute the eigenvalues of this eigenvalue matrix
- Use the eigenvalues to correct the naive version of the pearson statistic to a χ^2 statistic, and refer the value to the appropriate table.

The covariance matrix of the estimated model coefficients may be obtained from the expression

$$\hat{\Sigma}_{\hat{B}} \approx \left(X^T U X\right)^{-1} X^T V \left(\hat{\mathbf{\Sigma}}_{\mathbf{D}} + D_{\hat{\pi}} \hat{\mathbf{\Sigma}}_{\mathbf{P}} D_{\hat{\pi}} \right) V X \left(X^T U X\right)^{-1}$$

The covariance matrix of the model estimated mortality rates is given by

$$\hat{\Sigma}_{\hat{\pi}} \approx D_{h'} X \hat{\Sigma}_{\hat{B}} X^T D_{h'}$$

The covariance matrix of the repsonse residuals is given by

$$\hat{\Sigma}_{\hat{p}-\hat{\pi}} \approx A^T \left(\hat{\Sigma}_{\mathbf{D}} + D_{\hat{\pi}} \hat{\Sigma}_{\mathbf{P}} D_{\hat{\pi}} \right) A$$

where

$$A = \left(\mathcal{I} - D_{h'} X \left(X^T U X \right)^{-1} X^T V \right)$$

The diagonal of $\hat{\Sigma}_{\hat{p}-\hat{\pi}}$ is used to obtain standard errors for the response residuals, from which we obtain standardized residuals.

As an example, Table I contains the results of fitting a logistic model to US cancer mortality rates, estimated using the NMFS and NHIS data. The model contains terms for Sex (Male,Female), Race(White, Non White), Age(40-49,50-59,60-69, 70-79 year:) and Smoking Status(Never, Current, Former). Smoking information was available for the population at risk in the 1987 NHIS Cancer Epidemiology Supplement only, and that survey provided the denominator data. The model fit was a main effects model, with sum-to-zero constraints. The corrected Pearson goodness of fit statistic is 18.8 on 17 degrees of freedom, indicating a good fit of the model. The Sex, Age and Smoke coefficients are all statistically significant. If the model fits, the standardized residuals should have a standard normal distribution. Figure 3 shows the standardized residuals on a norrnal quantile-quantile plot. In such a plot normally distributed data appears as a straight line.

This method of analysis is made possible by the availablity of two compatible surveys. Although we have illustrated the results with the example of NMFS/NHIS data, its utility is not limited to mortality data. Further extentions of these methods are possible, including consideration of multiple outcomes and use of nested models.

Table 1: Coefficients and Odds Ratios for NMFS/NHIS Cancer Mortality Model

	Coefficient	Odds Ratios
Intercept	-5.60	
Sex		
Male	0.09	1.00
Female	-0.09	0.84
Race		
White	-0.06	1.00
Nonwhite	0.06	1.14
Age		
40-49	-1.53	1.00
50-59	-0.18	3.87
60-69	0.54	7.96
70-79	1.16	14.81
Smoke		
Never	-0.48	1.00
Current	0.74	3.42
Former	-0.26	1.25

References

[1] Eli Frenkel and Stanley M. Aronson. Family income and mortality rates: An analysis of the national mortality survey. ***Rhode Island Medical Journal,*** 69:165–170, April 1986.

[2] Frank Godley and David L. Krugel. Cigarette smoking and differential mortality: New estimates from representative national samples. In ***Population*** Associalion of ***America*** Annual Meeting, Seattle, Washington, April 1975.

[3] William Haenszel, Donald B. Loveland, and Monroe G. Sirkin. Lung cancer mortality as related to residence and smoking histories. i. white males. ***Journal of the*** National Cancer Institute, 28:947–1001, 1962.

[4] J.N.K. Rao and A. J. Scott: On chi-squared tests for multi-way tables with cell proportions estimated from survey data. ***Annals of Statistics***, 12:46–60, 1984.

[5] G. Roberts, J.N.K. Rao, and S. Kumar. Logistic regression analysis of sample survey data. ***Biometrika***, 74(1):1–12, 1987.

[6] Theodor D. Sterling, Wilfred L. Rosenbaum, and James J. Weinkam. Race,income and mortality. ***Journal of the National Medical*** Associalion, 00:???–??, 1993.

[7] James J. Weinkam, Wilfred L. Rosenbaum, and Theodor D. Sterling. Computation of relative risk based on simultaneous surveys: An alternative to cohort and case-control studies. ***American Journal of*** Epidemiology, 136(6):722–729, 1992.

Figure 1: 1986 NMFS/1986 NHIS and US Vital Statistics
Mortality Rates
White Males, All Causes

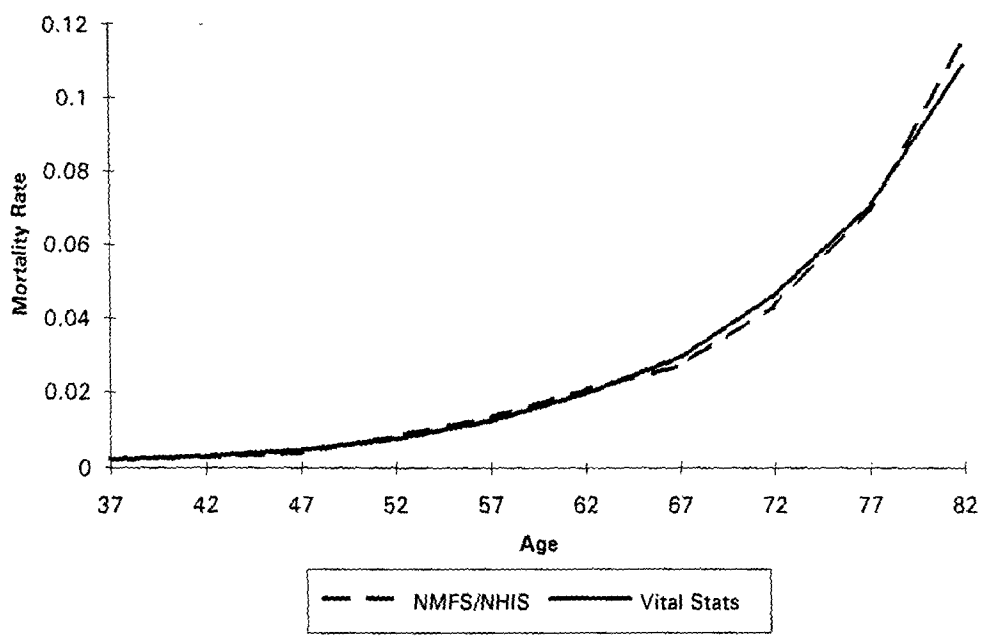


Figure 2: 1986 NMFS/NHIS and US Vital Stats Mortality Rates
White Female Lung

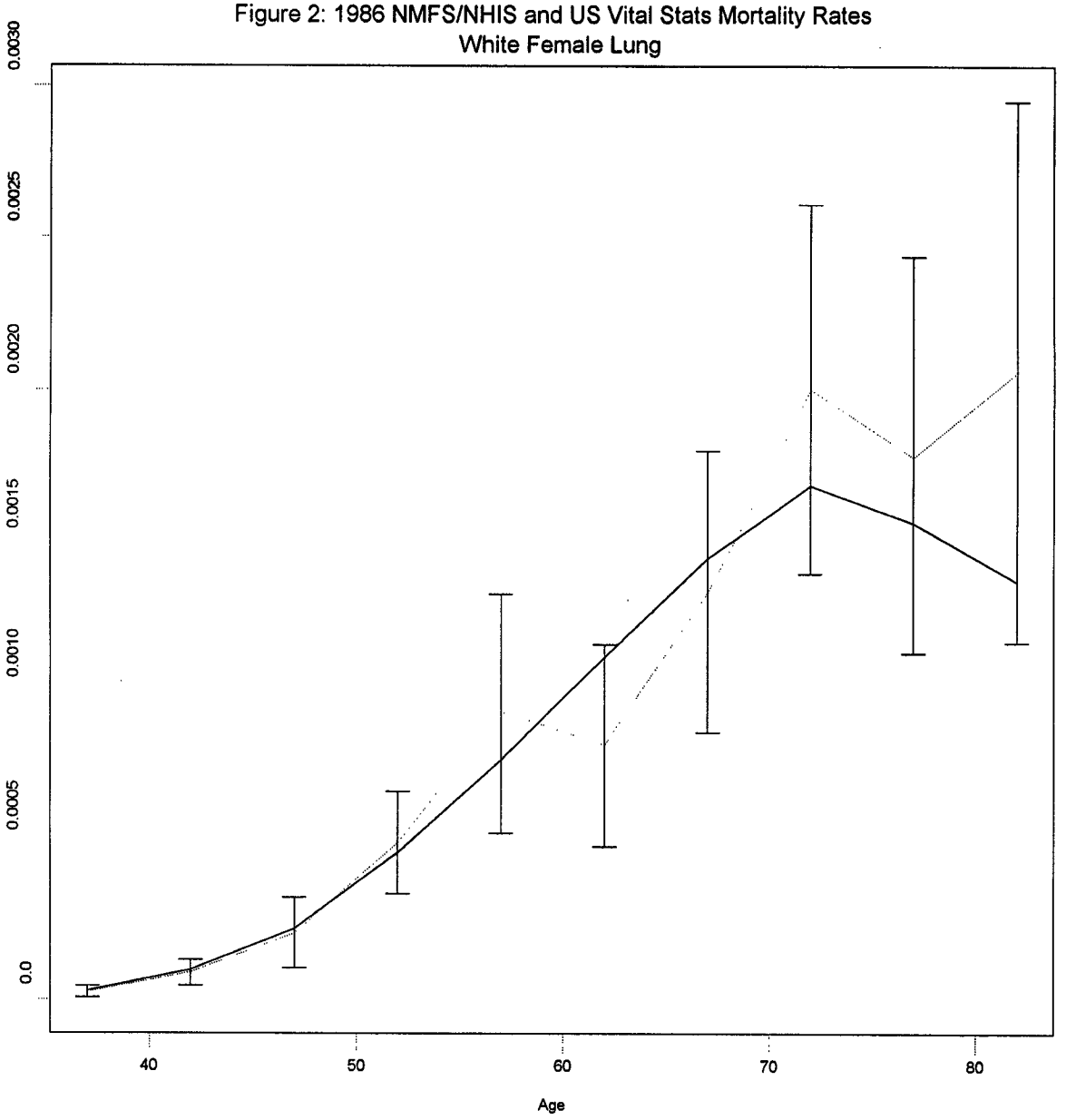
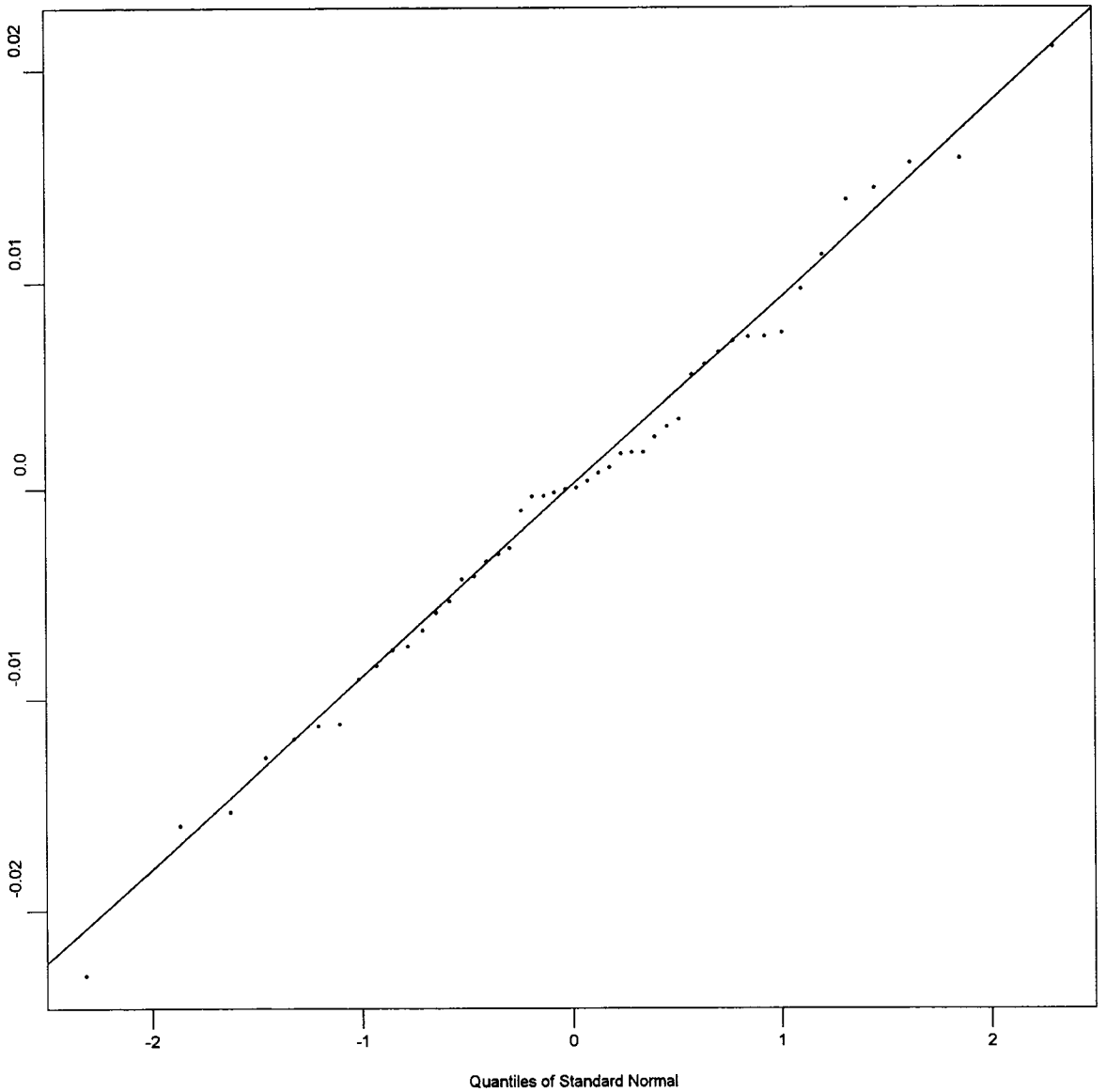


Figure 3: Standardized Residuals from
NMFS/NHIS Cancer Model



BENEFITS OF PROBABILISTIC LINKAGE OF ELECTRONIC PATIENT RECORDS FOR HIGHWAY SAFETY

William H. Walsh, Jr.
National Center for Statistics and Analysis

Introduction

Motor vehicle traffic crashes represent $137 billion dollars in costs to society from lost productivity, medical costs, and property damage. The National Highway Traffic Safety Administration (NHTSA) was created to implement programs and countermeasures to reduce this staggering toll. The National Center for Statistics and Analysis (NCSA) is NHTSA's office responsible for collecting data used to evaluate the effectiveness of existing or to identify the need for newly proposed safety countermeasures that help meet NHTSA's goal. NCSA's analytical activities focus on identifying causal factors, quantifying effectiveness, relating occupant behavior factors to injury and fatality rates, and presenting a clear and concise national picture of the characteristics of highway traffic crashes.

NHTSA Data Sources

Information to monitor the nature, severity, and injury outcome of motor vehicle traffic crashes is necessary to efficiently allocate highway safety resources and to develop effective countermeasures. Within NCSA, data for these analyses are obtained primarily from the following sources: the Fatal Accident Reporting System (FARS), the National Accident Sampling System's (NASS) Crashworthiness Data System (CDS), and the General Estimates System (GES).

FARS is a national census of all fatal motor vehicle traffic crashes. Data are obtained from the police accident report (PAR), driver and vehicle records, and some autopsy reports.

NASS is a national sample of approximately 5,000 police-reported crashes each year involving at least one towed passenger car, light truck, van, or utility vehicle. Detailed occupant injury and vehicle damage information are collected from scene and vehicle inspections, personal interviews, and medical records.

GES is a national sample of police-reported traffic crashes including all severity levels from which national estimates are generated. An annual sample of about 46,000 PARs are obtained. Data are obtained only from the PAR. NCSA also obtains and uses in its analyses the motor vehicle crash data files from 17 states. These are data coded by the state from PARs. We call these **State Data.**

CODES Project

A new source of data for highway safety analyses is being developed through NHTSA's Crash Outcome Data Evaluation System program (CODES). CODES starts with the police-reported occupant-specific crash data (state data) for all crash-involved occupants. Then, for those who are injured, the crash data are linked to injury data collected for medical, reimbursement, and epidemiological purposes. The CODES data provide NHTSA, at reasonable cost, access to a large volume of population-based, longitudinal crash outcome information. NHTSA in turn can use these data to evaluate the medical and financial consequences of specific vehicle, crash, and occupant characteristics. The addition of this information is accomplished without the delay normally required to implement new data collection systems on a large scale. Linkage makes it possible to relate specific medical and financial information to crash, vehicle, and occupant characteristics, an important upgrade in the usefulness of state data to highway safety analyses.

Grantees: NHTSA has provided grants to organizations in Hawaii, Maine, Missouri, New York, Pennsylvania, Utah, and Wisconsin to create CODES. These states were chosen based on the maturity and completeness of their existing state data systems.

ISTEA: Funding for CODES was provided through the Intermodal Surface Transportation Efficiency Act of 1991. What we call the ISTEA. The linked crash/injury outcome information compiled by the grantees will be provided to NHTSA and we will prepare an ISTEA mandated report to Congress about the benefits of safety belt and motorcycle helmet use.

Purpose: The ISTEA specifically requires that the benefits of belts and helmets be measured in the following areas: on injury type, body region injured, and injury severity; on acute, rehabilitation, and long term care consequences at 6 to 12 months; on the financial consequences; and, on mortality and morbidity outcomes. Only linked crash and injury data will generate this type of information.

Linkage: Occupant specific and environmental information describing the various aspects of a crash are useful for highway safety analyses. State Data from PARs have been used by NHTSA and many other highway safety analysts for

many years. CODES data will enable these researchers to follow the victims of crashes from the scene through the medical care system to final disposition, something they aren't able to do currently being restricted to police reported injury severity. CODES will do this by linking data collected at the following sources: crash, EMS, emergency department, hospital discharge, rehabilitation/long term care facilities, non-hospital outpatient, and claims. Data from death certificates, census, driver licensing, vehicle registration, and roadway design files also can be linked. In Figure 1, the data sources in boxes with double lines represent computerized, statewide data systems. Those in boxes with a single line represent data systems which are usually not statewide.

Concepts of Probabilistic Linkage: The linkage will be performed using a probability based linkage software. The purpose of any data linkage is to match records pertaining to the same person that are located in different files. The linkage of crash to injury records is complicated by the fact that it is uncertain which crash records are expected to match to an injury record and vice versa. The probabilistic linkage methodology assumes that each field has a measurable reliability and probabilities of chance agreement for its variable attributes. Linkage requires that sufficiently accurate information be available to discriminate among the events and among individuals involved in a specific event. Weights are assigned to indicate the likelihood that the records which match on specified identifiers do, in fact, represent the same person. Probabilistic linkage reduces the linkage problem to records included within small blocks. Because of this method of operation, fewer staff and computer resources are required to perform the linkage. Most of the CODES grantees are in fact using PCs to perform the match. As a result, this software makes the linkage of statewide databases feasible on a routine basis.

Identifiers: Linkage variables may be direct or indirect. Direct variables include information which directly identifies a person, that is, a social security number or a unique patient identifier. Direct identifiers on injury records are frequently not available because of patient confidentiality issues. When they are available, they are subject to the same potential for missing and inaccurate data as indirect identifiers. Indirect variables must be combined to identify a person. Included are age, sex, date, and location. Linkage is more effective when both direct and indirect variables are used together to compensate for potential missing or inaccurate information existing in each. Probabilistic linkage allows almost 100 variables to participate in the linkage, although efficiency decreases somewhat as the number of linkage variables increases.

File and Field Preparation: The linkage software requires ASCII files in which the variable lengths and file locations are fixed. Also, better linkage results are achieved when the data are accurate. Crash times should occur before arrival of the police at the scene. Calls to EMS before EMS arrives. Identification codes for the police, EMS, and hospitals must be consistent in both files being linked. Particular care must be taken to make the coding consistent. Coding systems used to identify towns, service areas, sex, etc., in one data file must match the codes used in the other file. A common problem is the use of a zero or 99 to designate unknown ages without distinguishing the newborns or 99 year-olds. Variable lengths may also be a problem. Age coded as a 2-digit field in one file must be a 2-digit field in the other file. Information which is non-uniform in the files being linked can be standardized into new variables useful to sort the files into blocks or link the records within a block. Creation of an occupant specific date of birth field facilitates linkage to the date of birth information on the injury records. Military times converted into 4 or 8 hour time blocks make it possible to use time as a blocking variable.

Data Elements: Each of the data sources provides unique information. The crash data include occupant identifiers and descriptions of the crash. The EMS data include identifiers, type of EMS response, and medical information collected at the scene and enroute. The hospital data include identifiers, medical diagnoses and severity, disposition, and financial information.

Linked Record: For example, a linked record could describe an event in which a 13 year-old male was an unrestrained passenger in a 1986 Buick involved in a single vehicle frontal collision with a fixed object. The crash occurred in a

rural area. The passenger suffered a fractured cervical spine, a broken arm and a brain injury. EMS was called 10 minutes after the crash occurred and provided advanced life support treatment at the scene. The patient was transported to the nearest local hospital within 2 hours of the crash and then transferred to the trauma center within 6 hours of the crash. Ten hours lapsed between the time of the crash and spinal surgery. The patient remained in the hospital for 8 days, incurred a hospital bill of $15,300 and total charges of $20,000. The information reported for this linked record represents only a small part of the computerized data available from each data system.

Obstacles To Data Linkage: Poor data quality, non-computerized data, interagency politics, unnecessarily restrictive data access policies, and insufficient resources are some of the many obstacles which jeopardize data linkage. Resolution of these obstacles requires the collaboration of the highway safety and medical communities.

CODES State Advisory Committee: To promote that collaboration, each CODES grantee was required to convene an advisory committee consisting of the users and owners of the statewide data. The purpose of the committee is to resolve problems associated with the obstacles and to establish guidelines for access, data security, privacy and confidentiality, and financial support. These committees have already been responsible for facilitating needed improvements to the crash and EMS data files among the CODES sites.

Uses of Linked Data

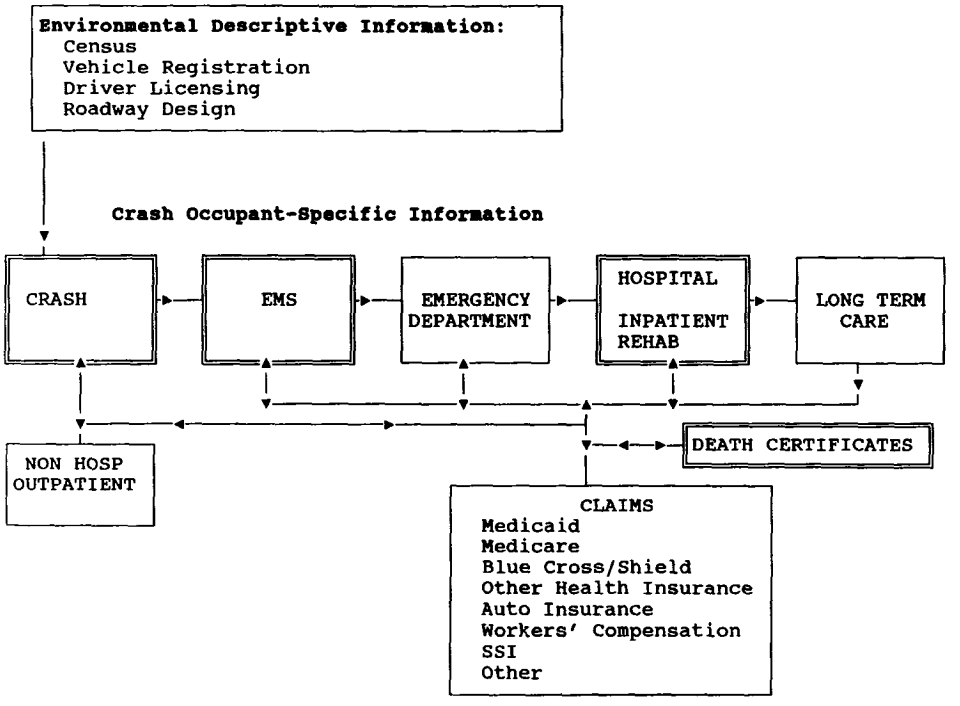
Linked data have many uses. First, as part of the CODES project, each CODES site will generate linked data describing the characteristics of the crash, vehicle, occupant, the severity and type of injury, the utilization of protective devices, and financial information for inclusion in a uniform model to research the benefits of safety belts and helmets. But the linked data can also be used for vehicle related analyses which focus for example on relating the type of injury to the type of restraint system or a specific vehicle characteristic. Linked data can be used for occupant related analyses to focus on medical outcome in terms of the utilization of protective devices or the absence of alcohol. Linked data can be used to target specific injury control initiatives for urban and rural areas. Linked data can be used to evaluate aspects of the EMS system such as the impact of the delay to EMS notification or the use of advanced life support on medical outcome.

Conclusion

There are trade offs among the various NCSA data sources related to timeliness, expense, and content. Crashworthiness and crash avoidance problems identified in the national data (NASS or FARS) can be investigated further to determine prevalence using the population-based CODES data. Patterns identified in the state data can be evaluated for national importance by using the national data.

NHTSA programs have resulted in increased protection of motor vehicle occupants -- through enhanced vehicle and roadway designs, increased occupant restraint use, and a decline in the incidence of drunk driving. To continue this progress, we must strengthen the data systems needed to guide the development of effective solutions for preventing further crashes and subsequent injuries. Probabilistic linkage projects have the advantage of improving state data capabilities. Improvement in state data over the long run improves the data available for analytical purposes both to the states and to the Federal Government. The numerous technical and program office safety issues require different types of data. CODES data compiled from routinely collected state data will complement the detailed research data from systems like NASS. They supplement and complement each other.

FIGURE 1: LINKED CRASH AND INJURY DATA SYSTEMS



Session N

International Comparisons

Toward the year

2000

Refining the measures

TRENDS IN MORTALITY IN EAST AND WEST GERMANY DURING THE 3 DECADES BEFORE UNIFICATION

Karl E. Bergmann, Federal Health Office

(Not available for publication)

Mortality in the United States and France: 1988.

Alvan O. Zarate, National Center for Health Statistics
Manning Feinleib

Prior to the long-term declines in mortality that have taken place in industrialized countries, the leading causes of death were the result of common health problems. A variety of epidemiological patterns have been responsible, however, at lower levels of mortality.

While both the United States and France had, by the late 1980s, achieved age-adjusted total mortality rates of approximately 10 per thousand, an examination of cause-specific mortality reveals important differences between the two countries.

This paper describes the two distinct cause-of-death profiles in the United States and France and discusses some of the reasons for the dissimilarities.

Background

Both the United States and France are known to have among the world's lowest mortality in certain respects. By 1988, the United States had achieved an age-adjusted death rate among males for stroke that was lower than that of any other industrialized country. Similarly, in France mortality from ischemic heart disease (IHD) was the lowest of any western nation - in spite of the relatively high prevalence in that country of health-related behaviors known to constitute risk factors in its development.

However, the striking nature of these observations often serves to overshadow other important differences in the mortality patterns of the two countries.

Methods

Data for this presentation are taken from the mortality data bank made available by the World Health Organization (WHO). For the United States this information is supplemented by data provided by the NCHS Division of Vital Statistics.

Age-adjusted rates were computed using age-specific deaths and population estimates also provided by WHO. The standard employed is the WHO 'European' standard population.

In view of space restrictions, this presentation will be limited to data for white U.S. and total French males. A similar analysis for females and U.S. blacks is in preparation.

Results

Although total mortality rates among U.S. white and French males are similar, there are marked dissimilarities in the specific causes of death that produce them (Figure 1).

In the United States heart disease accounts for well over a third of total white male mortality, whereas in France this category makes up less than one-fifth.

Cancer, on the other hand, accounts for around one-third of French male mortality, compared with about one-quarter of U.S. white male deaths.

Both stroke and accidents account for a greater proportion of French than U.S. mortality.

Table 1 provides a somewhat different perspective on how different causes of death contribute to total mortality. Those major causes of death for which U.S. rates exceed French rates had, in 1988, a combined rate among U.S. white males of all ages of about 475 per 100,000 - almost twice that among French males for the same causes.

The great bulk - 80 percent - of the difference between these combined rates is accounted for by the difference in mortality from heart disease alone (369.9 vs 186.0 in the U.S. and France, respectively).

For those major causes of death for which mortality is higher among French males, the combined rate for U.S. white males of about 359 per 100,000, is three-quarters of that for French males.

Slightly more than half of the difference is accounted for by the difference in mortality from cancer.

Age patterns of mortality

Figure 2 indicates that the tremendous advantage enjoyed by French males relative to U.S. white males with regard to heart disease is largely compensated for by their much higher rate of mortality from suicide, cirrhosis of the liver, cancer and stroke. This is true among all of the age groups shown and is especially notable at ages 45-64.

Similar though less sharp patterns can be seen among older males. Among the oldest, deaths coded to the ICD category "Signs, symptoms and ill-defined conditions" form a large proportion of deaths represented as "All other".

Nevertheless, among older age groups, the heart disease group continues to be more important among U.S. white males than the group headed by cancer and the reverse is true for French males.

In sum, while these two broad groups of cause of death together make up about 80 percent of all deaths among males over age 45 in both countries, the two separate groups contribute about equally to the total mortality of U.S. white males at these ages, whereas among their French counterparts, the group headed by cancer is by far the greater killer.

As a result, the relationships between the disease groups discussed here are virtually mirror images of each other; the group headed by heart disease having roughly the equivalent importance among U.S. white males that the group headed by cancer has among French males and the group headed by cancer being only slightly more important in the U.S. than is the group headed by heart disease in France.

Discussion

Having thus contrasted mortality among males in France and the United States, what can we say as to the factors underlying the different mortality patterns.

- Effects of known and suspected variations in diagnosis of cause of death and coding practices.

A frequently used indicator of the overall quality of cause-of-death statistics is the proportion of death certificates coded to "symptoms, signs, and ill-defined conditions (ICD-9 nos. 780-799). In this regard, while France has, for years coded many more deaths to this category than the United States, the difference has narrowed considerably.

However, as we have pointed out, the difference between the two countries is much larger at very advanced ages where most deaths take place and we should consider whether this might have affected our results.

Furthermore, researchers have reported underreporting of diabetes in both countries (1), of heart disease - particularly in France (2), overreporting of respiratory disease in France (3), and differences between the U.S. and France in the coding of cancer (4). As for the possible role such variations might play in bringing about our results, we offer the following considerations:

1. While generational or historical differences that may affect comparisons between older males cannot be ignored, an increase in deaths assigned to cancer at the oldest ages in France has coincided with declines in mortality from stroke and heart disease (the two other major causes of death at these ages) and a sharp reduction in deaths coded to Signs, symptoms and ill-defined conditions (Figure 3). It is likely, therefore that many deaths previously coded to Signs, symptoms, etc. are being increasingly coded to cancer. To the extent that reallocation of deaths now assigned to Signs, symptoms, etc. would tend to increase deaths assigned to cancer, the pattern among French males over 75 would more closely resemble that seen among younger men.
2. While there is evidence that the French tend to under-report deaths to heart disease, this must be balanced against a similar tendency in the U.S. Even if all French deaths classified in "Signs, symptoms, etc." were reallocated to "heart disease", our conclusions would not be seriously affected.
3. Adjustments for tendencies to over- or under-report cancer, respiratory disease and diabetes, would either result in wider U.S.-French differences than seen here or change our results only slightly due to the magnitude of the error or the level of the mortality rate involved.

It seems unlikely, therefore, that differences in data quality or coding and diagnostic practices account for the overall pattern we have observed.

- Effects of differences in risk factors.

If consideration of such differences does not substantially modify our conclusions, what other factors may be responsible?.

There is general agreement that recent changes in mortality in industrialized countries have been brought about by modifications in a number of factors including cigarette smoking, hypertension, overweight, dietary fat, amount of exercise, alcohol consumption, and the use of screening techniques (5).

We would like to assess some reasonably well-documented evidence with respect several of them.

The low mortality from Ischemic Heart Disease (IHD) despite the high fat content of the French diet, often referred to as the "French paradox", is illustrated in Figure 4.

IHD constitutes the great bulk of heart disease in both countries, and, as we have indicated, there is solid evidence that one of the principal risk factors in its development is amount of dietary fat. Based on the level of fat of animal origin consumed in France, its level of IHD mortality ought to approximate that of the United States. Instead, it is considerably lower.

As many researchers have pointed out, the outstanding difference between France and other countries seems to be the exceptional French consumption of wine (6).

To demonstrate how this factor might be responsible for the pattern seen in Figure 4, we have combined wine consumption and animal fat consumption in a regression analysis (Table 2).

Keeping in mind that inférences based upon ecological correlations are extremely hazardous and must be made with great care (7), we may observe that although the overall model is highly significant - as is the beta coefficient for wine consumption - the beta coefficient for fat consumption is not. The pattern we see in Figure 5 in which the position of France is no longer that of an outlier from the regression line, but very close to it, appears to be primarily the result of the association of IHD mortality with wine consumption. The predictive power of wine consumption alone is further demonstrated by Figure 6.

While the prophylactic effect of greater wine consumption among French males may account for much of their advantage over U.S. males with regard to heart disease (especially IHD), other evidence indicates that this gain has come at the price of higher mortality from other diseases.

When we examine those major causes for which French mortality exceeds that of the U.S. (Table 3), we find a predominance of causes that have been closely linked with excessive alcohol consumption: cirrhosis of the liver, suicide, accidents other than automobile and cancer.

Cancer, of course is not an etiologically homogeneous disease. While cause of death statistics by site are subject to a much greater degree of uncertainty than the broad cause of death categories previously presented, Tables 4 and 5 are nonetheless instructive.

There is strong empirical evidence of the causal role of alcohol consumption for cancers of the larynx, oral cavity, pharynx, liver, and the esophagus (10). These data suggest that the bulk of the cancer differential in mortality between the U.S. And France is attributable to those cancers.

The major cancer sites for which U.S. white male mortality exceeds that of French males are cancers of the pancreas, prostate, trachea, bronchus, lung, and colon.

Not all of the French excess in cancer mortality is due to alcohol, however, for, with the exception of rectal cancer, there is also evidence concerning the role of cigarette smoking -in addition to and in combination with alcohol- in the etiology of cancer at each of the sites in Table 4 (8,10).

In the case of respiratory cancers - i.e., cancer of the trachea, bronchus and lung - while age-adjusted rates for all ages are clearly higher in the U.S. than in France (Table 5), this relationship is apparently undergoing change.

Among men 45-64 while respiratory cancer has levelled off in the U.S., and perhaps even declined, it continues to rise among their counterparts in France (Figure 7).

Figure 8 throws additional light on possible changes in cigarette smoking in the two countries.

Secular trends in cigarette sales and consumption indicate that heavy smoking reached its peak during the 1960s in the U.S., whereas it is just beginning to level off in France.

Conclusions

From the vantage point provided by these data, it is clear that low or moderately low national mortality levels do not imply uniformity with respect to specific disease patterns.

Although the same diseases - heart disease, cancer and stroke - are the major killers in both countries, they do not contribute to total mortality in the same degree.

While differences in data quality, or coding and diagnostic practices undoubtedly have an impact on the comparisons made, we do not believe they are large enough or operate in such a way as to have brought about our results.

Indeed, these disease patterns are fully consistent with what is known concerning differences in dietary and other health habits - most notably, cigarette smoking and alcohol consumption.

Excessive animal fat consumption does not appear to account for the pattern described largely because of the protective role of alcohol.

However, the influence of alcohol consumption appears to extend beyond the protective and, when excessive amounts are consumed, by itself and in combination with cigarette smoking contributes to a cause of death configuration distinctly different in France and in the United States.

The data utilized in this report are neither as detailed nor reliable as we would prefer. Furthermore, our conclusions presume empirical relationships whose scientific acceptability is yet to be confirmed.

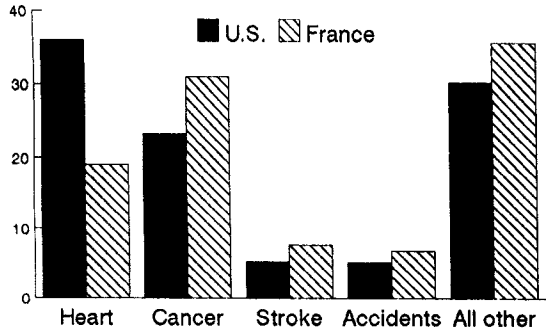
Nonetheless, our understanding of the factors underlying the broad patterns we have traced seems plausible and would, in fact be useful in the consideration of comparisons between the United States and other countries with similar levels of mortality.

Certainly, the proposition considered at the outset, that industrialized countries may pursue a variety of paths to low mortality has been, we think, well illustrated.

References

1. Jougla E, Papoz L, Balkau B, **et al**. Death certificate coding practices related to diabetes in European countries - the 'EURODIAB' Subarea C' study. **International Journal of Epidemiology** 1992; **21**:343-351.
2. Nuttens M C, Salomez J L, Tillard B, **et al** Validité des certificats de décès pur l'étude des cardiopathies ischémiques. **Presse-Meds.** 1990 **19**:1143-1146.
3. Cooreman J, Thom T J, Higgins M W. Mortality from chronic obstructive pulmonary diseases and asthma in France, 1969-1983. Comparisons with the United States and Canada. **CHEST** 1990; **97**:213-219.
4. Percy C, Muir C. The international comparability of cancer mortality data. Results of an international death certificate study. **American Journal of Epidemiology** 1989; **129**:934-946.
5. Hahn R, Teutsch S M, Rothenberg R B, Marks J S. Excess deaths from nine chronic diseases in the United States, 1986. **JAMA** 1990; **264**:2654-2659.
6. Tunstall-Pedoe H. Autres pays, autres moeurs. Theories on why the French have less heart disease than the British. **BMJ** 1988; **297**:1559-1560.
7. Feinleib M, Leaverton P.E. Ecological Fallacies in Epidemiology. In Leaverton P, Masse L, eds. **Health Information Systems.** New York, Praeger Publishers, 1984; 33-61.
8. U.S. Department of Health and Human Services. **Reducing the Health Consequences of Smoking: 25 Years of Progress. A Report of the Surgeon General.** DHHS Publication No. (CDC) 89-8411. 1989.
9. Food and Agriculture Organization of the United Nations. **FAO production yearbook.** Rome. 1987.
10. International Agency for Research on Cancer. **IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Volum 44. Alcohol Drinking.** Lyon. 1988.
11. in Ravenholt RT. Tobacco's global death march. **Population and Development Review** 1990; 16:213-240.
12. Ministére de la Solidarité, de la Santé et de la Protection Sociale. **La Santé en France. Faits majeurs grandes tendances.** Paris: La Documentation Française. 1990.

Figure 1. Selected causes of death as percent of total deaths: U.S. (white) and French males



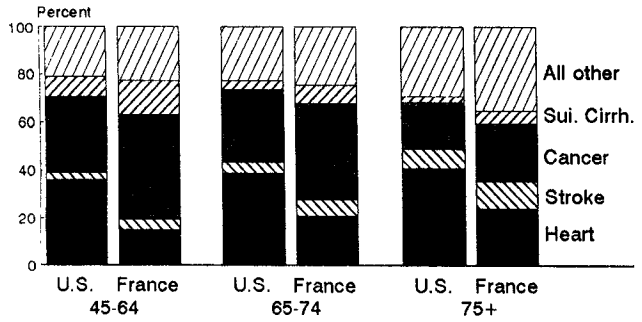
SOURCE CDC,NCHS,NVSS and World Health Organization, 1986

Table 1. Comparison of cause of death groups for U.S.(white) and French males

Cause	Age adjusted mortality rate	
	U.S.	France
U.S. exceeds France.		
Heart, diabetes, COPD pneumonia, influenza homicide	475.1	246.0
France exceeds U.S.		
Cancer, stroke, other accidents, suicide, cirrhosis	358.5	477.3

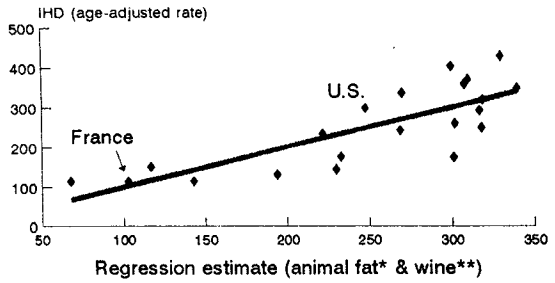
SOURCE CDC/NCHS,NVSS and World Health Organization 1986

Figure 2. Mortality from selected causes as percent of all deaths: U.S. (white) and French



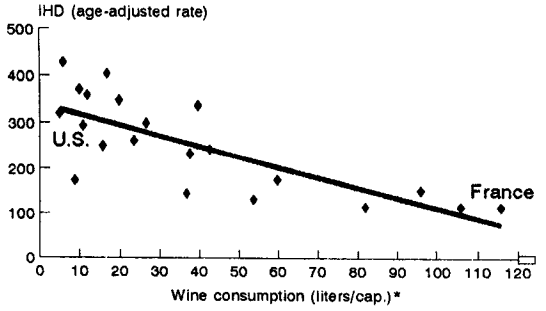
SOURCE CDC/NCHS NVSS & World Health Organization 1986

Figure 5. Regression of IHD mortality on animal fat and wine consumption:
21 industrialized countries



*FAO (9), 1983-85 **IARC (10), 1986
SOURCE World Health Organization 1985

Figure 6. Regression of IHD mortality on wine consumption: 21 industrialized countries



*IARC, 1986
SOURCE World Health Organization, 1985

Table 3. Major causes of death with mortality higher among French males

Cause	Age-adjusted rate*	
	U.S.	France
Cancer	239.9	303.3
Stroke	54.8	74.7
Other accidents	26.3	40.2
Suicide	21.6	29.9
Cirrhosis of the liver	15.9	29.2

*per 100,000 population
SOURCE CDC/NCHS, NVSS and World Health Organization, 1986

Table 4. Mortality from selected cancer sites: males age 35-74

Cancer site	Rate*		U.S. as % of France
	U.S.	France	
Larynx	4.8	22.8	21
Oral & pharynx	9.3	37.9	25
Liver	3.5	14.2	25
Esophagus	11.4	30.1	38
Bladder	6.7	13.1	51
Rectum	5.7	10.7	53
Stomach	11.4	19.2	59

*per 100 000 population
SOURCE World Health Organization, 1985-89

Table 5. Mortality from selected cancer sites: ages 35-74 (cont.)

Cancer site	Rate*		U.S. as % of France
	U.S.	France	
Pancreas	16.8	16.4	102
Prostate	23.5	22.1	106
Trachea, bronchus & lung	137.2	113.3	121
Colon & sm. int.	30.6	24.7	124

*per 100 000 population
SOURCE World Health Organization, 1985-89

Figure 3. Trends in selected causes of death at ages 75+: French males 1955-90

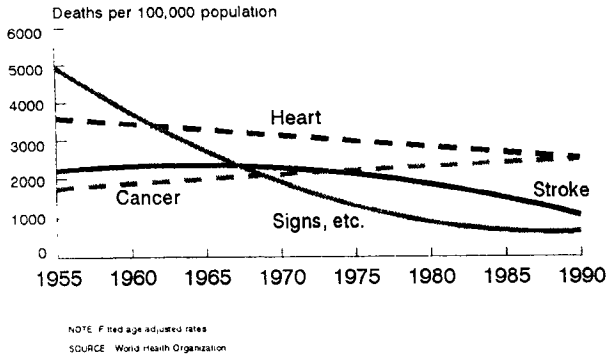


Figure 4. Regression of IHD mortality on animal fat consumption: 21 industrialized countries

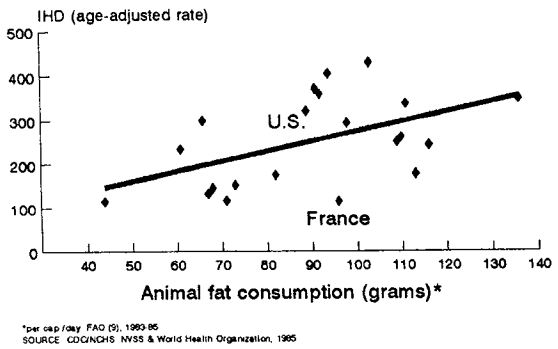


Table 2. Ecological regression analysis of IHD and national consumption of animal fat* and wine**

$$\widehat{IHD} = 232.84 + 1.09 \text{ fat} - 2.00 \text{ wine}$$

Variable	*df*	*Coefficient*	*Significance*
Fat consumption	1	1.09	0.14
Wine consumption	1	-2.00	0.0004
Intercept	1	232.84	0.006
Error	18	---	---
R-Square	0.6341	---	0.0001

*FAO (9), 1963-85 **IAR C(10), 1986
SOURCE World Health Organization, 1985

Figure 7. Mortality from respiratory cancer: males aged 45-64

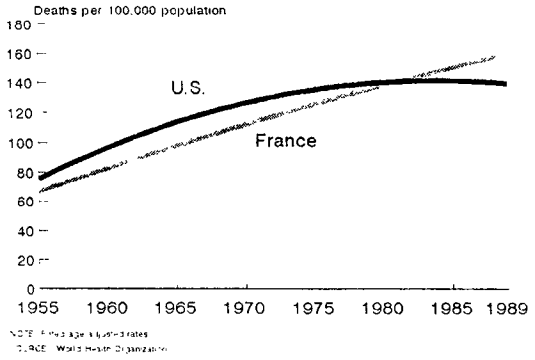
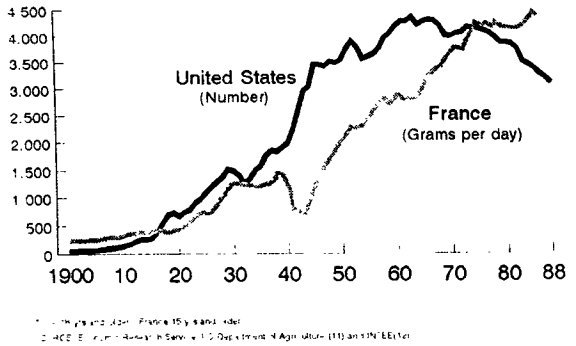


Figure 8. Annual per capita* cigarette consumption



THE INTERNATIONAL RANK OF THE UNITED STATES IN INFANT MORTALITY: ARE WE REALLY 24TH?

Robert B. Hartford, National Center for Health Statistics

(Not available for publication)

HEALTH EXPECTANCY OF THE ELDERLY IN XICHANG, CHINA AND CANADA: COMPARISONS BASED ON MEASURES OF AUTONOMY IN VARIOUS ACTIVITIES OF DAILY LIVING[1]

Edward Ng, Statistics Canada
Zong-kai Qiao, Russell Wilkins, Min Yang, Yajia Lan,
Xiaochuen Chen, and Yuanchang Xu

INTRODUCTION

Health expectancy indices such as disability-free life expectancy (DFLE) have now been estimated for many developed and lesser-developed countries (Robine and Romieu, 1991-1992). Calculations of health expectancy for the elderly in Xichang prefecture of Sichuan province (southwestern interior) of China were presented in an earlier paper (Qiao et al., 1992a), where the classification of persons into health states was based on a "severity of disability" scale.

This paper presents health expectancy calculations, where the classification of persons into health states was based on a simpler and more easily interpretable "degree of autonomy" scale, for the elderly (ages 65 and over) in Xichang. The similar calculations of health expectancy for all adults in Xichang were presented in another paper (Qiao et al., 1992b). The results for Xichang will be compared with those for Canada and for all of China.

MATERIALS AND METHODS

Xichang Health Quality Survey, 1989-90

From December 1989 to January 1990, the Xichang Health Quality Survey was conducted in Xichang prefecture, Sichuan province, China. The total population of Xichang prefecture was about 170,000 persons, 56% of whom lived in urban and 44% in rural areas. The survey used stratified random cluster sampling. The sample size was 21,812 persons, 17,179 of whom were 15 years of age or older, and 1,329 of whom were 65 years of age or older. The Xichang Survey excluded institutional residents and other members of collective households.

Canada Health and Activity Limitation Survey, 1986

The 1986 Health and Activity Limitation Survey (HALS) was a Statistics Canada post-censal survey that interviewed some 203,600 persons, 19,100 of whom were institutional residents.

China National Sampling Survey of the Handicapped, 1987

For China as a whole, calculations were based on the 1987 National Sampling Survey of the Handicapped (hereafter referred to as the Disability Survey of China). This was a multi-stage probability sampling survey that interviewed some 1,579,000 persons, chosen from about 369,000 households in China. The Disability Survey of China excluded institutional residents and other members of collective households.

Disability defined

The HALS in Canada used the "International Classification of Impairment, Disabilities and Handicaps" framework developed by the World Health Organization to measure the consequence of disease and operationalize the concept of "disability" through a series of questions concerning the "Activities of Daily Living" (ADL) and " Instrumental Activities of Daily Living" (IADL). For Canada, disability includes institutionalization and dependency or partial dependency in ADLs or IADLs.

The questions in the Xichang Survey concerned elementary activities of daily living (ADLs) and various instrumental activities of daily living (IADLs). For Xichang, disability includes any trouble with or inability to do ADLs or IADLs.

The definition of disability used in the Disability Survey of China included questions on visual, hearing and speech impairments as well as mental and physical disabilities (Grab et al., 1991).

Disability prevalence rates

For all China, only overall rates for any disability were calculated and given in Grab et al. (1991). In the Xichang Health Quality Survey, the ADL questions concerned bathing, dressing, going to the toilet, transferring, continence, feeding and grooming. The IADL questions concerned food preparation, housekeeping, shopping and handling money. An IADL question on travelling outside the home was abandoned as the essence of "local mobility outside the home" seemed to have been adequately captured in the shopping question.

For each question, the interviewer asked if the individual was able to do the activity without trouble, was having trouble with, or was completely unable to do the activity. That information was used to estimate the disability prevalence rates.

Respondents in Xichang were classified as unable to do at least one elementary activity of daily living ("ADL-unable"); able to do all elementary ADLs but unable to do at least one instrumental activity of daily living ("IADL-unable"); able to do all ADLs and

IADLs, but having trouble in at least one ADL or IADL ("ADL/IADL-trouble"); or able to do all ADLs and IADLs without trouble ("ADL/IADL-no trouble"). These categories were interpreted as a simple scale of "degree of autonomy" in the elementary and instrumental activities of daily living.

The Canadian results were based on direct questions about a person's dependency on other persons for assistance with essentially the same ADL and IADL items, instead of assuming dependency if a person was unable to do the activity. In the Canadian findings, the most severe category reported was "institutionalized or dependent for personal care or mobility within the home", which corresponds fairly closely to the "ADL-unable" category from the Xichang survey.

We also construct a combined "less severe" category including partial dependency for personal care or mobility within the home, and dependency or partial dependency with instrumental activities. That combined category roughly corresponds to the categories "IADL-unable" and "ADL/IADL-trouble" from the Xichang survey.

For each age group and sex and for each age group and place of residence (urban or rural) in Xichang and in Canada, population prevalence rates for each category of the autonomy scale were calculated by dividing the number of persons in each category by the relevant total subpopulation. Age-standardized disability rates were also calculated for both Xichang and Canada by various socio-economic dimensions using the total elderly population of Xichang as the reference.

Life table values

Abridged life tables were calculated by Chiang's (1984) method, using population and death data obtained from the 1990 Census of China. The population data were for 1 July 1990 while the death data were for the period 1 July 1989 to 30 June 1990. In Canada, the 1985-97 life tables specially tabulated by Statistics Canada (Wilkins and Adams, 1992, special tabulations) were used. For all China, life tables used were based on deaths for 1981 as shown in the 1982 Census of China (Calot and Caselli, 1988).

Health expectancy calculations

In Canada, health expectancy has been calculated for 1986-87 using a "degree of dependency" scale (Wilkins and Adams, 1992, special tabulations). In Xichang, for each age group, sex and urban/rural residence category, the number of life-years expected to be lived in each state of health or disability was obtained by applying the disability prevalence rates obtained from the survey to the life years derived from the life tables. The sum of this product from each age (x) onward was then divided by the number of life table survivors at each age, yielding the expected years of life in each health state.

RESULTS

For Xichang and Canada, the number of elderly survivors at different ages and the number of life years lived between successive ages are shown in Table 1. For ages 65 and over, the number of life table survivors and life years lived was always higher for females than for males and always higher for urban than for rural residents. The same pattern held for Canada, except that the urban and rural difference was minimal[2].

Table 1 Life table survivors (L) and life years lived (LL) by sex and by residence, Xichang, 1990 and Canada 1985-87

Place	Males		Females		Urban		Rural	
Age	L	LL	L	LL	L	LL	L	LL
Xichang								
65-74	67974	512748	74845	609823	73796	587872	66495	509315
75+	33369	225874	45428	293533	44265	328474	32436	185928
Canada								
65-74	77261	663379	87264	807267	83262	743417	81443	729515
75+	53384	486827	72296	868205	63904	691110	62919	678539

N.B. For Canada, more detailed information of L and LL were used for age 75+ in our calculation of health expectancy.

Table 2 Disability prevalence rates, by age, sex, and residence, Xichang, 1990 and Canada, 1986 (in %)

Place Age	ADL/IADL no trouble	ADL/IADL trouble	IADL unable	ADL unable	ADL/IADL no trouble	ADL/IADL trouble	IADL unable	ADL unable
	Males				Females			
Xichang								
65-74	63.81	25.70	5.35	5.14	50.64	35.68	8.55	5.13
75+	39.13	33.54	10.56	16.77	29.61	30.90	21.46	18.03
Canada								
65-74	79.95	5.30		4.72	75.32	19.65		4.90
75+	61.34	21.21		17.44	44.69	29.83		25.45
	Urban				Rural			
Xichang								
65-74	64.57	26.38	3.15	5.91	48.48	35.83	11.48	4.22
75+	42.99	31.31	9.35	16.36	22.22	32.78	26.11	18.89
Canada								
65-74	80.64	14.20		5.12	75.63	19.64		4.71
75+	50.08	28.42		21.49	52.37	25.87		21.73

The disability prevalence rates by age, sex, and urban/rural residence in Xichang are shown in Table 2. At ages 65 and over (where most severe disability is found), both males and females in Xichang and Canada were about equally likely to be unable to do ADLs, with elderly females aged 75+ in Canada having slightly higher disability rate. The percentage of population having any disability was higher for females than for males. This was true in both Xichang and Canada.

As shown in the lower portion of Table 2, elderly people living in rural areas in both Xichang and Canada were about equally likely as residents of urban areas to be unable to do ADLs, but were considerably more likely to be unable to do IADLs or to have trouble with ADLs or IADLs. The percentage of elderly population in Xichang having any disability was always higher for rural dwellers than for urban residents. However, the urban and rural rates in Canada were almost the same.

Figures 1 and 2 show age-standardized disability rates by sex for various socioeconomic dimensions in Xichang and in Canada. For both sexes in Xichang, the elderly with higher level of education tended to have lower disability rates (unable or trouble). In Canada, elderly people with higher income tended to have lower disability rates.

By place of residence, rates of disability (unable or trouble) among elderly in Xichang were in general higher in rural than in urban areas. In Canada, disability rates in rural areas are almost the same, if not slightly higher than those in urban areas.

Based on the combined ADL and IADL disability prevalence rates shown in Table 2, we calculated health expectancy in each state of health. The results are shown in Figure 3. In Xichang, while females at age 65 had higher life expectancy than did males, their disability-free life expectancy was lower. Females were expected to spend more years than males in various states of disability, although most of the additional years of disability for females were in the less serious categories (ADL/IADL-trouble and IADL-unable, rather than ADL-unable). Elderly people residing in urban areas had both higher life expectancy and disability-free life expectancy than did those living in rural areas.

Figure 1 AGE-STANDARDIZED DISABILITY RATES FEMALES, 65+, XICHANG, 1990 AND CANADA, 1986

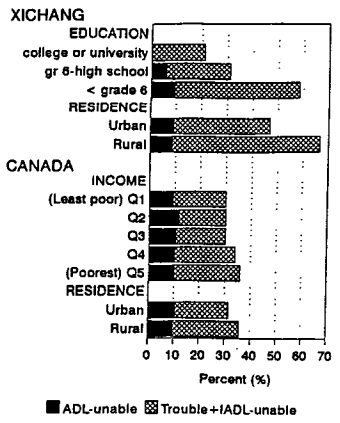
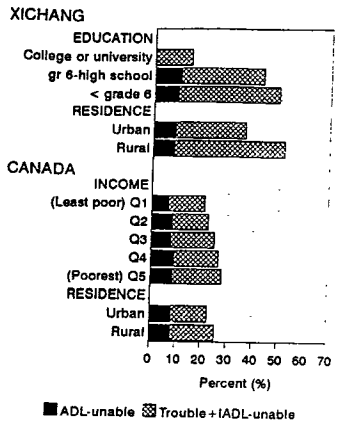


Figure 2 AGE-STANDARDIZED DISABILITY RATES MALES, 65+, XICHANG, 1990 AND CANADA, 1986



DISCUSSION

Since the additional years of life lived by females were usually accompanied by at least some degree of disability or loss of autonomy, the advantage to females in terms of life expectancy may become a disadvantage in terms of disability-free life expectancy. In Xichang, however, there was still a slight advantage to females in terms of expected years of life free of ADL-inability in Xichang.

As rates of disability in Xichang were higher and life expectancy was shorter in rural compared to urban areas, all of the health expectancy indices favoured the urban areas. This was not the case for Canada given the minimal rural-urban difference in both life expectancy and disability.

Questions requiring modification in Xichang for comparisons

Some of the interview questions borrowed from developed countries had to be modified to take account of the different socioeconomic and cultural context in China. In testing of the "toileting" question ("are you able to go to the toilet by yourself?"), it was found that many persons in older urban areas were unable or had trouble with this activity mainly because they had no toilet in the home. In such cases, "going to the toilet" implied leaving the house for a shared privy or toilet somewhere else in the neighborhood--so the ability to get around within the neighborhood was also required. Therefore, in case of reported difficulty in "going to the toilet", the interviewers were instructed to inquire whether the person was able to use a chamber pot, or had trouble doing so. Compilations of the "toileting" question were based on responses to the latter questions.

Another question which had to be restated by the interviewers in order to overcome cultural differences was the "bathing" question ("Are you able to bathe by yourself?"). Occasionally, some older persons in rural areas responded to the effect that they did not bathe. In such cases, the interviewer demonstrated the motions required, and asked if the person could do that (with or without difficulty), in order to determine if the person would be able to bathe, or would have trouble bathing.

Comparisons with findings for all China

In spite of the definitional difference in disability in the Disability Survey of China, to the extent that similar underlying problems are discerned by each survey, we would expect that the least severe categories of disability should differ the most, while the most severe categories should be more similar. Unfortunately, only overall rates for any disability were given Grab et al. (1991).

Comparing the Xichang and Canada findings, as shown in Figure 3, we see that life expectancy and disability-free life expectancy at age 65 for both sexes were consistently higher in Canada than in Xichang. For instance, in Canada, males at age 65 were expected to live for another 15 years out of which 11 years would be disability-free and 13 years would be autonomous (free of ADL dependency or institutionalization). In Xichang, males at age 65 were expected to live for another 11 more years, 6 years of which would be disability-free and 10 years of which would be autonomous.

In Canada, urban/rural differences in life expectancy, disability- free life expectancy and autonomous life expectancy were very small compared to the differences seen in Xichang. For both elderly males and females in urban and rural areas, expected years of non-autonomous life (total life expectancy less autonomous life expectancy) were much fewer in Xichang compared to Canada. This may indicate reduced survival for persons with severe disabilities in non-institutional settings in China, compared to their increased survival in institutional settings in Canada. By contrast, expected years with somewhat reduced autonomy (autonomous life expectancy less disability-free life expectancy) were similar in both Canada and Xichang.

Figure 3 HEALTH EXPECTANCY AT AGE 65 CANADA, 1986 AND XICHANG, 1990

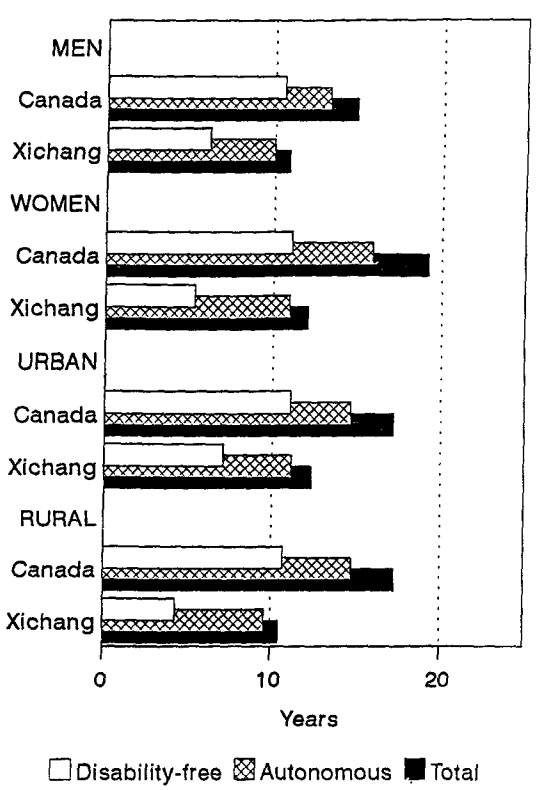


Table 3 Health expectancy for elderly by sex and by residence, Canada 1986, Xichang 1990, and all China 1987 (in years)

Dimension	Total (life expectancy)			Disability-free[a]			Autonomous[b]		
Age	Canada	Xichang	China	Canada	Xichang	China	Canada	Xichang	China
Males									
65	14.9	10.9	12.5	10.6	6.1	8.9	13.3	10.0	n/a
75	9.1	6.8	7.4	5.4	2.6	4.2	7.4	5.7	n/a
Females									
65	19.2	12.1	14.6	11.1	5.3	9.9	15.9	11.0	n/a
75	12.0	6.5	8.7	5.0	1.9	4.7	8.6	5.3	n/a
Urban									
65	17.2	12.4	12.8	11.1	7.1	n/a	14.7	11.2	n/a
75	10.8	7.4	7.5	5.1	3.2	n/a	8.1	6.2	n/a
Rural									
65	17.3	10.5	12.4	10.7	4.3	n/a	14.8	9.6	n/a
75	10.8	5.7	7.2	5.2	1.3	n/a	8.1	4.6	n/a

Note: a For Canada, disability includes institutionalization and dependency or partial dependency in ADLs or IADLs. For Xichang, disability includes any trouble with or inability to do ADLs or IADLs. For all China, disability includes visual, hearing and speech impairments, as well as mental and physical disabilities.

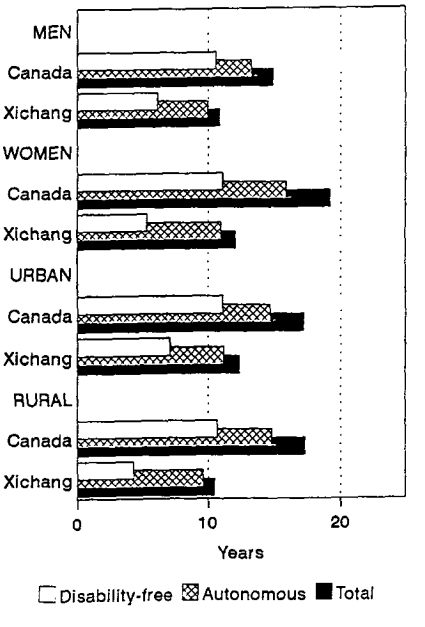
b For Canada, autonomous excludes years of institutionalization and dependency for personal care or mobility within the home. For Xichang, autonomous excludes years unable to do any ADL.

Health expectancy results based on the Disability Survey of China used mortality data from the 1982 National Census of China while the Xichang study had the benefit of the 1990 Census results. Health expectancy results would be influenced by any changes to the mortality conditions of the Chinese population in that 8 year period.

In Xichang, where data on institutionalization were available, we found that even at age 85, less than 0.4% of the population was living in institutions. Assuming that rates of institutionalization in other parts of China were similar to those in Xichang, exclusion of the institutional population would not have had much effect on the results.

While neither the life expectancy nor disability-free life expectancy results are strictly comparable for the reasons stated above, a comparison of the Xichang and country-wide data for China showed that both life expectancy and disability-free life expectancy at selected ages were consistently lower in Xichang than in China as a whole. Figure 4 and Table 3 show that, at age 65, remaining life expectancy was 11 years for males and 12 years for females in Xichang, compared to 13 years for males and 15 years for females in China as a whole. The lower life expectancy of Xichang might have been expected, given the generally higher than average mortality rates of Sichuan province (Calot and Caselli, 1988). At age 65, remaining disability-free life expectancy was 6 years for males and 5 years for females in Xichang, compared to 9 years for males and 10 years for females in China as a whole. Thus at age 65 in Xichang, males could expect to lose about 5 years of life to disability, compared to 4 years for males in china as a whole. Elderly females in Xichang could expect to lose about 7 years, compared to 5 years for females in China as a whole.

Figure 4 HEALTH EXPECTANCY AT AGE 65 CANADA, 1986 AND XICHANG, 1990



CONCLUSION

In the language of the World Health Organization's International Classification of Impairments, Disabilities and Handicaps (1980), our "degree of autonomy" scale is a measure of handicap. Reduced autonomy in the various activities of daily living is a possible consequence of health impairments (resulting from disease or injury or congenital anomalies) which are serious enough to result in functional disabilities (despite any prostheses or aids which may or may not be employed to overcome the effects of the impairment).

However, a reduction in autonomy is not a necessary result of disabilities--it depends on the environmental context in which the person is living. Moreover, impairments do not necessarily result in disability, but they are more apt to be disabling where prostheses and aids are not readily available. For example, in China, visual impairments which would not be disabling in a developed country will frequently result in disability and handicap, simply because corrective lenses are not readily available to those who could benefit from them.

In developed countries where "professional" home care and institutional care for the elderly are more widely available (and sought after), the service implications of ADL- and IADL-disability questions are more relevant. However, in a situation of few formal services for the disabled, and where almost all home and personal care services are provided informally by family members or neighbors, these results are less relevant for services planning. Nevertheless, they still provide good summary indices of population health status, which are reasonably comparable to those available from developed countries.

Compared to Canada, the most striking differences concerned expected years of life with ADL-inability, which were much fewer in Xichang. Urban/rural differences in health expectancy were also much greater in Xichang than in Canada.

The comparisons in Xichang and Canada, were relatively straightforward because the same ADL and IADL items were used. However, comparisons with the Disability Survey of China were much more difficult, due to major differences in what was measured. In order to make the comparisons more meaningful, we would have preferred, of course, to have had comparable measures of ADL and IADL dependency from the Disability Survey of China.

ENDNOTE

1 This project is a collaboration between researchers from Canada (E. Ng and R. Wilkins of Statistics Canada) and People's Republic of China (Z-K Qiao, M. Yang and Y. Lan of West China School of Medical Sciences; X. Chen of Hunan Provincial Anti-epidemic Centre; and Y. Xu of Xichang Anti-epidemic Centre). Dr. Z-K Qiao was the principal investigator of the Xichang Health Quality Survey.

2. In Canada, urban areas were defined as communities with population size over 1,000,000 while rural areas were communities with less than 9,000. The choice of the two extreme groups in the community size continuum was to maximize the rural-urban difference in disability in Canada.

REFERENCES

Calot G. and Caselli, G. (1988): La mortalité en Chine d'après le recensement de 1992. I. Analyse selon le sexe et l'âge au niveau national et provincial. II. Tables de mortalité par province. Dossiers et recherches no 16. Paris: Institut national d'études démographiques (INÉD).

Chiang, C.L. (1984): The Life Table and Its Applications. Malabar, California: Robert E. Krieger Publishing Company.

Grab B., Dowd, J.E, and Michel, J.-P. (1991): Estimates of disability-free life expectancy in China, 1987. Paper presented to the 4th Meeting of the International Network on Health Expectancy, Noordwijkerhout, June.

Qiao, Z.-K., Wilkins, R., Yang, M., Lan, Y., Chen, X. and Xu, Y. (1992a): Health expectancy of the elderly in Xichang, China. Paper presented to the 5th Meeting of the International Network on Health Expectancy, Ottawa, February.

Qiao, Z.-K., Wilkins, R., Yang, M., Lan, Y., Chen, X., Xu, Y. and Edward Ng (1992b): Health expectancy of adults in Xichang, China, 1990: autonomy in various activities of daily living. Paper presented to the 6th Meeting of the International Network on Health Expectancy, Montpellier, October.

Robine, J.-M. and Romieu, I. (1991-1992): Health Expectancy. Bibliography Series No. 1, 2, 2 supplement, 3. Montpellier: Institut National de la Santé et de la Recherche Médicale and National Institute on Aging, January 1991, April 1991, June 1991, February 1992.

Wilkins, R, and Adams, O. (1992): Independence and autonomy expectancy of Canada's seniors. Paper presented to the 5th Meeting of the International Network on Health Expectancy, Ottawa, February.

World Health Organization (1980): International Classification of Impairments, Disabilities, and Handicaps. A Manual of Classification Relating to the Consequences of Disease. Geneva: World Health Organization.

ACKNOWLEDGEMENTS

The National Natural Science Foundation of China funded the Xichang Health Quality Survey and provided support for related analyses of health expectancy.

Session O

Issues in Confidentiality

Toward the year

2000

Refining the measures

PRIVACY POLICY IN AN ELECTRONIC WORLD

Joan Turek-Brezina, Department of Health and Human Services

My presentation focused on the HHS Task Force on the Privacy of Private Sector Health Records, of which I am the chair. The task force was established in 1990 and plans to have a written report produced by late 1993. Because a written report and recommendations will be produced, I will only outline the tasks being pursued by the task force.

The task force was originally charged with:

- examining the extent to which there are problems with the use of personally identifiable medical and other health-related records in the private sector;
- identifying what needs for health information exist in the public and private sector;
- reviewing current laws and practice related to the privacy of private sector health records;
- recommending steps that the Federal government could appropriately pursue to protect nonfederal record systems, if problems were identified.

However, emerging events have refocused the task force's mission to concentrate on the move toward electronic records within the context of health care reform.

The current vision of future health record systems is:

- A comprehensive longitudinal computer-based patient record (clinical, financial and research data, including diagnostic images or pictures).
- A "national" electronic network for accessing this health record for a variety of purposes.
- Use of a "smart card" for purposes ranging from providing health insurance coverage information to providing a conception-to-death record of all health care.
- Use of unique patient-specific identifiers.

The task force held a series of meetings with representatives from professional associations, privacy advocates, the research community, the insurance community, etc. In February 1993 the task force sponsored a conference "Health Records: Social Needs and Personal Privacy". The following major policy issues/questions have been identified by the task force for consideration in developing its report and recommendations:

- What records should be governed by principles of health record privacy?
- Should "especially sensitive" records receive special treatment?
- What is the impact of automation on the privacy of health records?
- What concerns are raised by use of a unique identifier for health records?
- At what level should legislation be enacted to protect health information?
- What constitutes informed consent for disclosure of health information and under what circumstances may informed consent not be a sufficient basis of disclosure?
- Under what circumstances should record keepers be allowed to disclose records? How will policies and requirements be enforced?
- What type of structure is needed to oversee privacy policy, confidentiality and security matters and violations?
- What training, education and awareness programs should there be both for individuals regarding their own records and for those working with health records?

TOWARDS COMPILING BETTER HEALTH INDICATORS THROUGH DATA LINKAGE

Brian K. Gallagher, New York State Department of Health
Michael Sawyer

Overview

The New York State Department of Health (DOH) maintains one of the best sources of health indicators for the birthing population in the state, i.e., the birth certificate (BC) file. The data is more useful, however, when linked to other data sources. The files are linked taking special care to ensure that confidentiality is maintained. The linked files are used to analyze prenatal care and birth outcomes of participants in Medicaid and other programs, e.g. WIC.

I will describe how the DOH linked its BC file to the data base of a second state agency, and describe the benefits we derived from linking the data.

Background

The NYS DOH administers the Prenatal Care Assistance Program (PCAP), a Medicaid program which provides comprehensive prenatal care services to low income women.

Obstetricians who care for low income women receive a fee per visit from Medicaid. If an obstetrician offers the PCAP model of care, they receive an enhanced Medicaid fee. The New York State Department of Social Services (DSS) provides payment for these prenatal services, while the DOH sets standards of care for the PCAP program, recruits health care providers and monitors the quality of care.

To analyze the impact of the PCAP model of prenatal care, client specific data records from both agencies are combined. The BC file provides maternal demographic information, maternal medical obstetrical history, and outcome data (gestation, birthweight, and congenital malformation indicators). The DSS Medicaid claims file identifies the provider of prenatal care and also includes information on what services were billed during the prenatal period.

Why Match ?

Enhanced Research Possibilities

The DOH BC file is a robust set of demographic and medical risk data, but by combining it with Medicaid claims data, it is even more useful. The linked records are used to ascertain program participants' health status and quality of care, as well as to identify the determinants of poor birth outcome.

Program Evaluation

The linked file allows us to assess the effectiveness that program changes have had on pregnancy outcomes. Linked files are also used to augment the health status data of program participants and evaluate the effectiveness of specific interventions (i.e. smoking cessation).

Data quality

The two data sets have several fields which contain the same information. The combined records allow cross checking between the two separately maintained data bases. This allows us to detect data quality problems which might otherwise go undetected.

Ease of matching

With both data sets in an electronic format, a mainframe computer is used to perform the matching. We match our 300,000 record birth certificate file with the 25,000 record Medicaid file using an IBM 3090 running SAS on the MVS operating system. The decision criteria and software required to match a majority of the file (50 - 60%) are not very complex. A larger percent of the file can be matched (we achieve rates in the 90 - 93% range) with more complex decision criteria and a greater expenditure of programming and match verification effort.

Pre-match Considerations

Three major areas of concern had to be addressed before the records could be matched.

Commitment To Project

First, both agencies had to 'buy into' the project. This was not a trivial matter in our case, since DSS is primarily concerned with the cost effectiveness of the program, while the DOH was interested in the impact of the program on birth outcomes. Once both agencies had their project champions, the administrators of each agency were approached to secure a commitment for the project to ensure on-going support.

Institutional Barriers

Secondly, institutional barriers had to be addressed. The two agencies involved had never entered into an agreement to share data, and overcoming this reluctance was a lengthy process.

The barriers within each agency also

had to be addressed. In the DOH, the Bureau of Production Systems collects birth certificates and maintains the BC file. The unit within DOH which was conducting the analysis had to assure them that their file was not being used inappropriately. By the time commitments were secured from both agencies to support the project and institutional barriers were addressed, five years had elapsed.

Confidentiality

Ensuring the confidentiality of the two data sources was the final pre-match consideration.

The Department of Health and Department of Social Services both have extensive safeguards on data with personal identifiers. Both agencies have a process in place to review requests for using data with personal identifiable information. This internal review process is used to approve all access to confidential data whether the request is from within the agency or from a private party outside of the agency.

A memorandum of understanding (MOU) was drafted by the two agencies to ensure the confidentiality of the records. The owners of the data, the counsels office's for each agency and the program staff which was interested in analyzing the data were all involved in drafting the MOU between the two agencies. The MOU describes the seven step process used to ensure that confidentiality is maintained.

The process described below has two important features. First, personal identifiers are removed from the file as soon as the matching is completed. Also, the files do not contain any outcome or service utilization data while personal identifiers are still on the file.

MOU main points

1. DOH gets records with DSS id number, womens' names and other identifiers from DSS.

2. DOH matches women to BC records.

3. DOH eliminates the DSS personal identifiers from the file.

4. DOH adds an encrypted DOH id number to each matched record.

5. DOH sends DSS a file with
   - DSS id number
   - encrypted DOH BC number
   - DOH BC outcome fields of interest

6. DSS adds data to the file, and encrypts the DSS id number.

7. DSS sends back to DOH the same file with
   - encrypted DSS id number
   - encrypted DOH BC number
   - DOH BC outcome fields of interest
   - DSS service provider fields of interest

Precautions are taken during the implementation of the project to ensure that confidentiality is protected. Access to the data is limited to the project director and the data specialist. The electronic file is accessable only by the data specialist. These electronic files are protected through a mainframe computer security system of identification codes and passwords.

Any hardcopy which includes personal identifiers are accessible only to the project director, and are stored in a locked file cabinet when not in use. The hardcopy which includes personal identifiers is destroyed within a specified time frame after being generated, and the computer file abstracts are destroyed after all matching has been completed. The owners of the file also reserve the right to inspect the materials relating to the project at any time, to insure that appropriate safeguards are in place.

Matching Administrative Records

Both the DOH and DSS files are large files which were developed to fulfill administrative purposes. The DSS file was created to keep track of medicaid claims for individuals. The DOH birth certificate file was created to record each birth to a NYS resident. The files were not developed for matching purposes, so neither file contains a unique personal identifier. Therefore, any data elements that are common or comparable to the two files are used to match the files.

The principal behind matching is straight forward. Bring together the information which pertains to an individual from two separate sources. The mechanics of matching are more complicated. Below find a step by step description of how the DOH linked its BC file.

1. Gather the data. The DSS supplies the DOH with a list of women who were enrolled in Medicaid and had a hospital admission for delivery. The list includes the receipients name, social security number, mother's date of birth, county of residence, zip code of residence, hospital code were the birth took place, and the date the delivery was billed.

   The DOH BC file has several of the same fields, and several comparable fields that are useful for matching, including:
   mother's first, middle and last name
   mother's maiden name

mother's date of birth
mother's social security number
mother's county of residence
mother's locality of residence
mother's residence zip code
mother's mailing address zip code
infant's last name
infant's date of birth
infant's facility of birth id code
infant's facility of birth locality
infant's county of birth
father's last name
father's social security number

2. Get to know your data. The data in both the DSS file and the DOH BC file are analyzed to determine the ranges and frequencies of variables. Data quality and completeness are checked, so that erroneous characters or junk can be eliminated before matching begins. Getting to know your data is very important and will help you select appropriate variables for matching.

   Variables or combinations of variables which uniquely identify an individual record are best for matching. These variables include social security number, first and last name combined, or date of birth combined with last name.

3. Match the files. The actual matching of files is an iterative process which may take several combinations of linking variables before a sufficient number of records are matched.

   Match the files using your first combination of variables so that every record in your first data set is mathced to every record in the second data set where the linking field agrees.

4. Verify the matches. Compare the comparable fields which are available from each file, side by side, for every potential match (Table 1).

   Although manual verifiction is tedious and time consuming it is important, because each data set has errors. Just because two SSN's match, that does not mean you have a pair of records that is a true match. Numbers can be miscoded, and women can use their spouses SSN, so verification is necessary to ensure true matches.

   In my experience, social security number has been the most accurate field for matching. Unfortunately, not all of the records in the files I work with have the social security number. With our files, 30% are matched by using SSN alone.

5. Save the good matches. Once a pair of records is matched, the DSS ID number and the DOH ID number are recorded and written to a 'matched' file. These id numbers are then eliminated from the master files, and the remaining unmatched records are re-matched using the next linking variable.

6. Pick the next linking variable. In my experience, last name plus first name is a good linkage variable, with some caveats.

   Depending on the characteristics of the clients in your data base, name variables may be problematic. If your clients are homogeneous in ethnicity, then you may have thousands of Mary Smith's or Virginia Cruz's. Some ethnic groups use maiden name and last name interchangeably. Also, watch out for compound or hyphenated names.

   With our files, we also have an anomoly with last name registration which is attributable to the nature of the files. For example, when the mother registers herself for Medicaid services, she is likely to use either her maiden name or her married name. When the child is born, it is much more likely that the women's married name, or the father's last name is used.

   After accounting for these name field characterisitics, match the files using first and last name combined. Compare all the fields which are available from each file (Table 2) for every potential match, including mother's last, mother's maiden and

Table 1.

File	First Name	Middle Initial	Last Name	Maiden Name	Father's Last Name	Age	SSN
DSS	Nora	K	Xenops	Tricops		32	123-45-6789
DOH	Nora	K	Xenops	Smith		32	123-45-6789

Table 2.

File	First Name	Middle Initial	Last Name	Maiden Name	Father's Last Name	Age	SSN
DSS	Nora	K	Tricops			32	123-45-6789
DOH	Nora	K	Xenops	Smith	Tricops	32	

father's last name.

7. Next steps. Continue picking linkage variables and matching the files until you have matched a sufficient number of records.

Conclusion

The linkage project was difficult to initiate and ensuring the confidentiality of program participants is a rigourous exercise. But, the enhanced data set which we develop by combining data from seperate files is useful for many applications.

THE UNIQUE RECORD NUMBER: AN ALTERNATIVE TO TRADITIONAL PERSON IDENTIFIERS

Andrew Asher, Mathematica Policy Research
David Edson, Embry Howell, Karen Pence

To answer many key policy and programmatic questions, researchers and program analysts require person-level information that provides an accurate and complete assessment of conditions and service delivery patterns. To obtain such data, person identifiers are often necessary to link person and encounter records from multiple sources. In many instances, the use of traditional unique identifiers such as name, date of birth, or social security number is unacceptable because disclosure of such information may have negative consequences for the individuals being studied. The Unique Record Number (URN) is an alternative to traditional identifiers. Based on encrypted portions of traditional identifiers, the URN is a mechanism that links person-level data from multiple sources without the use of traditional identifiers that could ultimately disclose sensitive personal information.

The URN is a major component of the Uniform Reporting System (URS), a proposed system developed by the Health Resources and Services Administration (HRSA), with assistance from Mathematica Policy Research, Inc. Under the URS, providers will report person-level data using the URN in place of other identifiers. The URN will be assigned by service providers to each HIV-infected client receiving eligible services. Since all providers will be expected to use the same encryption algorithm, the URN allows for the generation of unduplicated counts of clients, many of whom may be served by more than one provider. The URN approach has potentially wide applicability, extending beyond HIV data collection and analysis efforts.

This paper discusses the importance of person identifiers and presents an overview of the URN system. Section I describes the role of person identifiers in research and program management, identifies criteria on which to judge them, and discusses the problems with traditional identifiers. Section II introduces the URN approach. The role of the URN in the Ryan White Comprehensive AIDS Resources Emergency (CARE) Act Uniform Reporting System is discussed in Section III, followed by a review of key URN-related operational issues in Section IV. Conclusions are presented in Section V.

I. The Role of Person Identifiers

Person identifiers are an important tool in health research and program administration. They can link data sets from multiple sources, providing researchers and program analysts with a more accurate representation of the characteristics of individuals or the services they receive. Identifiers can also be used to link data across several time periods to create a longitudinal file that may be used to examine how individuals or care delivery patterns change over time.

Identifiers are evaluated according to three criteria: accuracy, replicability, and confidentiality.

- **Accuracy.** Accurate identifiers have a minimum number of "false positives" and "false negatives" when two or more databases are linked. A false positive is the linking of records on different individuals. A false negative is the failure to link records from the same individual. Inaccuracies can occur for a number of reasons when information is collected, transcribed, or automated.

- **Replicability.** Identifiers should be easily replicated. They should be based on obvious characteristics of an individual that do not change over time and that are typically obtained in service settings, such as date of birth or gender. In addition, program personnel should be able to assign identifiers simply and conveniently. That is, identifiers should not increase the likelihood that personnel will either neglect to assign them or do so erroneously.

- **Confidentiality.** Identifiers should protect confidentiality in two ways. They should prevent unauthorized parties from being able to link them to specific people. They should also minimize the number of authorized individuals that need to have access to confidential information, such as name or social security number.

Implicit in traditional identifiers is the trade-off between confidentiality, and accuracy or replicability. For instance, relative to alternatives, traditional identifiers such as name and date of birth are accurate and easily replicable, but do not adequately protect confidentiality. Administrative I.D. numbers, such as Medicaid numbers, are generally adequate for protecting confidentiality, but cannot be replicated outside the context in which they are used.

II. The URN Approach

The need to protect highly sensitive information is not unique to research or program management. Cryptographic methods are used in a variety of commercial and military applications to protect sensitive communications, and to store and transmit data securely. The URN holds promise as an accurate, easily replicable, and highly confidential approach largely because of the cryptographic method on which it is based: an encrypted string (the URN) is generated from an unencrypted string based on components of traditional identifiers--ideally those that are the easiest to obtain and the most accurate. Unlike many cryptographic methods, this one uses a message digest algorithm with a one-way mathematical function.[1] This function generates URNs from unencrypted strings, but no related function can be used to generate unencrypted strings and identify individuals from a list of pre-existing URNs. Because of this feature, the URN offers a level of security that surpasses that of most traditional identifiers. Once generated, the URN becomes the sole mechanism by which records are linked, and all traditional identifiers are deleted from the record.

III. The Role of the URN in the Ryan White Uniform Reporting System

The Ryan White CARE Act of 1990 is a major source of funding for medical and support services for persons with AIDS and HIV disease. The Act provides grant funding to states and metropolitan areas that have been disproportionately affected by the AIDS epidemic. These funds are administered by the Bureau of Health Resources Development (BHRD) within HRSA, and the grantees are responsible for funding local service providers, state drug assistance programs, and state health insurance continuation programs.

The purpose of the URS is to help HRSA and Congress monitor these programs and assess the extent to which the goals of the act are being achieved. This will be accomplished by requiring providers to collect person-level data on individuals being served by the CARE Act. The data will be used to generate unduplicated counts of clients, identify the characteristics of persons receiving services, and review the distribution of and gaps in services provided to key client subgroups.

In the URS, service providers who receive CARE Act funds will forward client-level records to their grantee (either a state or local government agency). These records contain the URN and are stripped of traditional identifiers before they are sent to the grantee. The grantee will work with providers to ensure that these data are accurate before preparing summary reports and electronic files that are sent to HRSA. HRSA will use this information to create a set of national tables that provide counts of the number of recipients and identify their characteristics, and to prepare ad hoc analyses. It is expected that HRSA will provide grantees with information to assist them in program management, and that grantees will present this type of information to their providers.

Because clients are likely to receive services from more than one provider and would therefore be counted more than once, HRSA recognized the need to use a unique identifier that could link these records. Given the significant confidentiality concerns raised by person-level HIV reporting, traditional identifiers such as name or social security number could not be use in the URS. While an administrative case number system could have been used to assign numbers that are unrelated to the identify of the client, such a system was infeasible. It would have required each provider to obtain numbers from a central source, and each client to remember and use his or her number.

The precise construction of the proposed HRSA URN is shown in Figure 1. The unencrypted string is based on the first and third letters of the first and last names, all 6 digits of the date of birth, and a 1-digit code for gender. This 11-digit alphanumeric code is encrypted into a nearly unique[2] 9-digit alphanumeric string by the message digest algorithm. As the figure shows, the URN does not resemble the unencrypted input string. It is noteworthy that even a 1-digit change in the input string will generate an entirely new and unrelated output string. (For instance, when the date of birth for the example shown in Figure 1 is changed from 04/23/59 to 04/24/59, every digit of the URN changes.)

IV. Operational Issues Associated with the URN Approach

Three broad sets of operational issues are associated with the URN approach: those related to design, to the needs and capabilities of system participants, and to security. This section reviews these issues and provides relevant examples based on HRSA's experience in implementing the URN approach on a trial basis in the URS.

In designing the URN, the criteria of accuracy, replicability, and confidentiality must carefully be considered. Concerns about accuracy were the impetus for the decision to base the HRSA URN on portions of data elements that are unlikely to change and that were found to be accurately recorded in administrative databases.[3] Replicability was responsible for an interest in basing the URN on elements that already tended to be collected by providers. Concerns about confidentiality influenced the construction of the input string. While the input string alone was not intended to be a major safeguard, its use of only portions of the name increases its ability to protect an individual's identity.

The needs and capabilities of system participants must be considered by the agency overseeing the implementation of the URN. For HRSA, this involved gaining the trust of providers and grantees, offering training and technical assistance, and making appropriate technology available. To gain the trust of participants, HRSA needed to convince them that HRSA shared their commitment to protecting client confidentiality. The effort devoted to the development of the URN approach and related administrative procedures were important in persuading them. Once providers and grantees were satisfied with HRSA's commitment to protecting confidentiality, HRSA provided training and technical assistance for implementing the URN. To actually generate the URN on site, providers also needed access to technology options. To aid providers and grantees in the creation of the URN (and in the collection and processing of URS data), HRSA supported the development software systems, related training activities, and the acquisition of hardware.[4]

Finally, agencies interested in implementing the URN should consider whether it meets their security needs. Compared to traditional identifiers, the URN provides a significantly higher degree of security. While providers and grantees have been advised to handle URNs as if they were names, encryption eliminates the possibility that casual observations could lead to identification. Also, because no identifying information leaves the provider agencies, the URN limits access to identifying information. And as mentioned, the algorithm used to generate the URN is based on a one-way mathematical function that cannot be used to regenerate the unencrypted string and identify clients.

Though very promising, the URN approach by itself does not provide absolute security. Inadequate provider or grantee security procedures for storing, using, or transmitting person-level data could enable unauthorized parties to obtain URNs. If these URNs were obtained, unauthorized parties could initiate the following "attacks":

(1) an external database attack in which person-level administrative records, such as those found in employee databases, are used to generate URNs and then compared to a pre-existing list of URNs to identify HIV service recipients, and (2) a dictionary attack, in which all possible combinations of input strings and corresponding URNs are generated and compared to a pre-existing list of URNs. The probability of actually identifying an individual through these types of attacks is difficult to quantify and depends on several factors, including the extent to which the input and output strings are unique, and on the extent of errors in the data. HRSA is conducting several analyses to assess both the degree of risk posed by these attacks and the most effective ways to further reduce such risks.

V. Conclusion

The URN is a promising alternative to traditional identifiers. As part of the URS development effort, HRSA is field testing the system with 15 grantee agencies and their providers to assess the feasibility of the URN approach and, more broadly, of the URS. HRSA is also sponsoring complementary studies to examine the vulnerability of the URN and to identify ways to improve its construction or implementation. These efforts will help HRSA identify improvements to the URN and build a greater understanding of its strengths and limitations.

Acknowledgments

The authors would like to thank Robert Baitty and Wendell Pope of the Health Resources and Services Administration and Burton Kaliski of RSA Data Security, Inc. for their work on the URN approach.

Notes

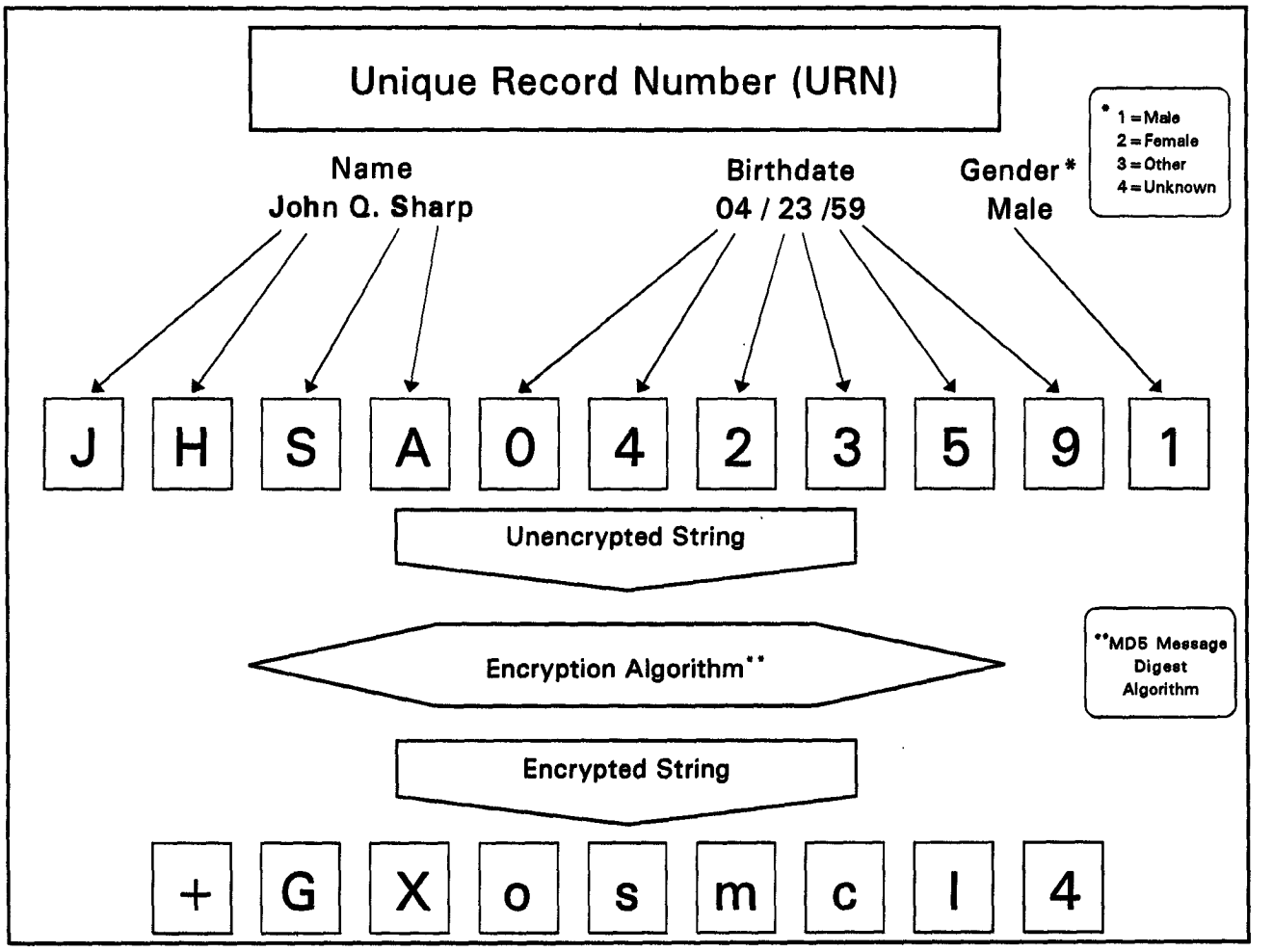
1. R. Rivest. "RFC 1321: The MD5 Message-Digest Algorithm." Internet Activities Board, April 1992.

2. We refer to the URN as being "nearly unique" because there is an extremely small probability that different input strings will result in the same output string.

3. Internal HRSA analysis.

4. This hardware and software was made available to help grantees and providers fulfill URS responsibilities, including the generation of URNs, and it provides them with the added benefit of helping them complete other agency functions.

Figure 1. Construction of the URN



CALIFORNIA'S EXPERIENCE: SSN AND CONFIDENTIALITY

Eleanor Meux, California Office of Statewide Health Planning & Development

The State of California instituted its hospital inpatient discharge data program via legislation in 1981. The program is administered by the Office of Statewide Health Planning and Development, which is a freestanding department within the California Health and Welfare Agency. The purpose of our program is to make data publicly available in order to answer aggregate statistical questions, particularly about hospitals, without allowing the identification of individual patients. Our data elements, which are established in state law, do not include the name and address of the patient. The data are reported from the medical record, not from a billing form, and follow definitions that we establish in regulation.

Our data elements did not include any form of personal identifier. After a few years, it was realized that lack of a unique identifier limited the usefulness of the data, due to the inability to do a number of things requiring a unique identifier, such as to identify readmissions, to follow patients between hospitals, or to link to mortality files in order to obtain post-discharge deaths. After considering possible unique identifiers, such as name, address, medical record number, and combinations of these, it was concluded that the social security number (SSN) is the only realistic identifier that is almost universally present and reliable in linking records for one individual across hospitals and years. Social security number was added to our data elements by legislation in 1988. The legislation also added this statement: "No person reporting data pursuant to this section shall be liable for damages in any action based on the use or misuse of patient-identifiable data which has been mailed or otherwise transmitted to the office pursuant to the requirements of this subdivision." This statement relieves hospitals of liability and leaves the responsibility on our Office. The legislation also added another statement: "Patient social security numbers and any other data elements that the office believes could be used to determine the identity of an individual patient shall be exempt from the disclosure requirements of the California Public Records Act." The law already said: "It is the expressed intent of the Legislature that the patient's rights of confidentiality shall not be violated in any manner." [All quotes are from section 443.31(g) of the California Health and Safety Code.]

California's constitution has an explicit constitutional right to privacy, added by voters in 1972. The California Information Practices Act of 1977 provides, among other things, restrictive conditions for the disclosure by a state agency of personal information in a manner that could link the information to an individual. The Act includes criteria for disclosure, such as legal authority, informed consent, and whether the use of the information is consistent with the purposes for which it was collected.

We do not publicly release data such as dates of birth, admission, discharge, and procedures, because they could possibly serve to identify an individual. We have established a policy under which we review requests for data that we do not release publicly. After close review, including review by a committee for the protection of human subjects, and with appropriate safeguards, we have provided a few requestors with the actual dates of birth, admission, discharge, and procedures on individual patient records. No SSN has been released to anyone.

The timing of our program is fixed in law: the reporting period is six months (calendar semiannual), and the data are not due until six months later. Then we are allowed another six months to edit the data. After all data are edited, they are publicly released. So, although the law adding SSN was passed in 1988 and was effective for July 1990 discharges, we have as of now released only 18 months of data containing "a unique identifier."

Let me describe what I mean by the "unique identifier" that we do release. It was obvious to us that we could never publicly release any SSN, given the confidentiality restrictions in California law. So, we had two choices in order to provide our data users the benefits of having a unique identifier: we could either make every data user come to us when they wanted to link records, count unduplicated individuals, etc., or we would have to find some way to "disguise" the SSN before public release. I developed a method of encrypting each SSN into what we call a record linkage number (RLN), which consists of nine alphanumerics. The SSN cannot be obtained from the RLN, not even by our computer programmer who has access to the encryption program. The RLN has no external utility in identifying an individual. It is exceedingly improbable that two SSNs would result in the same RLN (the probability value contains at least five zeros). And, the same SSN will always result in the same RLN. The RLNs can be released publicly without fear that anyone can derive the encryption technique from the released values. I have written up the method for publica-

tion, but it is still under review and so copies are not yet available.

We decided to encrypt the SSN into its RLN, and make the RLN publicly available. We considered replacing the SSN on our database with the RLN, so there never would be any chance of accidental (or even deliberate) release of an SSN. But it was decided that it would be a "waste" to throw away information, once collected.

When hospitals first heard about the new requirement to report SSN, they were very concerned, particularly psychiatric and substance abuse hospitals. But protests ceased after information was circulated about how we were encrypting the SSN. I am not aware of anyone who still has concern about how we are handling this confidential information.

We also had to resolve confidentiality questions in collecting, handling, and editing SSNs. Although only about two percent of our records are submitted on paper abstract forms, that still amounts to 37 thousand sheets of paper each six months, given the scope of our program (3.7 million discharge records a year, from almost 600 hospitals). We bought a shredder and now shred these onsite, after the information is keyentered and verified. The bulk of data is submitted on computer media (tapes and diskettes), and we developed security procedures for their secure transportation and storage.

We edit all our data very extensively. The encryption of the SSN is not done until editing is complete, so that we can edit the SSN itself. We started out with minimal edits on SSN: it had to consist of nine digits, but not nine identical digits; if it was reported on a patient under 60 days of age, it was probably an error; and, if it was not reported and the expected payer was Medicare, we would ask the hospital about it, because we thought it should have been available.

After about six months with these edits, we obtained definitive information from the Social Security Administration, and added more edits on SSN: the last four digits could not be 0000, and the first three digits could not be 000, 750, 751, greater than 755, or between 649 and 700. After looking at records for the first set of available RLNs, we added the edit that if the SSN had eight identical digits it almost certainly was in error, and if it had seven identical digits it still had a very high probability of being an error.

We print at least two days worth of SSNs for each hospital, so that we can look for patterns in errors and duplicates. Any page on which an SSN appears is shredded onsite. As you can see, we have made some compromises here and there, balancing strict confidentiality concern with our striving to find and correct every error we can, before releasing the data.

This has been an overview of California's experience with SSN and confidentiality. The story would be incomplete, though, without adding an account of our findings so far, regarding RLNs.

Table 1 presents the overall picture of our findings, using data for 1991.

Table 1
California, Jan.-Dec., 1991

Total discharges	3,725,059	
Newborn discharges	589,102	
Non-newborn discharges	3,135,957	
Discharges with an RLN	2,709,020	(86%)
Different RLNs	1,960,546	
RLN appears only once	1,527,077	(78%)
RLN appears more than once	433,469	

Some patients just do not have an SSN, so there is a maximum to the availability rate. How much it can be improved above the current 86 percent will have to be determined through experience. The real problem is that SSN availability is not evenly distributed. For example, of women admitted to deliver, with an expected payer of Medicaid, only 68 percent had an SSN reported in 1991. This limits the information obtainable through a unique identifier, on a patient population of special interest and probably at higher risk of negative outcome.

Table 2 addresses the issue of the number of readmissions per patient.

Table 2
Patients Admitted More than Once
California, Jan.-Dec., 1991

Admissions	Discharge Records	Patients (Different RLNs)
2	551,664	275,832
3	258,387	86,129
4	142,148	35,537
5	82,205	16,441
6	50,550	8,425
7	32,347	4,621
8	19,888	2,486
9	12,960	1,440
10	9,020	902
11	5,885	535
12	4,128	344
13	2,938	226
14	2,114	151
15	1,530	102
16	1,104	69
17	799	47
18	792	44
19	665	35
20-24	1,208	57
25-29	553	21
30-39	491	15
40-49	226	5
50-68	235	4
106	106	1
Total	1,181,943	433,469

Table 2 shows, for example, that in 1991, there were 8,425 patients (as iden-

tified by a unique RLN), each of whom was discharged 6 times, and the discharges for those 8,425 patients added to 50,550 discharges. And, there were 57 patients, each of whom was discharged 20-24 times, and their discharges totaled 1,208.

Three points can be made about these data: most patients are not readmitted in a year, those who are readmitted are usually readmitted only once, and some people are readmitted repeatedly, as many as 106 times in a year.

The first question we had when we saw these data was whether these high numbers of readmissions were artifacts produced by many people using the same SSN. We found that these readmissions were not artifacts. For example, the RLN associated with 106 discharges in 1991 does identify one individual patient. This same person had another 55 discharges in the last six months of 1990. This patient has an inherited blood disorder and comes to the hospital every few days for injection of gamma globulin. But there is an even more interesting issue raised by these 161 discharges -- this procedure is generally considered outpatient. In fact, the total days of stay for all 161 discharges for this patient is 2 days -- only on two occasions did the patient have to stay overnight. Most hospitals consider this as outpatient care and would not report a discharge record for each injection. But this one California hospital formally admits patients receiving prophylactic immunotherapy to inpatient care (although they never stay 24 hours), and insists that this patient and other such patients are inpatients and so these are inpatient discharges. Because there is no definitive definition distinguishing an inpatient from an outpatient, all discharge data programs are going to contain some records like this. (While most "outpatient" records will be distinguished by a stay of 0 days and a routine disposition, this same pattern is seen in "outpatient" deliveries -- uncomplicated births delivered in alternative birthing rooms and labor/delivery/recovery rooms, which need to be reported as inpatient discharges so that the hospital's C-section rate will be comparable to that of hospitals with no such "outpatient" births.)

We looked at 2,170 discharge records from the 158 patients with 11 or more discharges in the first half of 1991 (an average of 13.7 discharges each in just these six months). In examining these records, three patterns caught our attention, the first two of which are related to the two most frequent reasons for these multiple admissions. First, we noticed 17 patients with a total of 232 discharges, all with a diagnosis of sickle cell crisis. Their 232 hospital stays included 707 days and charges of $1.554 million. These patients were seldom admitted for any other reason. Second, we noticed 23 patients with a total of 199 discharges for chemotherapy. These patients had additional admissions, associated with other treatment of the underlying malignancy and with complications from the chemotherapy and other treatments. Their 199 chemotherapy stays alone totaled 570 days and charges of $1.471 million.

Third, we noticed that there are some patients who are constantly in and out of hospitals -- and almost every discharge is from a different hospital. They appear to have no fixed residence, because they report a wide variety of zipcodes, and their expected source of payment changes almost randomly from record to record. They seldom receive any surgical procedure, and they report a variety of urinary, kidney, stomach, and heart symptoms. For example, of the 17 patients with 17 or more stays in the first six months of 1991, nine were discharged from 11 to 38 different hospitals, each. These nine patients are listed in Table 3.

Table 3
Nine Patients
Discharged from Multiple Hospitals
California, Jan.-June, 1991

Patient	Stays	Days	Charges (thousand)	Hospitals
1	17	165	$208	11
2	19	107	128	11
3	20	44	115	17
4	24	101	222	24
5	24	77	111	16
6	25	119	223	14
7	29	65	181	23
8	30	44	126	30
9	38	109	262	38
Total	226	831	$1,576	

These nine patients were discharged a total of 226 times, stayed 831 days, with charges of $1.576 million, in the six month period. Over a year, such patients, however rare, represent a disproportionate burden on the health care system. When we complete the programming of a linked file covering all readmissions, we will be able to report more completely on the types of patients and conditions associated with multiple readmissions.

We are constructing a file which will link death certificate data with the latest hospital discharge record, if any, prior to death. The death certificate contains the SSN of the decedent, enabling the match. The resultant linked file will be available to data users, to identify post-discharge death rates. Having linked records on our large general patient population is going to be very useful.

Session P

Methodological Issues for National Minority Data

Toward the year

2000

Refining the measures

DESIGN OF THE NATIONAL HEALTH INTERVIEW SURVEY, 1995-2004

James T. Massey
National Center for Health Statistics

(Not available for publication)

ESTIMATION AND ANALYTIC METHODS FOR MINORITY STATISTICS BASED ON MULTIPLE-SOURCE SURVEY DATA

Lester R. Curtin, National Center for Health Statistics

(Not available for publication)

USE OF BEHAVIORAL RISK FACTOR SURVEILLANCE SYSTEM TO MONITOR RISK FACTORS IN MINORITY POPULATIONS

Emma L. Frazier, Centers for Disease Control and Prevention
David V. McQueen

Introduction

A priority for public health is to address the health gaps which exist between different sectors of the population. Health, U.S.[1] reports that individuals in minority groups have higher morbidity and mortality rates from major chronic diseases than that experienced by the rest of the general population. The rapid growth of minority groups in the United States population mandates the collection of minority specific risk factor data if health disparities are to be better understood and addressed.

Healthy People 2000 [2] has specified numerous goals and objectives to bring about greater equity in the health experience of the population. While there are many objectives and goals that are targeted towards specific populations, Healthy People 2000 Objective 22.4 calls for the development and implementation of a national process to identify significant gaps in the Nation's disease prevention and health promotion database, particularly with respect to disease inequities. Special target groups that have been identified are racial and ethnic groups, as well as people with low incomes and individuals with disabilities.

The Behavioral Risk Surveillance System (BRFSS) has become an invaluable tool for health professionals interested in monitoring risk factors at the state level. The system has been useful to state health planners in the identification of major risk behaviors in various groups, including individuals of minority populations. This paper presents a number of methodological issues related to the collection of risk factor data on minorities using BRFSS. Examples are cited of how BRFSS has been used to collect data on risk behaviors among various populations. Additionally, insights about strategies to enhance the data collection and analysis of data on minorities are discussed.

Background of BRFSS

Since 1981, CDC has helped states plan, design, and conduct surveys which ask adults by telephone about their health behaviors. Between 1981 and 1983, 28 states and the District of Columbia conducted one time telephone surveys. By 1984, 15 states constituted BRFSS and were collecting monthly survey data using this system. Today, 49 states and the District of Columbia are participating. Each state complete from 100 to more than 350 interviews per month during a 8- to 14-day period. The average and total number of interviews completed during each of the years are shown in Table 1.

BRFSS was designed with the idea of building a system that permitted states to monitor the prevalence of health behaviors, use of preventive services, and other health related issues which the state deems critical. The BRFSS has been instrumental as a capacity building mechanism, assisting the states in building the ability to conduct telephone surveys and monitor progress toward Year 2000 Objectives. Many states have used BRFSS data to develop baseline estimates for health behaviors

Table 1. Number of Participants and Sample Sizes, 1984-1993

Year	Number or Participants	Average Sample Size+	Total Sample Size
1984	15	817	12,258
1985	22	1,146	25,221
1986	26	1,323	34,395
1987	33	1,518	50,081
1988	37	1,526	56,448
1989	40	1,672	66,867
1990	45	1,812	81,557
1991	48	1,830	87,844
1992	49	1,950*	95,500*
1993	50	2,000*	100,000*

+ per year, per state *estimated

or problems that are relevant to their state. Use of the BRFSS has provided a mechanism that states have used to examine behavioral trends in their states and regions, as well as compare estimates between states. It is also notable that the capacity of BRFSS gives each state the flexibility to conduct special surveys of subpopulations at the state and local levels.

BRFSS Questionnaire Contents

The BRFSS questionnaire has three parts: the core, standard optional modules, and state added questions. The questionnaire is developed through a cooperative process between state representatives and the Centers for Disease Control. When feasible to our survey mode, questions are often taken or adapted from surveys like the National Health Interview Survey.

The core questionnaire is a set of questions asked by participating states. Topics covered in the standard core include health status and health care access; tobacco and alcohol use; cholesterol and hypertension awareness and screening; and use of preventive health services. Modules are optional and different modules may be chosen by a participating state. Module topics have included smokeless tobacco use, radon testing, dietary fats intake, fruits and vegetables consumption, and numerous other health related subjects. State added questions are developed at the state level and provide the flexibility to investigate specific behaviors or issues of interest to each state.

There are a number of demographic questions. Those pertinent to this paper are: "What is your race? Would you say : White; Black; Asian or Pacific Islander; Aleutian, Eskimo, or American Indian, or Other ?" This is followed by the question, "Are you of Hispanic origin such as Mexican American, Latin American, Puerto Rican or Cuban?" These two questions are combined to produce racial/ethnic groups such as : White, non Hispanic; White Hispanic; Black nonHispanic; Black Hispanic; Other Hispanic; Asian or Pacific Islander; Aleutian, Eskimo, or American Indian; and Other. While these are standard groupings used by BRFSS and some national surveys, modifications have been made when special populations have been surveyed.

BRFSS Design Characteristics

The BRFSS is a state based data collection tool. All interviews are collected by telephone. The system is a surveillance system, which collects data on a monthly basis at the state level. Most states use the Waksberg multi-cluster sampling strategy to collect data on non-institutionalized adults, age 18 and older. The survey is designed to give every adult who has household access to a telephone an equal probability of being selected in the sample. Additional technical information on the BRFSS is available elsewhere[3].

As in all telephone surveys, adults without telephones do not have a chance of being included in the sample. Individuals without telephones are usually of lower socioeconomic status and are underrepresented in surveys of this type.

Use of the BRFSS to monitor minority populations

We will present examples of the use of BRFSS to collect data on several ethnic groups and discuss methodologies and techniques used to analyze the data in each survey. For comparability purposes, the prevalence of smoking among men in each of these ethnic groups will be presented.

A study of Hawaiians was carried out by Chung et al[4] using the 1986 Hawaii BRFSS. Surveillance data was collected on 1,557 participants during 1986. Additionally, the survey was modified to reflect the most prevalent ethnic groups in Hawaii, that is, Hawaiian, Japanese, Chinese, Filipino, and others. The rate of smoking was highest in Hawaiian men-34%, followed by 27% in both White and Filipino men, 26% in Japanese men with the lowest smoking rate in Chinese men at 14%. This is an example of aggregating a year of BRFSS data with appropriate modifications to race/ethnicity questions to provide estimates on special populations at the state level.

Data were aggregated from 1987-1989 surveillance years by Sugarman et al[5] to produce estimates of risk factors for American Indians in the US. During this time period, 0.63% of BRFSS respondents identified themselves as American Indians; during that same time American Indians constituted 0.70% of the population of those 36 states participating. Four geographic regions were derived from the 36 states who participated in BRFSS during all of 1987-1989. Based on a total sample size of 1,055 American Indians, smoking rates for men in each of the regions ranged from 18% in the Southwest; 25% in the West; 38% in Other regions; with the highest reported smoking rate among American Indians residing in the Plains at 48%. This is an example of the aggregation of BRFSS data over multiple years to obtain more reliable estimates for a minority group.

In 1990, a special survey was conducted by the Asian Health Services in Oakland, California[6,7]. The standard BRFSS questionnaire was modified for

cultural appropriateness, translated into Chinese, backtranslated, and pretested. Two census tracts in Oakland with the largest percentage of Chinese residents were selected as a sampling frame. Among 296 adults, randomly selected from every second identified Chinese household, interviews were conducted in the home. The rate of smoking among Chinese men was estimated to be 28%. This is an example of the modification of BRFSS to a specific culture in a community, the adaptation of the instrument to do a face to face survey, and use of an independent one time survey to collect data.

In Monterey County, California, Hispanics constitute 34% of the population. There was a need for local data to plan appropriate and culturally sensitive programs in the Hispanic community[8]. This could only be achieved through the proper assessment of the risk factors related to chronic diseases in the Hispanic community. Using the 1990 BRFSS instrument, updates were made to make the current instrument more culturally specific and to assess the level of acculturation. Additionally, 16% of Hispanics in this community did not have telephones and a sizable number of Hispanics were in migrant camps. To include this group of individuals, the telephone survey was supplemented with face to face interviews in households without telephones and migrant camps. Hispanic men reported a smoking rate of 25%. This is an example of modification to the BRFSS instrument to become more culturally appropriate and the use of the augmentation of a telephone survey with face to face interviews in a population with high non-telephone coverage.

In 1990, states with at least 15% of African Americans (predominately states in the South and Southeast US) were aggregated to provide estimates of risk factors. African American men had the highest rate of smoking (28%)[9]. This is an example of the aggregation of data from several states to obtain more reliable estimates on a minority population.

In Figure 1, the percentage of women aged 40 and older who reported that they never had a mammogram is presented. Data from 45 states are aggregated in 1990 and show the following differences in the never use of mammography among African American, Hispanic and White women. Thirty six (36%) of African American women, 38% of Hispanic women, and 34% of White women reported that they never had a mammogram. Additional independent studies in 1990 showed that 34% of Chinese women[6,7] and 48% of Vietnamese women[10] in 1990 said they never had a mammogram. This is a brief illustration of the use of the BRFSS to assess another important public health issue which has minority health impact.

Sampling Strategies: Some options to capture data on minority populations

The first strategy is to increase the sample size. This will work for some states, but will not work be an effective mechanism in states with small numbers of minorities. States with a significant proportion of minorities in its population will benefit from a simple increase in sample size.

The use of stratification is another acceptable technique. The purpose of stratification is to formulate groups into strata that are homogenous with respect to characteristics of interest, such as racial and ethnic groups or income groups. Stratification is expected to improve the precision of those related variables.

Based on 10 years of experience of data collection in BRFSS, one consideration is the use of known telephone prefixes where target groups are aggregated. These identified telephone prefixes could then be used to oversample the group of interest. This will work best in a state or area where there is a geographic concentration of the group of interest. When the group is fairly dispersed in the sampling area, this method will probably not yield high benefits. For many surveys on special groups, an independent survey is the best option. In light of situations where the sample size is not sufficient, one needs to also consider aggregation of data over several states or over several years of data to obtain more stable estimates.

Conclusions

The BRFSS is an excellent survey tool which allows states to gather data on minority groups, particularly those identified in Healthy People 2000 Objective 22.4. While many national databases include data for minorities, the problem of small sample sizes become magnified when local and state data are needed. Many of the Healthy People 2000 Objectives have been measured at the state level and may also be monitored at this level for minority groups.

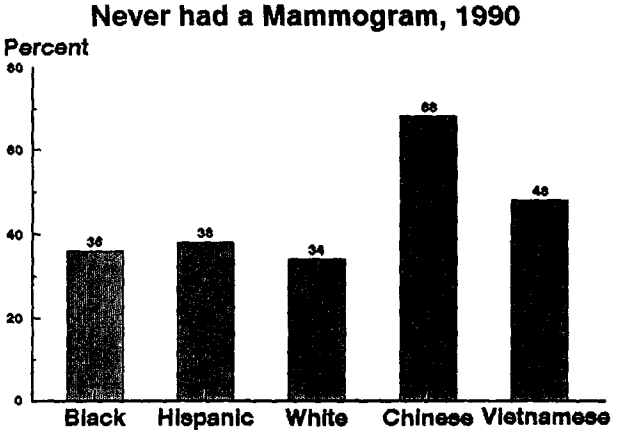
BRFSS provides and supports the mechanism to allow states to build the capacity to monitor risk factors in minority populations. The BRFSS has been successful by a number of states to identify needs of selected minority groups. States have been able to use this information to target prevention and intervention efforts, and make appropriate decisions about where these resources should be targeted. More emphasis is needed on designing surveys to collect data on minorities. More special studies are essential to assist

in program planning and policy making decisions about over diverse ethnic populations.

States and local areas are therefore challenged to respond to Healthy People 2000 objectives by assessing the needs of their minority communities. States should also consider and prioritize Healthy People 2000 objectives, in light of state specific needs which will enable them to establish intervention programs to meet the needs of special populations in their community, especially racial and ethnic groups and people of low income and low education.

In designing these special surveys, specific culturally sensitive materials must be adapted for each ethnic group. It is of importance to recognize that one survey method may not necessarily work for all groups. It is unlikely that one minority-specific strategy can be devised that will help us meet national objectives for all groups. Specific strategies must be examined, implemented and supported to collect data on minorities at the local, state and national levels.

Percent of Women Aged 40 and Older Who Never had a Mammogram, 1990



REFERENCES

1. National Center for Health Statistics. Health, United States, 1991. DHHS Publication No. (PHS) 92-1232. Hyattsville, Maryland: Public Health Service, 1992.

2. Public Health Service. Healthy People 2000: National Health Promotion and D
Disease Prevention Objectives. DHHS Publication No. (PHS) 91-50213. Washington, DC: U.S. Department of Health and Human Services, 1991.

3. Frazier EL, Franks AL, Sanderson LM. Using Behavioral Risk Factor Data. In: Using Chronic Disease data: a handbook for public health practitioners. Atlanta: Centers for Disease Control, 1992.

4. Chung CS, Tash E, Raymond J, et al. Health Risk Behaviours and Ethnicity in Hawaii. International Journal of Epidemiology. 1990, 19: 1011-1018.

5. Sugarman JR, Warren CW, Oge L, et al. Using the Behavioral Risk Factor Surveillance System to Monitor Year 2000 Objectives Among American Indians. Public Health Reports. 1992; 107:449-456.

6. Lew R, Chen A. Conducting a culturally-sensitive health survey in the Chinese community. Proceedings of the 1991 Public Health Conference on Records and Statistics, 1991:349-352.

7. Chen A, Lew R, Thai V et al. Behavioral Risk Factor Survey of Chinese-California, 1989. MMWR. 41(16):266-270.

8. Littlefield D, Snider JP, Melton RJ et al. Conducting a modified Behavioral Risk Factor Surveillance Survey in a Hispanic Community. Proceedings of the 1991 Public Health Conference on Records and Statistics, 1991:343-348.

9. Frazier EL, Blount S. Black-White Differences in Cardiovascular Disease Risk Factors, presented at the American Public Health Association meetings, November, 1991 (unpublished).

10. McPhee SJ, Jenkins CN, Hung S, et al. Behavioral Risk Factor Survey of Vietnamese - California, 1991. MMWR. 41(5):69-72.

OBTAINING DATA FOR ASIAN PACIFIC AMERICANS FROM NATIONAL SURVEYS

Larry Hajime Shinagawa, Asian American Health Forum

Background and Significance

Asian and Pacific Islanders (A/PI) lag the most behind all other groups in the execution, analysis, interpretation, and publication of health data. As a rare population, constituting less than 3% of the total population in the United States, there is a paucity of A/PI health data due to the aggregation and the small sample size of A/PI in national surveys. Numbering over 7 million with over 25 subgroups, A/PI are usually classified as one category, "Asian and Pacific Islanders," or put in a residual "other" or "etcetera" category. While this aggregation may result in statistics that may be statistically significant, it is of little social or policy value usually for assessing intervention models and health status among A/PI subgroups. Small or neglected A/PI groups such as Tongans, Cambodians, and children of interracial and interethnic marriages are among such populations. Without the identification and intervention-relevant classification and identification of these rare A/PI groups, strategies for improving the health of these groups will be neglected.

In light of the scarcity of A/PI health data, which is largely due to both aggregation and small sample size of specific A/PI populations in national surveys, this has fostered a de facto policy effect of the model health minority myth, where aggregate A/PI health statistics may hide or obfuscate specific health needs among particular populations, and where these aggregate figures may sometimes lead individuals to conditions and needs. However, by utilizing GIS map technology, we identify those specific census tracts or block groups that have a plurality or majority of a population of a particular group. We then attach to this geographical identification sociodemographic and health data. Based upon the geocoded link dataset, and by summing all census tracts with populations at a given percentage range with similar profile, we then infer the health and socio-economic status of the specific population. Thus, even though health data such as HIS may not have specific ethnicities such as Samoan Americans, we have accomplished an approximate assessment of the conditions of that particular group. As an example, lets look at several Asian and Pacific Islander population concentration maps for Hawaii.

Geocoding techniques linking Summary Tape File (STF) 3A[1] with TIGER topographic maps[2] are used to identify the characteristics and locations of specific A/PI subgroups with the "A/PI" category. The datasets that are analyzed are the U.S. Census Bureau 1990 STF3A, the 1990 TIGER geographic files, and the Hawaii Health Surveillance Survey (HHSS).[3]

Hypotheses

Our hypotheses are:

1. That it is feasible to use linked files of health and census to generate inferences about the health of rare A/PI populations;
2. That rare A/PI populations tend to cluster in small geographic areas;
3. That characteristics of a small geographic area approximate the characteristics of rare A/PI subgroups who reside in the area; and,
4. Certain A/PI subgroups are not captured by present surveys.

Specific Steps in Obtaining Data on Asian Pacific Americans

Mapping of the State of Hawaii Down to the Census Tract Level and Linking of the STF data to the TIGER dataset.

Since most health-information is geographically-specific, the ability to visualize and manipulate geographic-based linked health and sociodemographic data provides new insights into the relationship between sociodemographic factors, geographic concentrations, and health conditions. With desktop mapping Geographical Information Systems (GIS) programs and a wealth of both federal and commercial databases available for 486-level personal computers, the Health Forum was able to implement a cost-effective geographic research program.

As the official dissemination center on demographic data on Asian Pacific Americans, as authorized through the National Services Program with the U.S. Census Bureau, we were able to obtain two crucial datasets.

The first of these is called TIGER, which stands for Topographically Integrated Geographic Encoding and Referencing. It is a digital map of boundaries, street segments, landmarks, and features that has been developed by the U.S. Census Bureau. Included in this data are boundary lines for states, counties, places, tracts, block groups, and blocks.

The second major source of information comes from the 1990 Decennial Census. called STF 3A, it is short for summary Tape File 3A. This dataset holds most of the crucial sociodemographic data of specific ethnic and racial groups. It also has geographic indicators that, with

manipulation and programming, will allow for matching with geographic TIGER data.

unfortunately, the procedure of linking both the TIGER and STF data together is not simple. Many preliminary steps are necessary such as the conversion of the TIGER data to polygons, the creation of match characteristics in the STF data, and the purchase and use of a GIS package.

We first need to use a proprietary commercial package, such-as Colorado Map and Cad, for converting line segments in TIGER to objects that are called polygons. Once Map and Cad converts the TIGER information to polygons, then attributes can later be attached to each of these polygons. The Map and Cad program proceeds to create state, county, minor civil division, place tract, block group, and block polygons.

Once the files are polygonized, the next step is to process the STF files for use with the GIS program. This requires the use of a CD-ROM to read the STF data, and FOXPRO or some other dBase program that is capable to linking fields within the STF file and creating a new field that is the modified concatenation of other fields. One example would be to create an unique field for census tract, this requires concatenation of the state field, county field, and census tract field, with conditional branching to take care of anomalies in the census tract field. Once the FOXPRO program completes the concatenation and selection of the boundaries that one wishes, then we are ready for the third step.

The third step is the use of a GIS package to link the sociodemographic data to the geographic data using an unique and common match field.

Fortunately, the Health Forum received Atlas GIS, a GIS program, as a donation from Strategic Mapping. With this donation, we have utilized this package to graphically print and display geographically linked data files.

Among its many features, the package has the ability to thematically color-shade maps by characteristics calculated and created by the end-user. For example, if the user wants to see the percentage of Hawaiians by census tract in Hawaii in 1990, then the program can visually display and print the results by color-shading the different percentages. If health data was connected, it would be possible to see rates of influenza by census tract if such health data had a census tract field.

A second major feature, one that is crucial to the first feature, is the ability of the program to link census and geographic data together, once a unique identification field is created that can be the basis of match for each of these datasets. Using this second feature, used its merge command syntax to link and match the census and geographic to one another.

The Health Forum has completed processing of the TIGER and STF data for Hawaii and has linked them within ATLAS GIS.

Subsequent Steps

Since the procedures indicated above have accomplished the initial link of census and geographic data, the next step is the linkage of health data to the census-geographic data.

The HHSS survey provides a population-based approach to health assessment. It collects information which can describe a population's well-being or health status from three directions: 1) demographic characteristics; 2) morbidity; and 3) functional capacity. One of the strengths of the dataset is the ability to provide sociodemographic characteristics along with a representative picture of the incidence and prevalence of chronic and acute morbility, accidents, injuries, and disabilities/impairments in the general population.

In the near future, we will aggregate and average the health characteristics of individuals of specific groups within specific census tracts within the HHSS so as to have a composite picture of the health conditions of an ethnic group.

We will then proceed to select, using the sociodemographic census data in the GIS system, the matched areas that have predominantly one A/PI subgroup. We would then compare analyze the health indicators and health outcomes of the A/PI subgroups from this area using the HHSS.

These subsequent steps are still being devised as we speak at this conference. We are also contemplating the use of the coefficient of variation to evaluate to what degree 'the characteristics of selected places correspond to the characteristics of specific A/PI subgroups.

Conclusion

As this paper indicates, we are still in the process of evaluating this place-people methodology. While we have already shown that it is feasible to link files of health and census to generate inferences about the health of rare A/PI populations, we have not evaluated whether the characteristics of a small geographical area (place) can approximate the characteristics of specific rear A/PI populations. We hope to generate results in the next three months to answer such questions.

1. U.S. Bureau of the Census. *Summary Tape File* 3d *CD-ROM Technical Documentation*. Washington, D.C.: Government Printing Office, 1992.

2. U.S. Bureau of the Census. *TIGER Technical Documentation.* Washington, D.C.: Government Printing Office, 1992.

3. Neil Oyama and David B. Johnson, *Hawaii Health Surveillance Program Survey Methods* and *Procedures,* State of Hawaii: Hawaii State Department of Health Research and Statistics Office, 1990.

Session Q

Socioeconomic Indicators

Toward the year

2000

Refining the measures

MEASURING SOCIOECONOMIC INEQUALITY IN HEALTH: AN UPDATE ON MBTHODOLOGICAL ISSUES

Elsie R. Pamuk, Ph.D., National Center for Health Statistics
Harold R. Lentzner, Ph.D.
Robert Brackbill, Ph.D.

I. Introduction

By now most of us are well aware that a stated national goal for the year 2000 is the reduction of health disparities (1). We are also aware that a major contributor to disparities in health is inequality in socioeconomic resources. However, effectively monitoring the nation's progress toward reducing socioeconomic inequality in health has been hampered by a variety of methodological problems. This paper reviews work that has been done on some of these methodological issues in recent years, focusing on the problem of summarizing inequality in health so that the experience of a population can monitored over time, and compared with other populations at the same point in time.

II. The problems

In an article written for the World Health Organization in 1986 (2), Harald Hansluwka identified the issues that must to be addressed when choosing a summary indicator of socioeconomic inequality in health. He noted that, across populations and over time, the composition of social groups will differ, both definitionally, and in the proportion of the population they represent. These differences must somehow be taken into account.

The traditional way of summarizing the magnitude of social inequality in health has been to form a ratio of the experience of the lowest ranked social group to that of the highest ranked group. The inadequacy of this approach was brought into focus by the controversy surrounding *The Black Report* (3), the official report of the Working Group on Inequalities in Health established by the British government in 1977. Their conclusion, that social class disparities in health had widened in recent decades, was heavily reliant on the type of analysis typified by Table 1 (4), which shows Standardized Mortality Ratios (SMRs) in the Registrar General's 5 social classes. The social classes are based, and ranked, on occupational status, with Class I consisting of individuals in professional & higher administrative occupations, and Class v of unskilled manual laborers.

TABLE 1. ENGLAND AND WALES: MALE MORTALITY BY SOCIAL CLASS[a] (*National average = 100*)

Period	Social classes I	II	III	IV	V	Ratio of class V to class I
1921–1923[b]	82	94	95	101	125	1.5
1930–1932[c]	90	94	97	102	III	1.2
1949–1953[cd]	86	98	101	104	118	1.4
1959–1963[c]	76	81	loo	103	143	1.9
1970–1972[c]	77	81	104	114	137	1.8

Source: Office of Population Censuses and Surveys, *Occupational* Mortality: the *Registrar General's Decennial Supplement for England and Wales. 1970-1972, Series* DS. No. 1 (London. Her Majesty's Stationery Office, 1978), p. *174*.

[a] *Depending on the period. the* data refer to ages 15 or 20 to 64 or65 *years* end have not been adjusted to allow for changes in classification.
[b] Index calculated on the basis of the comparative mortality figure
[c] Index calculated on the basis of the standardized mortality rate.
[d] Results corrected by John Fox.

Reprinted from reference 4.

The most obvious flaw in summarizing this information using the ratio of mortality in Class V to that in Class I is the omission of information on Classes II through IV, which contain the majority of the population. This is particularly well-illustrated in comparing 1959-63 with 1970-72. Although there was a slight decline in the ratio of mortality in Class v to that in Class I, there were increases in the SMRs of Classes III and IV.

An additional problem, that characterizes not just the summary ratio, but the entire data presentation, is that class size is not taken into account. This poses a particular problem for assessing trends over time. As numerous observers have pointed out, shifts in the occupational structure have had a pronounced effect on the social class distribution of the population. In 1931, 18% of the adult male population was designated as belonging to Class V; but by 1971 it's share had declined to only 8% (3). If good health is related to upward mobility, it would be reasonable to expect an increase in the mortality ratio of the lowest ranked class as it becomes smaller and is made up of individuals in relatively poorer health.

III. Proposed solutions

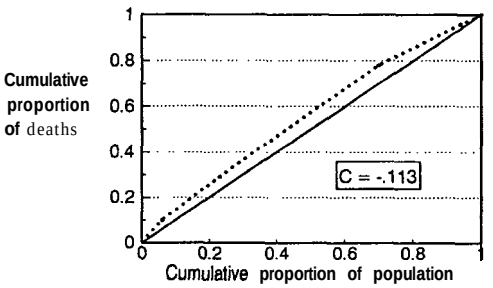
These criticisms prompted several efforts to come up with an alternative summary index that would both incorporate information from all of the groups and take into account their relative sizes. A critical appraisal of these efforts has recently been published by Wagstaff and his colleagues (5) and we will attempt only to briefly summarize their work here.

The most common approach was to try and adapt measures of income inequality used by economists. Most of these measures are based on the Lorens curve, a curve formed by ranking individuals according to their income, and plotting the cumulative proportion of income against the cumulative proportion of the population. Perfect equality is represented by the diagonal, where the cumulative distributions of income and population are identical. But if income is distributed unequally, the lower ranked percentiles of the population will have less than the corresponding percentile of income and the Lorenz curve will lie below the diagonal. Most summary measures of income inequality reflect some mathematical aspect of the difference between the Lorenz curve and the diagonal. The Gini Index, for example, is the the area between the Lorens curve and the diagonal expressed as a proportion of the total area under the diagonal.

Adapting income inequality measures to apply to the analysis of social inequality in health turned out to be far from straightforward. Measuring income inequality involves plotting the distribution of income over the population ordered by level of income. But measuring socioeconomic inequality in health involves plotting the distribution of health over the population ordered by socioeconomic status.

If the ordering of the groups by social status matches the ordering by level of health, the Lorenz curve and its derived indices work perfectly well. It even works with measures of ill health, such as death, where more death is concentrated among those with the least social status. In Figure 1, all points lie above--rather than below--the diagonal and the summary measures carry a negative sign.

Figure 1. Social class inequality in mortality, economically active men aged 20-64, England & Wales, 1979-83

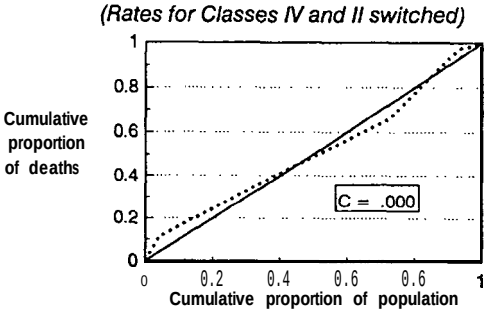


But if the ordering of the groups by social status differs from the ordering by level of health--if mortality is higher in Social Class II than in Social Classes III and IV, for example--the traditional Lorenz curve cannot be maintained except by arbitrarily switching the position of the classes so they remain ordered by level of health.

However, Wagstaff and his colleagues demonstrated that an adaptation of the Lorenz curve used in tax analysis can be applied to the situation where an external ordering is imposed. In the example presented in Figure 2, the calculation involves subtracting that portion of the area that lies on the other side of the diagonal. The summary index formed from this curve is called the concentration index (C), and is two times the "net" area between curve and the diagonal. If the death rates for Classes II and IV were switched, the ratio of mortality in Class V to that in Class I remains the same, but the concentration index goes to zero.

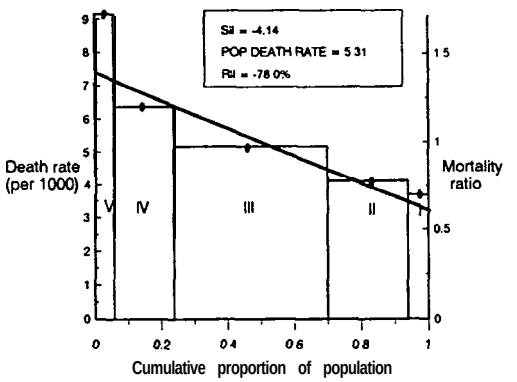
Figure 2. Social class inequality in mortality, economically active men aged 20-64, England & Wales, 1979-83

(Rates for Classes IV and II switched)



An alternative approach to the problem of a summary index was developed from the usual graphical presentation of the relationship between SES levels and health outcomes. SES levels are typically represented as equal-sized bars on a horizontal axis while the outcomes, such as age-adjusted death rates or ratios, are plotted on a vertical axis. In this alternative approach, the relative size of each group is taken into account by converting the horizontal dimension into an X-axis that reflects the cumulative proportion of the population represented by the ordered groups (Figure 3). [Note that this is identical to the bottom axis for the Lorenz and concentration curves]. The death rate for each group is then plotted at the mid-point of the bar representing that group's share of the total population.

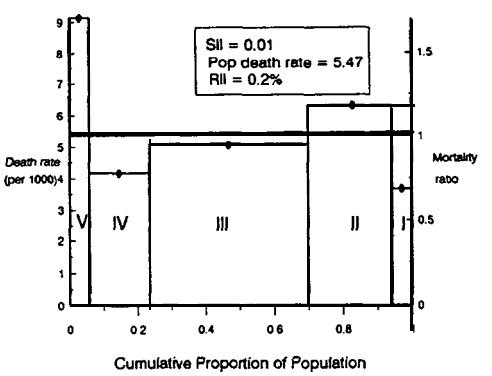
Figure 3. Social class inequality in mortality, economically active men aged 20-64, England & Wales, 1979-83



A summary index is formed by calculating the slope of a straight line fitted by weighted least-squares regression--using the proportion of the population in each class as the weights. This Slope Index of Inequality (or SII) can be interpreted as the average decline in the age-adjusted death rates over the entire population ranked from lowest to highest social status. It is, therefore, a measure of inequality in absolute terms. However, the SII is easily converted into a relative measure (which we call the Relative Index of Inequality or RII) simply by dividing it by the age-adjusted death rate for the population as a whole. This is equivalent to plotting mortality ratios, and calculating the slope in the same way. The sign of the slope will be negative when mortality declines with increasing social status, positive when it increases with increasing social status, and, like the concentration index, will go to zero when there is no consistent relationship with SES (Figure 4).

Figure 4. Social class inequality in mortality, economically active men age 20-64, England & Wales, 1979-83

(Rates for Classes IV and II switched)



The similarity between the concentration index and slope index is not fortuitous, although it may be serendipitous. As Wagstaff and his colleagues demonstrated, there is a clear mathematical relationship between the two. Both are based on the weighted covariance of x and y, where x is the ranked cumulative population and y is the health indicator, and multiplying the RII by twice the weighted variance of x produces the concentration index.

Therefore, much of what is developed from here on applies to either approach to measuring socioeconomic inequality in health. But we will limit most of our examples to the SII and RII, because audiences familiar with epidemiological research are used to beta coefficients, and because these indices may have "...more immediate graphical interpretation" (5, p.550). For example, Wilkinson (6) was able to show both the trend in overall mortality and the trend in social class ineauality bv. plotting the value of the SII centered-on the mean age-adjusted death rate (Figure 5)--the top of the bar is the intercept and the bottom is the predicted death rate at-the 100th percentile.

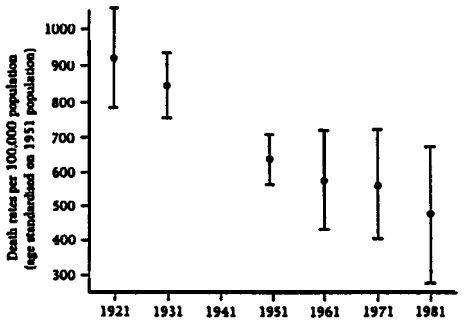


Figure 5. The decline in mortality and trends in social class mortality differentials a measured by slope indices of inequality. Men in England and Wales 1921–81.
● Mean age-standardised mortality rate.
| Slope index of inequality based on social classes and centred on the mean.
Source: R. Pamuk (1985), table 1.

Reprinted with permission of the author from reference 6.

Both approaches clearly go a long way toward addressing issues raised by Hansluwka. Both take the relative sizes of the social groups into account. They also address, at least partially, the issue of definitional consistency by treating the social groups as an ordinal scale, so that a group's rank order is more important than it's specific characteristics.

Hansluwka was also concerned that the issue of the number of groups be addressed, since "...(t)he more groups identified the smaller will be the within group heterogeneity and the larger the between group heterogeneity.* (2, p.1391 Without denying the truth of this statement, weighted summary indicators are remarkably robust to changes in the number of groups.

Table 2 shows educational inequality in mortality among white men aged 25-64 in the United States, based on Kittagawa and Hauser's study of death records linked to 1960 census records (7) and compares it to educational inequality in the early 1980s based on the National Longitudinal Mortality Study (8).

The distribution of the population according education level shifted dramatically in the intervening 20 years; by 1980, 43t of white non-Hispanic men were in the top category identified by Kittagawa & Hauser--one or more Years of college. Looking a more detailed educational classification, it is clear that this group is heterogeneous with respect to mortality; men completing l-3 years of college have 1.6 times the mortality of those with p&t-graduate training. Despite this, the SII, RII and concentration index based on 8 education levels are only slightly higher than the comparable indices based on the 5 levels identified by Kittagawa and Hauser.

Table 2. Inequality in mortality based on levvel of education: United States, white males in 1960 and white non-Hispanic males in 1979–81, aged 25–64

Years of school completed	1960 % of Pop	1960 DR (1000)	1960 Rel Rate
< 8	17%	9.7	I.2
8	17%	8.7	1.1
HS 1–3	20%	8.2	1.0
HS4	26%	7.4	0.9
C 1+	21%	6.3	0.8
INEQUALITY INDICES?		SII =	-4.0
		Rii =	–50%
		c-	-0.05

Years of school completed	1979–81 % of Pop	1979–81 DR (1000)	1979–81 Rel Rate
< 8	4%	9.4	1.5
8	5%	9.5	1.5
HS 1-S	12%	8.5	1.3
HS 4	88%	6.6	1.0
C 1+	43%	4.9	0.5
INEQUALITY INDICES:		SII =	-5.5
		Rii =	–86%
		C =	-0.19

Years of School completed	1979–81 % of Pop	1979–81 DR (1000)	1979–81 Rel Rate
5-7	3%	9.4	1.5
8	5%	9.5	1.5
HS 1–3	12%	8.5	1.3
HS4	36%	6.5	1.0
Cl-S	18%	6.1	1.0
C4	14%	4.5	0.7
C 5+	12%	3.8	0.6
INEQUALITY INDICES:		SII =	-5.7
		Rii =	-SO%
		C =	-0.14

• SII, Slop Index of Inequality;
Rii, Relative index of Inequality;
C, Concentration index (see text)

Sources: reference 7 end U.S. National Longitudinal Mortality Study

Table 3. Inequality in mortality among white non-Hispanic U.S. men aged 25–64, 1979–81 [based on 8 levels of education]

Mortality from:	Age-standardiied death rate (p e r	SII*	RII**	% of SII	
All causes	8.98	-5.87	-89%	100.0%	
Ischemic heart disease	1.86	-1.55	-83%	27.4%	43%
Cerebrovascular disease	0.22	-0.33	-146%	5.8%	
Other cardiovascular disease	0.65	-0.56	-85%	9.8%	
Smoking related cancer	0.84	-1.01	-121%	17.6%	22%
Prostate cancer	0.07	0.04	51%	-0.6%	
All other cancers	0.85	-0.29	-33%	5.0%	
Respiratory disease	0.29	-0.60	-266%	10.7%	
Diabetes	0.11	0.04	36%	-0.7%	
Cirrhosis & chronic liver disease	0.23	-0.20	-88%	3.5%	
Motor vehicle accidents	0.25	-0.39	-157%	6.9%	13%
Suicide	0.21	0.05	23%	-0.8%	
Homicide	0.05	-0.09	-165%	1.5%	
Other external causes	0.23	-0.33	-143%	5.9%	
All other causes	0.48	-0.44	-92%	7.8%	

* SII, Slope Index of Inequality (see text)

** RII, Relative Index of Inequality (see text)

Source: National Longitudinal Mortality Study, 1979–81 (5 year follow–up)

Hansluwka was also concerned that the implications of choosing to measure inequality on an absolute or relative scale be understood. We would argue that the best way to understand these implications is to examine both absolute and relative inequality, and that this is made easier by using absolute and relative indices that are easily derived from each other.

Table 3 shows the SII and RII for several categories of underlying cause of death, again for white non-Hispanic men aged 25-64 in the early 1980s. Each SII shows the average change in the cause-specific death rate over the population ranked by education level, and thus, the magnitude of the SII is dependent on the cause-specific death rate.

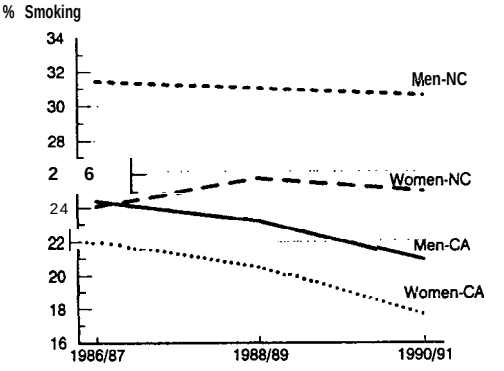
For each cause, the RII is formed by dividing the SII by the cause-specific rate, permitting a comparison of socioeconomic inequality by cause. The highest levels of relative inequality are found in deaths from respiratory diseases, external causes except suicide, stroke, and smoking-related cancers. These might be important candidates for prevention programs if our primary concern is to address causes where socioeconomic inequality is most severe.

However, if we are concerned with developing policies and programs aimed at reducing the overall level of inequality, we might do better to examine each cause's contribution to the SII for all causes combined--a calculation that takes into account each cause's contribution to overall mortality. Just as the cause-specific death rates sum to the total death rate, the cause-specific SIIs sum to the SII for total mortality; so each cause-specific SII can be converted to a percentage of the total SII. This shows that cardiovascular disease and smoking-related cancers together account for over 60% of the overall socioeconomic inequality in mortality in this population.

IV. Application to other health indicators

Until now we have focused on mortality, but these approaches to measuring inequality can be generalized to other health indicators. Figure 6 shows the trend in smoking prevalence in two states, North Carolina and California, for three periods from 1986/87 to 1990/91, using data from the Behavioral Risk Factors Surveillance System (9). While smoking prevalence declined substantially in California, especially after 1988/89, North Carolina was able to achieve only a slight reduction among men, and smoking prevalence actually increased slightly among women.

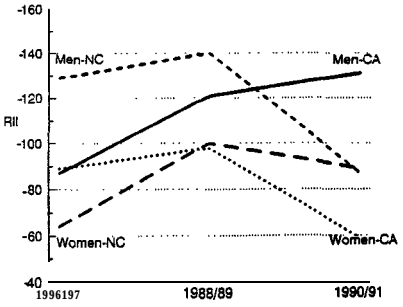
Figure 6. Smoking prevalence among adults 18 and older, California and North Carolina, 1986-1991



Source: Behavioral Risk Factor Surveillance System

However, the picture with respect to the distribution of smoking by socioeconomic status is quite different (Figure 7). Initially, educational inequality in smoking was greatest among men in North Carolina and least among women in that state, with inequality levels among California men and women falling in between and nearly identical to each other. Between 1986/87 and 1988/89, educational inequality in smoking increased in all four groups, very substantially for California men and North Carolina women, and only slightly for women in California and men in North Carolina.

Figure 7. Inequality in smoking prevalence among adults 18 and older, California and North Carolina, 1986-1991
[Education-based Relative Index of Inequality]



Source: Behavioral Risk Factor Surveillance System

Between 1988/89 and 1990/91, changes in the distribution of smoking by education level were more varied. Inequality fell sharply among North Carolina men and California women, but fell only slightly among women in North Carolina, and continued to increase among men in California. The overall comparison of these groups in 1990/91 is an interesting reversal of what was observed in 1986/87--socioeconomic inequality in smoking is now greatest among men in California, least among California women, with inequality among women and men in North Carolina located in between and nearly identical.

This tells us that the overall trend may mask much greater changes in smoking prevalence8 within socioeconomic categories. In California, for example, men and women experienced similar declines in smoking prevalence. However, this was achieved through greater relative reductions in smoking among men at higher educational levels and among women at lower levels of education. It is certainly arguable that information regarding changes in the socioeconomic distribution of smoking is crucial to formulating effective strategies for continuing the declines observed in California, and attempting to initiate a substantial decline in smoking prevalence in North Carolina.

Another example of the usefulness of this approach to examining socioeconomic disparities in health, is the comparison of relative levels of inequality across a broad range of health indicators. This is an abbreviated example for Alameda County, using data from the Human. Population Laboratory, with inequality again based on education (Table 4). It shows the SII and RII for 2 indicators of health status--the percent reporting their health as fair or poor and the average number of sick days over the past year; for 3 types of illness or health conditions--arthritis, diabetes, and high blood pressure: for 3 types of behavioral risk factors--the percent reporting smoking, being overweight, and exercising less than 3 times per week: and for 2 indicators of access to and need for care--having no health insurance, and the average number of doctors' visits in the past year (10). Clearly, these are very disparate measures, but by using the Relative Index of Inequality, the degree of concentration of any given indicator at one end of the social scale (or the other) can be examined and compared to that for other indicators.

Table 4. Absolute and relative indices of inequality baaed on level of education for various health indicators, romeo 18 years and older in Alameda County, California. 1990

	Population Average (% or mean)	SII*	RII**
HEALTH STATUS			
In fair or poor health	15.5	-21.0	-135.7
Days sick in bed last year	5.0	−2.7	−54.1
DIAGNOSED CONDITIONS			
Prevalence of arthritis	21.6	-8.7	-40.2
Prevalence of diabetes	5.0	-3.7	-74.4
Prevalence of hypertension	18.0	-12.3	-60.3
HEALTH-RELATED BEHAVIORS			
Current smokers	21.6	-13.1	-60.8
Overweight	19.2	-23.5	-122.3
Exercise < 3 times/wk	66.3	-28.1	-42.3
ACCESS/NEED INDICATORS			
Has no insurance	12.6	- 18.4	- 145.6
MD visits last year	3.8	-4.7	-124.7

* SII, Slope Index of Inequality (see text)
**RII, Relative Index of Inequality (see text)
Source: reference 10

Indicators can then be grouped according to their level of relative inequality. At the most cursory level of examination, this exercise shows that, although men and women in Alameda County have similar levels of inequality in overall health (as judged by the percent reporting being in fair or poor health), women experience substantial inequality in a greater range of health indicators-than do men (RII >= -50%). The differences shown here indicate that the strategies developed to reduce inequality may well need to be sex-specific.

Table 5. Distribution of health indicators by level of relative inequality (based on educational level), for men and women in Alameda County, California, 1990

Relative Index of Inequality	Men	Women
−100% or more	Fair/poor health Sick days Current smokers Has no insurance	Fair/poor health MD visits overweight Has no insurance
−50% to -99.9%	Arthritis Overweight	Sick days Diabetes High blood pressure Current smokers
0% to -49.9%	No regular exercise	Arthritis No regular exercise
0% − +49.9%	High blood pressure	
+50% to +99.9%	MD visits Diabetes	

Source: reference 10

V. Conclusion:

We do not mean to imply that all of the problems associated with monitoring inequality in health have been resolved. There are still a number of methodological issues to be addressed, including the appropriate basis for classifying individuals into SES categories and the extent of the population covered by any classification. But, we would argue that the remaining issues exist regardless of whether we simply report rates, or prevalences, or risk ratios by SES category--which we do all the time--or whether we attempt to quantify inequality. And given the need to address the goal of reducing health disparities, and the likelihood of specific policies and programs being instituted to achieve this goal (including reform in the health care system), a serious attempt at the regular monitoring of trends in socioeconomic inequality in health seems warranted.

REFERENCES

1. U.S. Department of Health and Human Services. (1991) Healthy *People 2000: National Health Promotion and Disease Prevention Objectives.* Washington: Public Health Service.

2. Hansluwka H. (1986) "Reflections of the measurement of social inequality of death," in New Developments in *the Analysis of Mortality and Causes of Death* (H Hansluwka, Y Porapakkham, AD Lopez, and P Prasartkul, editors). Bangkok, Thailand: Mahidol University and Geneva, Switzerland: World Health Organization.

3. Townsend P. and Davidson N. (1982) *Inequalities in Health: The Black Report.* Harmondsworth: Penguin.

4. Vallin J. (1980) "Socioeconomic determinants of mortality in industrialized countries." United Nations, Secretariat. *Population Bulletin of the United Nations. No.* 13:42-52.

5. Wagstaff A, Paci P, van Doorslaer E. (1991) *On the measurement of inequalities in health," *Social Science and Medicine* 33:545-57.

6. Wilkinson RG. (1989) "Class mortality differentials, income distribution, and trends in poverty." *Journal of Social Policy* 18:307-35.

7. Kitagawa EM, Hauser PM. (1973) *Differential mortality in the United States: a study in socioeconomic epidemiology.* Cambridge, MA: Harvard University Press.

8. Rogot E, Sorlie PD. Johnson NJ, Schmitt C. (1992) A *Mortality Study of 1.3 Million* Persons *by Demographic, Social, and Economic Factors: 1979-1985 Follow-Up, U.S. National Longitudinal Mortality Study.* NIH Publication No. 92-3297. Bethesda, MD: National Institutes of Health.

9. Remington PL, Smith MY, Williamson DF, Anda RF, Gentry EM, Hoglin GC. (1988) "Design, characteristics and usefulness of state-based Behavioral Risk Factor Surveillance: 1981-1987. *Public Health Reports* 103:366-75.

10 Pamuk E, Lentzner H, Mendlein J, Kaplan G. (1993) "Alternative indicators of socioeconomic inequality in health in Alameda County, California, 1989-1990" in *American Statistical Association: 1992 Proceedings of the Social* Statistics *Section.* Alexandria, VA: American Statistical Association.

UNEMPLOYMENT AND FERTILITY: A TRANSFER FUNCTION ANALYSIS WITH ADDITIONAL INTERVENTIONS

Andreas Muller, University of Arkansas at Little Rock

Introduction: Between 1929-1936, the U.S. fertility rate decreased by 15% (NCHS, 1990) suggesting that the Great Depression may have had a retarding effect on population growth. The following presentation will explore whether more recent, but less severe business cycles have had corresponding effects on U.S. fertility rates. Transfer functions analysis, a branch of dynamic regression models, will be used to examine the research question.

Theoretical Perspeotive: Unemployment has been viewed as an experience producing economic downward mobility and heightened psychic stress adversely affecting individuals and entire communities (Jahoda,1982). Perhaps best known is Harvey Brenner's finding (1971;1987) that business recessions are associated with increased heart disease mortality 2 to 6 years lagged. Largely due to Brenner's work, the impact of business cycles on various aspects of physical and mental health continues to be investigated (Kasl,1979; Iversen et al., 1989, Starrin et al., 1990; Dowdall et al.,1990).

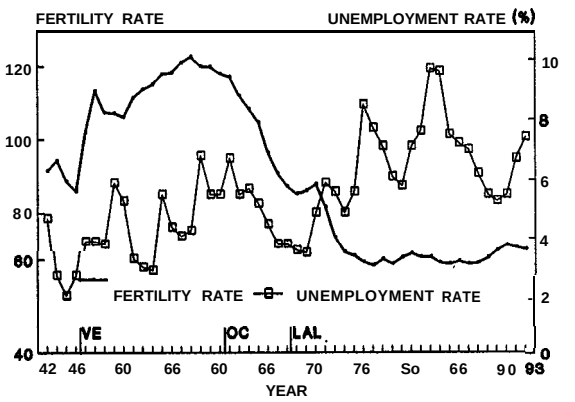
It is reasonable to suspect that business cycles are also related to fertility. In recent recessions, disposable income of the unemployed typically decreased by one-half or more (Economic Report of the President, 1982; Berki et al., 1985; Piacentini and Cerino, 1990). During periods of unemployment long-term commitments such as marriage, or having children may have to be postponed and unemployed persons can be expected to take measures to prevent pregnancies. Due to the length of human gestation, *increases in the unemployment rate are expected to be followed by decreases in the fertility rate primarily during the next year.*

However, unemployment is only *one* of several factors expected to explain variations in U.S. fertility rates for the last half century. World War II depressed marriage rates and the decision to have children was postponed by many until the return of the veterans (1945-46). In the early 1960s, the introduction of effective and convenient birth control methods, particularly oral contraceptives, provided greater choices'with regard to timing and the number of pregnancies and may have reduced fertility rates. In addition, liberalization of state abortion laws, beginning in 1967, may have further reduced fertility rates.

Data: U.S. fertility rates will be related to five variables: (1) unemployment rates, age and sex-specific, (2) marriage,' (3) legal abortion rates and (4) homecoming effect after World War II. Since no annual data on (5) oral contraceptive use could be found, the potential effect will be represented by an additional transfer function to be discussed. The time series analysis is based on annual data for the period 1942-1992 (n=51).

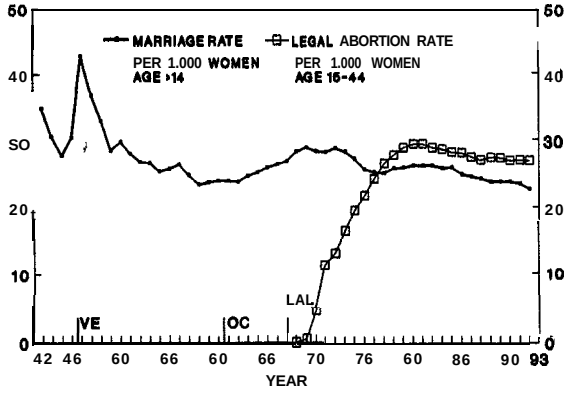
Figure 1 superimposes U.S. fertility rates as measured by the number of live births per 1,000 women, ages 15-44 years (NCHS, 1990, CDC, 1993) and national unemployment rates which represent the percentage of the U.S. civilian labor force 16 years and older out of work and looking for work within the last two weeks (U.S. DOL, 1992). Fertility rates show some sharp fluctuations during World War II and the immediate post-war period, an upward trend for most of the 1950s (baby boom) followed by a downward trend until the mid 1970s; thereafter the time series appears stationary. The unemployment rate shows seven sharp increases representing recessions. An underlying upward trend in unemployment is noticeable.

FIG. 1: U.S. FERTILITY AND UNEMPLOYMENT RATES, 1942-1992



NCHS AND U.S. DOL

FIG. 2: U.S. MARRIAGE AND LEGAL ABORTION RATES, 1942-1992



NCHS AND GUTTMACHER INSTITUTE

Figure 2 shows age and sex-specific marriage (NCHS, 1991; CDC,1993) and legal abortion rates. With the exception of high rates following World War II, marriage rates are fairly stationary over the last four decades. Legal abortion rates (Cates, 19823 Henshaw and Van Vort, 1990) increased exponentially during 1969-1971 and more slowly for the remaining years of the 1970s reaching their highest level in 1980-81. Thereafter, legal abortion rates are only slightly decreasing.[1]

Ordinary Least Squares (OLS) Distributed Lag Regression Analysis: The OLS model including one-year lagged variables is presented below; the time series context is denoted by subscript (t). The model is based on five assumptions: the residuals are (1) independent over time, (2) normally distributed, (3) homoscedastic, (4) sum to 0 and are (5) uncorrelated with the independent time series.

$$Y_t = b_0 + b_{it}*X_{it} + b_{it}*X_{it-1} + e_t;$$

Y_t = annual fertility rates;
b_0 = intercept;
b_{it} = regression coefficient of independent time series (i);
X_{it} = (1) annual unemployment rates, (2) age and sex-specific marriage rates, (3) legal abortion rates, (4) homecoming effect, dummy variable coded 0 prior to 1945 and 1 thereafter;
X_{it-1} = independent time series (1) (2) and (3) lagged one year;
e_t = residuals, $IN(0,\sigma)$;

Table 1 shows the result of the OLS estimation: all analysis were performed using *RATS 4.0* software (Doan, 1992). The regression model is statistically significant explaining 84.2% of the variation in fertility rates. However, individual regression coefficients are highly variable. The sign changes suggest multicollinearity often encountered with distributed lag models based on the original, untransformed data. The inspection of the correlation matrix of the independent variable time series (not shown) indicates that the correlations between the same variables, entered contemporaneously and lagged, range between .57 and .99.

The low Durbin-Watson statistic[2] indicates that the residuals are first-order autocorrelated. The sample autocorrelation coefficients exceed ±2 s.e. at lags 1,2,3,11 and 12. clearly violating the independence assumption. A high degree of autocorrelation means that the regression coefficients are inefficient, t-statistics are invalid, and may also indicate model misspecification (Maddala, 1988). Omitting variables (i.e. oral contraceptive use) correlated with the independent variables will bias the regression coefficients.[3]

TABLE 1. OLS DISTRIBUTED LAG REGRESSION ON U.S. FERTILITY RATES

VARIABLE	b_i	t
Unemployment Rate	1.63	1.30
Unemployment Rate(l)	**-.20**	**.21**
Homecoming Effect	13.63	2.53
Marriage Rate	-1.04	-1.91
Marriage Rate(l)	**.64**	1.19
Legal Abortion Rate	-4.18	-4.72
Legal Abortion Rate(l)	2.53	2.83
Constant	99.03	7.05

$F_{(7,43)}$ 39.01; P=.000; S.E.: 8.34; R^2 adj.: .842; DURBIN-WATSON: .46; Q(12): 87.3; P=.000

SAMPLE AUTOCORRELATION FUNCTION

LAG	1	2	3	4	5	6
r						
LAG	.77 7	.52 a	.39 9	.24 10	.13 11	.05 12
r	-.02	-.10	-.16	-.25	**-.35**	**-.42**

Note: Coefficients >|2 s.e.| are shown in bold type and lagged variables by number in parentheses.

FIG. 3: U.S. FERTILITY RATES AND OLS DISTRIBUTED LAG MODEL, 1942-1992

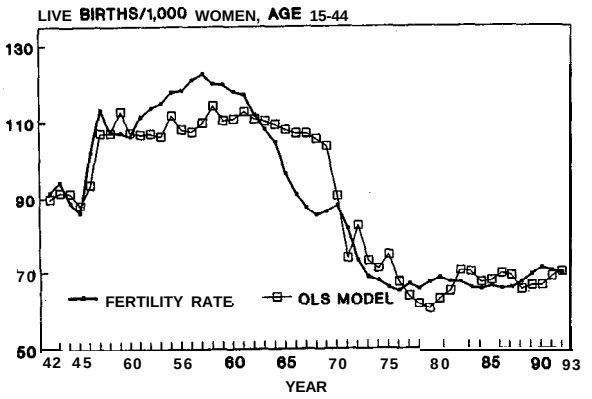


FIG. 4: OLS MODEL RESIDUALS
MEAN=0; SKEWN.=-.84, p=.02;
KURTOSIS=.78, p=.3

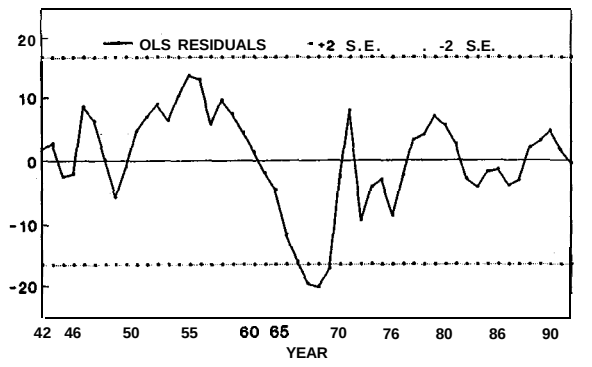


Figure 3 shows the fit of the OLS distributed lag model: it roughly fits the data. Long runs of positive and negative residuals, for the periods 1952-1962 and 1963-70, indicate the autocorrelation problem (see Figure 4). Residuals for the period 1967-1969 fall outside the ±2 s.e. range indicating heteroscedastity.

Transfer Function Analysis with Multiple Inputs: Box-Jenkins *ARIMA (autoregressive-integrated-moving average)* models can correct for autocorrelation in the OLS error term. That is, the error term e_t is decomposed into a systematic component modeled by a low-order autoregressive, or moving average process and a stochastic component a, which meets the standard residual assumptions $(IN(0,\sigma))$.

The multiple input transfer function model for a nonseasonal, stationary time series can be stated:

$$Y_t = c_0+(\omega_i(B)B^{bi}/\delta_i(B))*X_{i,t} + \phi^{-1}(B)\theta(B)*a_t;$$

Y_t = output time series;
c_0 = constant;
B = backshift operator:
$\omega_i(B)B^{bi}$ = rationally distributed effects of input series X_i;
$\delta_i(B)$ = decay rate of order (B);
$X_{i,t}$ = input time series(i):
$\theta(B)$ = MA operator $(1-\theta_1 B - \ldots -\theta_q B^q)$;
$\phi(B)$ = AR operator $(1-\phi_1 B - \ldots -\phi_p B^p)$;
a_t = residuals, $IN(0,\sigma)$.

In contrast to the OLS model, the transfer function model relates the fertility series (output series) also to lagged values of itself captured by the ARIMA model $(\phi^{-1}(B)\theta(B)*a_t)$.

The ARIMA modelling process and the transfer function extension are described in a number of sources (Box and Jenkins, 1976; McCleary and Hay, 1980; Pankratz, 1991; Pindyck and Rubinfeld, 1991). Using transfer function analysis with more than one input series can create significant difficulties at the identification and specification stages (Liu and Hanssens, 1982). Therefore, the procedure used in this paper is based on the *common filter* approach (Pankratz, 1991:173-184) which more readily allows for multiple inputs.

The model building process starts with an OLS free-form distributed lag model including several lags. Second and third-year lagged effect of the input variables were estimated; but no statistically significant lagged effects were found (models not shown).

Inspection of the sample autocorrelation function (see Table 1) indicates statistically significant coefficients at high lags suggesting the need for differencing. According to the common-filter approach, *the output series and all input time series* are differenced once and the model is reestimated by non-linear least squares with the constant term deleted from the model.'

TABLE 2. TRANSFER FUNCTION *ANALYSIS ON* **U.S. FERTILITY RATES:** PRELIMINARY MODEL WITH DIFFERENCED DATA

INPUT	ω_i	t
D-Unemployment Rate	-.03	-.14
D-Unemployment Rate(l)	-.41	-1.75
D-Homecoming Effect	11.04	2.93
D-Marriage Rate	.21	.88
D-Marriage Rate(l)	1.04	8.15
D-Legal Abortion Rate	.09	.27
D-Leg. Abortion Rate(l)	-.71	-2.14
NOISE MODEL: $\phi(1)$	.70	6.42

S.E.: 2.27: R^2 adj.: .988; DURBIN-WATSON:2.01; Q(12): 7.57; P=.75

SAMPLE AUTOCORRELATION FUNCTION

LAG	1	2	3	4	5	6
r	-.07	.12	.02	-.06	.11	-.05
LAG	7	8	9	10	11	12
r	.17	-.01	-.10	.01	-.17	-.11

The estimates of the preliminary transfer function are shown in Table 2. Incorporating first-order autocorrelation $\phi(1)$ greatly improves model fit: the adjusted R^2 value (.988) is substantially larger and the standard error of estimate (2.27) is much smaller than those for the OLS model. The Durbin-Watson and the Q-statistics indicate no significant first or higher-order autocorrelation. With the exception of the homecoming effect all contemporaneous effects are nonsignificant, but one-year lagged effects are statistically significant and in the expected direction. The lagged effect of the unemployment rate (ω=-.41; t=-1.75) is significant at the .05 level with direction predicted.

Additional models based on differenced data were tested including third-year lagged effects with first-order decay effects, but none of those proved superior to model M-l, Table 3. Model M-l deletes all non-significant contemporaneous and lagged effects resulting in a more parsimonious model.[6]

Figure 5 shows improved fit when compared with the OLS model (Fig. 3). The residuals of the transfer function model are presented in Figure 6. There is downward movement in residuals during the early 1960s culminating in an outlier in 1965. Since the downward movement coincides with the introduction of improved birth control methods, several transfer functions specifications were tried to capture the event. The best model fit was achieved by adding a dummy variable coded 1 for the period 1962-66 and specifying a first-order decay effect (δ).[7]

Table 4 shows transfer function model M-2 including the additional first-order

TABLE 3. **TRANSFER FUNCTION ANALYSIS ON U.S. FERTILITY RATES: MODEL M-l**

INPUT	ω_i	t
D-Unemployment Rate(l)	-.35	-1.59
D-Homecoming Effect	13.76	7.43
D-Marriage Rate(l)	.98	8.85
D-Legal Abortion rate(l)	-.71	-2.20
NOISE MODEL.: $\phi(1)$	.67	6.15

S.E.E.: 2.22; R^2 adj.: .989;
DURBIN-WATSON:1.99: Q(12): 6.63; P=.83

SAMPLE AUTOCORRELATION FUNCTION

LAG	1	2	3	4	5	6
r	-.07	.09	.05	-.03	.08	-.09
LAG	7	8	9	10	11	12
r	.14	.00	-.09	.01	-.15	-.15

FIG.5: U.S. FERTILITY RATE8 AND TRANSFER FUNCTION MODEL 1, 1942-1992

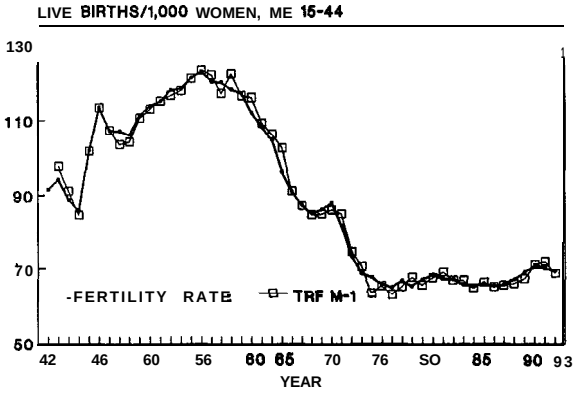
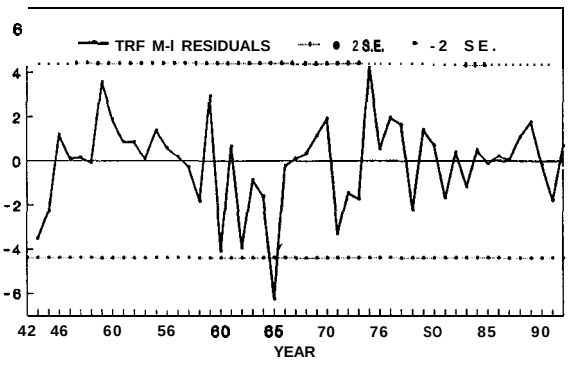


FIG. 6: TRANSFER FUNCTION RESIDUALS, M-I MEAN=-.08, pm.78; SKEWNESS=-.69, p=.05; KURTOSIS=1.54, p=.04



transfer function labeled "birth control". Model fit further improves (R adj.=.991) and the standard error of estimate decreases by an additional 12% to 2.01. The addition of the first-order transfer function has only a minor effect on other model coefficients with the exception of the substantial reduction in the autocorrelation coefficient ($\phi(1)$). The lagged unemployment effect improves somewhat (ω=-.40; t=1.91).

TABLE 4. TRANSFER FUNCTION ANALYSIS ON U.S. FERTILITY RATES: MODEL M-2

INPUT	ω_i	t
D-Unemployment Rate(l)	-.40	-1.91
D-Homecoming Effect	13.73	7.38
D-Marriage Rate(l)	.96	8.62
D-Legal Abortion Rate(l)	-.71	-2.62
Birth Control ω_0	-3.09	-2.83
Birth Control δ_1	.59	3.38
NOISE MODEL: $\phi(1)$	.42	3.03

S.E.E.: 2.01; R^2 adj.: .991
DURBIN-WATSON:1.86; Q(12) 4.03; P=.97

Figure 7 shows improved model fit for the 1960s and early 1970s. Figure 8 exhibits no outliers and the residuals do meet the criteria for residuals (see statistics Fig. 8). A run of positive residuals in the 1950s may indicate that the baby boom is not fully captured by the pulse variable specified for the year 1946.

FIG. 7:U.S. FERTILITY RATES AND TRANSFER FUNCTION MODEL 2, 1942-1992

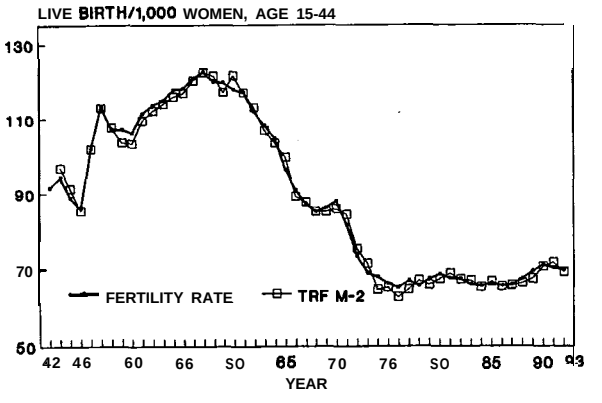
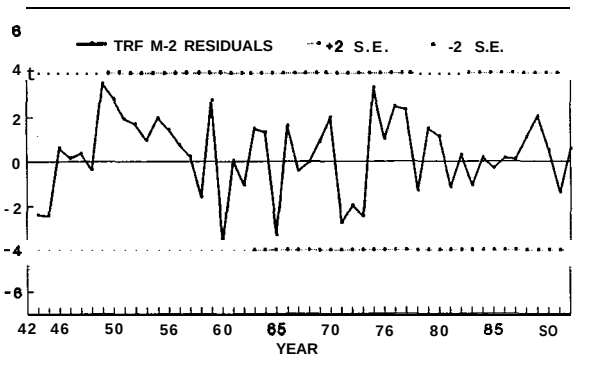


FIG. 8: TRANSFER FUNCTION RESIDUALS, M-2 MEAN=.3, p=.22; SKEWNESS=-.41, p-.25; KURTOSIS=-.38; p=.61



When the residuals of model M-2 are crosscorrelated with prewhitened[8] unemployment and marriage rates no significant correlations emerge within a ±9 year range. With the exception of the first-order transfer function whose numerator and denominator term are highly

correlated (r=.74), the remaining correlations between inputs are less than r=±.16. Since all transfer function assumptions are met, no further models were examined.

Sensitivity Analysis: The analysis was restricted to the period 1942-1988, to test for bias due to incorporating provisional and some estimated data. The transfer function coefficients changed little, the lagged unemployment effect remained at (ω=-.39; t-1.81; model not shown).

The effect of discontinuing *NCHS's* birth underregistration adjustment was tested by adding a pulse variable for the year 1960. The adjustment effect was marginally significant (ω=-3.36; t=1.80) and increased the lagged unemployment effect (ω=-.49; t=2.34; model not shown).

The third analysis examined the causal relationship of unemployment, marriage and fertility rates. Unemployment may effect fertility *indirectly* by postponing marriages. If this assumption is correct, then the unemployment rate and marriage rate must be *contemporaneously,* negatively related. The corresponding correlation coefficient of the differenced series is r=-.33.[2] To obtain the total effect of changes in unemployment rates on fertility rates, the model was reestimated deleting the one-year lagged marriage rates (model not shown). The reestimated lagged *total* unemployment effect increases to ω=-.71; (t=-2.77) controlling the birth under-registration adjustment and ω=-.62; (t=-2.45) without its control.

Discussion:The transfer function analysis indicates that a 1% increase in the unemployment rate, is estimated to reduce next year's fertility rate between -.7% and -.8%; roughly one-third of that effect may be *indirect* by postponing marriages. The effect of a short term recession, i.e. a one-year rise in the unemployment rate of 2%, is estimated to result in a reduction of the fertility rate by approximately 1.5%, or about 60,000 live births during the next year.[3] If the recession drags on for two, or three years, the effect will be cumulative resulting in 120,000 to 180,000 fewer live births.

The homecoming effect after World War II produced an estimated 15% increase in the fertility rate beginning in 1946. The effect may not fully capture the baby boom of the 1950s, but seems to be a reasonable approximation. The homecoming effect is independent of and controls for changes in marriage rates which fluctuated sharply after World War II.

Part of the downward trend in U.S. fertility rates in the *1960s* is attributed to the introduction of effective birth control methods, particularly oral contraceptives. *The National Fertility Study* and the *National Surveys of Family Growth* (Mosher, 1982:8) indicate that use of oral contraceptive increased from 15.1% to 25.1% of fertile females between 1965 and 1973 and then decreased to 22.5% by 1976. Simultaneously, other effective birth control methods, particularly sterilizations, were also increasingly used. Since the first-order transfer function cannot distinguish between different birth control methods, it is likely to represent their combined impact. The cumulative impact is an estimated reduction in the fertility rate of 37/1,000 for the period 1962-1976.

An increase in the legal abortion rate of 1/1,000 fertile females is estimated to be followed by a reduction in the fertility rate of .7/1,000 during the next year. One implication of this finding is that without the liberalization of state abortion laws, the U.S. fertility rates would have been in the 80-90/1,000 range rather the 60-70/1,000 range for the last two decades. Legalization of abortions in Japan, Czechoslovakia and Hungary was also followed by reductions in live births (Rude1 et al. 1973:220-226).

The preceding analyses presume correct model specification. Different models were examined including time series on female labor force participation rates, divorce rates and the age distribution of fertile females (proportion <30 years of age). These inputs neither improved model fit, nor did they eliminate the lagged unemployment effect.

In contrast to OLS multiple regression analysis which can lead to highly suspect findings in the time series context, transfer function analysis with multiple inputs is more likely to yield reasonable results.

NOTES

1. To obtain time series of equal length, female marriage rates for 1988-92 were estimated using linear regression with marriage rates as predictor (CDC, 1993). Likewise, legal abortion rates for 1990-1992 were forecasted by exponential smoothing based on the period 1979-1989.

2. The Durbin-Watson statistic tests the null hypothesis that regression residuals are independent against the alternative that the residuals follow a first-order autoregressive process.

3. Computer routines are available which correct OLS residuals for first-order autocorrelation. However, significant autocorrelation may remain in the "corrected" residuals when the time series is non-stationary, or the underlying autocorrelation process is of higher order. For instance, the autocorrelation function for the OLS model with first-order autoregressive correction resembled that shown in Table 1, but was somewhat

dampened. Autocorrelation coefficients at lags 1,2 and 12 remained statistically significant.

4. A second-order autoregressive noise model was also tried using the original time series. The autoregressive parameters approached stationarity suggesting the need for differencing.

5. When included in the model, the constant term turned out to be statistically insignificant.

6. There is a statistically insignificant, negative, second-year lagged effect for the legal abortion variable.

7. The five-year pulse and single decay effect use only 2 degrees of freedom and constrain the intervention to annual reductions and decay effects identical for each year (1962-1966). The cumulative effect of the first-order transfer function is reverse S-shaped and corresponds to the S-shaped curve of a typical product life cycle (Kotler and Clarke, 1987:336).

8. Prewhitening entails filtering of the original time series by its ARIMA process resulting in a new series of white noise residuals. The latter series is cross-correlated with residuals of model M-2.

9. The correlation coefficient is based on the period 1950-1992 for which the marriage rate time series is variance stationary.

10. The estimate assumes 4,000,000 live births per year.

REFERENCES

Berki, S.E., Wyszewianski, L., Lichtenstein R. et al. 1985. Health insurance of the unemployed. Medical *Care* 23(7):847-854.

Box, J.E.P., Jenkins, G.M. 1976. Time Series *Analysis:* Forecasting and *Control.* San Francisco, CA. Holden-Day.

Brenner, M.H. 1971. Economic changes and heart disease mortality. Am. *J. Publ.* Health 61:606-611.

Brenner, M.H.1987. Economic changes, alcohol consumption and heart disease mortality in nine industrialized countries. *Soc.* Sci. *and Medicine* 15(2):119-132.

Cates, W. 1982. Legal abortion: The public health record. *Science* 215(26):1586-1590.

Centers for Disease Control and Prevention, NCHS. 1993. *Monthly* Vital *Statistics* Report 42(9) Febr. 16.

Doan, T.A. 1992. RATS 4.0. Evanston,IL. Estima.

Dowdall, G.W., Marshall, J.R., Morra, W.A. 1990. Economic antecedents of mental hospitalization: A nineteenth-century time-series test. *J.* Health and *Soc.* Behavior 31(2):141-147.

Economic Report of the President, 1982. Washington, DC: Gov. Printing Office.

Henshaw, S.K., Van Vort, J. 1990. Abortion Services in the United States, 1987 and 1988. *Family Planning Perspectives* 22(3): 102-109.

Iversen, L., Sabroe, S., Damsgaard, M.T. 1989. Hospital admissions before and after shipyard closure. *Brit.* Med. *J.* 23(7):847-854.

Jahoda, M. 1982. *Employment and Unemployment: A* Social *-Psychological Analysis.* New York, N.Y. Cambridge University Press.

Kasl, S.V. 1979. Mortality and the business cycle: Some questions about research strategies when utilizing macro-social and ecological data. Am. *J. Publ. Health* 69(8):784-788.

Kotler, P., Clarke, R.N. 1987. Marketing for *Health Care Organizations.* Englewood Cliffs, N.J. Prentice-Hall.

Liu, L.M., Hanssens, D.M. 1982. Identification of multiple-input transfer function models. *Communic. in Statistics, Part A-Theory and Methods* 11(3):297-314.

Maddala, G.S. 1988. *Introduction to Econometrics. New* York, NY. Macmillan Publishing Company. 185-222.

McCleary, R., Hay R.A.1980. *Applied Time Series Analysis for the Social Sciences.* Beverly Hills, CA. Sage Publications.

Mosher, W.D.1982. *Trends in Contraceptive Practice, United States, 1965-76*,NCHS, Vital and Health Statistics, Series 23, No. 10. DHHS Pub No. (PHS) 82-1986.

NCHS: *Vital Statistics of the United States, 1987,* Vol III, Marriage and Divorce. DHHS Pub No. (PHS) 91-1103. PHS, Washington. U.S. Government Printing Office, 1991.

NCHS: *Vital Statistics of the United States, 1988,* Vol I, Natality. DHHS Pub No. (PHS) 90-1100. PHS, Washington. U.S. Government Printing Office, 1990.

Pankratz, A. 1991. *Forecasting with Dynamic Regression Models. New* York, NY. John Wiley.

Piacentini, J.S., Cerino, T. 1990. *EBRI Data Book on Employee Benefits.* Employee Benefit Research Institute Education and Research Fund, Washington.

Pindyck, R.S., Rubinfeld, D.L. 1991. *Econometric Models and Economic Forecasts* (3rd ed.) New York, NY. McGraw Hill.

Rudel, H.W., Kincl,F.A., Henzl, M.R.1973. *Birth Control; Contraception and Abortion. New* York, NY. Macmillan.

Starrin, B. Larsson, G., Brenner, S. et a1.1990. Structural changes, ill health and mortality in Sweden, 1963-1983: A macroaggregated study. *Intern. J. Health Services* 20(1):27-42.

U.S. Dept. of Labor, Bureau of Labor Statistics. 1992. *Employment and Earnings* 39(1):162.

A Comparison Of Socioeconomic Indicators Across Three Decades

Alan B. Humphrey, University of Rhode Island
Jay S. Buechner, Rhode Island Department of Sealth
Larry Manire, Databasics Inc.

INTRODUCTION

The objective of this study is to test a procedure used on the 1980 census data to create socioeconomic clusters of census tracts using the 1990 census data. In addition a comparison with 1970 socioeconomic clusters will be made.

Rhode Island is a completely census tracted state and has since 1970 been coding much of its health data to these census tracts. The census tracts contain approximately 5,000 people, are defined to be uniform demographically, and are hierarchically consistent with Cities and Towns. Aggregates of these census tracts can provide populations subsets that have a large enough population base to provide stable rates. If the aggregation is based on a common set of demographic variables they can be used to calculate utilization rates, vital statistics rates, and market shares.

The decennial Census of Population and Sousing provides the social, demographic and economic data necessary for this aggregation process. The data, available in computer compatible media, can be obtained at the block group, census tract, and city/town levels. This aggregate data is so large that variables need to be carefully selected and defined so that they are consistent with the analysis objectives and meaningful to those who may want to use the final geographic clusters.

In the early 1970's Sakoda and Karon (1) used the cenaua data to classify these tracts into four socioeconomic groups; high , middle, low, and poverty. Their methodology included using factor analysis to reduce 25 census variables to five meaningful factors. The factor with the highest loading was called socioeconomic (SES) which was used to classify the census tracts. The other factors were Group Quarters, Migration, Suburban, Non-ethnic, and Social Stability.

In 1984 Humphrey, Buechner, and Velicer (2) expanded on the work done by Sakoda and Karon and classified census tracts into four socioeconomic areas. The methodology used was a combination of factor analysis and cluster analysis. Factor analysis was used to reduce 60 census derived variables into four factors; Wealth, Education, Suburban, and Group. The first two factors were used as SES and on their factors scores the census tracts were clustered into four SES groups.

The methodology used in 1980 was repeated with the 1990 census data. A five factor solution was obtained with the factors being Socioeconomic Status, Ethnicity, Urban, Family, and Elderly.

All three analyses had the same objectives, namely reducing a number of census variables to a manageable number using multivariate techniques. The census tract was the unit of measurement and data was obtained from the Censua tapes. While the analysis used in all three cases was factor analysis, the 1970 study used a FORTRAN program written by Sakoda while the 1980 and 1990 studies used factor procedures from the Statistical Analysis System (SAS). The number of variable used in analysis increased frcnn 56 in 1970 to 60 in 1980 and 70 in 1990. Also, there were 213 census tracts for both the 1970 and 1980 analyses, but these were increased to 232 in 1990 by splitting some tracts.

VARIABLE SELECTION AND DEFINITION

In the 1984 study 60 variables *were* constructed from the census tabulation and used to create the SES clusters. In the 1990 census some of the variables used in 1980 were not included while other questions were added. Therefore, it is not possible to do a perfect emulation of the 1980 study. There were 70 variables used for the 1990 study, of which 50 were the same as those used previously, 10 variables were not included in 1990 due to unavailability, and 10 new variables were added. Most of the variables were in the form of ratios, though there were a few continuous variables such as family income and value of owner occupied house which were not ratios.

1980 and 1990 *FACTOR ANALYSES*

Principal components method was used to extract the factors and then rotated using varimax rotation. Scree plots were used to determine the number of factors to retain.

If a variable loaded on no components (unique variable) or on *more* than one component (complex variable) were excluded as far as computing factor scores. Unique variables have little variability, have low reliability, and are measuring a domain that is not consistent with the other variables in the data set. Complex variables, since they are related to more than one component, are difficult to interpret and increase the correlation between the scale scores.

Once the components are determined, the final step in the factor analysis is the calculation of scores for each census tract. Two accepted alternatives exist for creating scores for the new variables: component scores and scale scores. Component scores represent weighted linear composites of all the observed variables, where the weights are proportional to the factor loadings. The principle advantage of component scores is that they are completely uncorrelated. Scale scores, on the other hand, are unweighted linear composites of only those observed variables that are selected for inclusion in each component. The principle advantages of scale scores are simplicity of scoring, ease of interpretation, and stability in cross-validation. Scale scores, however, will generally be somewhat correlated.

1980 and 1990 CLUSTER ANALYSES

Ward's hierarchical clustering method was used to create clusters based on the factor scale scores. The cubic clustering criteria (3) was used to help determine the optimum number of clusters.

COMPARISON OF FACTORS: 1970 - 1990

In 1970 and 1990 one factor was identified as the SES factor. The 1980 analysis required the use of two factors to create SES. The similarity of these factors, based on the variables included (Table 1.) is interesting in that 1970 had no education variables included in SES and was primarily comprised of income related variables. The 1980 and 1990 SES was

Table 1. Variables included in the 1970, 1980, and 1990 Socioeconomic Factors

1970 SES

Dishwasher
Median Family Income
More than one Bathroom
Median Value of Owner Occupied House
Median Gross Rent
New Souse Coating $25,000 or *More*
Professional / Managerial Occupation
- Blue Collar Occupation

1980 SES

WEALTH

Single Housing Unit
White
Families
- Poverty
- Female Beaded Families in Poverty
- Families with Children in Poverty
- Public Assistance Income
- Divorced or Separated

EDUCATION

Completed College
Self Zmployed
Speak English
Central Heating
Professional / Managerial Occupation
- Completed High School
- Blue Collar Occupation
- Median Gross Rent
- *Foreign* Born

1990 SES

Completed College
Professional / Hanagerial Occupation
Percapita Income
Median *Family Income*
Median Value of Owner Occupied House
Median Gross Rent
Self Employed
Bouse with 4+ Bedrooms
Speak English
- Female Beaded Families with Children
- Families on Public Assistance Income
- Disabled Workers
- Completing less than Ninth Grade
- Completing less *than* High School
- Blue Collar Occupation

comprised of a combination of Income, Education, and Occupation variables.

There were 15 variables included in the 1980 analysis that were not present in the 1990 analysis. Thirty variables were added to the 1990 analysis that were not available in 1980. Of these, ten were not included in defining a factor. *Ten of* the twelve variables that were not used in defining a factor in 1980 were used as defining variables in 1990. Nine of the seventeen variables defining SES in 1990 were used to define the Wealth and Education factors in 1980. The complete distribution of variables across factors is presented in Table 2.

CLUSTER COMPARISONS: 1980 - 1990

The clusters formed in 1980 and 1990 used similar methodologies with different variables defining SZS. The 1980 clusters used two factors to cluster the census tracts (WEALTH and EDUCATION) while the 1990 clusters were obtained using the one SES factor. In addition there were 213 census tracts in 1980 and 230 census tracts in 1990. The dietribution of census tracts and population in the two years (Table 3) is quite for the HIGH and LOW clusters. About a quarter of the population is in the ZIGB 1980 cluster which drops to about 11% in 1990. Conversely, about a quarter of the population was in the LOW 1980 cluster which rose to about, 45% in 1990.

If 1980 and 1990 SES factors are measuring a common characteristic and there is no change in the population one would expect that the census tracts would be placed in the same group in 1990 as 1980. Any deviation from this indicates a difference in the measure or a change in the demographic composition of the population. The greatest disparity can be seen in the MIDDLE and LOW clusters (Table 4).

The scale scores used to define the clusters are somewhat different in that there was a greater range in 1990 than in 1980 (Table 5). However, the average income and percent completing college are very similar for the two periods. The percentages of those completing college are somewhat higher for all clusters in 1990 than in 1980 except for the poverty cluster.

To measure the internal variability of the clusters coefficients of variability were calculated for theIncome, College, and SES *scale* scores variables (Table 5). In both years the internal cluster variability was nearly the same. The 1990 clusters were less variable based on the scale scores except for the HIGH cluster, less variable for all clusters based on College, and more variable for all but the MIDDLE cluster for Income.

Table 2. Distribution of Variables across the 1980 and 1990 Factors.

	1990 FACTORS						
1980 FACOTRS	SES	MINORITY	URBAN	SUBURBAN	ELDERLY	1990 NOT PRESENT	TOTAL
WEALTH	3	3	2	1		1	10
EDUCATION	6		1			2	9
SUBURBAN			1	3	2		6
GROUP		2	4	1			7
NOT PRESENT IN 80	4	6	4	2	4	10	30
NOT INCLUDED IN 80	4	2		3	1	2	12
TOTAL	17	13	12	10	7	15	74

Table 3. 1980 and 1990 Cluster Distributions of Census Tracts and Population

	1980				1990			
	Census Tracts		Population		Census Tracts		Population	
	#	%	#	%	#	%	#	%
HIGH	51	24%	242,038	26%	24	10%	108,084	11%
MIDDLE	77	36%	359,216	38%	69	30%	308,723	31%
LOW	48	23%	213,908	23%	105	45%	469,978	47%
POVERTY	37	17%	131,992	14%	34	15%	116,679	12%
TOTAL	213		947,154		232		1,003,464	

Table 4. Number of Census Tracts per SES Cluster in 1980 and 1990

	1990 CLUSTERS				
1980 CLUSTERS	HIGH	MIDDLE	LOW	POVERTY	TOTAL
HIGH	19	35	5	0	59
MIDDLE	5	30	50	0	85
LOW	0	3	40	6	49
POVERTY	0	1	10	26	37
TOTAL	24	69	105	32	230

Table 5. 1990 Average Income, Peroent Completing College, 1980 SES Scores, and 1990 SES Scores for the 1980 and 1990 SES Clusters

	1990		COLLEGE - 1990		SES80	SES90
	CLUSTER 80	CLUSTER 90	CLUSTER 80	CLUSTER 90	CLUSTER 80	CLUSTER 90
HIGH	42,741	45,074	0.42	0.52	0.96	1.93
MIDDLR	34,692	37,876	0.29	0.36	0.37	0.61
LOW	27,300	30,680	0.16	0.21	-0.42	-0.37
POVERTY	18,591	18,110	0.12	0.10	-1.55	-1.43
AVERAGE	32,592	32,592	0.27	0.27	0.04	0.02

Table 6. Coefficients of Variability for Income, Percent completing College, 1980 SES Scores, and 1990 SES Scores for the 1980 and 1990 SES Clusters

	INCOME		COLLEGE		SES80	SES90
	CLUSTER 80	CLUSTER 90	CLUSTER 80	CLUSTER 90	CLUSTER 80	CLUSTER 90
HIGH	0.26	0.35	0.30	0.23	0.26	0.30
MIDDLE	0.20	0.18	0.32	0.20	0.59	0.51
LOW	0.19	0.24	0.31	0.27	0.88	0.80
POVERTY	0.29	0.33	0.52	0.37	0.25	0.22

Table 7. Hospital discharge rates (per 1,000 population) for injuries, Rhode Island C-unity Hospitals 1988 - 1991

	1990 CLUSTERS				
1980 CLUSTERS	HIGH	MIDDLE	LOW	POVERTY	TOTAL
HIGH	20.83	21.77	16.69		21.04
MIDDLE	44.71	24.69	24.98		25.74
LOW		19.78	32.71	33.17	31.84
POVERTY		24.37	37.18	44.14	41.52
TOTAL	24.52	22.99	28.56	42.02	27.95

Table 8. Cluster Comparieons for Mean of Average Length of Stay for injuries, Rhode Island Community Hospitals 1988 - 1991

1980 CLUSTERS	1990 CLUSTERS				
	HIGH	MIDDLE	LOW	POVERTY	TOTAL
HIGH	7.34	7.99	7.91	-	7.76
MIDDLE	9.12	7.87	8.72	-	8.46
LOW		8.85	8.56	8.10	8.52
POVERTY		12.04	7.74	8.60	8.43
TOTAL	7.84	8.03	8.52	8.53	8.33

COMPARISON OF HEALTH VARIABLES

Hospital discharge data from Rhode Island community hospitals for 1988 through 1991 was selected for those diagnoses of injury. The Discharge rate and the length of stay generally increases as the SES category goes from HIGH to POVERTY. The rates based on the 1980 clusters do follow a general increase with the POVERTY rate about twice that of the HIGH rate (Table 7). The pattern is not as clear for the rates based on the 1990 clusters, in fact the MIDDLE cluster has a slightly lower rate than the HIGH cluster. A comparison of the rates within the table follows the same pattern except for the MIDDLE clusters in both 1980 and 1990.

There is a similar pattern for length of stay (Table 8). However, the length of stay for the 1990 clusters follows the expected pattern better than the 1980 clusters

CONCLUSIONS

Factor analysis of census data for 1970, 1980, and 1990 has provided the basis for partitioning Rhode Island into Socioeconomic Status (SES) clusters. Each analysis generated different factors and hence different definitions of SES. While the definitions differ, they are similar and provide a common basis for aggregating census tracts into meaningful groups. These groups have been used by health planners to quantify health problems and issues. With new data available every ten years and hence a new definition of SES and new SES clusters, the question of comparability over time has been raised.

The results of this study indicate that the measurements of the extremes (HIGH versus POVERTY) is comparable over time and that the shifts from one cluster to an adjoining cluster may be due to changes in the population characteristics and not to changes in SES definitions. However, the shifts observed in the MIDDLE and LOW clusters indicate that there are some comparability problems arising from SES definitions.

These differences call for analysis of 1970, 1980, and 1990 data using the same variables in each year and the same methodology. This analysis would then show which variables are the most stable over time with respect to differentiating geographic areas, which factors are the most important for demographic constructs, which areas of the State are changing, and finally a measure of the magnitude of that chance.

References

1. Sakoda, J. M. and Karon, J. P. Factor Analysis of Rhode Island Cenmum Tracts, 1970. Providence: Rhode Island Health Services Research , Inc. 1974.
2. Humphrey, Alan B., Buechner, Jay S., and Velicer, Wayne F. Differentiating Geographic Areas by nSocioeconomic Characteristics. The Northeamt Journal of Business and Economics. 13:47-64. 1987.
3. Sarle, Warren S., Cubic Clustering Criteria. SAS Technical Report A-108. Cary NC: SAS Institute Inc., 1983.

Session **R**

Minority Health Research

Toward the year

Refining the measures

HEALTH DATA ON HISPANIC WOMEN: METHODOLOGICAL LIMITATIONS

***Hortensia* Amaro, Boston University**
school of Public Health

This presentation focuses on two issues pertinent to research on Hispanic women's health: 1) an assessment of the adequacy of current national health data systems for informing the health status of Hispanic women, and 2) models to support research on Hispanics and specifically on Hispanic women.

Relevance of National Data Systems

National health data compiled by federal government agencies are critical for understanding morbidity and mortality among Hispanic women. Annual health statistics and publications are used to identify major health problems needing legislative and programmatic attention. Problems of inadequate Hispanic ethnicity identifiers, small sample sizes, linguistic and cultural inappropriateness of measures in national health studies have resulted in a dearth of information on the health problems and conditions of Hispanics, including women.[1] The limitations of national data systems for providing data on the health status of Hispanics pose severe problems for understanding and addressing the health problems of Hispanic women in the United States. Without baseline data it is difficult to develop priorities for prevention programs and for health services to address the health problems of Hispanic women and to assess the impact of programs and services on the health of Hispanic women. Yet, progress in the availability of health data on Hispanics and other minority groups has been slow. Notwithstanding a few exceptions (eg., vital health records and the Hispanic Health and Nutrition Examination Survey), there has been little improvement in the nation's health data systems ability to provide data on Hispanics and other racial/ethnic groups.

Limitation of National Data Systems Procedures

In order to ascertain which national health data sets provide reliable information on the health status of Hispanic women, a review of the major national data systems was conducted. Fifteen data systems were identified based on their relevance to health problems of women and the importance of these data systems for informing public health efforts and health service needs. A survey questionnaire was sent to individuals identified as the contact person for each data set and a follow-up telephone interview was conducted to verify and clarify responses on the survey. The questionnaire asked for general background information on each data set and on the extent to which Hispanics were included in the study. More specifically, this presentation reports on findings regarding: 1) if and how Hispanic ethnicity was assessed; 2) whether Hispanics were included in the sample, if they were oversampled, and the sample size for the study, for Hispanics and their subgroups and for Hispanic women; 3) whether detailed gender x age analyses can be conducted for Hispanics and for Hispanic groups; and 4) if Spanish speaking respondents were included and how translation of instruments was handled.

Results

Ethnicity identifiers. Table 1 presents findings from the survey. Identifiers for Hispanic subgroups were collected in 10 of the 15 data sets reviewed. However, 5 data sets continue to use a global identifier (i.e., Hispanic) that does not allow for analyses of Hispanic groups.

Even when adequate Hispanic ethnicity identifiers have been included, serious problems in the collection and completion of such information continues to hamper the usefulness of the data. For example, the 1990 mortality data, which are based on Hispanic-origin population from 46 States and the District of Columbia, exclude Hispanics from New York City because more than 10 percent of death certificates had inadequate data for ethnicity.[2] The Puerto Rican population (58% covered) is grossly underrepresented. Since about half of deaths to Puerto Ricans are accounted for by New York City, the mortality data for Hispanics overall and for Puerto Ricans are underestimated.[2]

Sampling Strategies. Most federal population-based surveys employ a stratified probability sample and as such, obtain samples of Hispanics in proportion to the population size, about 9% or less of the national sample. This yields a sample size that is too small to allow for meaningful analyses of data for women of Hispanic subgroups. As a result, if analyses are conducted on Hispanics at all, they often group all Hispanics even when subgroup ethnicity identifiers are employed in the surveys. Of the general population surveys in this review, two oversampled Hispanics (ie., NMES and NHSDA) and one oversampled Mexican Americans (ie.,

NHANES III). One survey (ie., H-HANES) focused specifically on Hispanics and sampled the three largest Hispanic groups (Mexican Americans, Cuban Americans, and Puerto Ricans). Collection of data on Hispanics and Hispanic subgroups under the major surveys such as the NHANES, NMES and NMHIS has been sporadic. As a result, it has not been possible to document changes in diseases and conditions across time.

Most of the major national health surveys that document the health status of the U.S. population have not provided meaningful data on specific Hispanic subgroups due to small samples. For example, the National Health Interview Survey whose purpose is to assess the health status of the non-institutionalized population in the U.S. completed a 1990 fielding that included 11,580 Hispanics who comprised 10% of the sample. The 1990 sample of Hispanics is sufficiently large to conduct some general analyses for the two largest Hispanic groups, Mexican Americans and Puerto Ricans. However, detailed gender x age analyses, especially for Puerto Ricans and other Hispanic groups are not feasible.

The National Health and Nutrition Examination Survey conducted by NCHS has also traditionally not included sufficiently large samples to provide data on the health status of women in Hispanic subgroups. The most recent fielding (1988-1991) of this survey oversamples Mexican Americans who comprise 30 percent of the 15,884 person sample and provide the largest sample ever of any Mexican Americans in a national health survey. The large sample size allows for detailed gender x age analyses among Mexican Americans. However, the survey will not provide data on Puerto Ricans, who have the most jeopardized health status among Hispanics. National data systems that provide information on child and maternal health (eg., Child Health Supplement to the NHIS, NSFG, and NLSY) have been relatively uninformative on the health problems of Hispanic women and children. Similarly, the major data systems for assessing expenditures for health care and health insurance coverage (eg., NMCUES and NMES) also have small samples of Hispanics and yield very limited data on women in Hispanic subgroups.

The national study which has made the greatest contribution to informing Hispanic women's health is the Hispanic Health and Nutrition Examination Survey (H-HANES) conducted by NCHS between 1982 and 1984. The purpose of the H-HANES was to provide information about the health of Hispanic children and adults living in the United States. Since approximately half of the sample for each Hispanic subgroup was female, the H-HANES contains a sufficiently large sample of Hispanic women to conduct gender comparisons and gender-specific analyses of many major diseases and conditions and health care utilization patterns. The H-HANES continues to be the largest and most complete data set that provides information on Hispanic women and researchers continue to use it to assess the general health status of Hispanics and to plan for health care needs and programs. This single research effort resulted in the largest number of publications on Hispanic health and in a major advance in our knowledge of Hispanic health. Two major health journals, The American Journal of Public Health and the Journal of the American Medical Association published thematic issues based on articles employing the H-HANES data.[3,4] Despite the usefulness of the data produced by the H-HANES, there are no plans to repeat it now that the data are ten years old. Further, they are no plans to include all of the major Hispanic groups in the future NHANES.

Linguistic and Cultural Comparability. The inclusion of adequate samples of Hispanics in national health studies must also be accompanied by the use of appropriate methods that render the results representative and valid. Until recently, most surveys were conducted in English, a practice that introduced bias in the Hispanic sample through exclusion of Spanish dominant respondents. This represented a significant omission since two out of three Hispanics speak Spanish at home,' reflecting greater fluency and/or preference for Spanish. Further, findings from the H-HANES indicate that a significant minority of Hispanics have limited or no English language fluency.[6]

In recent years many national surveys have included Spanish dominant respondents through the use of translated questionnaire and interview instruments. In our review of ten major data systems that collect information directly from sample persons, all but two (ie., NHIS and the Child Health Supplement to the NHIS) reported that they are currently using Spanish translations of instruments. The two surveys that do not employ Spanish translations of the instruments rely on in-house translators who translate the instruments when a Spanish speaking respondent is encountered. If this cannot be accomplished, the interviewer terminates the interview. This practice introduces bias in the sample of Hispanics included in the study and presents threats to validity and reliability of the measurements. Using interviewers for simultaneous translation without the control

introduced by standardized procedures required for valid translations[7,8] brings into question the comparability of the data obtained from English and Spanish respondents.

In summary, while most major health data systems reviewed have incorporated ethnicity identifiers for Hispanic groups, serious limitations continue to exist in the ability of national data systems to inform Hispanic women's health. Most data systems continue to have small samples of Hispanic women and therefore do not allow for accurate estimates of diseases, health conditions, and health care needs in this population. Small samples generally do not allow for gender x age and other detailed analyses for subgroups of Hispanic women and therefore provide data of limited usefulness in planning for health care and public health programs. Finally, the issue of linguistic and cultural appropriateness of instruments used in many surveys needs attention since the lack of proven cross-cultural validity of such measures brings into question the interpretation of findings.

The Hispanic Health Research Consortium

The 1990 Disadvantaged Minority Health Improvement Act in its report language, recognized that in addition to changes in national data systems, there was a need to build a research infrastructure within the Hispanic community. Unlike, the African American community in the U.S., the Hispanic community lacks a system of higher education that can support a research infrastructure on health research. As a result, Hispanic researchers and students often lack important mentoring relationships that provide support and lend critical expertise to the advancement of their. research. Further, the lack of a research infrastructure often results in Hispanic researchers working in isolation without the support of peers with shared interests who can provide valuable input and collaboration.

Recognizing the need for developing a research infrastructure to support Hispanic health researchers, the 1990 Disadvantaged Minority Health Improvement Act, through a legislative mandate to NCHS, provided one year of funding to the National Coalition of Hispanic Health and Human Services Organizations (COSSMHO) to conduct secondary data analyses of existing national data sets through its Hispanic Health Research Consortium (HHRC) founded in 1982. With these funds COSSMHO developed an initiative that funded five university-based research teams to address research questions focused on Hispanic women's health. While modest in funding, this project represents the largest single effort to date to improve the knowledge base on Hispanic women's health.

The HHRC model incorporates the following components: 1) funding for a series of university-based research teams composed of a Senior Research Fellow (typically a junior to mid-level faculty member), one or more Junior Research Fellows (undergraduate or graduate students), and a statistical consultant; 2) mentoring to Senior Research Fellows by a Senior Research Scientist who directs the HHRC and mentoring to Junior Research Fellows by the Senior Research Fellow; 3) national level networking with other Hispanic researchers, which was facilitated through COSSMHO's extensive network; and 4) dissemination of findings through presentations at regional and national scientific meetings and through publication of presentations in scientific journals. The HHRC Model has been extremely successful and cost effective in promoting the development of Hispanic researchers and in providing a meaningful contribution to the scientific body of literature on Hispanic health. In the last year, the HHRC researchers produced over 12 research presentations on Hispanic women's health, which have been presented at scientific meetings and which will be published in scientific journals. Equally important in this model, is a supported team context, that facilitates skills enhancement, collaborations and national networking among Hispanic researchers. It is imperative that we look for ways to increase the representation of Hispanic women in health research, to foster the development of Hispanic health researchers and to train young scientists. In doing so, we must not only increase funding for individual investigators conducting research on Hispanic health but also fund mechanisms that support the development of a research infrastructure on Hispanic health. Models for funding research which incorporate mentoring, networking and team efforts hold great promise for the training and development of Hispanic researchers who can contribute to research on Hispanic women's health and to Hispanic health research in general.

Acknowledgements

Work for this presentation was conducted by the author while a visiting Senior Research Scholar at the National Coalition of Hispanic Health and Human Services Organization during which time she directed the HHRC. Work for this Presentation was supported with funding from the National Center for Health Statistics. Special recognition goes to Dr. Britt Ellis who provided invaluable assistance in the collection of data. I

also thank Marjorie Greenberg from NCHS for her useful comments on the survey and Dr. Ruth Zambrana for her suggestions on an earlier version of this presentation.

References

1. General Accounting Office. Hisoanic Access to Health Care: Significant Gaps Exist, Washington, D.C.: U.S. Government Printing Office; 1992.
2. National Center for Health Statistics. Advance Report of Final Mortality Statistic, 1990. Monthlv Vital Statistics Report, 1993, 40 (7) (Suppl.).
3. Trevino FM, Falcon AP, Stroup-Benham CA. (Eds.) Hispanic Health and Nutrition Examination Survey; 1982-1984: Findings on health status and health care needs. American Journal of Public Health 1990, 80 (suppl.):1-72.
4. Journal of the American Medical Association, 1991, 265.
5. Schick FL, Schick R. Statistical Handbook on U.S. Hispanics. New York: Oryx Press; 1991.
6. Amaro H, Whitaker R, Coffman G, Heeren T. Acculturation and marijuana and cocaine use: Findings from the H-HANES 1982-84. AJPH, 1990, 80 (suppl.): 54-60.
7. Marin G, Marin BV. Research with Hisoanic Populations. Applied Social Research Methods Series, London: Sage Publications; 1992: 92-99.
8. Zambrana RE. Cross-cultural methodological strategies in the study of low income racial ethnic populations. In Grady ML. (Ed.), Primarv Care Research: Theorv and Methods. Rockville, MD: Agency for Health Care Policy and Research, 1991: 102-114.

Table 1. Summary of Hispanic Ethnic Identifier, Sample Sii and Spanish Translation of Instruments on Selected Data Systems

Name Agency Year	Purpose	Hispanic Ethnic Identifier	Total Sample Sic Oversample: Hispanic N Hispanic Female N	Translated Instrument
National Vital Registration System/ NCHSI 1989	To Collect general purpose statistics regarding births including prenatal care, birthweight, birth rates, teenage and unmarried births, family formation and dissolution, and characteristic8 of the fetus, the pregnancy, and the mother	Mexican American Puerto Rican Cuban Central & South American Other Hispanic	Total N = 3,903,000 Oversample: NA N = 530,000 Hispanics	NA
National Mortality Statistics: General Mortality/ NCHSI 1989	To provide general mortality data including infant mortality, life expectancy, and causes of death	Mexican American Puerto Rican Cuban Central & South American Other Hispanic	Total N = 2,150,466 Oversample: NA N = 76,379 Hispanics N = 30,404 Hispanic women	NA
National Notifiable Diseases Surveillance System/ CDC/1991 (ongoing)	To determine the prevalence of 49 reportable diseases and conditions	Hispanic	Total N = 321,366 Oversample: NA N = 19,642 Hispanics N = 9,178 Hispanic women	NA
National Health Interview Survey/ NCHS/ 1990	To provide basic data on health status and disability, utilization ofhealth care, AIDS knowledge and attitudes, alcohol and drug use, and Year 2000 health prevention objectives	Mexican/Mexicano Chicano Puerto Rican Cuban Other Latin American Other Spanish Spanish- don't know type Multiple Hispanic category	Total N = 122,000 Oversample: No N = 11,580 Hispanics N = 5,928 Hispanic women	No
National Survey of Family Growth/ NCHSI 1988	To provide information on contraception and sterilization, adolescent sexual activity and pregnancy, family planning, adoption, breastfeeding, and infertility in women age 15 to 44	Mexican American Puerto Rican Cuban Central & south American Other Hispanic	Total N = 8,000 Oversample: No N = 641 Hispanic women	Yes

Hispanic Health and Nutrition Examination Survey/ NCHSI 1982-1984	To provide information about the health status and health cab utilization among Hispanic adults and children including: prevalence of disease/conditions, undetected diseases/conditions, use of cigarettes, alcohol, illicit drugs and depression diabetes	Mexican/Mexicano Chicano Hispano Puerto Rican Boricuan Cuba" American Cuban Latin American Other Spanish Other Hispanic	Hispanic focused study Examined N = 7,197 Mexican Americans N = 2,645 Puerto Ricans N = 1,291 Cuban Americans Women: N = 3,946 Mexican Americans N = 1,597 Puerto Ricans N = 721 Cuban Americans	Yes
National Health and Nutrition Examination Survey III/ NCHSI 1988-1991	To provide information about tbe health status and health care utilization among adults and children including: prevalence of disease/ conditions, undetected diseases/conditions, including heart disease, diabetes, osteoporosis. imn deficiency anemia and other nutritional disorders, lead exposure, & children's growth and development	Mexican Mexican American Other Latin American Other Spanish	Total N = 15,884 Oversampled: Mexican Americans N = 4,856 Mexican Americans N = 2,479 Mexican American Women	Yes
Child Health Supplement to the National Health Interview Survey/ NCHSI 1991	To provide data relevant to the healthy People 2000 objectives related to child health and immunization	Mexican/Mexicano Mexican American Chicano Puerto Rican Cuban Other Latin American Other Spanish Multiple Hispanic	Total N = 17,102 Oversample: No N = 2,116 Hispanic children N Hispanic girls not available	No
National Ambulatory Medical Can Survey/ NCHSI 1990	To estimate the number of office visits to office-based physicians and to describe patient demographics, reason for visit, diagnosis, services provided	Hispanic ethnicity noted if mentioned by provider	Total N = 48,065 Oversample: NA N = 2,235 Hispanics N = 1,274 Hispanic women	NA
National Hospital Ambulatory Medical Cam Survey/ NCHS/ 1992	To provide data on the U.S. population's experience in hospital emergency and outpatient medical care	Hispanic	Total N = 500 hospitals and 72,000 visits Oversample: No No N on Hispanics tabulated	NA
National Medical Care Utilization and Expenditure Survey/ NCHS & HCFA/ 1980	To provide data on use ad expenditures for health cam	Mexicano/Mexican American Puerto Rican Chicano Cuban Other Latin American Other Spanish	Total N = 17,123 Oversample: No N = 1,192 Hispanics N = 596 Hispanic women	Yes
National Medical Expenditure Survey/ NCHS/ 1987	To provide data on health can expenditures and health insurance coverage	Mexican/Mexicano/Mexican American Chicano American Puerto Rica" Cuban Other Latin American Other Spanish	Total N = 37,446 Oversampled: Yes N = 4,308 Hispanics N = 2,204 Hispanic women	Yes (household sample)
National Household Survey on Drug Abuse/ NIDA/ 1992	To develop estimates of drug use prevalence	Mexican American Puerto Rican Cuban Other Hispanic	Total N = 28,000 Oversample: Yes N = 7,916 Hispanics N = 4,386 Hispanic women	Yes
National Crime Victimization Survey/ BJS (1991) ongoing	To measure rate of crime victimization for rape, robbery, assault, personal theft, household theft, burglary and motor vehicle theft	Hispanic	Total N = 47,000 Oversample: No N = 3,760 Hispanics N = 1,880 Hispanic women	Yes
National Hospital Discharge Survey/ NCHSI 1991	To provide national estimates of shod-stay hospital use	Hispanic	Total N = 266,00 discharges in 474 hospitals Oversample: No 75 % of ethnicity data missing; no estimates available on size of Hispanic sample	

COOPERATIVE AGREEMENTS TO ADVANCE THE UNDERSTANDING OF THE HEALTH OF ASIAN AND PACIFIC ISLANDER AMERICANS

Dong Suh, Asian American Health Forum, Inc.

Introduction

During the 1980s, the Asian and Pacific Islander (A/PI) population nearly doubled from 3.1 million to 7.3 million. The 95% growth rate in a decade was larger than any other racial group. This growth has challenged the current health information system since the current system is inadequate in providing necessary and meaningful information for this emerging population. To begin to address the rising need for more health information, National Center for Health Statistics/Centers for Disease Control and Prevention has initiated an extramural cooperative agreement to advance the understanding of the Health of A/PI population. This paper presents an overview of the issues in health research for A/PI population and how these issues are being addressed through the cooperative agreement.

In conducting research in the A/PI and other populations, there are five inter-related problem areas. To gain the necessary information on the health of Asian and Pacific Islander population, it is critical to address these issues.

I. Lack **of ethnic A/PI specific data collection efforts and survey** instruments.

Most of the national, state and local health data do not have the racial breakdown necessary to gauge the health of A/PI population. In most cases, the **race/ethnicity** information is available for Blacks, Whites and Hispanics; A/PI population is often in the residual category with American Indian/Alaskan Natives. Even if information are available for the aggregate A/PI population, the diversity within the population renders much of the information inadequate for any one specific sub-population within Asian and Pacific Islanders.

11. Absence **of research and researchtoolsthat have been normalized for A/PI population.**

Despite the obvious lack of health information for A/PI population, there are very few research efforts. Moreover, the existing research efforts often disregard the cultural and behavioral differences among various Asian and Pacific Islander populations and other populations. Therefore, the validity of many of the research finding may be suspect.

III. Little **baseline data or research on A/PI population.**

Of the over 600 national health objectives for the Year 2000, only eight objectives are targeted specifically for Asian and Pacific Islanders. This absence of A/PI population in the document for national health priority is due to the fact that there are very iittle baseline data with which to establish and measure objectives. The impact of the lack of baseline data has resulted in considering health of A/PI population as a low priority.

IV. Lack of trust within A/PI community in research and data collection efforts.

Conducting research in any population needs the cooperation of that population in order to collect accurate and timely information. Due to a number of reasons from language problems to privacy issues, it is difficult to conduct research in the A/PI community. Moreover, many have not seen any benefit from research or data collection efforts conducted in their communities.

V. Methodological and cost issues in surveying A/PI population.

The small size of the A/PI population and geographic concentration presents major difficulties in undertaking of a nationwide survey. The current method of surveys are prohibitive in cost and alternative methodologies with lower cost are not available.

As a consequence of above problems, the health of A/PI population is not very well understood and at the same time, is not receiving much attention. Often, the health problems are not addressed until too late. The next section examine the possible consequences of lack of health information and offer alternatives.

Measles Morbidity and Mortality

In 1990, half of the measles contracted .in the state of Minnesota were children of Hmong immigrants. This was an astounding number since there are only 15,000 Hmong in the entire state. With state population of 4,375,000, Hmong population in Minnesota make up 0.34% of the state's population. By June, three Hmong children were succumbed by the disease.

From 1989 to 1990, there were 58 deaths in California from measles. Of that 17 deaths (about 30%) occurred in the Asian and Pacific Islander community. The Hmong, Samoan, Lao and Cambodian population combined make up less than seven tenth of one percent of the California population.

This type of ethnic specific information in invaluable in determining the population at risk and formulate effective method of intervention. If these deaths occurred in a state without sub-A/PI specific data, this type of information may not have been available.

Immunization

Health U.S., an annual publication detailing health status of the nation, has only two categories, "whites" and "all others" in a section on immunization status. This type of data do not provide any useful information for A/PI population. However, it is a typical presentation of health data and is a grave problem when trying to determine the health status of minority population, especially A/PI population.

On the other hand, a data collection effort that disaggregate A/PI population yield much more useful information. The 1992 Kindergarten Retrospective Survey, conducted by the Immunization Unit of California Department of Health Services, provide information not available nationally. This survey disaggregated A/PI population into two groups, Asians and Southeast Asians. As Table-I indicates, the survey found that although Asian children had second highest level of immunization, Southeast Asians had the lowest level of immunization. In fact, the immunization rate for Southeast Asians were 15% lower than the second lowest group, Hispanic children.

Table I - Percent of California'5 Children with Adequate Immunization by Age 2

Race/Ethnicity	% Immunized
All Races	48.2%
White	58.0%
Black	44.2%
Asian	55.6%
Southeast Asian	21.6%
Other	59.5%
Hispanic	36.9%

Source: 1992 Kindergarten Retrospective Survey, CA DHS.

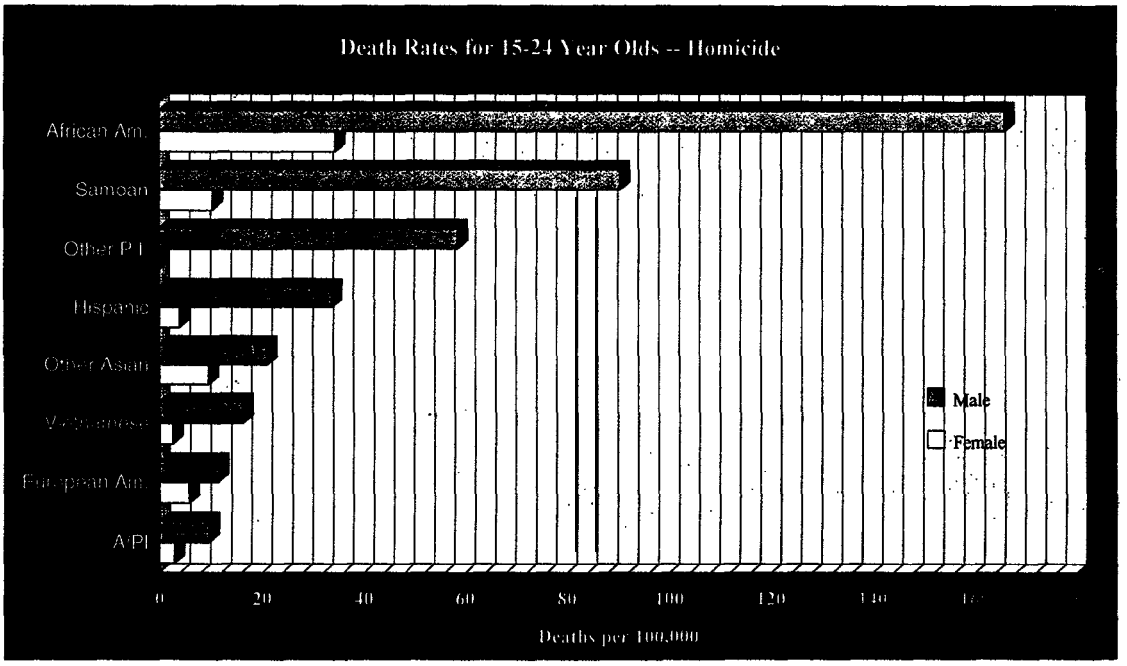
Vital Statistics

Currently, there are very little ethnic specific mortality analysis available for Asian and Pacific Islanders. Although the mortality rate for the A/PI population in general and some population such as Chinese and Japanese are available, data on other population are very rare. Analysis of California vital statistics performed by Asian American Health Forum provides information on other A/PI populations. Its findings show high rate of mortality for specific sub-groups of Asian and Pacific Islanders.

Figure-I illustrates death per 100,000 population due to homicide and legal intervention in California among 15 to 24 year old population using the 1986 to 1988 data. It indicates that Samoan5 and Other Pacific Islanders have highest mortality rate after African Americans.

Figure-I

California Vital Statistics. 1986-88



Survey and Research Program, Asian American Health Forum, Inc. 1992

Alternative Methods

Due to small size of samples or number of incidents, the rates or mortality or morbidity are often statistically insignificant. Since the validity of the findings for data analysis alone is in question, other means of supporting the data need to be utilized. By using supporting studies, community input and alternative technologies such as mapping, one can determine whether these findings support the findings from quantitative analyses.

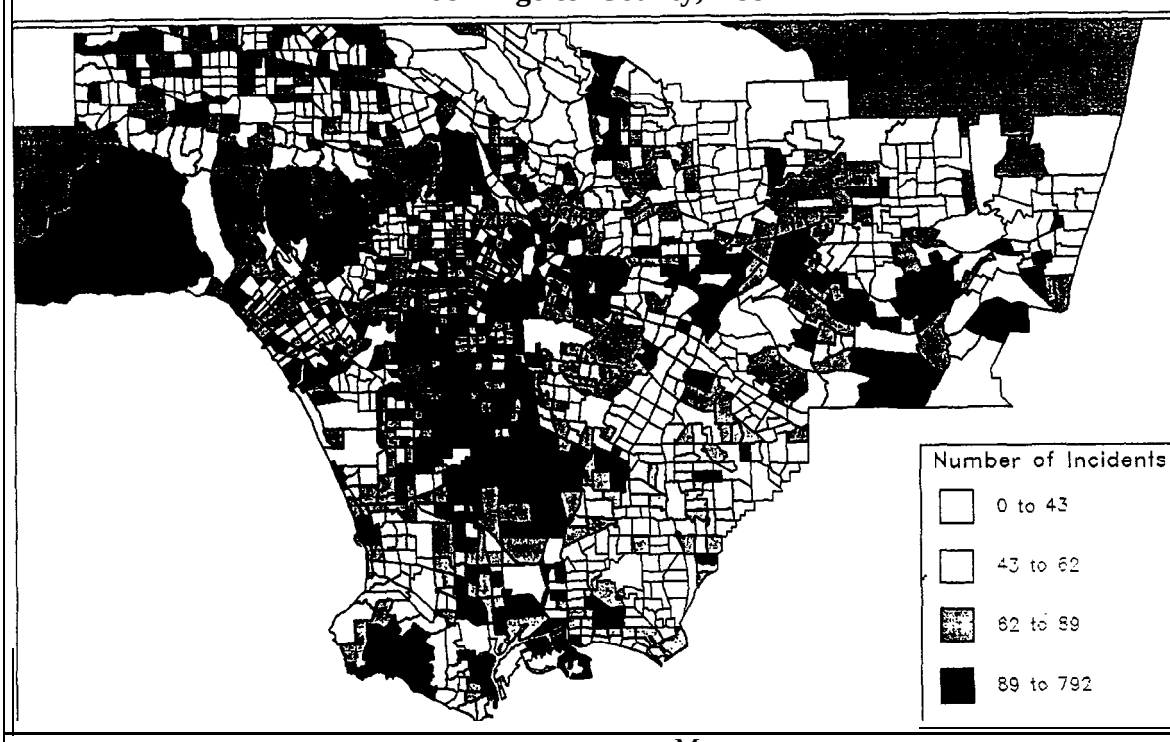
To illustrate this point, a simple map will be used to support high homicide rate for Samoans. The first map (Figure-II) illustrate the number of homicide in each census tract in Los Angeles County. The next map (Figure III) shows the concentration of Samoan population in Los Angeles County by census tract. There is an overlap between census tracts with high concentration of homicide and census tract with high concentration of Samoan population. Since Samoans reside in area with high number of homicides, the findings from the mortality analysis that Samoans have higher rate of deaths due to homicide and legal intervention can be supported by this method.

More traditional approach to mapping is data from existing literature. Many studies indicate that the Pacific Islanders, especially Samoans, are at high risk for diseases of the heart. This supports the findings from the mortality analysis that show Samoans have one of the highest rate of deaths due to diseases of the heart. California mortality analyses show death rates due to diseases of the heart for Samoans are second highest among 35-44 year olds, third highest among 45-54 year olds, and second highest among 65-74 year olds.

Figure-II

Number of Homicides by Census Tract

Los Angeles County, 1991

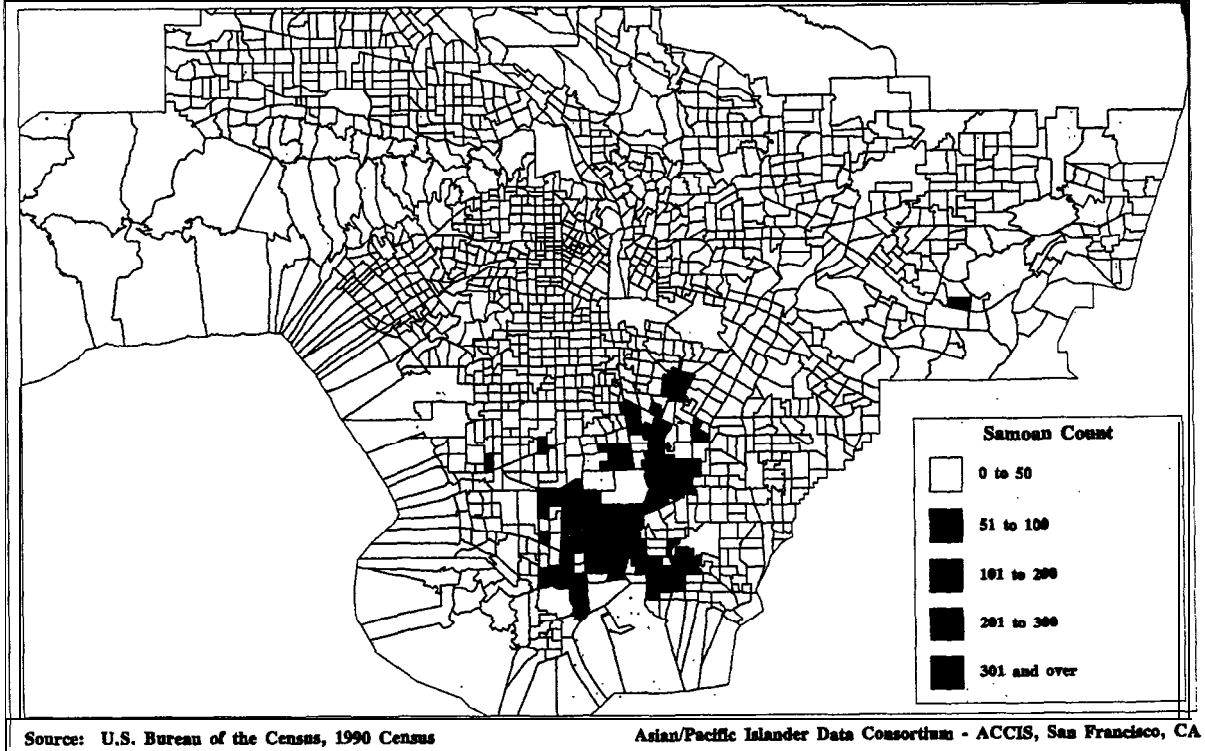


Source: U.S. Bureau of the Census, 1990 Decennial Census
Asian and Pacific Islander Data Consortium – ACCIS: San Francisco, CA

Map made with ATLAS GIS from Strategic Mapping, Inc.

Figure-III

Samoan Population in Los Angeles County by Census Tract, 1990



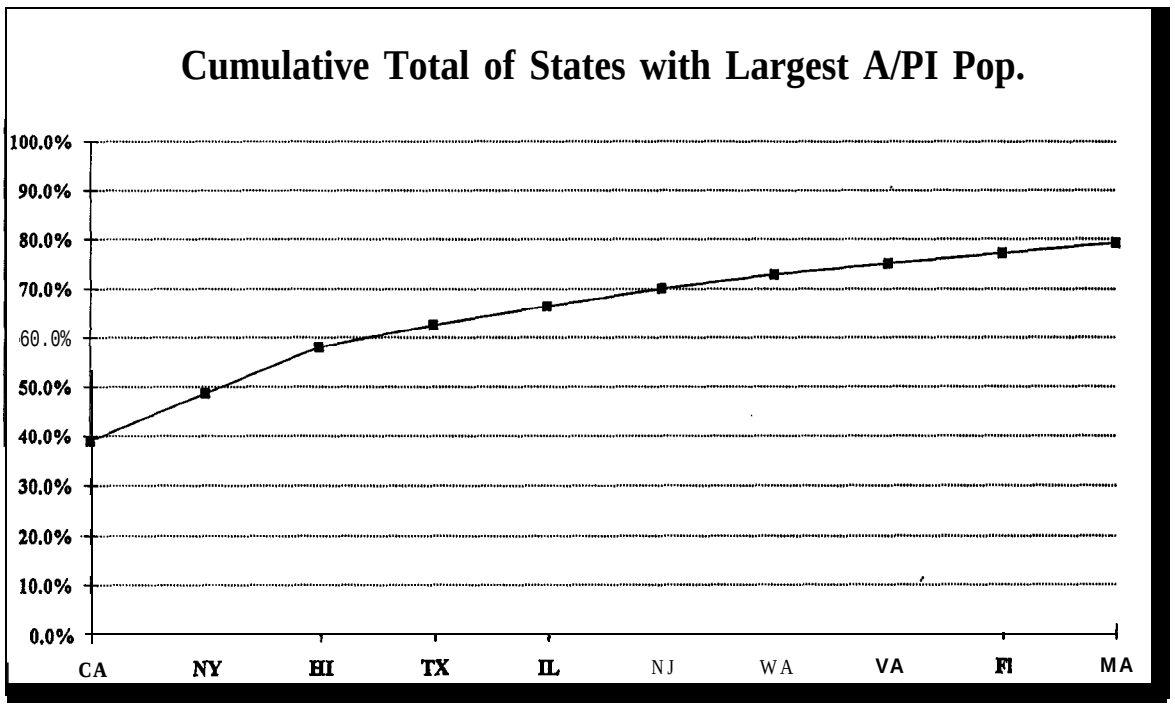
Source: U.S. Bureau of the Census, 1990 Census

Asian/Pacific Islander Data Consortium - ACCIS, San Francisco, CA

As demonstrated, without relevant and dieaggregated health information, the health care delivery system cannot deal with the problems faced by this rapidly growing population. The disaggregation of A/PI population given it's diversity and heterogeneity for the entire nation seems to be a daunting task. However, due to the high concentration of A/PI population in small number of states, large level of data can be obtained from concentrated efforts.

As Figure-IV indicates, two-thirds of the A/PI population reside in five states and about 80% of the A/PI reside in ten states. Asian and Pacific Islanders also reside in diverse *regions in* the United States. Given this diversity of geography along with concentration of the population in only a small number of states, a focused effort in several states can yield large amount of useful and needed data.

Figure-IV



Source: 1990 U.S. Census

The Extramural *Research*

To supplement the efforts of the Asian American Health Forum to further advance the knowledge of A/PI's health in other *areas,* AAHF has awarded four grants to community- and university-based researchers. These grants are designed to further address five problems mentioned above.

I. Sources of Bias in Health Statistics among Elderly Chinese and Asian Indian Americans in Flushing, New York by Zibin Guo. The objective of this research is to:

1. Obtain full inventory of types of health resources utilized,

2. Understand interaction between the elderly and the modern health care system,

3. Develop explanatory models of illness and health maintenance,

4. Determine difference in answers for culturally worded questions versus standard questions.

This study alleviates the need for dfsaggregated data, normalized ethnic research tool, and 'ethnic sensitive research. The research included key informant interviews and focus groups to eater to the characteristic and needs of the study population. Currently, survey has been conducted and data are being analyzed.

II. *Preventative Health Practices of American Samoans: Socio-Cultural Values Influencing Patterns of Child Immunization* by Gregory Loos. This study is designed to:

1. Determine socio-cultural values that influence health-related behavior.

2. Compare the socio-cultural values of Samoans residing in American Samoa and Hawai'i.

3. Develop a map of socio-cultural values influencing preventive health practices.

This study addresses the need for ethnic specific data, standardized research tool. The survey instruments were developed after large

number of interviews and surveys are being conducted in both American Samoa and Hawai'i.

III. *Assessing Health Status and health Risk Behaviors in the Korean* American community *in Los Angeles County* by Toshiaki Sasao. The objective of the study is to:

1. Determine the feasibility of using surname-based telephone survey versus random digit dialing survey.

2. Conduct community-based, community oriented survey.

This study is designed to meet the need for disaggregated data, standardized research tool, establishing baseline, community-oriented research and testing different methodologies. Extensive interviews and focus groups were conducted, as well as, extensive media campaign with public service announcements. Eighty percent of those surveyed were aware of the research. The completion rate was 88.6% with refusal rate of less than seven percent.

IV. Maternal *Characteristics, Newborn Growth* Parameter *and Low Birth Weight in Asian and Pacific Islander population in* San *Francisco* by Don Wong. This study aims to:

1. Develop growth parameter and ethnically appropriate reference graph.

2. Determine factors that influence low birth weight and growth parameters.

The study will obtain disaggregated data, standardize research tool and establish baseline data. Data will be collected for next few months. Thus far, information was collected for 500 hundred Chinese and Southeast Asians.

Conclusion

In response to rapidly increasing Asian and Pacific Islander population, more data are necessary. Few national surveys and databases have been responsive to the change in population and has expanded their capacity to provide more useful information. The National Health Interview Survey has disaggregated race information for Asian Indian, Chinese, Filipino, Guamanian, Hawaiian, Japanese, Korean, Samoan and Vietnamese. Vital records will also be disaggregated for states other than California and Hawai'i. In the future, it is hope that disaggregated data will be available for National Hospital Discharge Surveys, National Health and Nutrition Examination Survey, and ' National Medical Expenditure Survey to provide information not available in National Health Interview Survey or vital statistics, such as morbidity and health care financing information.

The health research in the Asian and Pacific Islander population is just beginning and has a long way to go. The theme of this conference is "Refining the Measures." Without overcoming the problems in health research for Asian and Pacific Islanders and other minority groups, this goal cannot be achieved.

Reference

Asian American Health Forum, Inc. *Dispelling the Myth of a Health Minority.* Asian American Policy Review. Boston, Vol. III. 1993.

Asian American Health Forum, Inc. *Teen and Young Adult Mortality in California: Analysis of State of California Mortality Data 1986-1988. San* Francisco, 1992.

Asian & Pacific Islander Center for Census Information and Services. Our Ten Years of Growth: A Demographic Analysis on Asian and Pacific Islander Americans. San Francisco, 1992.

California Department of Health Services. 1992 Kinderaarten Retrospective Survey. Berkeley, 1992.

Isaacs, Mareasa R. *The Complexity of Minority Health: An Overview of Key Issues.* Presented at the Setting A Research Agenda: Challenges for Minority Health Grants Program sponsored by the National Center for Health Statistics. December 4-6, 1991; Besthesda, MD.

Measles Mystery Newsweek June 4, 1990. p4

National Center for Health Statistics, DHHS, PHS. Health United States and Prevention Profile, 1991. p. 186.

Ponce, Ninez. Advancing the Understandina of the Health of Asian and Pacific Islander in the United States: the Pressina Need for Subpopulation Data Development. Asian American Health Forum: San Francisco. 1991.

SOCIAL COGNITION AND RESPONSES TO HEALTH SURVEY QUESTIONS AMONG MINORITY POPULATIONS: PRELIMINARY EVIDENCE

Timothy Johnson, University of Illinois
Diane O'Rourke, Noel Chavez, Loretta Lacey, Seymour Sudman,
Richard Warnecke, John Horm

INTRODUCTION

Despite earlier assertions to the contrary (Freeman, 1969; Welch, Comer, & Steinman, 1973), most researchers now recognize the pitfalls of uncritically applying "standard" survey research technology to studies of minority populations (Aday, Chiu, & Anderson, 1980; Flaherty, 1987; Milburn, Gary, Booth, & Brown, 1991; Montero, 1977; Pareek & Rao, 1980). In particular, it is now understood that culturally mediated differences in cognition and interpretation are likely responsible for many of the systematic differentials that have been observed in cross-cultural surveys (Marin & Marin, 1989; Rogler, 1989). Acknowledged for years as a problem in achievement testing (Flaugher, 1978), the uncritical adoption of questionnaire measures developed by and for representatives of the dominant White middle class to surveys of minority group members may misrepresent the thoughts, feelings, and behaviors of those individuals (Berry, 1969). This phenomenon, referred to as "category fallacy" (Kleinman, 1977), is a problem in much of the health-related survey data collected in this country and is usually a consequence of the assumption that certain concepts or constructs are universally applicable to various cultures and contexts.

This problem can be alternately viewed through the distinction in cross-cultural psychology between "etic" and "emic" constructs (Triandis, '1972). Etic constructs are those assumed to be universal in that they are equally understood across multiple cultural groups. Emic constructs are culture specific, in that they have important meaning within a single cultural group and are either interpreted differently or completely ignored by persons from other cultures. When emic constructs are treated as if they are etic, they are described as *psuedoetic,* and result in the "category fallacy." The widespread use of pseudoetic constructs in survey research is in no small part responsible for the criticisms this methodology has received in the past (Cicourel, 1964). While it is commonly accepted today that many survey questions are likely emic, these items nonetheless continue to be, used in a pseudoetic manner.

Triandis (1972) has proposed a process that may be used in the development of *survey* instruments that emphasizes identification of survey items that are culturally specific (emic) and universally relevant (etic). This approach involves examination of survey questions believed to be etic, as well as other items thought to be emic, in order to assess and understand which aspects of the culture in question are unique and which are universal. Research by Triandis and Marin (1983) has tested this model, concluding that this "emic + etic" methodology is superior to the pseudoetic approach commonly used as part of traditional survey methodology.

The research to be presented in this paper applied cognitive survey methodology to the identification and evaluation of etic and pseudoetic components of health-related survey questions commonly asked in national surveys. Identification of pseudoetic questions and an understanding of why they are evaluated differently across cultures is a crucial first step in addressing the issue of cultural sensitivity and comparability in the collection of health survey data.

Cognitive Aspects of Survey Responses

It is now generally understood that there are four steps in the survey response process. (Tourangeau & Rasinski, 1988). The first step isquestion interpretion. It is likely that cultural perceptions moderate the meanings and interpretations respondents assign to many of even the most generic health survey questions. Questions dealing with various disease names, medical procedures, and health behaviors, for instance, may well be subject to both linguistic and interpretative differences.

The next step is retrieving either the answer to the question or relevant information that can be used to construct the answer. There has been substantial work reported recently about the effects of different question wording that act as cues to improve retrieval (Schwarz & Sudman, forthcoming). While some activities are typically recalled as individual events, others are recalled as or through schema or generalizations. The accuracy of information obtained from schema depends on the regularity of the behavior (Menon, forthcoming). Researchers may influence response accuracy by asking questions that deliberately evoke or discourage the use of schema. However, because most of the research in this area has been conducted with White populations, it is not clear whether results can be generalized. There is currently no evidence that information retrieval strategies are used in consistent manners cross-culturally.

The third step in the response process, forming judgments based on data retrieved from memory, is an important

aspect of attitude formation and the reporting of behavioral frequencies. Some important research has been conducted regarding cultural variations in norms for responding to survey questions. For example, it appears that there are cross-cultural differences in the processes used to map personal judgments onto survey rating scales. Available information suggests that both African-American and Hispanic survey respondents are more likely to qualify their answers (Bachman & O'Malley, 1984; Hui & Triandis, 1989). These findings have been interpreted as being a consequence of variations in perceptions of what constitutes honesty, with some cultures believing that extreme responses are more honest (Marin & Marin, 1989). Other research suggests that variations in probabilistic thinking across cultures may also influence response decisions (Wright et al., 1978).

Finally, respondents edit responses if they see certain answers as being more socially desirable than others (Bradburn, Sudman, & Associates, 1979). However, because personal definitions of social desirability are likely to be culturally mediated, the extent of bias due to response editing may be quite variable, depending *on* the topic and cultural tolerance for the behavior (Groves, 1989; Marin, Triandis, Betancourt, & Kashima, 1983). Several cultural factors, including cultural scripts for interaction (Triandis, Marin, Lisansky, & Betancourt, 1984) and ethnic affirmation (Yang & Bond, 1980), may also influence the editing of survey responses differently across cultural groups. A substantial literature also exists suggesting that respondents may edit answers on a selective basis as a consequence of social or cultural distance between respondent and interviewer (Campbell, 1981; Hatchett & Schuman, 1975-76; Weeks & Moore, 1981).

Cognitive Methods and Survey Research

It has only been in the last decade that survey researchers have begun to use the cognitive methods of psychologists in an effort to improve the design of survey questionnaires (Jobe & Mingay, 1991). Cognitive methods are used to identify sources of response error that are often missed in field pretests and to gather information about the possible reasons for that error (see, for example, Lessler, Tourangeau, & Salter, 1989). Field testing alone is not adequate because, while respondents may answer a question with seemingly little difficulty, it often remains uncertain whether they and the researcher share a common understanding of the question or whether they are in fact answering a *different* question, unless clarification is specifically requested. In order to identify problems such as lack of comprehension and differing interpretation and to better understand issues such as information retrieval and judgment formation, respondents must be queried with "special probes" (Oksenberg, Cannell, & Kalton, 1991). These probes are designed to identify the underlying thought processes used by respondents when answering survey questions. Procedures that incorporate this approach into the traditional interview process have been referred to using a variety of terms. For simplicity, we will refer to this process as "think-alouds."

METHODOLOGY

Research Design

The work currently being undertaken by our research team is concerned with evaluating the cognitive processes underlying responses to health survey questions among three minority groups found in the Chicago metropolitan area: African-Americans, Mexican-Americans, and Puerto Ricans. The two largest Hispanic communities in Chicago were separately -selected in recognition that although there is a core culture common to all Hispanics, there is also considerable heterogeneity, due in part to differences in migration experiences and acculturation (Flaherty, 1987; Marin & Marin, 1989). A stratum of White, non-Hispanic respondents is also being examined for purposes of comparison. A total of 400 adults aged 18 through 50 (100 from each of these four groups) will be asked to participate in think-alouds as part of this study. In recognition that some psychological processes may also vary across cultural groups by gender, age, and educational achievement, respondents within each of the four ethnic groups will be further stratified by these characteristics so their influence can also be investigated. Interviewing for this study is currently in progress and is expected to be completed in early 1994.

This paper will present preliminary data that were obtained from 30 pretest think-alouds: 8 each with African-American, Mexican-American and White respondents, .and 6 with Puerto Rican respondents. All interviewing was conducted by our research team during late April and May 1993. Interviews averaged approximately one hour in length.

Instrument Development

A three-step process was used to develop a survey instrument to address these questions. Initially, a large pool of health questions previously used in national health surveys was selected for consideration. A committee of six investigators next reviewed these items and selected approximately 40 questions for inclusion in the pretest instrument. Items were selected with consideration for producing variation in terms of substantive topics, question types (e.g.,

opinion/attitude, behavior, knowledge), and question formats, (e.g., open vs. closed-ended, Likert scale vs. yes/no response options, numerical recall). During the third stage, sets of specific probes were developed for use with each of the health questions selected in stage one. These probes each represented one of the four phases of the question response process described above (interpretation, information retrieval, judgment formation, or editing; examples are shown in Figure 1). On average, two or three probes were developed for each health question, with the number per question ranging from one to five.

PRELIMINARY FINDINGS

In reviewing the information presented in this paper, it is important to recognize that it is based upon the pretest only and should therefore be considered as illustrative of the method, rather than as a set of findings. Although some of the results are quantitative, formal statistical tests of significance were not conducted. The results are intended to highlight each of the four stages of the survey response process.

Question Interpretation

Our pretest data suggest evidence of both etic and emic concepts among the health questions examined. For example, questions about sexual orientation may be etic, as there was considerable agreement across ethnic groups regarding definitions of heterosexual, homosexual, and bisexual. Only a small number of the definitions reported were judged to be incorrect (less than 8%). Most incorrect definitions were provided by Puerto Rican respondents. This suggests that while there is considerable agreement across all cultures regarding these concepts, what confusion exists is concentrated within a single group, leaving the issue unresolved.

A question commonly asked as part of the CES-D Scale (Radloff, 1977) appears to be more emic: "During the past week, how often have you felt that you could not shake off the blues, even with help from family or friends?" While large proportions of African-American and White respondents indicated a belief that most people would have *little* difficulty understanding this question, persons of Hispanic origin (both Mexican-American and Puerto Rican) felt that people would have difficulty with this concept.

White respondents as a group differed from the three minority groups examined. Whites were less likely to believe people would have difficulty answering the question, "Please tell me which of these you think causes more cancer in the United States: personal behaviors or factors you have little control over?" Whites also differed from the other ethnic groups in the definitions they applied and the terminology they used when answering questions. When probed regarding this question: "During the last year, did anyone in'the family go to a dentist?" White respondents included in their definition of "family" people who did not live with them. In contrast, majorities within each of the three minority groups included only household members in their definition of "family." However, whether this is an issue of interpretation or reality is yet to be resolved, as these responses may be a consequence of the greater likelihood for extended families to be living together within minority households. This is an example of a question that was developed by White researchers with White respondents in mind: It can be answered by all respondents, even if they do not share the same meaning. Additional questions have been added to the main study questionnaire to further examine this issue.

When shown a card and asked which term they would apply to a common disease (diabetes), almost all White respondents endorsed the term "diabetes." Minority respondents, in contrast, exhibited considerable variability: Four different terms for this disease (diabetes, sugar diabetes, sugar, and high sugar) were selected by the eight Mexican-American respondents, and eight African-American respondents endorsed three of these terms (diabetes, sugar diabetes, and sugar).

Finally, African-American respondents were unique in reporting that they felt people would have difficulty understanding the following question: "How much control do you think you have over your future health?" African-Americans were also notable for indicating that they had not previously developed an opinion about this question. These findings are important, as African-Americans' responses to other questions suggest that they alone, of the four groups examined, do not feel they have much control over the likelihood that they will develop diseases such as cancer in the future. These responses suggest that the concept of control over future health is one that African-Americans may be less likely to consider when responding to health questions.

Information Retrieval

Preliminary information suggests similar recall strategies may be employed across the four ethnic groups. All four groups were consistent in reporting use of counting, or episodic, strategies for remembering the numbers of sexual partners they had in the previous five years. There was a clear tendency, though, for all persons reporting greater numbers of partners to rely on schematic recall. Respondents within each ethnic group also reported using similar methods for estimating other activities, such as the number of times certain foods are consumed and frequency of social inter-

actions. In each instance, there was a general preference for averaging experiences over a number of weeks, rather than thinking specifically about the most recent week, in developing an answer.

On the other hand, while majorities of African-American, Mexican-American, and White respondents reported relying on a counting strategy to retrieve information regarding the number of medical visits they had in the last year, most Puerto Ricans interviewed indicated that they had used some type of estimation strategy to retrieve this information, suggesting that there may also be some cultural differences in the methods of recall used for this item.

Judgment Formation

Several probes were used to examine consistency in response mapping across groups. One probe that we used was to ask respondents to provide numeric values for ordinal response categories. Respondents were asked to report how many days *were* represented by several of the ordinal response options commonly used with the CES-D Scale. With only a single exception, the estimates provided by each group of respondents were monotonic. On average, the response option "most of the time" was reported to represent 4.9 days of the last week; "occasionally" represented 3.5 days; "a little of the time" represented 2.3 days; and "only rarely" represented 1.3 days. These preliminary findings suggest respondents are interpreting and using these various response categories in a similar manner for these items. Respondents performed a similar exercise using the perceived health rating question, "Would you say your health is excellent, very good, good, fair, or poor?" Here, although there was considerable agreement regarding numerical ratings assigned to the scale endpoints ("excellent" and "poor"), some problems with the middle categories were noted. Most importantly, members of all three minority groups tended to assign higher numerical rating scores to the "fair" rating than to the "good" rating. Other studies have identified the term "fair" as representing a more positive rating among respondents selecting it than its placement in the set of response options suggests researchers may have intended it to have (Hougland, Johnson, & Wolf, forthcoming).

Another probe focused on attitude crystallization, or the likelihood that various groups had previously thought about or formed opinions regarding health-related topics. African-Americans were more likely to report having thought about several questions relating to various ways that one might contract the AIDS virus. In contrast, African-Americans were the only group of respondents to report *not* already having an opinion about personal "control over future health" prior to being asked the survey question. This is likely related to the finding reported above that African-Americans felt this question would be one that was difficult to understand, as persons not familiar with a specific concept would not be expected to have an opinion regarding it.

Probes were also developed to assess judgment certainty. No differences were observed in the degree of confidence that respondents expressed in their reports of the number of different days they had consumed alcohol during the last month. Whites, however, were less certain than the three minority groups regarding the behavioral estimate of their lifetime use of marijuana. Puerto Rican respondents expressed less certainty regarding their answers to a knowledge-based question regarding the possibility of contracting AIDS by sharing eating utensils with an HIV-infected person.

Response Editing

The pretest also yielded considerable information regarding possible respondent editing of survey answers. There was considerable agreement across ethnic groups that questions regarding anal sex practices would be perceived as embarrassing to most people. There was also a consensus across these groups that most people would underreport their lifetime use of marijuana. Some cultural differences were noted, however, as persons of Hispanic origin were much more likely to indicate a belief that questions regarding the seeking of help for personal or emotional problems would be embarrassing.

There was also evidence of selective response editing depending on the match between respondent and interviewer ethnicity. A majority of African-Americans felt that respondents would *not* be comfortable talking to an interviewer from another ethnic group regarding alcohol consumption (e.g., number of drinks consumed per day), whereas White and Hispanic respondents felt this would not be a problem. When discussing marijuana use, however, African-American *and* Hispanic (both Mexican-American and Puerto Rican) respondents felt that revealing this information to interviewers of a different ethnicity would produce discomfort. Virtually all White respondents continued to report that they would be comfortable in this situation. Similar findings were observed when discussing sexual practices (e.g., anal sex) and sexual orientation.

DISCUSSION

These preliminary findings have identified several examples of both etic and emic concepts among the health survey questions examined. Because the questions included in this study were mostly drawn from prior national health surveys, it would appear that those concepts identified as emic have been previously used

in a pseudoetic manner. Clearly, additional research will be necessary to assess the scope of this problem.

An important limitation of the findings reported here, of course, is the fact that they are based on fewer than 10 think-alouds with representatives of each ethnic group being examined. By the end of this project, we plan to have completed cognitive interviews with approximately 400 individuals, permitting an analysis of cultural differences that will enable us to control for the potentially confounding effects of gender, age, and educational attainment. We will also be able to assess and control for the effects of acculturation, migration status, and social desirability on responses.

This research demonstrates the considerable potential of applying the methods of cognitive psychology to the identification and improved understanding of cultural differences in the perception of and response to traditional health survey questions. This understanding, we believe, is a critical step in the design of health surveys that are recognized as valid across cultures.

REFERENCES

Aday, L. A., Chiu, G. Y., & Anderson, R. (1980). Methodological issues in health care surveys of the Spanish heritage population. *American Journal of Public Health, 70, 367-374.*

Bachman, J. G., & O'Malley, P. M. (1984). Yea-saying, nay-saying, and going to extremes: Black-white differences in response styles. *Public Opinion Quarterly, 48,* 491-509.

Berry, J. W. (1969). On cross-cultural comparability. *International Journal of Psychology, 4, 207-229.*

Bradburn, N. M., Sudman, S., & Associates (1979). *Improving interview method and questionnaire design: Response effects to threatening questions in survey research.* San Francisco: Jossey-Bass.

Campbell, B. A. (1981). Race-of-interviewer effects among Southern adolescents. *Public Opinion Quarterly, 45, 231-244.*

Cicourel, A. V. (1964). *Method and measurement in sociology.* Glencoe, IL: Free Press.

Flaherty, J. A. (1987). Appropriate and inappropriate research methodologies for Hispanic mental health. In M. Gaviria & J. D. Arana (Eds.), *Health and behavior: Research agenda for Hispanics.* (The Simon Bolivar Research Monograph Series No. 1.) Chicago: The University of Illinois.

Flaugher, R. L. (1978). The many definitions of test bias. *American Psychologist, 3, 671-679.*

Freeman, D. M. (1969). A note on interviewing Mexican-Americans. *Social Science Quarterly, 49,* 909-918.

Groves, R. M. (1989). A total survey error approach to AIDS-related survey research. In F. J. Fowler (Ed.), *Conference proceedings on health survey research methods* (DHHS Publication No. PHS 89-3447, pp. 265-270). Rockville, MD: Public Health Service.

Hatchett, S., & Schuman, H. (1975-76). White respondents and race-of-interviewer effects. *Public Opinion Quarterly, 39, 523-528.*

Hougland, J. G., Johnson, T. P., & Wolf, J. G. (forthcoming). A fairly common abmiguity: Comparing rating and approval measures in public opinion polling. *Sociological* Focus.

Hui, C. H., & Triandis, H. C. (1989). Effects of culture and response format on extreme response style. *Journal of Cross-Cultural Psychology, 20, 296-309.*

Jobe, J. B., & Mingay, D. B. (1991). Cognition and survey measurement: History and overview. *Applied Cognitive Psychology, 5,* 175-192.

Kleinman, A. (1977). Depression, somatization, and the "new cross-cultural psychiatry." *Social Science and Medicine, 11, 3-10.*

Lessler, J., Tourangeau, R., & Salter, W. (1989). *Questionnaire design in the cognitive research laboratory.* National Center for Health Statistics, *Vital and Health Statistics,* 6(1) (DHHS Publication No. PHS 89-1501). Rockville, MD: Public Health Service.

Marin, G., & Marin, B. V. (1989). *Research with Hispanic populations.* Newbury Park, CA: Sage Publications.

Marin, G., Triandis, H. C., Betancourt, H., & Kashima, Y. (1983). Ethnic affirmation versus social desirability: Explaining discrepancies in bilinguals' responses to a questionnaire. *Journal of Cross-Cultural Psychology, 14, 173-18.6.*

Menon, G. (forthcoming). Judgments of behavioral frequencies: Memory search and retrieval strategies. In N. Schwarz & S. Sudman (Eds.), *Autobiographical memory and the validity of retrospective reports* . New York: Springer-Verlag.

Milburn, N. G., Gary, L. E., Booth, J. A., & Brown, D. R. (1991). Conducting epidemiologic research in a minority community: Methodological considerations. *Journal of Community Psychology, 19, 3-12.*

Montero, D. (1977). Research among racial and cultural minorities: An overview. *Journal of Social Issues,* 33, l-10.

Oksenberg, L., Cannell, C., & Kalton, G. (1991). New strategies for pretesting survey questions. *Journal of Official Statistics, 7, 349-365.*

Pareek, V., & Rao, T. V. (1980). Cross-cultural surveys and interviewing. In H. Triandis & J. Berry (Eds.), *Handbook of cross-cultural psychology*

(vol. 2, pp. 127-129). Boston: Allyn and Bacon.

Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 3, 385-401.*

Rogler, L. H. (1989). The meaning of culturally sensitive research in mental health. *American Journal of Psychiatry, 146, 296-303.*

Schwarz, N., & Sudman, S. (Eds.) (forthcoming). *Autobiographical memory and the validity of retrospective reports. New* York: Springer-Verlag.

Tourangeau, R., & Rasinski, K. A. (1988). Cognitive processes underlying context effects in attitude measurement. *Psychological Bulletin, 103,* 299-314.

Triandis, H. C. (1972). *The analysis of subjective culture.* New York: Wiley-Interscience.

Triandis, H. C., & Marin, G. (1983). Etic plus emic versus pseudoetic: A test of a basic assumption of contemporary cross-cultural psychology. *Journal of Cross-Cultural Psychology, 14, 489-500.*

Triandis, H. C., Marin, G., Lisansky, J., & Betancourt, H. (1984). Simpatia as a cultural script for Hispanics. *Journal of Personality and Social Psychology, 47, 1363-1375.*

Weeks, M. F., & Moore, R. P. (1981). Ethnicity-of-interviewer effects on ethnic respondents. *Public Opinion Quarterly, 45, 245-249.*

Welch, S., Comer, J., & Steinman, M. (1973). Interviewing in a Mexican American community: An investigation of some potential sources of response bias. *Public Opinion Quarterly, 37, 115-126.*

Wright, G. N., Phillips, L. D., Whalley, P. C., Choo, G. T., Ng, K.-O., Tan, I., & Wisudha, A. (1978). Cultural differences in probabilistic thinking. *Journal of Cross-Cultural Psychology, 9, 285-299.*

Yang, K. S., & Bond, M. H. (1980). Ethnic affirmation by Chinese bilinguals. *Journal of Cross-Cultural Psychology, 11, 411-425.*

FIGURE 1
Examples of Think-Aloud Probes

Question Interpretation

A. QUESTION INTERPRETATION - "Please tell me in your own words what you think this question is asking about."

B. RESPONSE INTERPRETATION - "What does [SELECTED RESPONSE] mean to you?"

C. INTERPRETATIVE DIFFICULTY - "Do you feel this is a question that people would or would not have difficulty understanding?"

Information Retrieval

D. RECALL STRATEGY I - "How did you remember the number of times you did this?"

E. RECALL STRATEGY II - "In recalling this information, did you think *specifically* about the most recent week, did you *average* together experiences you've had over a number of weeks, or did you think about it in some other way?"

F. RECALL ORDER - "Did you answer this question by first thinking about visits that happened most recently, about visits that happened at the beginning of the year, or by just thinking about visits in whatever order you could remember them?"

Judgment Formation

G. RESPONSE MAPPING - "On a scale of 0 to 10, with 0 being the worst possible health and 10 being the best possible health, what number best represents (SELECTED RESPONSE) to you?"

H. JUDGMENT CERTAINTY - "How much confidence do you have in the accuracy of your answer to that question?"

I. ATTITUDE CRYSTALLIZATION - "Is this something you had already made up your mind about before we asked this question?"

Response Editing

J. EDIT FOR EMBARRASSMENT - "Do you feel this is a question that people would or would not be embarrassed to answer?"

K. EDIT FOR SOCIAL DESIRABILITY - "Do you feel that people might purposely say they used marijuana more times, fewer times, or would they try to answer accurately?"

L. INTERVIEWER ETHNICITY EFFECT - "Do you feel that this is a question that African-American respondents would be comfortable or uncomfortable talking about with a non-African-American survey interviewer?"

HEALTH STATUS AND ACCESS TO CARE AMONG RURAL MINORITIES: WHAT DO WE KNOW?

Keith Mueller
Nebraska Center for Rural Health Research
Suzanne Ortega
Ahuva Askenazi
Kashinath Patil
Keith Parker

Minority populations are routinely considered separately when health: statistics are reported. There are many reasons for doing so, among the more obvious being that the prevalence of adverse health conditions varies between races and ethnic groups for reasons related to the characteristics of those populations (e.g., prevalence of sickle cell anemia) and the conditions in which they typically live (e.g., lead paint poisoning among Black children because of poor housing conditions). There are also differences in access to medical are services across racial and ethnic groups; differences associated with economic conditions, availability of medical care providers in the area, and cultural barriers between providers and citizens. Unique problems of health status and access to services, then, justify separate attention to minority populations in formulating the nation's health objectives for the year 2000 (USDHHS 1990: 596-604). For the same reasons, the research community should pay particular attention to the problems of minority populations.

The problems faced by minority populations are compounded in rural areas for two reasons. First, health care resources are in short supply in rural areas, which places all rural residents at a comparative disadvantage for access to care. This is particularly a problem in more remote rural areas, the very places most likely to be populated by minorities such as migrant farm workers, and Native Americans. Second, in many regions of the U.S. minorities are small percentages of the persons living in rural areas. This fact results in any special needs they may have being ignored, both by policy makers focused on state-wide. or region-wide problems, and by researchers investigating rural problems by designing projects based on regions and aggregate populations.

This paper will provide an overview of the health care and access problems of minorities in rural areas in the U.S., as reported in the scholarly and applied literatures. Each of four principal minority groups will be considered--Native Americans, Afro-Americans, Hispanic Americans, and Asian Americans. Within each group special problems related to particular categories of diseases will be considered, as will special problems in accessing health care services. The location Of the research reported will be noted, demonstrating that there is less known about rural minorities in certain regions, such as the West North Central U.S. Census Region (Kansas, Missouri, Nebraska, Iowa, North Dakota, South Dakota, and Minnesota). Special consideration is given to the various methodologies used in previous research.

GENERAL BACKGROUND

Rural populations in the U.S. include the same minority groups found in the general population, e.g., African Americans, Native Americans, Black and Caucasian Hispanics, and Asian and Pacific Islanders. However, the concentrations of those groups will vary depending on the region of the country. For example, African Americans are more likely to be found in the rural South than other regions, and Hispanics in the rural Southwest and migrant farmer stream. Studies of the interaction of minorities with the health care delivery system can be challenging in the regions where particular minority groups may not represent a large percent of the total population, but where they nevertheless face serious problems related to health conditions, financial access to medical care, and geographic access to services.

In one region, the West North Central Census Region, the most prevalent rural (defined as non-metropolitan statistical area) minorities are Native American, Hispanic, and Asian. In Missouri and Iowa, there are substantial numbers of African Americans in rural counties. All of these minority residents of rural areas tend to be concentrated in certain counties, where they may represent as much as 15 percent of the total population. Table 1 shows the numbers of each of four minority groups in the seven West North Central states, and the percent of the total population they represent.

Table 1
Minority Populations

State	Total Rural	White	Black	AmInd	Asian	Hisp
MO	1,737,782	1,656,027	56,019	6,990	6,712	12,034
%		95.8	3.2	.4	.4	.7
IA	1,561,583	1,527,577	8,576	2,786	9,539	13,105
		98.3	.55	.2	.6	.8
KS	1,162,828	1,068,932	30,325	9,897	10,190	43,484
		93.4	2.65	.9	2.2	3.8
MN	1,422,295	1,379,355	2,480	20,907	6,936	12,617
		97.5	.2	1.5	.5	.9
NE	822,317	792,861	1,707	7,982	2,607	17,160
		97.6	.2	.98	.3	2.1
ND	517,460	483,688	2,770	24,237	2,748	4,017
		93.9	.5	4.7	.5	.8
SD	440,859	389,001	1,453	46,371	970	3,064
		88.7	.3	10.6	.2	.7
Total	7,665,924	7,297,441	103,330	119,970	39,702	105,481
		95.9	1.4	1.6	.5	1.4

Source: U. S. Census, 1990

A review of the data available through the 1988 National Health Interview Survey shows the most frequently mentioned health problems for rural minorities in the West North Central Census Region, as reported in Table 2. As can be seen in those data, the types of conditions of most concern vary across the different groups. These different prevalences imply different efforts in targeting services, and different expectations for researchers.

Table 2
Most Frequently Mentioned Health Problems
Rural Minorities in the Midwest

Problem	American Indian	Asian	Black	Hisp.
Impairment	14.8%	36.4%	28.1%	16.7%
Digestive	15.9	27.3	9.4	16.8
Respiratory	23.9	0	18.8	16.6
Skin & Musculi	10.2	9.1	6.3	16.4
Circulatory	11.4	9.1	28.1	16.3

Source: National Health Interview Survey, 1988

More detailed information about the needs of rural minorities is available from the comprehensive review of literature that follows in this paper. As will be evident, a great deal remains to be learned about the particular problems of minority residents in rural counties. We can gain further knowledge by asking different questions in large surveys using weighted samples. However, the best promise for improved knowledge and application is in increasing the volume of work being completed by researchers and state and local health officials.

METHODS

The first step in this comprehensive review of studies related to the health of rural minorities was to search relevant data bases and collect abstracts for those articles identified by using certain key words in the search. The second step was to review all abstracts in order to determine which articles are relevant to the objectives of this study. Those articles have been acquired and retained for further analysis. Additional research studies, appearing in reports and other documents not published as journal manuscripts or books, have also been acquired. Information was also requested of state offices of rural health, centers for minority health research. Very little actual information was received from sources other than published literature, hence only published articles are used in this summary.

Most of the articles included in this analysis were found using one of five data bases:

- Medline: The bibliographic data base of the National Library of Medicine.
- Sociofile: A cumulative index to English language periodicals, particularly in the fields of anthropology, area studies, community health and medical care, economics, geography, gerontology, international relations, law and criminology, minority studies, planning and public administration, police science and corrections, social work and public welfare, sociology, urban studies and related subjects.
- ERIC: A bibliographic data base covering the journals and technical literature in the field of education.
- Grateful Med: Access to the computer files of the National Library of Medicine which contains approximately 12 million records covering its holdings of books, journal articles and more.
- PsychLit: A comprehensive resource for information in psychology and the behavioral sciences.

Each literature search yielded abstracts of articles, over 660 were collected in all. Each abstract was read by each one of the four faculty investigators involved in this effort (Mueller, Patil, Parker, and Ortega). Each investigator indicated one of three preferences related to retaining the abstract in a data base and obtaining the manuscript: yes, no, perhaps. If any one investigator indicated a yes preference the manuscript was obtained. If two or more indicated a preference of perhaps and the others no, the manuscript was obtained. For all other combinations (principally all no) the abstract was not retained for further consideration. This process yielded 158 manuscripts for this analysis.

All manuscripts have been read by research assistants working on the project (Askenazi) and coded to display information concerning the group or groups investigated, the methodology employed, the sample size for empirical studies, the state or states in which the research was completed, the community selection process (random or non random), the health conditions investigated, and other comments. The coding has been verified through a random check completed by one of the investigators. The categorization of the 158 articles forms the basis of this paper.

FINDINGS

Due to more than one group being considered in some articles, the counts for each group used below will sum to 185. The articles address a variety of issues and use varying methods, as discussed below.

Health Care Issues Addressed

Native Americans. There are 59 articles investigating issues among rural Americans in the data. The most common type of articles were those addressing utilization of, and access to medical services--l2 articles did so. These studies examined utilization in the context either of general patterns or utilization for specific conditions, such as access after an injury. Two categories of medical conditions are studied most frequently--mental health and substance abuse. The 10 articles dealing with mental health focused on depression, stress, and suicide. The 11 articles, researching substance (alcohol, tobacco and illegal drugs) abuse focused on drug use and prevention. Other topics addressed by less than 6 articles each included: accidental injuries and death,

AIDS and AIDS knowledge, cancer, and birth and prenatal care. While most research was not focused on any particular age cohort, 14 were focused only on youth or teen health, while only 1 focused exclusively on the elderly.

African Americans. Seventy-one articles focusing on African Americans were retained. As was the case for Native American research, the single most common focus for research is health care utilization and access broadly defined--26 articles dealt with this subject matter. No other characterization of the research concerning rural African Americans accounts for more than 10 articles. Unlike Native American research though, cancer is a more common focus--l0 articles on cancer status and conditions and 3 on pap smears for women. As might be expected, another difference is attention to hypertension and cardiovascular ailments--8 articles. Drug use appears again as an important category--9 articles. The age focus is different for this population; 8 studies focused on the elderly and 14 on youths. Mental health was once again a topic of interest, but for only 6 of the articles. Other topics that were the focus of fewer than 5 studies include AIDS knowledge and risk, and infant mortality.

Hispanics. Forty-eight articles on the health of rural Latino or Hispanic Americans were retained. As was the case with other populations, the most common categorization of these articles is a focus on utilization and access to care--22 studies did so. Substance abuse and mental health emerge as important topics--8 and 6 studies, respectively. A variety of topics accounted for less than 5 studies each, some of them different from studies of other groups: cancer, hypertension or cardiovascular illness, eating or nutritional disorders, and oral health.

Asian and Pacific Islanders. The literature searches yielded only 7 articles focused on rural populations from this minority group. No topic was covered in more than one article, and the seven were: AIDS knowledge among teens, drug use among youth, dental utilization, fertilization, naso-pharynx cancer, maternal risks, health status, and Hepatitis B during pregnancy.

Reaional Bias

The data base created from the 158 articles includes information about the site at which the research was conducted. Some of the studies used national surveys, so for them there was no particular regional focus. Prior to the literature search the expectation was that there would be very little research focused on rural minorities in regions containing small number of rural minorities. More specifically, the West North Central Region is not expected to be included in very many of the studies.

Among the studies of Native Americans, most (23) were conducted in the Navajo nation, in the southwest (particularly New Mexico and Arizona). Another 14 were based on national surveys, and 3 were completed in Alaska. Only 5 studies specifically included populations from the West North Central region, and one of those is a three-state study including Arizona and Idaho.

Among studies of African Americans, 25 used national samples. Most of the others were, as one would expect, studies of African Americans in the South--4 "southern" samples, 6 from Georgia, 14 from the Carolinas, 3 from Florida, 3 from Alabama, 2 from Tennessee, and 2 from Mississippi. Only one study was in a West North Central state--Missouri, which has the highest number of rural African Americans among the seven states in the region.

None of the studies of Hispanics of Asian and Pacific Islanders were conducted principally in states in the West North Central Census Region. Two studies of migrant farm workers included persons who migrate through the region, particularly in Nebraska. However, most studies of the health of rural Hispanics, as one would expect, are focused in Southwestern states. Several studies of Hispanics use the Hispanic Health and Nutrition Study as a national sample. Three of the seven studies of Asian and Pacific Islanders also use national samples, and two focus on populations in California.

Methodoloaies

The purpose of this paper is to assess the current state-of-the-art research concerning the health of rural minorities. Thus far, the assessment has been based on quantity--number of articles published about various groups, health conditions, and regions of the country. The final consideration is one of quality, loosely considered. This is accomplished by summarizing the methods used by the studies, assuming multivariate research with large sample sizes is an ideal standard.

The largest samples sizes, and often the most sophisticated research methods, were found in studies using large national studies. For example, national surveys of Native Americans included samples of up to 800,000. Multivariate analysis was in 15 of the 59 articles focused on Native Americans, 29 of the 71 articles focused

on African Americans, 14 of the 48 Hispanics, and 2 of the 7 on Asian and Pacific Islanders. The combination of descriptive studies and literature reviews characterized many of the articles in this data base: 26 of the Native American articles, 24 of the African American, 21 of the Hispanic, and 3 of the Asian and Pacific Islander.

CONCLUSION

This paper has reported on the results of an extensive search of literature reporting studies of health care among rural minorities in the U.S. After four scholars read through over 660 abstracts, 158 articles were obtained for use in this analysis. This total reflects a cup that is half full or half empty. The number of articles was somewhat higher than expected, given that they had to be researched-oriented, not descriptions of programs or editorial opinions. On the other hand, the number is low since less than half approach a level of sophistication that would permit discriminating among various factors affecting minority health and access to care.

Why be concerned about rural minorities as a distinct subject for research, versus the health of minorities more generally? The interaction of natural cultural variables and the special conditions of health services in rural areas may pose particular problems for rural minorities. The populations themselves may differ from their urban counterparts, particularly among such groups as migrant workers and reservation residents. Rural minorities may face special problems because they must use the same health care system as others, with the exceptions of migrant health clinics and Indian Health Services facilities. There may no be providers for whom minorities are a large percentage of their practices, as is the case for some inner city providers. Certain areas face comparative disadvantages recruiting minority providers.

Why be concerned that any particular census region is understudied, as is the West North Central? The interaction of environmental variables and health care problems may vary by region, affecting such conditions as cancer. Minorities in one census region may be engaged in different daily activities, such as meat processing or working with particular pesticides, that affects their health differently than counterparts in another region. Minorities may also face peculiar problems accessing the health care system when they are a tiny fraction of the total population being served.

In light of these concerns, a great deal of work remains before we can appreciate fully the health care problems of minorities in rural areas. Future studies should build on previous work by improvingmethodologicalapproaches and by gaining more information from different regions of the country. Continued use of national surveys is also recommended. Those surveys need to be informed, though, by the findings of regional studies in order to ask the questions of greatest interest in a manner likely to yield the most relevant information.

IMPROVING HEALTH STATISTICS AMONG AMERICAN INDIANS BY DATA LINEAGES WITH TRIBAL ENROLLMENT REGISTRIES

Jonathan R. Sugar-man, Portland Area Indian Health Service;
Martha Holliday; Kerri Lopes; Doni Wilder

Approximately 1.9 million people reported their race to be American Indian or Alaska Native in the 1990 Census. These persons include members of over 500 federally-recognized tribes which are widely dispersed across the United States and characterized by substantial cultural, social, and biological heterogeneity. The diversity of these nations is an important part of American culture. This diversity, however, presents a number of challenges to public health statistics. In this paper, we will discuss two of these challenges, and will describe a specific method which can be brought to bear in meeting them.

Several papers in this volume address issues in health statistics germane to minority populations in the U.S. Although there is some overlap of issues related to public health statistics regarding other special populations, a number of features distinguish American Indian populations from other minority groups in the U.S. Many of the most important of these features stem from the special political relationship which exists between Indian tribes and the Federal government. Indian tribes are self-governing, sovereign nations located within the U.S. which have a government-to-government relationship with the Federal government. This relationship has been repeatedly acknowledged in treaties, federal law, and by executive order. Among the responsibilities of the Federal government is an obligation to provide health services to American Indian people. In Public Law 94-437, the 1992 Amendments to the Indian Health Care Improvement Act, Congress reiterated the Federal obligation to "provide . . . health services which will permit the health status of Indians to be -raised to the highest possible level, and to encourage the maximum participation of Indians in the planning and management of those services."

Much of the disparity in health status among racial and ethnic minorities in the U.S. is widely considered to be associated with socioeconomic status rather than with biological characteristics presumably associated with race, or cultural practices and beliefs associated with ethnicity (1,2). Indeed, it is frequently argued that consistent and scientifically valid definitions of race have not been applied in public health surveillance systems (3). However, in addition to the biological and cultural determinants of race and ethnicitv. however defined, there is a political dimension to American Indian race which has profound importance with regard to the collection and interpretation of statistics describing the health of American Indians. Even if, for instance, differences in health status among American Indians compared to other groups are attributable to differences in social class and socioeconomic status rather than to determinants of race and ethnicity, there will continue to be strong political justifications for reporting public health statistics specifically for American Indians because of the Federal responsibility to deliver health services and evaluate health care among this population.

The Indian Health Service (IHS), an agency of the Public Health Service, is the lead Federal agency charged with addressing the health needs of American Indians and Alaska Natives. IHS regularly prepares report0 and analyses describing the health status of American Indian based on vital records and on patient care provided or reimbursed by IHS (4,5). IHS reports data for aggregations of states comprising IHS Areas. rather than for individual states or reservations. These annual reports are focused primarily on Indian people served by the IHS. Thenatalityandmortalitystatistics arederived from tapes prepared by National Center for Health Statistics (NCHS), which originate with each state.

At least two features of dataderived from the vital records systems maintained by states and the NCHS pose difficulties in the description of Indian health status. The first issue is that misclassification of American Indian race can have a profound effect on the estimation of morbidity andmortalityrates among Indian people, almost always resulting in underestimation of rates. The second issue is that few data systems include information about tribal affiliation, andmanytribes are intensely interested in obtaining tribal-specific health data.

Racial Misclassification of American Indiansin Vital Records and Morbidity Registries:

Every rate consists of a numerator and a denominator, anditis axiomaticthatnumerators and denominators must be appropriately matched in order for rates to reflect accurately the experience of the population. In the case of mortality rates for American Indians, the denominator usually consists of census estimates of the Indian population. The numerator consists of deaths coded as American Indian on death certificates. In the case of the census, racial identification is self-selected by the respondent. On death certificates, race is recorded by an observer (such as a funeral director) who may not attempt to verify racial or ethnic identity. There is reason to believe that somepersonswhoareidentifiedas Indianduring their lifetimes are listed as non-Indianon death certificates.

For the years 1983-1985, among infants classified on birth certificates as American Indian (accordingtothe algorithm in use at the time), about 36.6% of infants who died within the first year of life were classified as another race on the death certificate. Using a consistent definition of infant race at birth and death (i.e., mother's race) resulted in an infant mortality rate for AI 46.9% higher than that calculated using standard methods (6).

Similarly, 36.4% of personswhoverereportedas American Indian in Current Population Surveys conducted by the Bureau of the Census and who weretatchedtodeathcertificates for 1979-1985 using the National Death Index were listed as non-Indian on death certificates in the National Longitudinal Mortality Survey. The authors concluded that national mortality rates among American Indians are underestimated by 22 - 30% (7).

Similar findings have been described in morbidity registries, particularly in the Pacific Northwest. For instance, almost 40% of IHS-registered Indian5 in a Surveillance, Epidemiology, and End Results (SEER) tumor registry in Western Washington were listed in the registry as non-Indian (8). End-stage renal disease rates among Indians in Washington, Oregon, and Idaho were underestimated by 15 - 20% in a database of dialysis patient5 in these three states (9). In a statewide injury surveillance system in Oregon, 35% of persons in the registry who were American Indian accordingto IHS records were listed as white (10).

In its annual statistical reports, IHS excludes data for the Portland Area (which includes Washington, Oregon, and Idaho), from California, and from Oklahoma in presenting national mortality data for American Indians because of a presumption that rates of racial misclassification on death certificates in these states are high (5). To the extent that resources for health service8 are at least partially allocated on the basis of mortality data, racial misclassification presents a dilemma for policy makers in appropriately directing these resources.

Lack of Tribal-Specific Data:

The second challenge which emerges from a study of public health data systems is the relative inability to derive tribal specific data from most data sets. National data describing Indian health obscure some extremely important inter-tribal variations. For instance, mortality from lung cancer among American Indians in the United States is, in aggregate, similar to that among the general population. However, mortality rates from lung cancer among American Indian women vary greatly among geographic regions served by the IHS. In some areas, such as Alaska and the Northern Plains, the lung cancer mortality rate is more than double that for the U.S. (11). In parts of the Southwest, lung cancer mortality rates among Indian women are lees than one-sixth those for the U.S.

Even within regions, there is substantial variability of disease prevalence for some conditions. Diabetes is more common among American Indians than among the general population. Although the prevalence of non-insulin-dependent diabetes mellitus among American Indians in Washington, Oregon, and Idaho is three times greater than that for the all races U.S. population, there are significant differences within and among cultural groups consisting of related tribes. (12) (Although no classification system is ideal, anthropologists have described cultural groups as groupings of tribes or band5 that shared significant traditional, religious, and language trait5 prior to European contact. In the Northwest, there are three distinct cultural groups: Coastal, Great Basin, and Plateau.)In a study of diagnosed diabetes which included ten reservations in the Northwest, the prevalence of diabetes on tribal reservations from the Great Basin cultural group was 3.6 time5 that for the U.S. Among the nearby Northwest Coast reservations, diabetes prevalence was only 1.9 times that for the U.S. Within the Plateau cultural group, the prevalence of diabetes among the Nez Perce Tribe was twice that among the Colville Tribe. Therefore, diabetes prevalence data aggregated by state or region in the Pacific Northwest would fail to illuminate important differences among tribes in the region. With the notable exception of New Mexico, state vital records system8 rarely include tribal identifiers (13).

There are significant implications to the absenceof tribal-specificdata. Healthcareis delivered to American Indian8 in the Pacific Northwest in a variety of setting8 operated by many agencies. Sovereign tribes have health department8 and ambulatory care clinics for the purpose of planning and providing health services. Under the provisions of Public Law 93-638, tribes can contract with the Federal governmenttodeliver services currently provided by the IHS. However, some tribes have been reluctant to exercise this option because of their inability to obtain accurate data describing the health status of their members. From a resource allocation standpoint, the lack of tribal specific data confounds decisions regarding the targeting of services to address specific health problems among particularly high risk groups.

Linking Tribal Enrollment Registries with Health Data Sets:

In ordertoaddress theproblemsof racial misclassification of American Indians in vital record8 andmorbidity registries, and to obtain tribal specific health data, the Northwest Portland Area Indian Health Board (NPAIHB) and the Portland Area IHS are collaborating on a study to linktribalenrollment registries with other data systems. The findingsdescribedbelow are based on the first nine months of a three-year collaborative agreement between the NCHS and the NPAIHB. The NPAIHB is a non-profit tribal advisory board which represents the forty federally-recognized tribes in Oregon, Washington, and Idaho to address health care needs and concern5 of American Indian8 in the region.

Indian tribes define their membership accordingto standards set by eachtribe. These standards, for instance, may include a specific *blood quantum", or proportion of tribal heritage, necessary to qualify for membership. Tribal rolls are maintained by each tribe, and essentially serve as enumeration of a defined population.

The approach taken in the study is conceptually straightforward-- to link tribal enrollment registries with vital record8 and disease registriesmaintainedbystates and other entities in order to correct for inconsistent classification in health data sets andtoobtain tribal specific data. Previous approaches to similar data linkage8 have relied on the IHS patient registration file, which only includes persons eligible for IHS services who actively register for health services. While the population of persons registered with IHS has relevance, in terms of the delivery of health services, the number of American Indians registered with IHS is only about half the number of person8 who identified themselves to be American Indian in the 1990 census. Tribal enrollment registries define specific populations based on tribal membership, rather than on persons who have registered to receive health services.

For the purposes of this research project, enrollment registries forparticipatingtribee in Washington, Oregon, and Idaho are being consolidated into a single electronic database includingpereonalidentifiere (suchasnameand birth date) which can be used for record linkages. In some cases, existing computerized databases are converted into a standard file format. In others, project staff keypunch data from manually-maintaitied registries into a

computerized database. (For tribes whose database is not currently computerized, an Epi-Info file with the registry database is returned to the tribe along with the software and documentation for the database). Various computer andmanual record linkage techniques, including computer-assisted probabilistic matching, can be conducted using the tribal database. Information regarding tribal-specific morbidity and mortality rates are provided to tribal government0 with appropriate attention to individual confidentiality.

The implementation of the approach requires substantial time, resources, and persistence. Release of the tribal enrollment registries must be approved by each tribe's governing body-- the tribal council. Based on previous experience, many tribes are cautious about participating in research projects. Tribal council0 have carefully scrutinized issues of confidentiality and data ownership, and have required that no data which can be used to identify a tribe be releaaed without prior approval from the council.

To date, 11 of the forty federally-recognized Northwesttribee have passed formal resolutions agreeing to participate in the research project, and over 9,000 names for eight tribes have been entered into an electronic database. Two pilot studies have demonstrated the feasibility of conducting data linkage0 on a small scale.

Results of Pilot Studies:

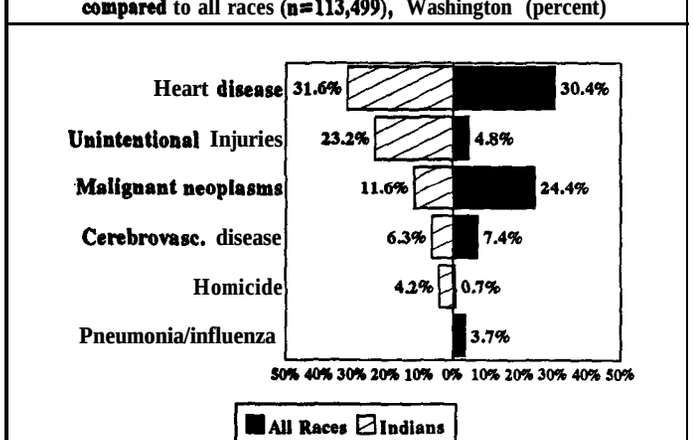
Two tribes located in Washington State have maintained a registry of deceased tribal member0 along with their tribal rolls. A search was conducted ueing Russell Soundex code6 to screen for potential matches among Washington etate death certificates. Matches were verified using identifying data on tribal rolls and death certificates. Racial coding and underlying cause of death were recorded.

The distribution of underlying cause of death based on the most recent 95 deaths for 2 tribes is shown in the slide. As shown in Figure 1, a much higher proportion of deaths among Indians in these communities are associated with injuries than Fe the case with the all race8 population for the state. Approximately. 15% of deceased tribal member0 were coded as non-Indian on death certificates.

The second pilot study is intended to provide information for tribal health services plannera. In the current climate of health care reform, information about the utilization of health service0 in a population is crucial for communities or organizations considering participation in capitatedplana. Administrative data bases maintained by the Health Care Financing Agency (HCFA) whichincludeexpenditure anddiagnosisdata for Medicare beneficiaries do not currently maintain a separate code identifying American Indianbeneficiaries. (The racial codes are white, black, andother.) Tribal groups wishing to contract for the delivery of health care to their population0 have been concerned that a low proportion of potential Medicarebeneficiaries are enrolled in Part Bof Medicare, although population-based data to supportthis assumptionhavenotbeenavailable. Because some tribes intending to assume responsibility for the administration of health care for theirmembers are consideringpurchasing Part B coverage eligible tribal members who have not already enrolled, documentation of the current level of enrollment of tribal members is important.

In collaboration with the Division of Health Standards and Quality, Region 10, HCFA we conducted a linkage study to identify thenumber of AI beneficiaries in a small tribe in a Northwest state, and to assess Medicare expenditures for these beneficiaries. Using a complex set of steps, 13 of the 18 persons listed by the tribe as being 65 years of age or older were definitively linked with the national Medicare beneficiary Enrollment Database. Four additional persons were linked. Twelve of the 17 persons (71)% werelisted as "other race" in the Medicare files. Among the five persons who were listed as white, 4 were at least 5/8 Indianblood quantum (percent Indian heritage), and the fifth was 1/2 blood quantum. When the list of beneficiarieswas linkedtothepayment file for Medicare Part A (hospital) expenditure0 for 1991, diagnostic and expenditure data for the cohort were identified. Linkage with a subset of the national HCFA denominator file including enrollee8 recorded as residing in the Northwest state during 1991 was used to determine the proportion of beneficiaries enrolled in Part B. Although the small size of the study precludes any generalizable conclusion0 regarding the specific characteristic0 of American Indian Medicarebeneficiaries, the success of the pilot study suggests that useful information may be obtained by linkages conducted on a larger scale.

Figure 1: Leading causes of death, members of two Indian tribes (n=95), compared to all races (n=113,499), Washington (percent)



All race data for 1989.1991; Tribe A data for 1988-1991; Tribe B data for 1973.1991

Limitations:

There are both practical and conceptual limitations to this method. From a conceptual standpoint, these studies are limited by the fact that all persons who identify themselves as American Indian are not included in tribal rolls. Specifically, many American Indian residents of Northwest states are members of tribes indigenous to other states (14). Therefore, registered tribal members may not be representative of all American Indians in the regions. In addition, the small populations in many of the tribes are problematic for analysts because of unstable morbidity and mortality rates.

The practical limitations are substantial. The process of collecting tribal rolls is extremely resource intensive. Obtaining tribal approval to conduct the study has often required several on-site vieite. Even after approval hae been obtained, compiling a computerized database from manual registries and from a plethora of proprietary computer software packages has proven to be a difficult exercise. Reaching agreements to assure all parties that confidentiality concerns will be appropriately addressed can be time-consuming.

Although this project is still in ite initial phase, it has already demonstrated that the data linkage approach is feaeible, and that a method can be developed to meet the needs of Northweet tribes to obtain tribal-specific data. Direct measurement of the rate of racial mieclaseification among deceased tribal members in Northwest states hae been achieved for several tribes. Tribal-specific information on enrollment of Indians for programs such as Medicare appears to be obtainable, and expenditure data can be generated as well. This project reyresents the first demonstration of which we are aware regarding the feasibility of reporting of data for a population-based sample of American Indians enrolled in Parts A and B of Medicare, which is a significant achievement in view of the importance of Medicare in the financing of health care in the U.S.

Application of the methode developed in the pilot studies to a larger number of Pacific Northwest tribes ie currently underway. If these efforts are eucceeeful, linkage of tribal enrollment registries with health data eets in other regions may be an effective means of gathering useful data to describe the health of American Indians. However, the true measure of success of these methods will be the application of the data by tribes and other agencies for the implementation of programs to improve the health statue of American Indians.

Acknowledgmenter

We gratefully acknowledge the tribal councils and staff who have cooperated in the project, staff at the Center for Health statistics, Waehington State Department of Health for assistance with retrieval of death certificates, and Joseph K. Lowery, Ph.D., for conducting the linkages with H.C.F.A. data systems. "The opinions-expressed in this paper are those of the authors. and do not necessarily reflect the views of the Indian Health Service. This research wae supported through a cooperative agreement from the NCHS, CDC (U83/CCU008678-01).

References:

1. Rogers RG. Living and dying in the U.S.A.: Sociodemdgraphic determinants of death among blacks andwhitee. Demography 1992;29:287-303.

2. Oeborn NG, Feit RD. The use of race in medical research. JAMA 1992;267:275-279.

3. Hahn RA. The etate of federal health statistic6 on racial and ethnic groups. JAMA 1992;267:268-271.

4. Indian Health Service. Trends in Indian Health 1992. Rockville, MD: US Department of Health and Human Services, 1992.

5. Indian Health Service. Regional Differences in Indian Health 1992. Rockville, MD: US Department of Health and Human Services, 1992.

6. Hahn RA, Mulinare J, Teutech SM. Inconeietenciee in coding of race and ethnicity between birth and death in US infants: Anew look at infant mortality, 1983-1985. JAMA. 1992;267:259-263.

7. Sorlie PD, Rogot E, Johnson NJ. Validity of demographic characteristics on the death certificate. Epidemiology 1992;3:181-184.

8. Frost F, Taylor V, Fries E. Racial misclaeeification of Native Americans in a Surveillance, Epidemiology, and End Results cancer regietry. JNCI 1992;84:957-962.

9. Sugarman JR, Lawson L. The effect of racial misclassificationoneetimatee of end-stage renal diseaseamong American Indians and Alaska Natives in the Pacific Northwest, 1988 through 1990. American J Kidney Dieeasee 1993;21:383-386.

10. Sugarman JR, Soderberg R, Gordon JE, Rivara FP. Effects of racial mieclaseification of American Indians on injury rates in Oregon, 1989-1990. American J Public Health 1993;83:681-684.

11. Valway S, Ed. Cancer mortality among Native Americans in the United Statee: Regional Differences in Indian Health, 1984-1988 and Trends Over Time, 1968-1987. Rockville, MD: Indian Health Service, US Department of Health and Human Services, 1991.

12. Freeman WL., Hosey, GM., Diehr P., Gohdes D. Diabetes in American Indians of Washington, Oregon, and Idaho. Diabetes Care 1989;12:282-88.

13. Halaean C, Totkamachi P, Towles L, Mueller A, Ortiz T. 1990-1991 New Mexico tribe specific vital statistics. New Mexico Department of Health, Public Health Division, Bureau of Vital Records and Health Statistics. Santa Fe, NM, November 1992.

14. Paisano E, Cowles J, Carroll D, Robinson A, Green R. American Indianopopulationbytribe for the United Statee, regions, divisions, and states:1990. Washington, DC: Racial Statistics Branch, Population- Division, Bureau of the Census, October 1992. (1990 CPH-L-99).

MINORITY HEALTH RESEARCH

Discussants:

Ruth Zambrana
Agency for Health Care Policy and Research

P. Douglas Williams
National Center for Health Statistics

(Not available for publication)

Session S

Intercensal Population Estimates

Toward the year

2000

Refining the measures

ESTIMATING THE POPULATION OF RURAL COMMUNITIES BY AGE AND GENDER: A CASE STUDY OF THE LOCAL EXPERT PROCEDURE

David A. Swanson, University of Arkansas
John F. Carlson, Linda K. Roe, and Christopher A. Williams

ABSTRACT

This paper examines the suitability of a survey-based procedure for estimating population in small, rural areas. The procedure is a variation of the Housing Unit Method. It employs the use of local experts enlisted to provide information about the demographic characteristics of households randomly selected from residential unit sample frames developed from utility records. The procedure is nonintrusive and less costly than traditional survey data collection efforts. Because the procedure is based on random sampling, confidence intervals can be constructed around the population estimated by the technique. The results of a case study are provided in which the total population, and its age and gender distribution, is estimated for three unincorporated communities in rural, southern Nevada and tested for accuracy against 1990 census data.

INTRODUCTION

This paper describes the methods used in the development of a demographic data base established in support of the Yucca Mountain Site Characterization Project Radiological Monitoring Program (RadMP)[1]. These data are required in the analysis of the RadMP study area. Presented in this paper are the results of an evaluation of the accuracy of the "Local Expert Procedure," a variation of the housing unit method (HUM) for estimating the age and gender distribution of populations in small rural communities.

The Local Expert Procedure. combines analytical techniques that have largely been developed in isolation from one another. These procedures are: (1) the HUM of population estimation: (2) random sampling; and (3) key informant ethnography (local experts).[2] These procedures, in combination with one another, may lead to a very powerful means of obtaining the population size and household composition information which are required for most impact assessment projects. Since the Local Expert Procedure relies primarily on people (the local experts) and administrative records that are available for any geographic area, it can in principle be applied to any small, rural area for the development of population estimates.

One of the best sources available for determining the number of households in place at a given point in time is the administrative record system of public utilities.3 Rives and Serow suggest this approach, as part of the HUM, to obtain information on housing types and location by relating each residential structure to an electric meter.

The HUM has been found by the U.S. Bureau of the Census to be used by 89 percent of state and local agencies responsible for the development of population estimates under the auspices of the Federal/State Cooperative Program for population estimates.' The principal advantage of the RUM for estimating the population of small, rural areas is that the researcher can accurately code the location and type of housing for all current residences within the study area. Once the total numbers for each housing type in a given community have been established, it is then possible to compute the sample sizes necessary to obtain representative information for eaah housing type, and to employ a random sampling technique for the selection of households to be included in each sample.

The practice of employing "key informants" in community studies has also been found to be a useful data-collection technique-particularly in areas experiencing rapid demographic and sociocultural change. Poggie[5] found that when the questions asked in the field are "noncontroversial concrete," and about "directly observable public phenomena," key informants have been shown to be a "highly reliable and precise" source of information (p.29). The informant method is generally acknowledged as 'reliance on a small number of knowledgeable participants, who observe and articulate social relationships for the researcher."[6]

The analysis presented in this paper is a follow-up to an earlier study[2] in which the Local Expert Procedure was found to produce accurate estimates of the total household population. In this paper, the accuracy of the Local Expert Procedure is examined for producing estimates of the gender and age composition of household population.

STUDY AREA

This research is part of a comprehensive program to assess the socioeconomic characteristics of communities located near Yucca Mountain, the site currently being evaluated for the nation's first geological repository for spent nuclear fuel and high level radioactive waste. During 1909, a Socioeconomic Plan for the Yucca Mountain Project7 was developed to ensure that adverse socioeconomic impacts that may result in Southern Nevada from the scientific investigations program are avoided or minimized to the maximum extent practicable. The data collected from this study are required to monitor the communities nearest Yucca Mountain for any changes in the size or composition of the population that may occur as the scientific investigations continue.

The communities included in this study-Beatty, Amargosa Valley, and Pahrump-are located in southern Nye County, Nevada. Each community is distinct, both geographically and demographically.[8] For example, the bulk of Beatty's 1,600 peoples is contained within a one square mile area. By contrast, Amargosa Valley is a community of nearly 900 peoples scattered over 500 square miles.10 Pahrump has the largest

population of the three communities with over 7,400 residents9 living in a 280 square mile area.10 Differences in age composition also exist. According to the 1990 census, nearly a quarter of Pahrump's population (21.5%) is 65 years of age and older.9 By contrast, only 6.3 percent of Beatty's population and 10.2 percent of Amargosa Valley's residents fall into this age category.9 Pahrump is the only one of three communities believed to have experienced significant growth since 1990. From 1990 to 1992, Pahrump experienced an increased of 702 households.11 Using the standard nonlinear method for calculating average annual rates of change and the 1990 household count of 3,555, the increase of households in Pahrump represents a 9.5 percent average annual rate of change. Beatty, on the other hand, experienced an increase of only 85 households-a 5.5 percent average annual rate of change. Amargosa Valley showed an increase of only 2 occupied units (0.3 percent average annual rate of change) during the same period.10

METHODOLOGY

The sample selection and data collection techniques employed in this research were designed to ensure the development of a demographic data base which accurately represents each community. During a preliminary phase of the research, several trips were made to the study area to develop a familiarity with the communities. Visits with community leaders and representatives of local businesses resulted in a network of contacts which later facilitated the collection of data. Field notes were taken during each of the trips describing the general layout of the communities, including the types and locations of business and residential areas. Four types of structures were identified and operationally defined, including mobile homes, permanent single family homes, multiple family structures, and travel trailers. Group quarters were not included in this study.

The next phase of the study involved mapping and coding the location and type of structures for all current residential electrical connections in each community. Using computer-generated plots, representatives from the electric company servicing southern Nye County identified the location and type of all households for the study area. Whenever questions arose regarding the location, type of housing, or the number of households serviced by a single meter, the location was visited to verify this information. This mapping process is part of the larger study. However, it is not a required feature of the local expert procedure.

A random selection of households, based on the coding system that identified each element by location and housing type, was drawn separately for each community. Sample sizes were based on the total number of households in the community. The formula used to calculate each sample assumed a 5% margin of error, an alpha level of .05, and a dichotomous variable with an estimated proportion of 0.5. This, of course, is a conservative strategy that accepts the "cost" of a sample size that is likely to be larger than required for the precision we desire. Once the sample size was determined, an additional 15% was added to allow for the possibility of having to exclude households from the sample for which data is unknown.

Finally, local experts were chosen from each community based on their experience in community activities and their familiarity with local residents. Each individual was interviewed and asked to complete a questionnaire which was designed to assess his or her qualifications.2

During the actual data collection process, the local experts were instructed not to identify households by name to ensure the anonymity of community members, although from the utility record the local experts know the householder's name under which the sampled unit's account was maintained. For each household, local experts were asked to record, or to communicate to the researcher, the total number of persons residing in the household as of July 15, 1990, the age category and gender of each household member, and the retirement status of persons fifty years and older. Eight age categories were used. The first two distinguished pre-school (O-4) from school aged (5-18) children. The next three accounted for persons normally considered to be in the work force. These categories include the age 19-29, 30-39, and 40-49. Retirement status was added as a variable for each of the last three categories 50-59, 60-64, and 65 and over.

Data were collected on a form that listed and identified the sample households by an attribute number (designated according to location on a map), the electrical meter number assigned to the unit, and the type of housing. All structures, including those identified as "burned down" or otherwise destroyed, unoccupied, or "removed from the pad" were considered as part of the final sample. Structures identified as "not a residence" or structures for which data were unknown were not included in the final sample. More detailed information on this part of the local expert procedure is found elsewhere.2

DATA ANALYSIS

The data collected for this study represent the population residing in each of the communities on July 15, 1990. This date makes it feasible to compare the study results with the data collected during the 1990 census.2 The data representing the age and gender distribution for each household were compiled from the original data-collection forms. Data for age were then collapsed into four categories: O-18 (pre-school and school age), 19-39 (labor force and primary child bearing years), 40-64 years (labor force and preretirement), and 65 and over (retirement) years. These age groups are comparable to the age groupings reported in the 1990 census data.

The primary goal in the evaluation was to construct confidence intervals for each of the study variables. The procedure used to compute the confidence intervals is illustrated below. Age is used as an example.

First, the arithmetic mean and standard error were computed for a given age category

Second, once the standard error was computed, lower and upper confidence intervals were estimated using the following formulas:

upper limit $=(N)\cdot(x+(t_{n-2},a/2)\cdot(s.e.))$

lower limit $=(N)\cdot(x-(t_{n-2},a/2)\cdot(s.e.))$

where : N = total households in the community
x = mean number per sampled household for the age category
s.e. = standard error
tn-2 = (a/2) 100th percentile of the t distribution, with (n-2) degrees of freedom

FINDINGS

Table 1 provides a review of the results from the previous study2 which found the Local Expert Procedure capable of providing accurate estimates of the total household population. Included in the table are: the total number of households in each community, the estimated persons per household (PPH), (generated from the sample data), the standard error of the estimated PPH, the estimated total household population, the 1990 Census, and the confidence intervals for the total household population estimates. These data show that the population counted in the 1990 census for all three communities fit within the 95 percent confidence intervals generated around the estimates.

Table 2 shows the estimated average number of males per household by community and the confidence interval for each of the estimates. These estimates are compared with the census data[9] in terms of the number of males in the total household population. The data show that in both Amargosa Valley and Pahrump the census count for males is within the confidence intervals. In Beatty, the census count is slightly higher (6 males) than the high end of the confidence interval.

The computed confidence limits of the estimates by age category are presented in Table 3 for each community. The actual 1990 census count is indicated to the right of the confidence limits. Overall, the figures show that the estimates are more accurate for the lower age groups than the higher age groups. The census counts for Amargosa Valley and Beatty, the two smallest communities, are within the confidence intervals for the two lowest age groups (0-18 and 19-39). In the highest age group (65 and over) the estimates for both communities are too low for the census count to fall within the confidence intervals. In Pahrump, the confidence interval for the highest age group (65 and over) is also too low to include the census count, and too high for the age group 40-64. Unlike Amargosa Valley and Beatty, the census count for the lowest age group (0-18) in Pahrump does not fall within the confide&e intervals. the only instance where the census count falls within the confidence interval in Pahrump is for age group 19-39.

SUMMARY AND DISCUSSION OF THE FINDINGS

Table 4 provides a summary of the findings along with the number of local experts used in each community and the size of the community. When compared to the total population and number of local experts, the findings point to a number of factors which could potentially influence the accuracy of the Local Expert Procedure for estimating age and gender.

The first factor presumed to have affected the accuracy of the estimates in this study is the size of the local expert "team" relative to the size of the community. In the smallest communities, where more than one local expert was employed (Amargosa Valley and Beatty), the estimates were generally more accurate than the estimates generated for Pahrump, where only one local expert was ultimately used. The highest number of local experts (4) was employed in Beatty which generated accurate age estimates for three of the four age categories. The estimated number of males in Beatty was also very close to the census which reported six more males than the upper limit of the confidence interval. In Amargosa Valley, data collected from two local experts produced accurate estimates for two of the four age groups and for the distribution of males. In Pahrump, only one age group (19-39) and the distribution of males were estimated accurately.

TABLE 1.
Sample Characteristics and Results of the Accuracy Test For The Estimated Total Population.

Community	Households	1990 Sample Estimates PPH	SE	Population	1990 Census Count	Confidence Interval Low	High
Amargosa Valley	326	2.58	0.11	841	853	771	911
Beatty	672	2.43	0.10	1,633	1,623	1,501	1,765
Pahrump	1,224	2.21	0.06	7,190	1,424	6,810	1.569

Table 2.
Estimated Number of Males by Community, the 95% Confidence Intervals for the Estimates, and 1990 Census Counts of Males.

Community	Households	1990 Sample Estimates Male PPH	SE	Males	1990 Census Count	Confidence Interval Low	High
Amargosa Beatty Valley	326 672	1.23 1.32	0.05 0.06	430 827	914 436	756 389	475 988
Pahrump	3,224	1.09	0.03	3,514	3,700	3,316	3,713

TABLE 3
Estimated Population by Age and Community, the 95% Confidence Intervals for the Estimates, and 1990 Census Counts by Age.

Age Group	Amargosa Valley Confidence Interval of Estimate Low	High	1990 Census Count	Beatty Confidence Interval of Estimate Low	High	1990 Census Count	Pahrump Confidence Intervals of Low	High	1990 Census Count
0-18	198	324	267	367	571	424	789	1,409	1,663
19-39	171	255	245	601	162	609	1,277	1,811	1,571
40-64	293	378	252	360	512	487	3,690	4,190	2,597
65+	17	53	89	18	73	103	459	812	1,593

TABLE 4
Summary Results of the 1990 Age and Male (Gender) Estimates and Number of Local Experts by Community

	AGE GROUP (Census count is within 95% confidence interval of the estimate)		MALES (Census count is within 95% confidence interval of the estimate)	SIZE OF THE COMMUNITY (1990 Census EXPcount)	NUMBER OF LOCAL
Amargosa Valley	0-18	Yes			
	19-39	Yes	Yes	853	2
	40-64	No			
	65t	No			
Beatty	0-18	Yes			
	19-39	Yes	Yes	1,623	4
	40-64	Yes			
	65t	No			
Pahrump	0-18	No			
	19-39	Yes	Yes	1,424	1
	40-64	No			
	65+	No			

Research on decision-making in groups indicates that accuracy increases when there is variance in the information held by group members and all information is shared and processed.[13] This is believed to reflect the experience of the local expert group in Beatty. There appeared to be less variance in information held by the two local experts in the Amargosa Valley, although they shared their information openly.

A second factor that could explain why the estimates are generally more accurate in Beatty is that compared to Amargosa Valley and Pahrump, Beatty appears to be a more "tightly knit" community. Since the community is small in land area and households are closer together, the opportunity for residents to interact is greater. This factor is presumed to influence extent and nature of each local expert's knowledge of households in the community. It fits within a long tradition of sociological research and theory regarding social intergration.[12]

The presence of identifiers associated with preschool and school-age children appears to have been an important factor in estimating the number of persons in the household 0-4 and 5-18 years of age. In all three communities, the estimates of the age distribution show a pattern of greater accuracy in the lower age categories. The absence of a strong identifier to differentiate the number of persons 65 years of age and older from those 40-64 may account for the fact that the census counts for the oldest age fall outside of the confidence intervals.

At this point, "judgment" was used by the local experts to develop the demographic characteristics of each sampled household. Research indicates that the development of "rule-based" procedures may improve accuracy.14 In this regard, follow-up work with the local experts used in this study would be useful. Accuracy may also improve if the subjective level of uncertainty that the local experts have is better understood, particularly in regard to estimating the age of those who are 65 and over. Research has shown that "desirability" increases positively with uncertainty.15 Since there was a tendency to underestimate the age of those 65 and over and at the same time overestimate those aged 40-64, it may be that social values concerning aging may be leading to the underestimation of the age of those in the initial years of retirement age since "youthfulness" is more socially desirable than "agedness."

The results of this study suggest that the extension of the Local Expert Procedure to the estimation of population composition (age, gender) holds promise. With a careful selection of local experts and good knowledge of the community in question, accurate, cost effective estimates are possible that are not based on intrusive data-collection techniques. The results also show that, at least for certain age groups, accurate estimates are more difficult for rapidly growing communities. Further analysis of the factors that are believed to affect the accuracy of the estimates is suggested. In this regard, ethnographic research may be particularly salient as well as the development of "rule-based" procedures. The experience of working with range of "local expert" group sizes, coupled with research on the accuracy of group decision-making, suggests that local experts be recruited with an eye toward maintaining "variance." That is, assembling a group in which each member has a high level of community knowledge, but in which their individual knowledge is based on different community experiences and overlaps only partially with that of others in the group. Once such a group is assembled, the research by Sniezek[13] suggests that the members receive training in the open sharing and processing of information as well is in the "rules" developed for decision-making[14] once a workable set of rules is created. It may even be the case that these extensions would lead to the accurate application of the local expert procedure to larger communities. For example, an exhaustive and mutually exclusive neighborhood by neighborhood application could lead to accurate estimates for a large city.

We conclude by noting that low costs and a nonexistent response burden are inherent features of the Local Expert Procedure. Coupled with its ability to produce accurate estimates as well as measures of estimation uncertainty, these features make the Local Expert Procedure particularly attractive.

REFERENCES

1. U.S.Department of Energy (DOE), 1990 Yucca Mountain Site Characterization Project Radiological Monitoring Plan. Las Vegas, Nevada: U.S. Department of Energy, Yucca Mountain Project Office.

2. Roe. L.K., J.F. Carlson, and D.A. Swanson 1992. "A Variation of the Housing Unit Method For Estimating the Population of Small Rural Areas: A Case Study of the Local Expert Procedure." Survey Methodology 18 June: 155-163.

3. Rives, N.W. and W.J.Serow 1984. Introduction to Applied Demography: Data Sources and Estimation Techniques. Newbury Park, California:Sage Publications.

4. Byerly, E. 1990. "State and local agencies preparing population and housing estimates." Current Population Report Series P-75. No.1063, Washington, D.C.: U.S. Bureau of the Census.

5. Poggie, J.J. Jr.1972."Toward Quality Control in Key Informant Data." Human Organization 31(1):23-30.

6. Seidler, John 1974."On using Informants: A Technique for Collecting Quantitative data and Controlling Measurement Error in Organization Analysis." American Sociological Review 39:816-831.

7. U.S. Department of Energy (DOE), 1991,a. Yucca Mountain Site Characteristics Project Socioeconomic Plan: Revision O. Washington, D.C.: Office of Civilian Radioactive Waste.

8. Carlson, J.F., D.A. Swanson, and C.A. Williams 1990. 'The development of Small Area Socioeconomic Data to be Utilized for Impact Analysis: Rural Southern Nevada." pp.985-990 in High Level Radioactive Waste Management Proceedings of the 1990 International Conference. La Grange Park, Illinois: American Nuclear Society.

9. U.S. Department of Commerce (DOC), 1991,. 1990 Census of Population and Housing, General Population Characteristics. Washington D.C.:Bureau of the Census.

10. U.S. Department of Energy (DOE), 1991,b. Draft Summary of Socioeconomic Data Analyses Conducted in Support of the Radiological Monitoring Program During Calendar Year 1990. Washington, D.C.: Office of Civilian Radioactive Waste.

11. U.S. Department of Energy (DOE), 1993. Summary of Socioeconomic Data Analyses Conducted in Support of the Radiological Monitoring Program During Calendar 1992. Washington, D.C.: Office of Civilian Radioactive Waste.

12. Brown, R., H.R. Geertsen, and R. Krannich 1989 "Community Satisfaction and Social Integration in a Boom town: A Longitudinal Analysis." Rural Sociology 54 Winter: 568-586.

13. Sniezek, J. 1992 "Groups under Uncertainty: An Examination of Confidence in Group Decision Making." Orgranizational Behavior and Human Decision Processes 52;124-155.

14. Dawes,R., D. Faust, P. Meehl 1989. "Clinical Versus Actuarial Judgment." Science 243:1668-1674.

15. Weinstein, N. 1980 "Unrealistic Optimism about Future Life Events" Journal Personality and Social Psychology 39:806-820.

16. Buckley, T. and J. Sniezek 1992 "Passion, Preference and Predictability in Judgment Forecasting." Psychological Reports 70:1022.

INTERCENSAL ESTIMATES OF THE POPULATION OF STATES AND COUNTIES by AGE, SEX, RACE, AND HISPANIC ORIGIN: 1980 TO 1989

Sam T. Davis, U.S. Bureau of the Census

INTRODUCTION

After appropriate data from the 1990 Census became available, the Population Division at the Bureau of the Census produced intercensal estimates for several levels of geography and, depending on that level, by different age, sex, race and Hispanic origin detail. Generally, these estimates were for July 1 of each year from 1981 through 1989. (We did make some estimates for July 1, 1980, but the controls were not the same as for the estimates we discuss below.)

Terminology

Before proceeding, it is necessary to define exactly what we mean here by "intercensal". It appears from the title of this session that intercensal refers to any estimates prepared between two censuses, whether both censuses have occurred or not. Estimates made for 1991 or 1992, for example, could be called intercensal because they are between the 1990 and 2000 censuses. This usage is not uncommon, but leads to some confusion when this usage is mingled with the stricter usage, which is the one employed for this presentation, where we will use intercensal to refer only to estimates which are produced using the results from two censuses. In this case the two censuses would be those for 1980 and 1990. The intercensal characteristic thus is due to methodology rather than location of the date for which the estimates are made. Estimates made after a census, using data only from that census are called "postcensal" estimates. Since there are no results yet from the 2000 census, any estimates we make for 1991, 1992, etc., are by necessity postcensal. Estimates for years during the 1980's can be postcensal or intercensal, depending on the type of methodology we use to make them.

GEOGRAPHY

In making intercensal estimates for the 1980s, we produced estimates on the National, State, and County levels. At the National level, we made intercensal estimates for single years of age, Sex, Race (4 groups - White, Black, American Indian and Alaska Native (AIAN), and Asian and Pacific Islander (API)), and Hispanic origin. The State estimates were produced for Single years of age and sex. County estimates were created by 5-year age group, sex, and 3 race groups (White, Black, Other Races, including AIAN and API). As a side product of these county estimates, we also made State estimates by Age, Sex, Race and Hispanic origin, along with some special products which will be discussed later.

SPONSORSHIP AND PREVIOUS ESTIMATES

These estimates were produced at the request of both the National Cancer Institute and the National Center for Health Statistics. The two agencies requested different sets of estimates, but due to their similar nature were combine into one overall project. Several other government agencies have used these numbers, as well as national private organizations and state and local agencies.

The first county estimates by age, sex, and race were produced during the 1970s and we have continued, with changes in methodology through the middle of the 1980s. we did not continue postcensal estimates through the later years of the 1980s, which leads to the following discussion of the different types of intercensal estimates.

"TRADITIONAL" INTERCBNSAL ESTIMATES

*Traditional" intercensal estimates are obtained using postcensal estimates carried through to the next census date (in our case April 1, 1990) and the actual Census data. The postcensal estimates rely only on the 1980 census and the various measures between the 1980 census base and the estimates date. When this procedure is carried through to the April 1, 1990 estimate date, the numbers obtained by the postcensal procedure are different from those given by the April 1, 1990 Census. The main question of how to make the intercensal estimates lies in how to distribute that difference to each of the intercensal years so that the revised, intercensal estimates arrive at the April 1, 1990 Census figures rather than the postcensal numbers. There are several ways of carrying out this allocation of the differences, six of which are presented in a memorandum written in the early 1980s by Prithwis Das Gupta of the Population Division's Population Analysis Staff. The differences among the methods is in how to distribute the discrepancy between the postcensal

estimates and census counts. The simplest method would allocate an equal proportion of the discrepancy to each year. Another method depends on the sire of the estimate, another puts more of the difference to years later in the decade, and other methods use various combinations of these and other criteria. The method we chose is referred to as "Number 6" and was so chosen by examining trial results obtained using the various methods, this method showed the quality of preserving much of the pattern of change through the decade given by the postcensal estimates. This was chosen over another similar method which was arithmetic rather than geometric as is method number 6. The geometric version tends to dampen extreme fluctuations. The formula for this procedure is

$$Pt = Qt(P10/Q10)t/10$$

where

t = time (in years) elapsed since April 1, 1980
Pt = Population Estimate at time t
Qt = Postcensal Estimate at time t
P10 = April 1, 1990 Census Count
Q10 = April 1, 1990 Postcensal Estimate

We used this procedure to obtain all the estimates at the national, state and county levels for which we had April 1, 1990 postcensal estimates. This included national estimates by age, sex, race, and Hispanic origin, State estimates by Age and Sex, and County total estimates. For estimates which we wanted but for which we did not have 1990 postcensals, including the county age, sex, race estimates, we had to use a different technique.

"SPECIAL" INTERCRNSALS

County Estimates

Bearing in mind we already had traditional county total intercensal estimates, we needed to obtain consistent intercensal estimates by age, sex, and race. While considering other techniques, we decided to employ one of the simplest, which was to use linear interpolation on the 1980 and 1990 age, sex, and race proportions of each county to obtain an age, sex, race structure for each county for each intercensal year. We then applied these interpolated proportions to the county totals.
Through several raking procedures, we controlled these numbers both to the intercensal state age-sex totals and to the national age, sex, race, totals.

The basic assumption of this procedure is that the age, sex, race structure of a county change in a linear fashion between 1980 and 1990, but the total numbers for all age, race, sex cells rose and fell as the county total did, and that each age and sex cell fluctuated consistent with its State total. In addition, we are assuming that all counties were affected by the overall national change in age, sex, and race structure. Unfortunately, other than anecdotal, single county cases, there are no real data available that we could use to test these assumptions. One alternate procedure we developed, discussed, but did not use for lack of any solid basis, was one based on the assumption that change in age structure is related to the change in total county population. If, therefore, a county had a tremendous increase between, say, 1984 and 1985, an equally large change in the age, sex, race structure would take place, yet still within the 1980-90 limits. This has an intuitive appeal, and we may possibly implement it in the future. Its effect, however, would be limited to counties with erratic changes in total population and substantial changes in age, race, sex, structure.

State Estimates

For states we wanted more race detail and a split for Hispanic origin, which we did not obtain for counties. To obtain what we wanted, we started by aggregating our county numbers to the state level. Next we carried out an interpolation similar to the one we did for county age, sex, race proportions, but for a different set of proportions. We interpolated the proportions we needed to split the state age, race, sex totals into the desired detail. These were then the proportion of other races that were American Indian/Alaska Native (AIAN) and Asian and Pacific Islander (API), and the proportion Hispanic origin/non-liispanic origin for all four race groups. We then applied these proportions to the appropriate state aggregates, and then carried out controls to state age-sex totals and national age-sex-race-Hispanic origin totals.

Special Areas Estimates

The NCHS needed estimates for 3 substate areas in the same detail as the state estimates we were providing them. These areas were defined by NCHS and closely approximated but were not identical to the metropolitan areas of Los Angeles and San Francisco in California and Seattle in Washington. Fortunately, they were the aggregates of a selection of whole counties. To obtain these

estimates,, we proceeded much as we did for the states, but we did this not for just the counties in the estimate areas; but for all the counties in California and Washington. This was in order to be able to control our county numbers to the state totals we had obtained for the desired age, sex, race and Hispanic Origin detail.

PRODUCTS AND AVAILABILITY

The file containing all 10 years of intercensal county estimates is available on a reel of tape from our Population Information Staff in the Population Division, (301) 763-5002. We also distributed state extracts of the county files to member agencies of our Federal State Cooperative Program for Population Estimates. The state numbers are available on floppy disk also from our Population Information Staff.

PREPARING POPULATION ESTIMATES FOR THE POST-1990 PERIOD: THE COMBINATION OF DATA FROM VARIOUS SOURCES

Signe Wetrogan, U.S. Bureau of the Census

(Not available for publication)

Session T

Special Focus Surveys

Toward the year

2000

Refining the measures

COMPARING TWO METHODS OF SURVEYING SPARSE MINORITY POPULATIONS

Jay S. Buechner, Rhode Island Department of Health
Josefina Reynes, Jana E. Hesser

Introduction. In Healthy People 2000: National Health Promotion and Disease Prevention Objectives (1), Objective 22.5 in the priority area Surveillance and Data Systems states:

> "Implement in all States periodic analysis and publication of data needed to measure progress toward objectives for at least 10 of the priority areas of the national health objectives."

Sub-objective 22.5a poses the additional requirement of:

> "Periodic analysis and publication of State progress toward the national objectives for each racial or ethnic group that makes up at least 10 percent of the State population."

Data from the 1990 United States Census reveal that a majority of states had no minority groups meeting the 10 percent cutoff of Objective 22.5a. The 10 percent criterion excludes many states with numerically significant minority populations. For example, only 19 states and the District of Columbia have a Black population which comprises 10 percent or more of the total population. California, with a Black population of 2 million, the second largest black population in the U.S., is excluded. Likewise, only 8 states have a Hispanic population comprising 10 percent or more of the total population. States like Illinois with a Hispanic population of nearly one million are excluded. Only one state has an Asian and Pacific Islander population that exceeds the cutoff, and no state has a Native American population comprising 10 percent of its population.

Rhode Island is typical of most states by having a number of numerically significant minority populations which comprise less than 10 percent of the total population. No single minority group in Rhode Island is as large as 5 percent of the total population, and the White, non-Hispanic population comprises 89.3 percent of the total (Table 1). However, the size of the state's minority populations has grown rapidly in recent years, especially its Asian and Pacific Islander and Hispanic populations (Table 2). Census data indicate that the rates of increase for these two groups from 1980 to 1990 were higher in Rhode Island than in any other state (2). The dramatic increase in these groups is due primarily to the arrival of recent migrants from South East Asia and Latin and Middle America. The linguistic, cultural, social, and economic barriers these groups must overcome as they adjust to their new environment pose significant health risks for them, and limit their access to health care. Data on the health status, health risk behaviors, health care utilization, and access to health care for these high-risk minority groups is very limited. However, the need for such data is critical if Year 2000 Objectives are to be met for these and other minorities in Rhode Island and the nation.

Table 1. Percentage Distribution of Population by Race and Ethnicity, Rhode Island, 1990

Race and Ethnicity	Percentage
White, not Hispanic	89.3
White Hispanic*	4.0
Black	3.9
Native American	0.4
Asian and Pacific islander	1.8
Other, not Hispanic	0.6

*Includes Hispanic persons of "Other" races.

Table 2. Percentage Increase in Population, by Race and Ethnicity, Rhode Island, 1980-1990

Race and Ethnicity	Percentage Increase
White	2
Black	41
Native American	41
Asian and Pacific Islander	246
Other	90
Hispanic*	132
Not Hispanic	3

*Persons of Hispanic ethnicity may be of any race.

Collecting health related information on populations that comprise a small proportion of a total population poses certain problems, and these vary depending on the nature of the data source. Ascertainment and comparability are significant problems for sources such as vital records and hospital discharge data that provide complete coverage of events in a population. Such problems are especially acute when

two or more databases are to be used together in rate studies or data linkage projects. For these databases, the problems associated with small numbers of cases can be addressed most readily by increasing the time span included in analysis.

When data are collected on a sample basis, such as in health-related surveys, the small proportion of minority residents in a population translates into very small sample sizes and associated confidence intervals which are too wide to support meaningful analysis for these groups. The problems of small sample size can be overcome during a survey of the general population by oversampling minority residents in numbers large enough to allow separate analysis of the minority data. Another approach is to conduct independent surveys of minority groups. This paper discusses the relative merits of these two approaches for collecting survey samples of minority populations. Both have been tried in recent years in Rhode Island.

The Rhode Island Department of Health conducts two health surveys of samples of the general population on a regular basis. The Rhode Island Health Interview Survey (HIS) is performed every five years, and the Behavioral Risk Factor Survey (BRFS) is performed continuously on monthly samples. In 1990, the standard sampling methods used in the two surveys yielded proportions of minority respondents that were lower than the minority proportions in the Census, with no group having enough respondents for reliable analysis (Table 3).

Table 3. Respondents to General Health Surveys, by Race and Ethnicity, Rhode Island, 1990

Race and Ethnicity	HIS	BRFS
White, not Hispanic	91.3	92.5
White Hispanic	3.3	2.8
Black	3.1	2.7
All Other	1.9	1.6
Missing	0.4	0.4
Number of Respondents	6223	1805

HIS = Health Interview Survey; BRFS = Behavioral Risk Factor Survey

Two approaches were adopted to address this problem. For the 1990 HIS, the sampling method was altered so that two minority groups, Blacks and Hispanics, were oversampled relative to other residents. When analysis of the entire population was performed, the differential sampling was adjusted for by assigning different weights to respondents in the oversampled groups. In order to augment the BRFS, a one-time independent survey was conducted only with Black and Hispanic respondents. This survey, called the Minority Lifestyle Survey (MLS), collected information only on members of the two targeted minority groups living in census tracts with large numbers of minority residents.

The relative merits of these two survey methods and data were assessed by addressing the following questions:

(1) How representative are the samples of their underlying populations?

(2) How different is the information produced by the two methods for those questions asked on both surveys?

(3) What are the comparative costs and benefits of the two survey designs?

Methods. The Rhode Island Health Interview Survey is a general health interview survey. Households are sampled and information collected on all household members, including children, from one knowledgeable respondent. During the most recent iteration of the survey, covering November 1990 through April 1991 (designated the "1990" HIS), interviewing was performed by telephone under contract to a professional survey research company. Sampling was performed by random-digit dialing under a stratification system that targeted a minimum of 100 responding households in each of 21 sub-areas of the state. Oversampling of Black and Hispanic households was performed by random-digit dialing within telephone exchanges that had yielded relatively high proportions (7.5% or greater) of minority residents in previous surveys, and by screening for race and ethnicity on contact. The survey included measures of health status, health risk behavior, health care utilization, and access to health care. The survey collected information on a total of 2,588 households comprised of 6,536 persons, including 104 households with 313 persons obtained through the Black and Hispanic oversample.

The Minority Lifestyle Survey (MLS) was a one-time telephone survey conducted from August 1991 through February 1992. The MLS collected information from a sample of Black and Hispanic adults on health risk behaviors corresponding to those included on the BRFS. The survey was conducted under contract to a different survey research

organization than the one used for the HIS. The contractor purchased a sample of residential telephone numbers for the eleven census tracts in the state with the largest numbers of Black and/or Hispanic residents, all located in two urban areas in the state. One adult per household was interviewed, selected according to a randomization protocol and screened for Black or Hispanic race and/or ethnicity. The survey collected information on 840 Black and Hispanic adults.

Demographic items from the two surveys were compared to 1990 Census data for Rhode Island to determine how well each sample represented its underlying population. Comparisons were made separately for Black and Hispanic respondents, ages 18 and older. (Because of the manner in which summary Census data are aggregated, Black Hispanics were included in both groups for these comparisons.) The HIS sample was compared to statewide data from the Census, while the MLS sample was compared to Census data aggregated over the eleven census tracts covered by the survey. The following measures were compared: (1) mean age in years; (2) percentage male; (3) percentage married (not available by race and ethnicity at the census tract level); (4) percentage with any college education; (5) percentage employed; and (6) percentage in households with incomes below the federal poverty level.

The two survey samples were compared with each other using the demographic measures listed above (excluding percentage married), and eight health risk measures collected in both surveys. The risk measures included: (1) current cigarette smoking [smoked at least 100 cigarettes in lifetime and currently smoking]; (2) moderate or heavy alcohol consumption [averaged one or more drinks per day during the past two weeks]; (3) not using auto safety belts [use safety belts only sometimes, seldom or never); (4) obesity [for males, having a body-mass index of 27.8 or greater (25.8 under age 20); for females having a body-mass index of 27.3 (25.7 under age 20)]; (5) high blood pressure [diagnosed by a physician or nurse]; (6) high serum cholesterol [diagnosed by a physician or nurse]; (7) no mammogram [none in the past two years, for women ages 40 and older]; and (8) no Pap test [none in the past two years, for women ages 18 and older].

Statistical significance was determined using the standard normal distribution to test differences of proportions and Student's t to test differences of means.

Results. Some important differences between the two surveys should be noted. The numbers of Black and Hispanic residents included in the Health Interview Survey and the Minority Lifestyle Survey are similar, but the Health Interview Survey obtained information on multiple household members through a single respondent while the MLS obtained information only on the selected adult respondent (Table 4). The Health Interview Survey sample

Table 4. Number of Respondents to Health Surveys of Minority Populations, by Race and Ethnicitv. Rhode island, 1990-1992

Race and Ethnicitv	HIS HH	HIS Indiv	MLS Indiv
Black, not Hispanic	121	331	370
Black Hispanic	15	49	16
White Hispanic	96	330	454
Total	232	710	840

HIS = Health Interview Survey; MLS = Minority Lifestyle Survey; HH = household; Indiv = individual

was designed to be representative of the Black and Hispanic populations statewide, whereas the MLS sample was drawn from census tracts containing approximately one-third of the state's Black and Hispanic populations.

Representativeness of samples. Comparison of demographic characteristics of the HIS Black and Hispanic samples with the underlying Census populations, for those ages l.8 and older, suggests that the samples were biased in several respects (Table 5). The Black sample included a significantly smaller number of males and a larger number of married persons than were expected. The Hispanic sample included a smaller number of males, of persons with any college education, and of employed persons than were expected. The difference in the proportion of employed persons was particularly large.

Comparison of demographic characteristics of the MLS samples of Black and Hispanic adults with the underlying Census populations of the eleven selected census tracts shows patterns of bias that are similar between the two groups but different from the patterns seen in the HIS (Table 6). For both samples, the number of persons with any college education and of persons with incomes below the federal poverty level were significantly higher than expected. The proportion in poverty was particularly high for the Hispanic sample.

Table 5. Selected Demographic Characteristics of Respondents to the 1990 Health interview Survey and of the 1990 Census Population, by Race and Ethnicity, Ages 18 and Older, Rhode Island

Demographic Characteristic	Black Population Census	Black Population HIS	Hispanic Population Census	Hispanic Population HIS
Mean Age in Years	38.8	38.4	35.9	38.9
Percent Male	49.5	43.0'	49.5	38.4*
Percent Married	35.4	43.5*	45.2	42.3
Percent with Any College	34.5	29.9	24.2	17.6*
Percent Employed	64.0	61.1	69.1	51.8*
Percent in Poverty	25.7	20.9	24.2	25.1

HIS = Health Interview Survey
*Significantly Different ($p<0.05$) from Census Population

Table 6. Selected Demographic Characteristics of Respondents to the 1991-I 992 Minority Lifestyle Survey and of the 1990 Census Population, by Race and Ethnicity, Ages 18 and Older, Rhode Island (11 Selected Census Tracts)

Demographic Characteristic	Black Population Census	Black Population MLS	Hispanic Population Census	Hispanic Population MLS
Mean Age in Years	39.7	41 .o	35.6	36.1
Percent Male	46.1	42.3	48.3	44.3
Percent with Any College	27.7	34.2'	13.8	21.5*
Percent Employed	58.8	58.4	62:0	63.0
Percent in Poverty	21.9	27.9'	28.5	41.6'

MLS = Minority Lifestyle Survey
*Significantly Different ($p< 0.05$) from Census Population

Table 7. Selected Health Risk Factors of Respondents to the 1990 Health Interview Survey and the 1991-1992 Minority Lifestyle Survey, by Race and Ethnicity, Ages 18 and Older, Rhode Island

Risk Factor	Black Population HIS	Black Population MLS	Hispanic Population HIS	Hispanic Population MLS
Current Cigarette Smoker	27.7	26.9	25.0	19.5
Moderate or Heavy Drinker	10.3	16.5	9.6	17.6
Non-Use of Auto Safety Belt	51.9	55.5	55.7	56.8
Overweight	30.8	39.0	27.7	27.4
High Blood Pressure	23.3	28.6	16.4	20.7
High Cholesterol	10.3	19.0	10.0	27.5
No Mammogram in Past 2 Years*	9.7	19.0	13.4	27.2
No Pap Smear in Past 2 Years**	17.5	6.7	15.8	13.2

HIS = Health Interview Survey; MLS = Minority Lifestyle Survey
*Women Ages 40 and Older
**Women Ages 18 and Older

Comparability of samples with one another. When demographic characteristics of the Black samples from the two surveys are compared with each other, no significant differences are found. However, the two Hispanic samples differ substantially from one another in two areas -- percentage employed and percentage in poverty. It should be noted that for both measures the differences between the two samples are larger than the differences between the two underlying populations.

There were substantial differences between the two Black samples and between the two Hispanic samples when compared for the eight behavioral risk factors included in both surveys (Table 7). Black respondents to the MLS were more likely than Black HIS respondents to report six of the eight risk factors. These included drinking moderately or heavily, not using auto safety belts, being obese, having high blood pressure, having high cholesterol, and not having a mammogram in the past two years (women ages 40 and older). The prevalence of cigarette smoking in the two samples was similar. Black women in the HIS sample were more likely not to have had a Pap smear in the past two years.

Risk factor patterns in the two Hispanic samples also differed from one another. The MLS sample was more likely than the HIS sample to report four of the eight risk factors including drinking moderately or heavily, having high blood pressure, having high cholesterol, and not having a mammogram in the past two years (women ages 40 and older). Prevalence for three factors was similar in the two samples. These included safety belt non-use, obesity, and not having a Pap smear in the past two years. Hispanics in the HIS were found to have a higher rate of cigarette smoking than in the MLS.

Survey Costs.The costs of collecting data using the two methodologies differed considerably. The Health Interview Survey interviewed Black and Hispanic households that were identified in two ways -- 128 minority households with 397 members were identified as part of the base sample in which no screening on race and ethnicity was performed, and 104 households with 313 members were identified in the minority oversample. Based on the survey contractor's billed costs for the two components, the costs per interview (household) were $29.19 in the base sample and $46.74 in the minority oversample, for an average cost per minority interview of $37.06. The costs per interview incurred by the contractor for the Minority Lifestyle Survey, in which all interviews were performed under a methodology that included,screening on race and ethnicity, were $59.52, or over 60 percent higher than in the HIS. The mean time to complete an interview was similar in both surveys, and 7 out of 8 contacted households were eliminated by screening on race and ethnicity in the HIS oversampling process, compared to only 3 out of 4 in screening for inclusion in the MLS.

Discussion. The two surveys of Rhode Island's Black and Hispanic populations differ substantially in the details of their methodology. Differences included the underlying populations from which the survey samples were drawn, the method of sampling telephone numbers and identifying eligible minority respondents, and the interviewing procedures. Nevertheless, comparison of the costs and benefits of the two methodologies can provide guidance in the design of future surveys of sparse minority populations, in Rhode Island and elsewhere.

The advantages of the Health Interview Survey, in which minority households were oversampled as part of a general survey, are the following:

The base survey includes respondents of all racial and ethnic backgrounds, providing information on the White, non-Hispanic population for comparison.

The Black and Hispanic samples provide estimates for the underlying statewide populations when weighted to account for the differential sampling probabilities created by the oversampling process.

The use of a single knowledgeable respondent to provide information on the entire household allows the collection of information on children, allows analysis on a household basis as well as an individual basis, and provides a larger sample at lower cost'compared to individual interviews.

The costs per interview for the minority sample were lower than those found in the "dedicated" survey of Black and Hispanic residents.

The advantages of the Minority Lifestyle Survey, in which only Black and Hispanic residents were interviewed, are the following:

The collection of information only on the interviewed person facilitates the inclusion of questions on knowledge and attitudes and minimizes

the difficulties and potential inaccuracies involved in interviewing by "proxy."

Interviewing within a small geographic area is useful for targeting public health programmatic efforts towards high risk populations.

The total cost of a survey including only minority residents is lower than the total cost of a general survey that oversamples minority residents.

In summary, the different sampling strategies used with the two surveys produced samples which differed in selected demographic characteristics. Not surprisingly, different patterns of risk factor behavior also emerged. The inclusion of an oversampling protocol within a survey of a general population was found to be a very cost-effective approach to the collection of data on sparse sub-populations and should be considered whenever a general survey is scheduled. If no general survey is planned, or if a more targeted survey is necessary for program considerations, a dedicated survey of sparse sub-populations that excludes other population groups is likely to be the alternative with the lowest total cost.

References

(1) U.S. Department of Health and Human Services, Public Health Service. Healthy People 2000: National Health Promotion and Disease Prevention Objectives. DHHS Publication No. (PHS) 91-50212. Washington, DC: U.S. Government Printing Office, 1991.

(2) U.S. Bureau of the Census. Statistical Abstract of the United States: 1992. Washington, DC: U.S. Government Printing Office, 1992.

REDUCING SAMPLING COSTS: A MULTI-STAGE DESIGN USING REPLICATED QUOTA SAMPLING TO YIELD A SAMPLE OF HISPANIC AND AFRICAN AMERICAN HOUSEHOLDS IN LOS ANGELES

Laurie M. Anderson, Centers for Disease Control and Prevention
Naihua Duan, Judith F. Perlman, and David L. Wood

Introduction

In 1989 and 1990, measles transmission reached epidemic proportions in the United States. Disease rates were particularly high among preschool children in urban settings. In Los Angeles, the measles epidemic clustered in low-income Hispanic and African American neighborhoods. For an understanding of the population characteristics that might decrease a family's demand for immunizations in neighborhoods where measles attack rates were highest, three inner-city Los Angeles County Health districts became the geographic focus of a community survey. According to the 1990 Census, the Northeast and the East Los Angeles Public Health Districts (abbreviated below as ELA) are contiguous areas that encompass 36 census tracts and have a combined total population of 179,891. The South Central Los Angeles Health District (SCLA) is composed of 55 census tracts and has a population of 293,377.

The 1990 Census described the ELA area as 84% Hispanics, of which 4.6% were children aged one or two. The racial and ethnic composition of SCLA was fairly equally distributed between Hispanics and African Americans: Hispanics comprised 40% of the population, and African Americans comprised 37%. Among the African American populations in SCLA, 3.8% were children aged one or two, while 5.6% of the Hispanics were children in this age group.

Method of sample Selection

Since population-based public health surveys entail considerable sampling costs, we sought an efficient sampling design that would achieve a balance among precision, bias, and cost. A simple random sample of households within ELA and SCLA would have been prohibitively expensive if we were to achieve a meaningful level of precision. A multi-stage, cluster design provided a functional approach in terms of cost and practicality of field organization and operations. The multi-stage cluster design was combined with a replicated quota sampling design to obtain an economic sample that also provided potential to detect bias in the sample. We wanted to examine the issue of bias inherent in a quota sample; therefore each selected census block was approached three times, so that comparisons could be made between interviews that were easy to obtain and those that took several attempts. Equal size samples of 400 households were sought among three separate sampling frames: ELA Hispanics, SCLA Hispanics, and SCLA African Americans. Each census block was sampled with probability proportional to the predicted number of eligible households in the block. At the household level, all eligible children were counted and one eligible child was selected at random. The interview was conducted with the adult who was mainly responsible for the care of the selected child.

Stratification of Sampling Frame

The 1990 Census Standard Tape File 1B served as the sampling frame, restricted to Northeast and East Los Angeles (ELA) and South-Central Los Angeles (SCLA). The main demographic characteristics about those districts are given in Table 1.

In ELA, Hispanic children aged one or two and in SCLA, both African American and Hispanic children aged one or two were deemed eligible. In SCLA, 115 census blocks were excluded by the interviewers in the field, because of lack of safety and other reasons. The demographic characteristics for SCLA were changed by the exclusions about 13% of the eligible African American children and about 2% of the eligible Hispanic children were not included in the survey (Table 1).

Blocks were sampled from each of the two areas; then households within the sampled blocks were sampled. Examination of the block level data revealed that some blocks had only a few households, thus eligible children might not be available for interviewing. Furthermore, some blocks had low proportions of African Americans or Hispanics, so too much screening would be required to identify eligible children. Because of these considerations, we stratified the sampling frame, using the following variables:

(1) Block Size = Number of households in block * Eligible Ratio.

(2) Eligible Ratio = $\frac{\text{Eligible population in block}}{\text{Total population in block}}$

Block Size served as a crude estimate of the number of households in the eligible population. If Block Size was small, the interviewers would be unlikely to obtain an interview. Eligible Ratio determined the amount of screening required to obtain an interview. If the Eligible Ratio was low, the interviewers might need to screen too many households to obtain an interview. Using Block Size and Eligible Ratio, we stratified the sampling frame into three Primary Strata:

* Stratum l--Block Size of at least 20, and Eligible Ratio of at least 50%.
* Stratum 2--Block Size of at least 10, and Eligible Ratio of at least 50%.
* Stratum 3--All other blocks.

In Stratum 1 an interviewer would likely succeed in obtaining at least one interview each time a visit was made to one of these blocks. We included Stratum 1 for sampling in both ELA and SCLA. For ELA, Stratum 1 covered 93% of the eligible Hispanic children. We considered this coverage to be satisfactory, so we restricted our sample in ELA to Stratum 1. For SCLA, Stratum 1 covered only 83% of the eligible African American children and 77% of the eligible Hispanic children. 'This coverage was unsatisfactory; therefore, we supplemented Stratum 1 with Stratum 2. Stratum 2 was more difficult to work with. Since there were fewer eligible households on each block, an interviewer might screen all households on the block and not obtain an interview. We decided to cluster adjacent blocks in Stratum 2 and consider each cluster as a primary sampling unit (PSU). The 314 blocks in Stratum 2 in SCLA were grouped into 139 clusters, among which 103 consisted'of two blocks and 36 consisted of three blocks. By combining Strata 1 and 2 in SCLA, we achieved coverage for 92% of the eligible African American children and 86% of the eligible Hispanic eligible children. Finally, to obtain a balanced sample of African American and Hispanic interviews in SCLA where the ethnic composition was quite heterogeneous, we further stratified Primary Stratum 1 in SCLA into three substrata:

* Substratum i.1 (Mostly Hispanic)--the Hispanic population in the block was at least twice the African American population.
* substratum 1.2 (Mostly African American)--the African American population in the block was at least twice the Hispanic population.
* Substratum 1.3 (Mixed)--all other blocks in Primary Stratum 1.

In summary, five strata were defined for sampling: Stratum 1 in ELA, Substratum 1.1 in SCLA, Substratum 1.2 in SCLA, Substratum 1.3 in SCLA, and Stratum 2 in SCLA.

Field Procedures

Three trips were assigned for each PSU, and one interview was targeted for each trip. A trip could be made either weekday daytime, weekday evening, or weekend. Each of the three trips was pre-designated randomly to one of those three times, so that all three times were covered across the three trips. A randomly pre-designated starting point in the PSU was assigned for the first trip to the PSU to begin screening for eligible children for interview. Each household approached was listed sequentially on an enumeration form. The disposition for each household was recorded according to the following major categories: approached, accessed, successfully screened, screened to be eligible, and successfully interviewed. These dispositions were used later to estimate the weights used in the data analysis. On subsequent trips, the interviewer used the enumerations obtained earlier to approach the remaining households (excluding those found earlier to be ineligible, those who refused, and those already interviewed) to screen for eligible children for interview. The dispositions were recorded on the enumeration form. Each trip finished either when an interview was successfully obtained, or when the interviewers had exhausted all households in the PSU. In the former case, the interviewers concluded by estimating the total number of households in the PSU without approaching the remaining households.

Response Rates

A total of 1,206 households in our sample were approached, accessed, screened, found eligible, and interviewed (Table 2). Very few of the accessed households declined screening. Among those successfully screened, we found a substantial variation in the number of eligible households was found across the strata: the predominately Hispanic strata had much higher proportion of eligibles among screened households than did the other strata.

Conclusion

The replicated quota sampling design was successful in obtaining access to approximately 80% of the households approached. In terms of survey cost per interview, the quota approach was substantially less than the traditional approach in which the household to be interviewed are pre-designated and the interviewers must make repeated visits to the sampled PSU until they access the pre-designated household.

To overcome potential bias in a quota sample, we modified the usual quota sampling approach into a *replicated* quota sample. Then, on the basis of the enumeration and disposition history for each household, interviews could be classified into those that were easy to obtain versus those that were difficult to obtain. In analysis of the data, this information allowed us to compare these interviews to evaluate the presence of bias by creating a variable that represented the number of trips required to access the household. This variable did not affect our multivariant model with respect to the outcome of interest: the child's immunization status.

Table 1.
Demographic Characteristics of the East Los Angeles
and south Central LO8 Angeles Sampling Area6

District	East LA	SCLA	SCLA After Exclusions
Total Population	179,891	293,317	271,676
African American	2,798	107,169	92,071
Hispanic	151,691	117,163	114,876
Children aged 1 or 2	7,746	12,624	11,890
African American	114	4,046	3,537
Hispanic	6,942	6,565	6,441
Total number of household6	44,682	79,794	73,408
Total number of census blocks	913	2,027	1,912
Total number of census tracts	36	55	53

Table 2
Number of Households in Sample

District	8 tratum	#(app)	#(acc)	#(scr)	#(elig)	#(int)
ELA	1	5,277	3,956	3,920	530	387
SCLA	1 + 2	18,066	14,307	14,138	1,455	819
SCLA	1	16,867	13,299	13,138	1,372	772
SCLA	1.1	2,809	2,092	2,053	397	232
	1.2	8,918	7,147	7,067	592	315
	1.3	5,140	4,060	4,018	383	225
SCLA	2	1,199	1,008	1,000	83	47,
Total		23,343	18,263	18,058	1,985	1,206

ISSUES OF METHODOLOGY FOR AGRICULTURAL HEALTH AND HAZARD SURVEYS

Lorraine L. Cameron
National Institute for Occupational Safety and Health

(Not available for publication)

EVALUATION OF METHODS EMPLOYED IN THE ASSESSMENT OF HEALTH CARE NEEDS AND ACCESS TO CARE IN DADE COUNTY FLORIDA FOLLOWING HURRICANE ANDREW

Nancy D. Barker, Centers for Disease Control and Prevention
N.E. Stroup, G.M. Lopez, J.T. Massey

On August 24, 1992, Hurricane Andrew hit south Florida causing extensive damage to south Dade county. More than 28,000 houses, mobile homes and apartment buildings were destroyed, and approximately 107,000 additional dwellings sustained major damage (1). An estimated 180,000 persons were left homeless; total damages were estimated at more than $30 billion.

Immediately after the storm, the Florida Department of Health and Rehabilitative Services (HRS) and the Centers for Disease Control and Prevention (CDC) coconducted a rapid needs assessment in the areas most affected by the hurricane. Information regarding the availability of water, food, and medical care was provided to decision-makers within 24 hours of completion of the survey to guide emergency response. Two days later, Hurricane Andrew struck Louisiana, and similar assessments were conducted there. These assessments were the first in which survey methods were used to obtain information for immediate use on acute health-care needs among affected populations following sudden-impact disasters (2).

Two months later, the Dade County public health unit requested federal assistance to assess health-care needs and access to health care in all of Dade county. This assessment was funded by the Federal Emergency Management Agency (FEMA) and was used as a basis for planning health services in the recovery phase. On this occasion three weeks rather than 24 hours were available to complete the survey and provide results to decision-makers. The surveys conducted during both the acute phase and the recovery phase were based on the 'rapid epidemiologic assessment' or REA method - a 2-stage cluster survey developed from the Expanded Programme on Immunizations (EPI) survey methodology (3).

The characteristics of 3 surveys relevant to this disaster assessment are to be discussed: the original EPI method, the modifications used in the REA conducted in Florida during the acute post-hurricane phase, and additional modifications used in the recovery phase for the follow-up health-needs assessment conducted in Dade County 2 months after the hurricane.

As originally formulated, the EPI survey is a 2-stage cluster survey designed to assess immunization status in developing countries (3). This survey includes 210 households in 30 clusters of 7 each. The sample size was determined by estimating the number of households needed in a simple random survey to obtain estimates with 10% precision and 95% confidence. Ninety-six was the desired sample size. Assuming a design effect of two for the cluster survey, this number was then doubled. To maintain normal theory, the sample was divided into 30 clusters of 7 households each. In the first stage, clusters are selected with probability proportional to size, as determined by the census. The survey is self-weighted. In the second stage, a constant number of 7 households are selected from each village by moving from the center of the village in a random direction to selecting the next consecutive household. If no one is home, or someone refuses to participate, the next house is selected until a total of seven are included in the survey. Thus, replacements are selected for non-resaonders.

The REA conducted in Florida immediately following Hurricane Andrew was a modification of that EPI survey methodology (2). In the RNA, unlike the EPI surveys, the basis for the sampling frame was destroyed (i.e., many households were destroyed or severely damaged to render them uninhabitable). The survey was to be done in only the high-impact areas to assess acute effects 'to target emergency response. The sampling frame used for the REA was developed by placing a grid over the map of South Dade County. Thirty grids were randomly selected in the first stage of sampling. In the second stage, seven households were selected in each grid by moving from the center of the grid and in a random dire&ion, then selecting the next closest household. This REA also used replacements for non-responders. The design was not self-weighting since grids were selected randomly and a constant number of households were selected within each grid. Since estimates were not weighted, an assumption was made that each grid contained the same number of households.

Results of the Florida surveys conducted 3 and 10 days postimpact indicated that the number of households without running water, electricity, or telephone service decreased significantly over the 7 day period (2).

Two months after the hurricane, another study, the follow-up needs assessment was conducted to evaluate health care needs and access to care in all of Dade County during the immediate recovery phase (4). The objectives of this study were to determine the distribution of displaced persons in the county, to assess health needs, to assess access to health and social services and to assess evacuation behaviors.

In this follow-up study, staff from the health department divided the county into 6 zones according to the extent of hurricane damage to assess county-wide needs. Impact from the hurricane was greatest in the Cutler Ridge and Homestead areas in the southern part of the county. Damage from the storm was slightly less severe in the Coral Gables/ South Miami and Kendall areas in the central part of the county, and was least severe in the Hialeah and Miami/Miami Beach areas in the northern part of the county.

Because information was not needed as quickly as in the immediate REAs, the REA method used in the acute phase was enhanced. The basis for the sampling frame was still destroyed. A census of remaining households had not been conducted, and aerial photographs, which may have helped to evaluate the distribution of remaining households, were not available for the entire county. To sample with probability proportional to size, census blocks were selected as clusters according to the number of households listed in the 1990 census. Therefore, the assumption was made that the relative distribution of households in the census blocks in 1990 was the same as the relative distribution of households remaining after the storm. Unfortunately, information is not available to evaluate this assumption.

In the first stage, 30 clusters or census blocks in each of the six zones were selected. In the second stage, 10 households per block were selected. Using a random start, after selecting the first inhabited household, interviewers proceeded in a randomly-selected direction to select the next nine consecutive households. Ten rather than seven households were selected because replacements for non-responding households were not selected, and a 70% response rate was assumed. Replacing nonresponders may be appropriate for rapid surveys conducted in developing countries or immediately after disasters, but it was decided that 2 months after this disaster, many people were probably getting back to their normal daily routines (i.e., jobs, etc.), and would be less likely be found at home during the day when the

surveys were being done. Attempts were made to increase response rates by leaving call-back letters explaining the importance of the survey and making follow-up visits on weekends and a holiday.

This survey was conducted by approximately 100 trained interviewers over a 2-week period. The data were analyzed using a recently-developed module for version 6 of the Hpi-Info software package that weighs stratum-specific estimates and adjusts variances for cluster surveys (5). Preliminary findings were presented to Dade County health officials 1 week after data collection was completed. Although data were collected on various health care issues, three of the key findings are presented:

Overcrowding, which was defined as at least one new person living in the household since the storm, was greatest in the Homestead zone (38%) and decreased progressively with distance from the storm track (Table 1).

The proportion of households in which at least one person had symptoms of stress or anxiety, defined by trouble sleeping, crying, thoughts of suicide, or excessive irritability, was also highest in the Homestead zone (53%) and decreased progressively to 18% in the northernmost zones (Table 2). In the Homestead zone, 12% of households reported that at least one person had lost health insurance because of the hurricane, compared with 5% and 6% in the other zones (Table 3).

This information has been used to target health services more effectively, particularly in areas with a-high degree of dependence on public programs. Based in part on these findings, mental health services have been expanded, public clinics for primary care will be rebuilt in Homestead and Cutler Ridge, and access to public programs is being improved through community health teams that provide vaccinations, counseling, information on financial assistance, and health and social services (4).

Conducting REAs immediately following disasters in the acute phase raises both scientific and logistic difficulties. There is no easy way to choose a sample when the basis for the sampling frame has been destroyed. The problems faced by residents -- no water, electricity, or telephone service -- and impassible roads, are also faced by the survey team. In spite of these difficulties, decision-makers need information immediately to target emergency response. We believe that the teams dispatched to Florida and Louisiana did an extraordinary job conducting these surveys and providing critical information to decision-makers within hours. Nonetheless, we do have one recommendation that could improve the validity of estimates from REAs following disasters.

We recommend that aerial photographs be used to obtain weights based on the relative amount of remaining households in the grid when a SRS of grids are selected. These weights could be used to adjust the estimates. Unadjusted estimates could be used initially until aerial photographs of the remaining households are available. At that time, crude estimates could be adjusted.

The study team that conducted the follow-up needs assessment 2 months after Hurricane Andrew also faced both scientific and logistic difficulties. Makeshift offices were used, and few computers, photocopy, or fax machines were available. Cellular phones were used because regular phone service had not been installed at the survey site. Missing street signs made it difficult for survey teams to locate census blocks. However, on the basis of our experience, we recommend the following measures be instituted to improve needs assessment in the recovery period following a disaster:

First, if there is not enough time to conduct a census of remaining households in the cluster (either a census block or grid), we suggest implementing the following measures: select the clusters with probability proportional to size, using a census conducted before the disaster then choose a constant number of households in each cluster. Choose the initial household randomly, and then select subsequent households systematically, for instance, every 5th or 10th household.

The second alternative should be followed when adequate time is available for conducting a census of remaining households in selected clusters. We suggest the following: select clusters using simple random sampling. Conduct a census of-remaining households in the selected cluster. Select either a constant number of households or a constant proportion of households to be interviewed from each cluster, and select households systematically after a random start. Additional consultation with survey experts and more practical experience conducting surveys following disasters will help to further refine these suggestions. Regardless of the sampling method used, care should be taken to assure that multi-unit dwellings are not under-represented. In our study, apartments were underrepresented because interviewers were instructed to select only one household from multi-unit dwellings. We attempted to assess the impact on our findings by comparing estimates from single and multiple-unit dwellings. Fortunately, we found few differences for most of our key findings.

REAs provided information that was critical for targeting emergency response in the acute phase following Hurricane Andrew (2). They also provided information needed to direct rebuilding efforts and enhance the delivery of health and social services during the recovery phase following this disaster (4). Our experience suggests that rapid surveys can be done well, although procedures for conducting these studies could be refined further. Each situation should be evaluated for improvements whenever possible. If more systematic guidelines are developed, rapid survey teams will still be required to make scientific decisions quickly with uncertain implications under difficult and rapidly-changing circumstances.

References

1. Governor's Disaster Planning and Response Review Committee. Final report, Hurricane Andrew. Tallahassee, Florida: State of Florida, January 15, 1993.

2. Centers for Disease Control and Prevention. Rapid health needs assessment following Hurricane Andrew - Florida and Louisiana, 1992. MMWR 1992;41:685-8.

3. Henderson RH, Sundaresan T. Cluster sampling to assess immunization coverage: a review of experience with a simplified sampling method. Bulletin of the World Health Organization, 1982;60:253-60.

4. Centers for Disease Control and Prevention. Comprehensive assessment of health needs 2 months after Hurricane Andrew - Dade County, Florida, 1993, MMWR 1993;42:434-7.

5. Dean AD, Dean JA, Burton AH, Dicker RC. Epi-Info, version 5; a word processing, database, and statistics program for epidemiology on microcomputers. Atlanta: US department of Health and Human Services, Public Health Service, CDC, 1990.

Table1

Follow-up needs assessment

Households with new member(s) since storm

Zone	%	(95% CI)*
Hialeah	I7	(11-23)
Miami/Miami Reach	15	(10-20)
Coral Gables/South Miami	23	(17-28)
Kendall	23	(18-28)
Cutler Ridge	26	(20-32)
Homestead	38	(30-45)
Dade county	19	(17-22)

*Variances adjusted for complex survey design

Table 2

Follow-up needs assessment

Households with at least one person with indicators of stress or anxiety

Zone	4	(95% CI)*
Hialeah	18	(12-25)
Miami/Miami Reach	18	(14-22)
Coral Gables/South Miami	22	(16-27)
Kendall	39	(32-46)
Cutler Ridge	46	(39-53)
Homestead	53	(46-60)
Dade county	24	(21-26)

*Variances adjusted for complex survey design

Table 3

Follow-up needs assessment

Households with at least one person who lost health insurance because of storm

Zone	%	(95% CI)*
Hialeah	5	(2- 8)
Miami/Miami Reach	5	(3- 8)
Coral Gables/South Miami	5	(2- 8)
Kendall	6	(3- 9)
Cutler Ridge	6	(3- 9)
Homestead	12	(8-16)
Dade county	6	(4- 7)

*Variances adjusted for complex survey design

DUAL FRAME SAMPLE SURVEYS BASED ON NHIS AND STATE RDD SURVEYS

Monroe G. Sirken, National Center for Health Statistics
David A. Marker, Westat

Introduction

Establishing nationwide information systems capable of satisfying the needs for public health data at all administration levels is one of today's most pressing and challenging problems in managing the nation's public health programs. For instance, a recent memorandum establishing a CDC steering committee for public health information and surveillance system development begins as follows: "One of the key strategies identified in the Center's for Disease Control and Prevention's (CDC) strategic thinking for the 21st Century is the development of an integrated, comprehensive, nationwide public health information and surveillance system to gather, synthesize, and communicate information to monitor and promote health."

No doubt any nationwide system for providing both national and subnational health statistics will involve several data systems. Some like the vital registration system will be based on administrative record files and others such as the National Health Interview Survey (NHIS) will be based on sample surveys. The latter is the focus of this paper. In particular, the paper addresses the long standing problem of designing nationwide household sample surveys that are capable of meeting health statistics needs at both the national and State levels and to do so at reasonable overall survey costs.

For example, the staff at the National Center for Health Statistics (NCHS) has for years wrestled with the problem of redesigning the NHIS sample so that NHIS could serve effectively as a nationwide household sample survey. The sample design challenge is to convert existing national sample data systems that have limited capabilities for producing reliable subnational statistics, such as the NHIS, into nationwide data systems with capabilities of producing highly reliable national statistics and subnational statistics. Unfortunately, there is a conflict in optimum sample allocation in designing national and subnational household surveys. Sampling proportional to population size is optimum allocation for national surveys, but distributing the sample uniformly over the States is the optimum allocation for State survey designs.

This design dilemma coupled with the relatively high per unit data collection costs of face-to-face interviewing make it prohibitively expensive to meet both State and national statistics needs simply by increasing the NHIS overall sample size. This paper proposes a sample design for a nationwide household sample survey that appears to be far more efficient than either redesigning the NHIS or executing any other sample design option currently available or under consideration.

There are now essentially two types of independently designed household sample surveys running concurrently. Random digit dial surveys (RDDS), such as the Behavioral Risk Factor Sample Survey, are becoming the survey design of choice to collect household information for State and local health surveys while the NHIS continues to serve as the principal source for national health statistics. At present, RDDS and NHIS samples are not commingled. State estimates are based exclusively on RDDS samples and national estimates are based exclusively on the NHIS sample. This paper proposes a design for a single nationwide household sample survey system that would commingle the RDDS and NHIS samples and thereby improve the quality of both national and subnational statistics and reduce the combined costs of conducting both types of surveys. The proposed survey is called the dual frame sample survey (DFSS). The DFSS estimator is defined in the Appendix.

This paper evaluates the potential gains and possible losses in data quality and costs that would be realized if State estimates were based on DFSS instead of RDDS, and if national estimates were based on DFSS instead of NHIS. The methodology for carrying out this evaluation is described in the next section and the findings with respect to the DFSS effects on the quality of State and national statistics respectively are presented in subsequent sections. Since these findings clearly demonstrate that DFSS has the potential for substantially improving the data quality and decreasing survey costs, the concluding remarks propose that demonstration studies be conducted in collaboration with selected States to investigate the infrastructural issues of the proposed nationwide data system.

Methodology

Three alternative DFSS sample design options were evaluated. For each option, the annual national NHIS sample of about 58,000 households is supplemented by RDDS State samples but the options differ with respect to their algorisms for determining the sixes and distributions of the RDDS State sample supplements. The options are described in Table 1.

DFSS design option 1 is based on supplementary samples of 1000 RDDS completed interviews per State. The RDDS samples in Options 2A nd 2B vary in size from State to State. In Option 2A, each State RDDS sample size represents the minimal supplemental sample which when added to the NHIS State sample would assure that the variance of the resulting DFSS State estimate would be equal to or less than the variance of the RDDS estimate based solely on a 1000 RDDS completed interviews. Option 2B assures that the mean squared errors (MSE) of DFSS State estimates would not exceed those for RDDS estimates based solely on 1000 RDD interviews per State. The bias component of the MSEs are based on current

cigarette smoking rates as reported in the 1985 NHIS.

Three indexes are used in comparing the quality of national estimates based on DFSS and NHIS, and of State estimates based on DFSS and RDDS: sampling errors and two surrogate measures of bias, namely nonresponse rates and nontelephone household coverage rates. The MSE is also used. Bear in mind, however, that the MSEs shown in this paper only reflects those biases in RDDS estimates that are due to differences in the reporting of cigarette smoking rates between persons living in telephone and nontelephone households.

The error model for evaluating the relative quality of DFSS estimates is based on the NHIS and RDDS inputs as shown in Table 2.

Comparison of DFSS and RDDS State Estimates

Tables 3 summarizes the error model's findings with respect to the relative effects that DFSS sample design option 1 has on the quality of RDDS State survey statistics. The 50 States and the District of Columbia are listed in order by State population size which is virtually tantamount to listing the States by NHIS samples sixes since, as Table 1 notes, the NHIS sample is distributed approximately proportional to population size.

Compared to estimates based solely on 1000 completed RDD interviews per State, Table 3 shows the improvements in quality that could be realized if the NHIS State samples were supplemented by these RDD interviews. DFSS sample design option 1 improves the quality of every State's RDDS estimates, without exception, for each of the 4 quality indexes. On the average, State sampling errors and MSEs are improved 35% and 55% respectively, and response rates and telephone coverage rates respectively are improved by 6 and 8 percentage points.

There is considerable variation among the States in the amount of quality improvement in the RDDS estimates that are realized by DFSS design option 1. Relative improvements in sampling variances and MSEs are about proportional to the State population sixes and relative improvements in response rates almost likewise. Thus the gains in these indexes are greatest for the most populous States like California and Texas which have the largest NHIS State samples, and least for Vermont and Wyoming which the smallest NHIS State samples. Improvements in State telephone coverage rates are greatest for States like Alaska and Mississippi that have the lowest telephone coverage rates and least for States like Wisconsin and Connecticut that have the highest telephone coverage rates.

Table 4 summarizes the findings with respect to the quality effects that DFSS sample design options 2A and 2B have on RDDS State survey estimates based on 1000 completed RDD interviews per State. It will be recalled that DFSS design options 2A and 2B involve supplementing the NHIS State sample sixes by smaller RDDS samples than 1000 completed interviews per State assumed by option 1. On the average, these options are based on only 460 and 230 RDD supplementary State interviews respectively. Even so they improve response and coverage rates in every State, and improve the sampling variances in about a dozen of the most populous States, and the MSEs in more than half the States.

Here again there is considerable variation among the States in the quality effects of DFSS. The gains are greatest in the 12 most populous States, each having an annual NHIS sample in excess of 1000 interviews. However, none of the quality indexes of any of the State estimates would be adversely affected by DFSS options 2A or 2B, and the quality of every State's estimates would be improved for one or more indexes.

Comparison of National Estimates

DFSS has a mixed effect on the quality of NHIS national estimates. It reduces NHIS sampling errors but increases nonresponse rates. The telephone coverage rates are 100% whether based in NHIS or DFSS national estimates.

Using supplementary samples of 1000 RDD completed interviews per state, DFSS design option 1 reduces the NHIS sampling variance by about 44%, and increases the NHIS nonresponse rate by about 7 percentage points. Based on a supplementary sample of 11,732 RDD interviews, DFSS design option 2B, decreases the NHIS sampling variance by about 8% and increases the NHIS nonresponse rate by 3%.

Substantially greater improvements in the sampling variances and MSEs of the national survey estimates would have been realized, if the supplementary RDD samples in DFSS options 1 and 2 had been distributed proportional to State population sizes instead of the ways they were. If, for example, the 51,000 supplementary RDD interviews in design Option 1 had been distributed proportionally rather than uniformly, the NHIS sampling variances would have been reduced by 56% rather than 44 %. However, distributing the State samples in that manner, would have resulted, in smaller improvements in the variances of the State survey estimates.

Concluding Remarks

The research findings in this report demonstrate that substantial gains could be realized by designing a nationwide household sample survey health information system as a dual frame sample survey (DFSS) that merges the sample designs of NHIS and RDD State surveys (RDDS). Compared to the current usage of NHIS and RDDS, the proposed DFSS would substantially improve the quality of State and national statistics and simultaneously reduce the current combined costs of conducting both types of surveys independently.

Some design research, currently underway, remains to be completed. This includes investigating the effects DFSS on nonresponse and response biases. On the other hand, this paper has not addressed at all the effects of DFSS on the infrastructure of a nationwide information system. In view of the substantial potential DFSS gains, it is timely if not urgent that the infrastructural issues also be addressed. Establishing a nationwide health information system based on the sample design proposed in this paper is likely to substantially change the ways in which State and national household sample surveys are currently being collected and

dramatically change CDC's relationships with State and Federal Agencies. Demonstration studies should he undertaken in collaboration with selected States to explore the infrastructure issues.

Appendix: The DFSS Estimator

The DFSS estimate (est) is the weighted average of the NHIS est and RDDS est for the population living in households (hhs) with telephones, plus the NHIS est for the population in hhs without telephones.

Thus,

DFSS (est) = t x (weighted average of NHIS est and RDDS est for telephone hhs) +

(1-t) x (NHIS est for nom&phone hhs).

The DFSS est is unbiased if the fraction

t = proportion of hhs with telephones.

Table 1 - Description of DPSS Design Options

	Design Options		
	1	2	3
NHIS			
Total sample size	57,640	57,640	57,640
State sample distribution	Inverse to State population size		
RDDS			
Total supplementary sample size	51, 000	23,500	11,732
State sample distribution	1000/State	Inverse to NHIS State sample size	Inverse to NHIS State sample size & coverage rate

Table 2 - NHIS and RDDS Input8 to the Error Model for Evaluating the Quality of DFSS Estimates

Input	NHIS	RDDS
Sampling errors	1995 sample redesign*	Waksberg's sampling procedure
Response rate	95%	80%
Telephone coverage rate	100%	States rates**
Cigarette smoking rate***	30.1%	28.8%

* 1995 NHIS sample redesign feature8 some change8 that favor DFSS including stratification by State and met/nonmet area and increasing the number of psus from 200 to 400.

** Based on the 1980 Census.

*** Based on the 1985 NHIS, current cigarette smoking rates were 49.6% and 28.8% respectively for nontelephone and telephone households.

Table 3. Improvements in State RDD survey estimates due to dual frame estimation: 1,000 completed RDD interviews per State (Option 1).

Popult'r Rank	State	NHIS Completed Sample Size	Improvements: Reduced Variance (percent)	Improvements: Improved coverage Rate*	State Estimates: Improved Response Rate*	State Estimates: Reduced MSE (percent)
1	CALIFORNIA	9,394	85.8%	5	13	90.3%
2	TEXAS	5,613	78.6%	9	14	91.4%
3	NEW YORK	4.427	75.1%	7	8	87.2%
4'	FLORIDA	3,529	71.1%	IO	12	89.0%
5	ILLINOIS	2,577	64.7%	5,	9	75.7%
6	PENNSYLVANIA	2,016	61.7%	4	11	70.1%
7	OHIO	1.876	60.0%	6	IO	74.6%
8	MICHIGAN	1,761	57.5%	4	10	66.6%
9	NEW JERSEY	1,906	57.4%	5	7	68.8%
10	GEORGIA	1,618	55.5%	12	8	86.3%
11	NORTH CAROLINA	1.478	53.0%	11	8	83.9%
12	VIRGINIA	1,366	50.8%	8	10	76.4%
13	MASSACHUSETTS	1,091	46.0%	4	8	57.7%
14	INDIANA	952	43.4%	7	8	66.0%
15	MISSOURI	936	43.0%	5	9	60.6%
16	MARYLAND	1,149	45.7%	4	6	57.3%
17	TENNESSEE	970	43.2%	IO	8	77.1%
18	WASHINGTON	849	40.5%	6	6	59.8%
19	WISCONSIN	792	39.4%	3	8	47.4%
20	ARIZONA	1,041	41.6%	II	8	78.8%
21	LOUISIANA	1,021	42.6%	11	8	79.4%
22	MINNESOTA	686	31.2%	3	7	45.7%
23	ALABAMA	899	40.4%	13	8	82.7%
24	SOUTH CAROLINA	838	38.3%	13	6	81.2%
25	COLORAW	797	36.2%	6	4	59.2%
26	KENTUCKY	610	33.1%	12	6	77.5%
27	CONNECTICUT	672	33.4%	3	6	42.0%
28	OKLAHOMA	579	31.8%	8	5	63.3%
29	MISSISSIPPI	480	28.1%	17	5	83.5%
30	OREGON	654	32.4%	7	6	57.3%
31	IOWA	394	25.2%	4	5	35.1%
32	ARKANSAS	461	26.7%	13	5	75.0%
33	KANSAS	445	26.0%	5	5	43.1%
34	UTAH	330	21.2%	5	4	39.2%
35'	NEW MEXICO	587	26.8%	14	7	78.3%
36	WEST VIRGINIA	260	18.5%	II	4	61.6%
37	NEBRASKA	254	17.4%	4	3	23.8%
38	HAWAII	239	15.7%	5	4	27.2%
39	NEW HAMPSHIRE	198	14.8%	6	3	30.6%
40	NEVADA	271	16.3%	10	4	54.7%
41	MAINE	186	14.2%	7	2	38.7%
42	RHODE ISLAND	183	12.7%	5	3	19.0%
43	IDAHO	178	12.5%	7	3	32.5%
44	MONTANA	120	9.4%	8	2	25.8%
45	DELAWARE	148	10.0%	5	2	11.0%
46	SOUTH DAKOTA	105	8.5%	6	2	11.8%
47	ALASKA	112	8.5%	17	I	57.8%
48	DIST. OF COLUMBIA	205	13.3%	5	2	20.8%
49	NORTH DAKOTA	93	7.6%	4	1	5.8%
50	VERMONT	a7	7.1%	7	1	5.2%
51	WYOMING	84	6.3%	8	1	7.2%

* Improvements in coverage rate and response rate are shown in number of percentage points.

Table 4. Improvements in Stare RDD survey estimates due to dual frame estimation: Variable completed RDD interviews per State.

Populat' Rank	State	RDD Completed Sample Size	Improvements in State Estimates: Option 2A: mproved Response Rate*	Option 2A: Reduced Variance* (percent)	Improved Coverage Rue*	Option 1: Improved Response Rate*	Option 1: Reduced MSE (percent)	RDD Completed Sample Size
1	CALIFORNIA	0	15	83.4%	5	15	88.8%	0
2	TEXAS	0	15	72.7%	9	15	89.2%	0
3	NEW YORK	0	15	66.8%	7	15	83.1%	0
4	FLORIDA	0	15	39.4%	10	15	84.7%	0
5	ILLINOIS	0	15	45.4%	5	15	62.8%	0
6	PENNSYLVANIA	0	15	37.9%	4	15	52.1%	0
7	OHIO	0	15	33.2%	6	15	58.2%	0
8	MICHIGAN	0	15	26.1%	4	15	42.7%	0
9	NEW JERSEY	0	15	25.8%	5	15	46.4%	0
10	GEORGIA	0	15	19.9%	12	15	76.3%	0
11	NORTH CAROLINA	0	15	11.2%	11	15	71.0%	0
12	VIRGINIA	0	15	3.1%	8	15	55.4%	0
13	MASSACHUSETTS	148	12	0.0%	4	15	10.6%	0
14	INDIANA	234	11	0.0%	7	15	25.5%	0
15	MISSOURI	247	11	0.0%	5	15	11.9%	0
16	MARYLAND	157	12	0.0%	4	15	9.2%	0
17	. TENNESSEE	239	II	0.0%	10	15	50.9%	0
18	WASHINGTON	318	10	0.0%	6	15	5.1%	0
79	WISCONSIN	349	9	0.0%	3	11	0.0%	193
20	ARRONA	287	10	0.0%	II	15	53.3%	0
21	LOUISIANA	257	11	0.0%	11	15	55.5%	0
22	MINNESOTA	408	8	0.0%	3	10	0.0%	246
23	ALABAMA	321	10	0.0%	13	15	61.8%	0
24	SOUTH CAROLINA	378	9	0.0%	13	15	56.7%	0
25	COLORADO	432	8	0.0%	6	14	0.0%	29
26	KENTUCKY	492	7	0.0%	12	15	42.4%	0
27	CONNECTICUT	498	7	0.0%	3	8	0.0%	339
28	OKLAHOMA	534	6	0.0%	8	14	0.0%	15
29	MISSISSIPPI	610	5	0.0%	17	15	55.4%	0
30	OREGON	521	'7	0.0%	7	12	0.0%	100
31	IOWA	663	4	0.0%	4	5	0.0%	488
32	ARKANSAS	637	5	0.0%	13	15	25.1%	0
33	KANSAS	648	5	0.0%	5	7	0.0%	361
34	UTAH	731	4	0.0%	5	5	0.0%	433
3s	NEW MEXICO	635	5	0.0%	14	15	37.2%	0
36	WEST VIRGINIA	774	3	0.0%	11	9	0.0%	119
37	NERRASKA	790	3	0.0%	4	3	0.0%	665
38	HAWAII	814	2	0.0%	5	3	0.0%	607
39	NEW HAMPSHIRE	826	2	0.0%	6	3	0.0%	548
40	NEVADA	805	3	0.0%	10	7	0.0%	206
41	MAINE	835	2	0.0%	7	4	0.0%	419
42	RHODE ISLAND	855	2	0.0%	5	2	0.0%	717
43	IDAHO	857	2	0.0%	7	3	0.0%	497
44	MONTANA	896	1	0.0%	8	2	0.0%	554
45	DELAWARE	889	1	0.0%	5	2	0.0%	824
46	SOUTH DAKOTA	907	1	0.0%	6	1	0.0%	784
47	ALASKA	907	1	0.0%	17	7	0.0%	97
48	DIST. OF COLUMBIA	847	2	0.0%	5	2	0.0%	693
49	NORTH DAKOTA	918	1	0.0%	4	1	0.0%	1,083
50	VERMONT	923	1	0.0%	7	1	0.0%	891
51	WYOMING	933	1	0.0%	8	1	0.0%	824

* Improvements in coverage rate and response rate are shown in number of percentage points.

Third Plenary Session

Assessing Preventive and Primary Health Services

Toward the year

Refining the measures

ASSESSING PREVENTIVE AND PRIMARY HEALTH SERVICES

Kenneth E. Thorpe
Department of Health and Human Services

(Not available for publication)

ACCESS TO WHAT AND WHY?
TOWARD A NEW GENERATION OF ACCESS INDICATORS

Lu Ann Aday, The University of Texas School of Public Health

INTRODUCTION

Reflections on new generations bring to mind the following observations and expectations:

first, their identity is rooted and formed by their progenitors with whom they are linked both by biologic pedigree and cultural persuasion;

second, they are nonetheless uniquely shaped and influenced by the new and evolving era and experiences of which they are a part; and

third, they are likely to raise irreverent and provocative questions that challenge the current or contemporary ways of thinking and doing.

In contemplating a "new generation of access indicators", this paper explores (1) their origins and identities in earlier generations of indicators; (2) the unique properties and insights they add or extend to these prior incarnations; and (3) the important and provocative challenges they pose to the assumptions under which we have been comfortably operating in the design of access measures.

TO WHAT EXTENT DOES MEDICAL CARE CONTRIBUTE TO PEOPLE'S HEALTH?

The question that is implicit, if not explicit, in the norms underlying the design of a new generation of access indicators, is, "To what extent does medical care contribute to people's health?".

This is not a new question. Radical critics of the medical care system, such as Ivan Illich (1975) and Ronald Carlson (1975), writing in the mid-1970s, argued, in fact, that the limits of Western medicine in improving people's health had been reached, and that it now had the potential to do as much harm as good to those who came within its compass. During that same period, drawing upon patterns of mortality in Western nations over nearly a century, Thomas McKeown (1976) similarly argued that most reductions in death rates could be better attributed to improvements in nutrition and hygiene, than to medical care per se.

Canadian Marc Lalonde's (1975) field concept, that formed the conceptual basis for the U.S. Surgeon General's Healthy People report, in the mid-1970s, and was subsequently made manifest programmatically in the Year 1990 and Year 2000 Objectives for the Nation's Health, to which we particularly attend in this week's conference, gave further expression to these simultaneously irreverent, provocative, and important inquiries (Public Health Service, 1979; 1990).

These investigations have, in fact, highlighted the greater importance of lifestyle and environmental, relative to biologic factors or medical care, in ameliorating many of the major health problems that plague contemporary Western (including U.S.) society. Other social and medical care critics, such as Sylvia Tesh (1988), in her book, Hidden Arguments: Political Ideology and Disease Prevention Policy, and Dan Beauchamp (1988) in the U.S., Robert Evans and Gregory Stoddart (1990) of Canada, and the White and Black reports in England (Day and Klein, 1989; Gray, 1982), among others (Aday, 1993), have looked more deeply into the social-structural origins of the variations in health risks and related health outcomes associated with varying access to social and economic resources as a function of age, gender, race, or social class in U.S., as well as other, societies.

This population-oriented view of the determinants of people's health has been most widely heard and adopted in policy and practice in the public health field, rather than the medical care and associated health services research communities, in the United States.

The current emphasis on medical outcomes research has highlighted the question of the extent to which medical care is effective. It may be seen however, to largely fit within the context of a clinically-oriented view of the role of medical care in promoting individuals' health and well-being (Aday, et al., 1993).

Some of the most recent work on developing empirical indicators of access to personal medical care services, embodied in the deliberations of the Institute of Medicine (IOM) Committee on Monitoring Access to Personal Health Services similarly may be seen as fitting more directly into the clinically-oriented view of the role of personal health services in enhancing individuals' medical care outcomes -- not the broader public health-oriented perspective on the determinants of a population's health (Institute of Medicine, 1993).

Nonetheless, the IOM committee's work does, much more so than was the case in the past, seek to surface and illuminate clinical effectiveness norms in the formulation of recommended access to medical care indicators.

The discussion that follows explores the conceptual, normative, and empirical foundations for an earlier generation of access indicators, and how these indicators served both to originate and evolve into a new generation of measures of access to medical care.

HOW HAS ACCESS TO MEDICAL CARE BEEN CONCEPTUALIZED?

Over the past twenty-five years, there has been a considerable interest in the multiplicity of factors reported to be defining of access, such as the availability and organization of medical care providers and resources; the geographic, attitudinal, financial, and organizational barriers that individuals confront in seeking services; and the rates of actual utilization and satisfaction descriptive of the "outcomes" of the care-seeking experience.

The Behavioral Model of Health Services Utilization and associated Access Framework developed by Ronald Andersen, myself and others at The University of Chicago during the 1970s and 1980s, attempted to capture and examine the relative importance of an array of empirical predictors and indicators of access to medical care (Aday, et al., 1980).

Health Policy

Access has been most often considered in a political context. Improved access is an important goal of many health policy efforts. Health policy is then characterized as the starting point for the consideration of access, as it is the effect of health policy on altering access to medical care that policymakers at the federal, state, and local levels are concerned with affecting.

potential Access

Potential access may be described in terms of structural indicators, such as the characteristics of the delivery system, or in terms of the wants, needs, and resources that people themselves bring to the care-seeking process.

Characteristics of delivery system

The delivery system component Of the model refers-to-those arrangements for the potential rendering of care to consumers. It includes service availability (volume and distribution of services), organization (mechanisms for entry and movement through the system), and financing (the allocation and distribution of resources for paying for them). These characteristics are aggregate, structural properties. The community, or a particular delivery organization, is the unit of analysis, rather than the individual.

Characteristics of population at risk

The characteristics of the population at risk are- the predisposing, enabling, and need components of Ronald Andersen's original Behavioral Model of Health Services Utilization.

Predisposing variables include those that describe the propensity of individuals to use services -- including basic demographic characteristics (e.g., age, sex), social structural variables (e.g., race and ethnicity, education, employment status, and occupation), and beliefs (e.g., general beliefs and attitudes about the value of health services, and/or knowledge of disease).

Enabling variables include the means individuals have available to them for the use of services. Both financial resources (such as family income or insurance coverage), as well as organizational resources (such as having a regular source or place to go for care) specific to the individuals and their families are relevant here.

Need refers to health status or illness as a predictor of health service use. The need for care may be perceived by the individual and reflected in reported disability days or symptoms, for example, or evaluated by a medical provider, in terms of actual diagnoses or presenting complaints.

Realized Access

Realized access indicators refer to objective and subjective indicators of the actual process of care-seeking. These are, in effect, indicators of the extent to which the system and population characteristics predict whether or not or how much care is sought, and how satisfied potential or actual consumers are with the medical care system.

Utilization of medical care services

Utilization has been characterized in terms of the type, site, purpose, or time interval of use, that is by type of provider (physician, nurse, dentist, etc.); where care is sought (hospital inpatient, outpatient, emergency room; doctor's office; public health clinic, etc.); the reason for the visit (prevention-related, illness-related); and whether care was obtained or how many times in a given period of time (e.g., the percentage seeing a doctor and mean number of visits in the year).

Satisfaction with medical care services

Satisfaction may be evaluated by patients at a specific facility or in

terms of their particular experience of seeking care, or in general, by the public with respect to their perceptions of the performance of the medical care system as a whole. Both have been used as subjective assessments of the performance of the medical care system with respect to access.

Types of Access Research

This framework has been used in three principal types of research related to the access objective: descriptive, analytic, and evaluative research.

Descriptive research focuses on profiling access in terms of the potential and realized access indicators (such as the number and distribution of medical providers, percentage of the population with insurance coverage or a regular source of care, the proportion having seen a doctor or dentist in the year and the mean number of visits for those who did) and in documenting and examining subgroup variation according to these indicators.

Analytic research is more directed toward understanding why some groups have better access than others through examining the role of potential access measures (system and population characteristics) in predicting their actual rates of utilization and/or satisfaction with medical care.

Evaluative research is concerned with assessing how well policies, programs, and services that have been developed and implemented, based on previous descriptive and analytic research, have done in accomplishing the equity of access objective. These studies rely primarily on quasi-experimental evaluation designs to assess program or policy outcomes (Aday, et al., 1993).

HOW HAS EQUITY OF ACCESS TO MEDICAL CARE BEEN EVALUATED?

The access framework just reviewed has guided a great deal of the descriptive and analytic research surrounding equity of access to medical care. Underlying the application of that framework in assessing equity is the value judgment that the system would be deemed to be fair or equitable if need-based criteria, rather than resources (such as insurance coverage or income), are the main determinants of whether or not or how much care is sought.

subgroup disparities (by income, race or ethnicity, or insurance coverage, for example) in the use of services relative to underlying need would, therefore, be minimized in a fair and equitable system.

HOW HAS EQUITY OF ACCESS TO MEDICAL CARE BEEN MEASURED?

Access indicators developed within the context of this perspective on equity have focused on overall estimates and subgroup variations in the use of what were deemed to be necessary primary prevention-oriented, as well as illness-related, services: "the utilization of medical care services relative to need" (Aday, et al., 1980).

Primary prevention-oriented measures include indicators of the trimester in which prenatal care was sought or the number of visits during pregnancy, as well as having a general physical exam, pap smear, or breast exam in a given period of time.

Illness-related indicators that have been used directly in applying the Aday and Andersen framework to evaluating equity of access to medical care included the use-disability and symptoms-response ratios (Aday, et al., 1980).

The use-disability ratio summarized the number of total physician visits reported by subgroups of interest to the total number of days of limited activity due to illness they reported during the year. Widely varying rates of overall utilization relative to disability were deemed to be indicative of an inequitable distribution of services relative to experienced need.

Similarly, the symptoms-response ratio compared the proportion of people in a given age group who actually contacted a physician for a set of symptoms with the proportion a panel of physicians thought should see them, to assess the extent to which what were deemed needed services were actually obtained.

HOW MIGHT ACCESS TO MEDICAL CARE BE CONCEPTUALIZED?

The current and evolving concept of access to medical care builds upon, extends, and modifies the perspective provided by the Aday and Andersen access framework, and associated use-relative-to-need approach to measuring equity (Aday, et al., 1993).

In the original formulation, health status was viewed principally as a predictor of health services utilization. In the emerging conceptualization, it is, importantly and in addition, viewed as an implicit or explicit outcome of the medical care-seeking process.

Further, the choice of indicators of the use of medical care services in general or relative to some generalized conception of underlying "need" has also been tempered and focused in the context of the question of, "To what might

access make a difference?", that is, to what types of medical care does it make the most sense to enhance utilization, because those services are most likely to be effective in improving health?

In this reconceptualization of access, there is more of an explicit focus on the bottom line of whether the health and quality of life of individuals are enhanced by gaining access to specific medical care services. Further, the goal of equity of access to medical care becomes a means toward this end, not simply an end in itself. for which the efficacious outcomes of seeking and receiving care are theoretically assumed rather than empirically evaluated.

Further, subjective measures of realized access, such as patient satisfaction, remain important evaluative yardsticks in assessing the performance of the medical care system. However, this subjective dimension is more directly expressed in effectiveness-oriented considerations of patient preferences (or utilities) for different functional or longevity outcomes that are likely to be associated with alternative treatment procedures. There is an effort in both medical care outcomes research and applied clinical decision-making sciences to more explicitly invite patients to consider the risks and benefits associated with treatment options and to be full partners in choosing among these options.

Effectiveness, efficiency, and equity then all become guiding norms or prerequisites in ultimately ensuring the health and well-being of those to whom medical care services are directed, where,

effectiveness refers to the benefits of medical care measured by improvements in health;

efficiency relates these health improvements to the resources required to produce them; and

equity assesses whether the benefits and burdens of medical care are fairly distributed (Aday, et al., 1993).

HOW MIGHT EQUITY OF ACCESS TO MEDICAL CARE BE EVALUATED?

Contemporary considerations of equity of access to medical care explicitly embrace both the effectiveness and efficiency norms.

The definition of equity focuses on "the utilization of effective and appropriate medical care services to improve health outcomes", as well as the minimization of subgroup disparities in the extent to which this goal is achieved.

"Effective" care refers to services that are likely to result in health improvements. "Appropriate" care refers to those services that are most timely or fitting for achieving this objective. For example, penicillin may be an effective intervention to treat syphilis. Appropriate care to deal with congenital syphilis (that is syphilis in newborns) would, however, concentrate on the administration of penicillin to affected mothers prior to delivery to prevent its transmission in utero or during delivery, rather than after delivery to an affected newborn, for whom it could have been averted through appropriate prenatal care.

Efficiency research attempts to quantify the cost-effectiveness or cost-benefit of these.or other procedures in assessing the extent to which finite public, private, or personal resources should be invested in assuring access to those procedures or services.

HOW MIGHT EQUITY OF ACCESS TO MEDICAL CARE BE MEASURED?

Following is an overview of a number of the access indicators that were the focus of the Institute of Medicine Committee on Monitoring Access to Personal Health Care Services, summarized in the committee's final report, cc ēs to Health Care in America (Institute of Medicine, 1993).

That committee pointed out that the indicators selected to measure explicit access-related objectives should consider both (1) what health outcomes access was intended to affect, and (2) what types of utilization would be most appropriate to 'enhance those outcomes. Their aim was to specify a core set of indicators that could comprise a monitoring system for evaluating the state of the nation's performance with respect to access, much as the Year 2000 guidelines provide for monitoring the nation's health. A number of the access indicators recommended by the IOM committee on access do, in fact, mirror those used to measure. the Year 2000 Health Objectives as well (Public Health Service, 1990).

The committee considered what may be once again broadly characterized as primary prevention-oriented or illness-related access indicators and objectives. The major objectives and utilization and health outcome indicators for each are summarized here.

Primary Prevention-Oriented

A number of the more prevention-oriented utilization indicators are similar to those highlighted in earlier generations of access studies. However, in selecting both relevant utilization and outcomes-oriented measures, the IOM committee identified those preventive

services that were most likely to be effective in achieving specific health outcome objectives.

Objective 1. Promoting successful birth outcomes

Utilization indicators
Adequacy of prenatal care -- percentage of pregnant women obtaining adequate care, based on the Kessner index of trimester in which care first sought, number of visits and gestational age

Health outcomes
Infant mortality -- number of children who die before their first birthday per 1,000 live births

Low birthweight -- percentage of infants born weighing less than 2,500 grams

Congenital syphilis -- cases per 100,000 population

Objective 2. Reducina the incidence of vaccine-preventable childhood diseases

Utilization indicators
Immunization rates -- percentage of preschool children vaccinated

Health outcomes
Incidence of preventable childhood communicable diseases (diphtheria, measles, mumps, pertussis, polio, rubella, and tetanus) -- cases per 1,000 population

Objective 3. Early detection and diaonosis of treatable diseases

Utilization indicators
Breast and cervical cancer screening -- percentage of women undergoing selected procedures (clinical breast exam, mammogram, pap test) in a given period

Health outcomes
Incidence of late-stage breast and cervical cancers -- percentage of (breast, cervical cancer) tumors diagnosed at late stages

Illness-Related

Objective 4. Reducina the effects of chronic diseases and prolonging life

Utilization indicators
Chronic disease follow-up care -- average number of physician contacts annually by those in poor health; proportion with no physician contacts in the year

Use of high-cost discretionary care -- admissions for referral-sensitive surgeries (those that rely on the judgment of the physicians who provide first-contact care, who may or may not refer the patient for specialized procedures, such as hip/joint replacement, breast reconstruction after mastectomy, pacemaker insertion, coronary artery bypass surgery, or coronary angioplasty)

Health outcomes
Avoidable hospitalizations for chronic diseases -- admissions for ambulatory-care-sensitive chronic conditions (admissions that might not have occurred had the patient received effective, timely, and continuous ambulatory medical care for the condition, such as angina, asthma, grand mal status, chronic obstructive pulmonary disease, congestive heart failure, convulsions, diabetes, hypoglycemia, or hypertension)

Access-related excess mortality -- Number of deaths per 100,000 population estimated to be due to access problems (that is differences in death rates for diseases that can be managed by medical care between groups, such as blacks and whites, that remain after statistical adjustments for differing physiological and behavioral risk factors are made)

Objective 5. Reducina morbidity and pain through timely and appropriate treatment

Utilization indicators
Acute illness-related medical care -- percentage of individuals with acute illness who have no physician contact

Health outcomes
Avoidable hospitalizations for acute conditions -- admissions for ambulatory-sensitive acute conditions (such as bacterial pneumonia, cellulitis, dehydration as primary diagnosis, gastroenteritis, kidney/urinary infection, severe ear, nose, and throat infections, skin graft with cellulitis)

More detail is provided in the full Institute of Medicine report on methods of constructing these indicators and the most recent national data for the U.S. as a whole and age, sex, race, income and education subgroups for which information is available.

CONCLUSION

In summary, the emerging conceptualization of access explicitly views health status as a desired outcome as well as predictor of access to medical care. The new generation of access measures...

(1) have their progenitors in those indicators that assumed the use of services relative to need was desirable in improving health outcomes;

(2) have spun offspring that

examine, rather than presuppose, these assumptions; and

(3) persist in raising the important and provocative questions of, "To what should access be assured and [perhaps most importantly] why?"

REFERENCES

Aday, L. A.
1993. At Risk in America. The Health and Health Care Needs of Vulnerable Ponulations in the United States, San Francisco: Jossey-Bass Publishers, Inc.

Aday, L. A., R. M. Andersen, and G. V. Fleming
1980. Health Care in the U.S., Equitable for Whom? Beverly Hills: Sage Publications.

Aday, L. A., C. E. Begley, D. R. Lairson, and C. H. Slater
1993. Evaluating the Medical Care System: Effectiveness, Efficiency, and Equity. Ann Arbor: Health Administration Press.

Beauchamp, D. E.
1988. The Health of the Republic: Epidemiology as Chal[illegible]ne, and Mor[illegible]ism Philadelphia: [illegible]racy. Temple University Press.

Carlson, R.
1975. The End of Medicine,. New York: Wiley..

Day, P. and R. Klein
1989. "The politics of modernization: Britain's National Health Service in the 1980s." The Milbank Quarterly, 67, 1-34.

Evans, R. and G. Stoddart
1990. "Producing health, consuming health care." Social Science and Medicine, 31, 1347-1363.

Gray, A. M.
1982. "Inequalities in health: the Black Report. A summary and comment." International Journal of Health Services, 12, 349-380.

Illich, I.
1975. Medical Nemesis. London: M. Boyars.

Institute of Medicine
1993. Access to Health Care in America. Washington, D.C.: National Academy Press.

Lalonde, M.
1975. A New Perspective on the Health of Canadians. Ottawa: Information Canada.

McKeown, T.
1976. The Role of Medicine: Dream, Mirage, or Nemesis? London: Nuffield Provincial Hospitals Trust.

Public Health Service
1979. Healthy People: The Surgeon General's Report on Health Promotion and Disease Prevention (DHEW-PHS Publication No. 79-55071). Washington, D.C.: U.S. Government Printing Office.

Public Health Service
1990. Healthy People 2000: National Health Promotion and Disease Prevention Objectives: Full Report, with Commentary (DHHS Publication No. PHS 91-50212). Washington, D.C.: U.S. Government Printing Office.

Tesh, S.
1988. Hidden Arguments: Political Ideology and Disease Prevention Policy. New Brunswick, NJ: Rutgers University Press.

Session U

Cost-Benefit Analyses

Toward the year

2000

Refining the measures

A MODEL OF COST-EFFECTIVENESS AND INJURY PREVENTION

David P. Hamburger, Children's National Medical Center
Catherine S. Gotschall, Falah Shamsa, Martin R. Eichelberger

INTRODUCTION

Cost effectiveness analysis requires that the cost characteristics of both a disease state and the means to ameliorate the disease state are estimatable. The case of infectious disease preventable by vaccines such as congenital cytomegalic disease (29), *Haemophilus influenzae* (14, 15), Hepatitis B (1, 21), and Measles (5, 23) have been approached in this manner.
This approach has been recently applied to the field of injury control, where burn injuries have been analyzed with respect to the cost effectiveness of automatic sprinklers and smoke detectors (4). Budnick provided an estimate of the effectiveness of automatic sprinklers as an intervention for preventing loss-of-life in residential fires. Ruegg and Fuller presented an extended cost-benefit model for home fire sprinkler systems.

In these situations the cost of the injury without intervention is generally known or estimatable. If, however, an intervention is already in place to a variable degree, it becomes more difficult to partition the costs between those injured while using the intervention and those injured while not using the intervention. Complexity is further increased when there are multiple levels of interventions.

We present a model for the analysis of cost of injury data where multiple levels of intervention usage exist.

METHODS

We consider the situation where both the incidence of injury and the cost-of-injury data are presented without regard to intervention usage. We further consider interventions that may reduce both the number and severity of injury. Reductions in injury severity are measured by comparing the cost of caring for persons injured while using an intervention to that of persons injured while not using the intervention. We present a model for estimating cost-of-injury for a mechanism of injury with k existing levels of intervention.

We consider the following general case for any given type of injury:

LET I_0 = Overall incidence rate of injury, regardless of intervention
I_i = Incidence rate of injury with intervention i,

SUCH AS, $I_i = \dfrac{\text{Total \# injuries to persons using intervention } i}{\text{Total population at risk using intervention } i}$;

where $i = 0, 1, \ldots, k$ interventions and $i = 0$ represents no intervention.

ALSO N = Total population at risk for injury, regardless of intervention;

AND RR_i = Relative Risk of injury using intervention i, such that

$$RR_i = \frac{\text{\# injured using intervention } i \;/\; \text{\# at risk using intervention } i}{\text{\# injured with no intervention } (i = 0) \;/\; \text{\# at risk with no intervention } (i = 0)},$$

THEREFORE, $RR_i = I_i / I_0$; for all $i = 0, 1, \ldots, k$. (1)

CONSIDER
E_i = Effectiveness of intervention i as defined by Partyka (28), and substituting intervention for restraint, such that:

$$E_i = 1 - \frac{\text{(injury rate, with intervention)}}{\text{(injury rate, without intervention)}} \quad \text{for all } i = 1, 2, \ldots, k.$$

THEREFORE, E_i and RR_i are related as follows: $RR_i = 1 - E_i$. (2)

ALSO P_i = Probability of person using intervention i; for all $i = 0, 1, \ldots, k$ and $i = 0$ represents no intervention;

WHERE $\sum_{i=0}^{k} P_i = 1$.

CONSIDER $I_G = \sum_{i=0}^{k} I_i P_i$. OR $I_G = I_0 P_0 + \sum_{i=1}^{k} I_i P_i$. (3)

SUBSTITUTING EQUATION (1) IN EQUATION (3) ABOVE, $I_G = I_0 P_0 + \sum_{i=1}^{k} I_0 RR_i P_i$,

OR $I_G = I_0 \left(P_0 + \sum_{i=1}^{k} RR_i P_i \right)$. (4)

LET T_i = Total # injuries to persons using intervention i for population of size N; for all $i = 0, 1, \ldots, k$;

SUCH THAT $T_i = I_0 P_i N$. (5)

LET TC = Total cost for injury under consideration, for population N,

AND TC_i = Total cost of injury for persons using intervention *i*; for all $i = 0,1,\ldots,k$.

LET $\overline{C}_i$ = Mean cost per injury for persons using intervention i, for $i = 1,2,\ldots,k$,

C_0 = *Mean* cost per injury for persons using $i = 0$, which represents no intervention;

SUCH TNAT $\overline{C}_i = TC_i / T_i$. **(6)**

CONSIDER. *Mi* = cost reduction factor, i.e. cost of injuries incurred by persons using intervention *i vs.* cost of injuries incurred by persons not using intervention: for $i = 0,1,\ldots,k$ and where i = 0 represents no intervention;

SUCH **THAT** $M_i = \overline{C}_i / \overline{C}_0$ for all $i = 0,1,\ldots,k$. **(7)**

THEREFORE, FROM EQUATION **(6),**

$$TC = \sum_{i=0}^{k} TC_i = \sum_{i=0}^{k} T_i \overline{C}_i ;$$

AND SUBSTITUTING FOR T, PROM *EQUATION* (5),

$$TC = \sum_{i=0}^{k} C_i \overline{I}_i P_i N ,$$

OR $$TC / N = \sum_{i=0}^{k} \overline{C}_i I_i P_i . \quad (8)$$

SUBSTITUTING EQUATION (1) INTO EQUATION (8) ABOVE,

$$\overline{TC} = \sum_{i=0}^{k} \overline{C}_i I_i RR_i P_i ; \quad (9)$$

WHERE $\overline{TC}$ = TC / N which is the mean total cost of injury.

SUBSTITUTING EQUATION (7) INTO EQUATION (9) ABOVE,

$$\overline{TC} = \sum_{i=0}^{k} \overline{C}_0 M_i I_0 RR_i P_i ,$$

OR $$\overline{TC} = I_i \overline{C}_0 \sum_{i=0}^{k} M_i RR_i P_i .$$

FURTHER,

$$\overline{TC} = I_i \overline{C}_0 (M_0 RR_i P_0 + \sum_{i=1}^{k} M_i RR_i P_i).$$

THEREFORE:

$$\overline{TC} = I_i \overline{C}_0 (P_0 + \sum_{i=1}^{k} M_i RR_i P_i); \quad (10)$$

WHERE $M_0 = 1$ and $RR_i = 1$ by definition.

Illustrating the Model

The cost of motor vehicle injury (MVI) in children is substantial. It has been estimated that the cost of MVI for children less than five years of age was over 1 billion dollars in 1985 (31). Private health insurance covers the medical costs for some children. Medicaid provides reimbursement for health care costs for impoverished children injured in motor vehicle crashes. Medical costs avoided by reducing the frequency and severity of MVI are savings to Medicaid. To the extent that cost is a barrier to using safety seats among economically disadvantaged families, by providing safety seats as a Medicaid benefit, the prevented and reduced medical costs of MVI in excess of the cost of this intervention are potential net savings for Medicaid.

"The direct cost [of injury] is the actual dollar expenditure related to illness. The indirect cost represents the value of lost output due to reduced productivity caused by illness and disability and losses due to premature death" (31). The total economic cost of injury is the sum of the direct and indirect costs. Injuries to the very young typically have increased indirect costs as the value of the lost output accrues over a longer portion of a person's lifespan.

We posit that it would be cost-effective for Medicaid to provide safety seats to its infant beneficiaries, i.e., that the cost of providing safety seats is less than the cost of MVI prevented and reduced by their use. We use the model above to analyze this position.

Rice and associates give MVI incidence data by age-specific groups. Cost of injury data are given both as aggregate data and as a cost per injured person. These data are not given with regard to the status of the injured as concerns restraint usage during the crash event. We considered the case for two types of intervention (i.e., $k = 2$): safety seat (SS) (i=1) and safety belt (SB) (i=2). It is necessary to disaggregate these data by restraint use status, i.e., unrestrained (UR) (i=0), safety seat, or safety belt. We can then reconstruct the population for any combination of restraint use and predict the aggregate cost of injury under the new conditions.

We modelled a hypothetical cohort of 100,000 Medicaid eligible children all born on the same day and provided with safety seats by Medicaid prior to discharge from the hospital. We estimated the annual incidence of MVI and cost of injury for the cohort using our model for each of the first four years of life. Four years was chosen to represent the 'lifespan" of a safety seat.

The size of the cohort was not adjusted for mortality over time. The correction would be approximately 1% in the first year and less than 0.1% in subsequent years (16). considerably less than the uncertainty inherent in the data elements.

AS the cohort is aged, we varied the proportion of children unrestrained, in safety seats and safety belts to approximate the distribution for 0-l, 1-2, 2-3, and 3-4 year olds for 1985 (Table 1). We then summed the costs of injury over four years and compared this to the cost of the intervention to determine cost effectiveness. Discounting and inflation of medical costs were ignored for the four year summation.

INTERVENTION	PROBABILITIES			
	YEAR 1	YEAR 2	YEAR 3	YEAR 4
UNRESTRAINED	.44	.45	.50	.70
SAFETY SEAT	.55	.45	.30	.10
SEAT BELT	.01	.10	.20	.20

TABLE 1: Patterns of restraint use for the first 4 years of life. Partially misused safety seats are combined with correctly used safety seats. Children in grossly misused safety seats are considered unrestrained.

Data Elements

To illustrate the model we have used rough estimates from the literature for many of the data elements. Moderate values corresponding to a plausible intermediate estimate were chosen. Values were also chosen to demonstrate plausible extreme boundaries of cost-effectiveness. A high estimate was chosen to illustrate a putative maximal degree of a cost-effectiveness and a low estimate showing a putative minimally cost-effective situation. The values selected for the intermediate, high, and low estimates are shown in Table 2.

DATA ELEMENT	INTERMEDIATE	HIGH	LOW
INCIDENCE OF MVI (I_o)*	623	989	257
TOTAL DIRECT COST of INJURY*	$1,164,387	$1,848,441	$480,333
GRAND TOTAL COST OF INJURY*	$5,583,949	$8,864,407	$2,303,491
RELATIVE RISK (RR-SS)	0.4	0.15	0.65
RELATIVE RISK (RR-SB)	0.5	0.3	0.7
COST FACTOR (M-SS)	0.5	0.33	0.99
COST FACTOR (M-SB)	0.5	0.33	0.99
COST of INTERVENTION	$3,500,000	$3,150,000	$3,850,000

* These figures are calculated per 100,000 population.

TABLE 2: Data elements used in illustrating the cost-effectiveness of Medicaid providing safety seats at birth for a cohort of 100,000 infants for three plausible cases.

Incidence of MVI (I_o)

Rice and associates cite the incidence of MVI per 100,000 children O-4 years of age as 622.7, with a standard error of "greater than 30%'. No upper range of error is given. Assuming incidence to be normally distributed, we estimated a 95% C.I. (257, 989) incidence using a standard error of 30% for our high and low bounds. Given the large standard error, incidence figures were rounded to the nearest whole number.

cost Of Injury

In 1985 direct costs of MVI in children 0 to 4 years of age were $1869 per injured person and the total costs (direct plus indirect) for MVI were $8963 per injured person (31). These values were multiplied by the estimated incidence of MVI for the intermediate, high, and low estimates to yield a total direct cost of injury and a grand total cost of injury for the respective estimates.

Relative Risk (RR,, RR_{SB})

A wide range for effectiveness of injury reduction for safety seats and safety belt is available for a variety of crash scenarios that involve fatalities only, towed cars, patients seekinghospitalcare, etc. Effectiveness is given variably for 'restraint use' and specifically for lap belts, shoulder-lap combinations, and car safety seats. Several authors (6, 7, 11, 18, 19, 39) report values as low as 25-35% for restraint effectiveness; while values as high as 60-75% are also reported (6, 11, 27, 28), a middle range of 40-55% is also in the literature (6, 7, 11, 12, 30).

Using our model, we have converted restraint effectiveness to relative risk of injury using Equation 2 and have chosen RR_{SS} = 0.4 and RR, = 0.5 for the intermediate estimate of restraint effectiveness since we are applying them across the wide spectrum of crash severities and the cost-of-injury data are given per injury. We used RR,, = 0.15 and RR,, = 0.3 for the high estimate and RR,, = 0.65 and RR,, = 0.7 for the low estimate.

cost Factor (M_{SS}, M_{SB})

Reath and Orsay have both shown that the medical cost for adults unrestrained by safety belts in MVI are 2-to-3 times higher than the costs for those restrained by safety belts. Kaplan & Cowley also found costs doubled for injured, unbelted adults. Nelson has found factors as high as 3.86, 95% C.I. (2.28,6.53) for males age 6-14 injured at less than 55 mph in Iowa. The values for M_{SS} and M_{SB} in Table 2 for the intermediate and high estimate reflect this range.

M_{SS} and M_{SB} for the low estimate were chosen as 0.99. This reflects the hypothetical estimate where the medical costs for injured restrained passengers approach the costs for injured unrestrained passengers. There is no evidence that the medical costs for restrained passengers ever exceed the costs 'for unrestrained passengers. M_{SS} and M_{SB} were chosen not equal to 1.00 to facilitate calculation.

Cost Of Intervention

We used a cost of $31.50 per safety seat ($3,150,000 for the cohort) based on telephone queries to manufacturers of safety seats. Manufactures were asked for the best unit price available to a large purchaser of safety seats for a seat suitable from birth to four years of age. It was envisioned that Medicaid would distribute sole-use vouchers redeemable for the purchase of safety seats. Vouchers would be redeemed by mail directly from the manufacturer. A similar program has been used by the League General Insurance Company of Southfield, Michigan (34, 35) Approximately 10% was added as the combined cost of administering the program and shipping and handling (34). Thus, a total cost of $3,500,000 was figured for a cohort of 100,000. The cost of intervention was varied by ± 10% for the high and low estimates to see the impact of changing key model parameters (5, 14, 15). This is shown in Table 2.

A break-even safety seat cost was calculated for the intermediate, high, and low estimates for both total and direct costs to illustrate the sensitivity of the model to the cost of safety seats. Break-even was defined as the cost per safety seat where the cost of intervention equalled the anticipated savings for any safety seat and safety belt usage while the sum of the usage rates was 100%. Cost-effectiveness increases proportionally as safety seat cost falls below break-even cost.

Hypothetical Probabilities of Restraint Usage (P_o, P_{SS}, P_{SB})

Parents may place children in safety seats correctly or incorrectly. Incorrect use may be partial or gross. Kahane presents data on safety seat usage from a variety of sources. The age-specific data were corrected for gross misuse of safety seats as noted in 1984 and the correctly used safety seats were combined with partially misused safety seats to yield P_{SS} as shown in Table 1. P_{SB}, also shown in Table 1, was obtained from 1983-84 North Carolina MVI data as presented by Kahane. Numbers were rounded to the nearest 5% except in the estimate of P_{SB} for year 1. Passengers in grossly misused safety seats and passengers not in safety seats or safety belts were considered unrestrained as reflected in P_o in Table 1.

RESULTS

Our results indicate that it is cost-effective for Medicaid to provide safety seats to infants for certain sets of conditions that may reasonably occur. These result are highly sensitive to which aspect of cost is being considered.

Intermediate Estimate

For the intermediate estimate when only direct costs (Fig. 1) are considered, our results indicate that it would not be cost effective at any level of restraint use for Medicaid to provide safety seats to its beneficiaries. The break-even safety seat cost was $17.58 where 100% of child passengers from 0-4 years of age were restrained in either safety seats or safety belts in any combination. At this level of restraint use, Medicaid would lose $1,566,184 by providing safety seats to the cohort.

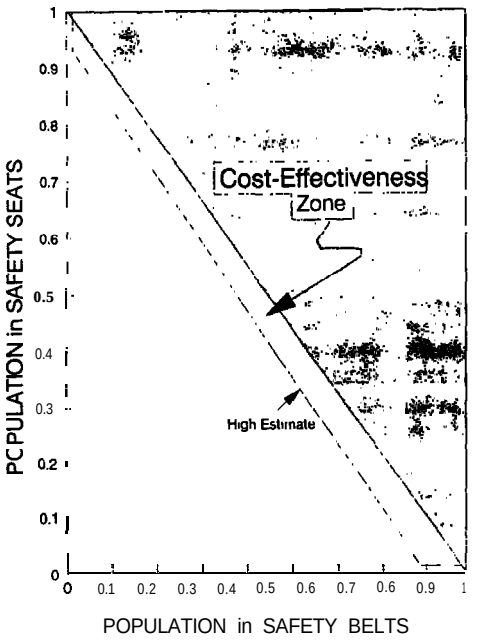


FIGURE 1: Cost-effective patterns of restraint usage when Direct Costs are considered. Medicaid purchase of safety seats is cost-effective only at high levels of restraint usage, high incidence of MVI, high cost of injury, and low costs of safety seats.

When total costs (Fig. 2) are considered, if child passengers are restrained in any combination of restraint use such that the sum of the safety seat and safety belt usage rates equals 75%. the intervention is cost effective. In fact, the break-even cost of a safety seat could be as high as $84.31 for the intervention to be cost effective if 100% of child passengers are restrained. A profit of $5,773,831 which is a return of $1.65 for each $1 invested, can be realized at 100% restraint use.

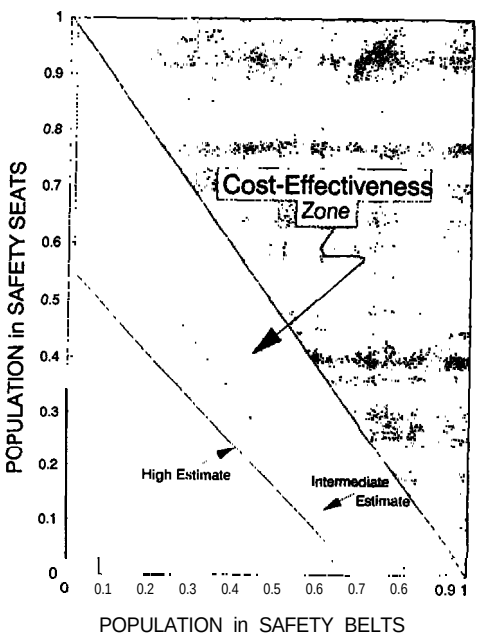


FIGURE 2: Cost-effective patterns of restraint use when Total Costs - Direct Costs plus Indirect Costs - are considered. Medicaid purchase of safety seats is cost-effective at moderate levels of restraint usage, moderate incident of MVI, moderate cost of injury, and median costs of safety seats.

Nigh Estimate

For the high estimate and direct costs (Fig. 1), 94% of child passengers must be restrained, in any combination, for Medicaid to cost effectively provide safety seats to its patients. If 100% of children are restrained in any manner, a $1,523,995 profit results and a safety seat cost of $42.49 or less makes the intervention cost-effective.

When total costs (Fig. 2) are considered, the range of restraint use for cost effectiveness is larger. In this situation any combination of restraint usage ≥ 11% makes the intervention cost effective. If restraint usage approaches 100%, the break-even cost of a safety seat is as high as $203.77 and profit approaches $19,264,670.

Low Estimate

For the low estimate there is no situation of restraint usage that makes the proposed intervention cost effective. Even if 100% of children were restrained, a safety seat would need to cost $0.11 to be cost-effective considering direct costs. Total cost considerations raise the break-even cost to $0.52. Losses vary from $3,838,077 for direct costs to $3,792,822 for total costs.

DISCUSSION

Costs of W I

Motor vehicle injury is the leading cause of death in childhood in the United States. The number of children injured but surviving is many times greater still (2). Restraint systems such as safety belts and safety seats reduce both the frequency and the severity of injury in a crash event (6, 30).

Survey data suggest that safety seat usage has increased in recent years (26). However, barriers remain to the routine use of safety seat by well-intentioned parents. These barriers include misperception of risk (10), the mistaken belief that infants are safe when held in a parent's arm (24) inconvenience, cost, and perceptions of child misbehavior in safety seats (20). Furthermore, it has been suggested that economically disadvantaged populations are less likely to use safety seats (32, 36). One likely reason for this is cost.

Medicaid is responsible for the direct costs of MVI for its recipients. If medical costs averted by the use of safety seats were greater or equal to the cost of providing them to the Medicaid population, the choice would be self evident. This is not the case under current restraint usage rates - however, other considerations may apply.

The total costs of MVI are, by definition, the economic cost of MVI to society taken as a whole. If total costs averted by the use of safety seats are greater or equal to the cost of providing them to the Medicaid population which may not be able to afford their purchase, it makes economic sense for a governmental program to invest in safety seats so that society may reap the anticipated savings. Under some of our assumptions this may be the case. Under conditions of increased incidence of MVI, increased cost of injury, or decreased cost of a safety seat, or any sufficient combination of these, an investment in safety seats makes sense on an economic basis alone.

Although Rice and associates account for rehabilitation costs, they are generated from total national 'vocational' rehabilitation costs. This would seem to give little weight to educational rehabilitation. Public school systems are currently tasked with providing an education for all children including those whose developmental or physical disabilities require extensive services. The economic cost of those services necessitated by MVI in Medicaid eligible children can be partially averted by safety seat usage. Society, though, considers more than economics.

Society places a value on avoiding pain and suffering and risk aversion in general. These are not considered in either the direct or total costs noted above. Rice and associates estimate that society is willing to pay considerably more than economic costs alone to avert injury. When economics are considered along with compassion and equity, having Medicaid provide safety seats becomes more acceptable.

Assumptions

Our assumption that crash risk is independent of restraint use tends to underestimate the costs attributable to the unrestrained occupant and, thereby, negatively bias estimates of cost effectiveness. The crash rate for unrestrained drivers is higher than that for restrained drivers (8, 17) and drivers who use restraints themselves are more likely to restrain their child passengers (20). If previously unrestrained drivers increase their restraint use concomitantly with using restraints for their child passengers and if their own restraint use translates to a decreased crash rate, the savings of the intervention is magnified.

Our model was illustrated with a scenario that took a cohort with a pattern of safety seat and safety belt use similar to the pattern of use for each of the first 4 years of life as in 1985 as illustrated in Table 1. We calculated costs and savings for different hypothetical levels of restraint use. These new levels of restraint use were held constant for each of the first 4 years of life. For example, costs and savings were calculated presuming that safety seat use was raised to 60% and safety belt use was raised to 30% and these rates were maintained during each of the four years. This was done for ease of calculation and analysis. Our model is applicable for any combination of restraint use in any year.

Assumptions of restraint effectiveness are somewhat dependent on the circumstances of crash events. More severe crashes are more likely to result in injuries unpreventable by any known intervention. Restraint effectiveness is greatest in those crashes of magnitude such that a severe injury is reduced to a moderate or mild injury or entirely averted with restraint use. In milder crashes, where events would suggest that the likelihood of severe injury is low regardless of restraint usage, restraint effectiveness might appear to be greater than it actually is. Given, however, that direct costs are driven by the high numbers of minor injuries, *vide infra,* restraints of modest effectiveness which prevent minor injuries entirely may greatly reduce the *cost* of MVI in this age group.

In the age group O-4, Rice and associates reported 102,000 children representing 91% of MVI (age O-4) suffering injury resulting in neither fatality nor hospitalization. This figure had a standard error exceeding 30%. This uncertainty resulted in a similar uncertainty for the overall incidence figures for MVI in that age group. Presuming that nonhospitalized patients had less serious injuries suggested that a intermediate case using modest figures for restraint effectiveness was appropriately conservative.

Implications of the Model

Although we have illustrated the model using MVI, it is clearly applicable to other situations. Any disease state with known or estimatable incidence and cost data and any proposed number of interventions characterized by a known or estimatable effectiveness in reducing the frequency and severity of the disease state may be analyzed with this method.

Figure 3 illustrates the relationship between savings and changing intervention effectiveness and restraint use for the intermediate estimate with respect to total costs. Relative risks for safety seat and safety belt ± 25% compared to values in Table 2 were taken as a measure of changing intervention effectiveness. In this illustration we did not require that overall MVI incidence be unvarying at 623 per 100,000 as we did for Figures 1 and 2. Incidence was allowed to vary in accordance with intervention effectiveness which emphasizes the financial impact of the purchase of safety seats. When 100% of passengers are restrained, savings would vary from $13.4 million to $12.0 million (depending on the mix of safety seat and safety belt use) for restraint effectiveness 25% greater than the intermediate estimate. These savings would vary from $9.7 million to $7.5 million for safety seat and safety belt effectiveness 25% lower than the intermediate estimate.

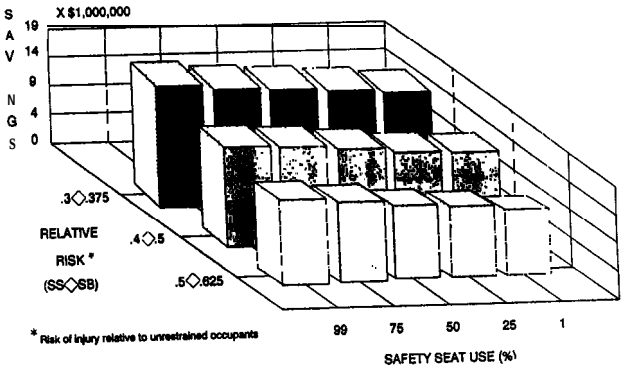


FIGURE 3: Savings associated with Medicaid purchase of safety seats for three levels of restraint effectiveness (expressed as relative risk) and five combinations of restraint usage. For the sake of illustration, all children are restrained, i.e. if safety seat use is 99%, then safety belt use is 1%; if safety seat use is 1%, then safety belt use is 99%.

Our model also predicts that as an intervention has an enhanced effectiveness for reducing the frequency of a disease state., costs attributable to that state are shifted away from the population using that intervention. As an intervention's effectiveness increases, further cost reductions will occur from higher intervention usage rather than improvements in intervention effectiveness.

Further, as the usage of interventions effective in reducing disease frequency increases, costs attributable to the disease state are increasingly spread over populations already using interventions. Further cost reductions will be dependent on newer interventions. Considering MVI leads one to the conclusion that as increasing usage of effective restraints decreases the risk of injury once a crash has occurred further reductions in the cost of injury will require interventions that involve attempts to avoid the crash event itself.

Limitations

Interventions are not always exclusive of each other such that a portion of the population is using one intervention and not another intervention. The model requires such independence to allow attributing costs to one intervention group only. Although it is possible to bypass this problem by subdividing the population into groups using one, another, or both interventions (in the case of two interventions), the model becomes more dependent on the population estimates for the subdivisions. This will increase the uncertainty of results unless intervention effectiveness estimates are relatively well known.

strengths

Our model is appropriate to wide variety of applications. It is only dependent on reasonably accurate incidence data, cost data, and effectiveness data. As we have modelled national data, a nursing home could easily model a program

to prevent falls resulting in broken hips and determine a range of effectiveness that an intervention must have to be cost effective. Applications to other areas of risk analysis are also possible.

SUMMARY

We have presented a model for the mathematical analysis of cost of injury data and cost-effectiveness of safety interventions capable of reducing the frequency and severity of injury. This model has potential application to any type of injury and any number of separate interventions for that injury where the effectiveness of the interventions, their costs, and the cost-of-injury are estimatable.

We have illustrated our model with the example of MVI using a hypothetical intervention of safety seat distribution to a cohort of Medicaid beneficiaries. We have suggested the cost-effectiveness of such an intervention for a plausible scenario of restraint usage and effectiveness when considering the total cost of injury.

ACKNOWLEDGEMENT

Dr. Hamburger's sabbatical year salary as Injury Control Fellow at Children's National Medical Center was supported by Group Health Association, Washington, D.C.

1. Arevalo JA, Washington AE. Cost-effectiveness of Prenatal Screening and Immunization for Hepatitis B Virus. JAMA 259:365-369; 1988.

2. Baker SB, Waller AE. Childhood Injury: State-By-State Mortality Facts. Washington, D.C.: The Johns Hopkins Injury Prevention Center, National Maternal and Child Health Clearing House; 1989.

3. Budnick EK. Estimating Effectiveness of State-of-the-Art Detectors and Automatic Sprinklers on Life Safety in Residential Occupancies. U.S.Department of Commerce, National Bureau of Standards report # NBSIR 84-2819: 1984.

4. Council on Scientific Affairs. Preventing Death and Injury from Fires with Automatic Sprinklers and Smoke Detectors. JAMA 257:1618-1620; 1987.

5. Davis RA, Markowitz LE. et al. A Cost-Effectiveness Analysis of Measles Outbreak Control Strategies. Am J Epidemiol 126:450-459; 1987.

6. Decker MD, Dewey MJ, et al. The Use and Efficacy of Child Restraint Devices: The Tennnessee Experience, 1982 and 1983. JAMA 252:2571-2575; 1984.

7. Evans L. Estimating Fatality Reductions from Increased Safety Belt Use. Risk Analysis 7:49-57; 1987a.

8. Evans L. Belted and Unbelted Driver Accident Involvement Rate Compared. J Safety Res 18:57-64; 1987b.

9. Evans L. Motorized two-point safety belt effectiveness in preventing fatalities. Accid Anal Prev 23:165-174; 1991.

10. Foss RD. Sociocultural Perspective on Child Occupant Protection. Pediatrics 80:886-893; 1987.

11. Graham JD. Preventing Automobile Injury. Dover, MA: Auburn House Publishing Company; 1988.

12. Graham JD, Henrion M. A Probabilistic Analysis of the Passive-Restraint Question. Risk Analysis 4:25-40; 1984.

13. Guyer B, Ellers B. Childhood Injuries in the United States: Mortality, Morbidity, and Cost. Am J Dis Chiild 144:649-652; 1990.

14. Hay JW, Daum RS. Cost-Benefit Analysis of Two Strategies for Prevention of *Haemophilus influenzae* Type b Infection. Pediatrics 80:319-329; 1987.

15. Hay JW, Daum RS. Cost-benefit analysis of *Haemophilus influenzae type* b prevention: conjugate vaccination at eighteen months of age. Pediatr Infect Dis J 9:246-252; 1990.

16. Horn A. Personal Communication. Division of Statistical Resources, National center for Health Statistics. Hyattsville, MD; 1991.

17. Hunter WW, Stutts JC, et al. Characteristics of seat belt users and non-user* in a state with a mandatory belt law. Health Education Research 5:161-173; 1990.

18. Kahane CJ. An Evaluation of Child Passenger Safety: The Effectiveness and Benefits of Safety Seats. US Department of Transportation, National Highway Traffic Safety Administration. report no. DOT-HS-806-890; 1986.

19. Kaplan BH, Cowley RA. Seatbelt Effectiveness and Cost of Noncompliance Among Drivers Admitted to a Trauma center. Am J Emerg Med 9:4-10; 1991.

20. Kernish R, London L. Strategies to Increase the Use of Child Safety seats: An Assessment of current Knowledge. US Department of Transportation, National Highway Traffic Safety Administration. report no. DOT-HS-807-116; 1986.

21. Leonard J, Holtgrave DR. Johnson RP Cost-Effectiveness of Hepatitis B Screening in a Mental Health Institution. J Fam Pract 32:45-48; 1991.

22. McLoughlin EM, McGuire A. The Causes, Cost, and Prevention of Childhood Burn Injuries. Am J Dis Child 1444:677-683; 1990.

23. Mast EE, Berg JL, et al. Risk Factors for Measles in a Previously Vaccinated Population and Cost-Effectiveness of Revaccination Strategies. JAMA 264:2529-2533; 1990.

24. Mohan D, Schneider LW. An Evaluation of Adult Clasping Strength for Restraining Lap-Held Infants. Human Factors 21(6):635-645; 1979.

25. Nelson D. Personal Communication. Epidemiology Branch, Division of Injury Control, Centers for Disease Control; Atlanta, GA: 1991.

26. National Highway Traffic Safety Administration. Occupant Protection Trends in 19 Cities. Washington, DC: US Department of Transportation, National Highway Traffic Safety Administration, 1990.

27. Orsay EM, Turnbull TL, et al. Prospective Study of the Effect of Safety Belts on Morbidity and Health care Costs in Motor-Vehicle Accidents. JAMA 260:3598-3603; 1988.

28. Partyka SC. Papers on Child Restraint: Effectiveness and use. US Department of Transportation, National Highway Traffic Safety Administration. report no. DOT-HS-807-286; 1988.

29. Porath A, McNutt A, et al. Effectiveness and Cost Benefit of a Proposed Live Cytomegalovirus Vaccine in the Prevention of Congenital Disease. Rev Infect Dis 12:31-40; 1990.

30. Reath DB, Kirby J, Lynch M, et al. Injury and Cost Comparison of restrained and Unrestrained Motor Vehicle Crash Victims. J Trauma 29:1173-1177; 1989.

31. Rice DP, MacKenzie EJ, et al. Cost of Injury in the United States: A Report to Congress. San Francisco, CA: Institute for Health & Aging, University of California and Injury Prevention center, The Johns Hopkins University; 1989.

32. Robitaille Y, Legault J, et al. Evaluation of an Infant Seat Program in a Low-Income Community. Am J Dis Child 144:74-78; 1990.

33. Ruegg RT, Fuller SK. A Benefit-cost Model of Residential Fire Sprinkler Systems. U.S Department of Commerce, National Bureau of Standard*. report no. NBS/TN 1203: 1984.

34. Saalberg JH. Second Evaluation of the League General Insurance Company Child Safety Seat Distribution Program. US Department of Transportation, National Highway Traffic safety Administration. contract no. DTNH22-80-C-05169; 1985.

35. Saalberg JH, Morrison AJ. Evaluation of the League General Insurance Company Child Safety Seat Distribution Program. US Department of Transportation, National Highway Traffic safety Administration. report no. DOT-HS-806-253; 1982.

36. Santer LJ, Stocking CB. Safety Practices and Living Conditions of Low-Income Urban Families. Pediatrics. 88:1112-1118; 1991.

37. Stulginskas JV, Pless IB. Effects Of a Seat Belt Law on Child Restraint Use. Am J Dis Child 137:582-585; 1983.

38. Vegega ME, Klein TM. Safety-Belt Use Among Drivers Involved in Alcohol-Related Fatal Motor-Vehicle Crashes--United States, 1982-1989. MMWR 40:397-400; 1991.

39. Williams AF, Zador P. Injuries to Children in Automobiles in Relation to Seating Location and Restraint Use. Accid Anal Prev 9:69-76; 1977.

EXPANDING MEDICAID ELIGIBILITY FOR FAMILY PLANNING SERVICES FROM 25 TO 133 PERCENT OF POVERTY: A COST BENEFIT ANALYSIS

Douglas R. Murray, Arkansas Center for Health Statistics

The purpose of this study was to determine that if Medicaid eligibility for family planning services were expanded from 25 percent to 133 percent of the federal poverty level, could the costs of providing those services be recouped within five years by direct savings in Medicaid maternity costs, WIC, and AFDC[1].

Most of the data used in this study come from readily available published articles. Much of the data in those articles are from the National Surveys of Family Growth (NSFG) sponsored by the National Center for Health Statistics. The analysis was conducted using Lotus 123. The flexibility of this software facilitated the use of about 50 separate variables, each of which could affect the outcomes[2].

To conduct this study, three separate analyses are undertaken. The first is to estimate the number of women in need of family planning services and who might be potentially eligible for Medicaid. The second is to estimate the probable rates of utilization of contraceptive services, their efficacy, and their impact on births. The third step is to estimate the costs of providing these services and the direct savings that might accrue from delaying or preventing unintended births; this is presented in the form of a benefit to cost ratio.

As noted, the first step is to estimate the number of women who could conceivably be affected by an expansion of Medicaid eligibility. According to the 1990 Census, there were 519,852 women in Arkansas ages 15 through 44. The next problem was to estimate how many of these women would be potentially eligible for family planning services if Medicaid were expanded from 25 to 133 percent of poverty. Unfortunately, detailed age/sex/poverty level data are not yet available from the Census.

Therefore, unpublished data from the 1990 Arkansas Health Insurance Survey[3] of almost 8,000 households were used to develop these estimates. Based on the survey respondents reports of their household income and household size, it was possible to calculate their poverty level. According to the survey data, 1,311 women ages 15 through 44 fell into the range of 25 percent to 133 percent of the poverty level. This represented 21.7 percent of all (6,037) women in that age category. This percent when applied to the 519,852 women is 112,891, which is the estimate of the total number of women in Arkansas aged 15-44 with incomes between 25 and 133 percent of the federal poverty level. This is the population that would be eligible for family planning services if Medicaid eligibility were expanded.

The next step is to estimate the impact that expanded Medicaid family planning services could have on births in this population. The question is, how many of these women are likely candidates for family planning services?

One approach is to look at low income women who have recently given birth and determine how many were unintended. Using data provided by Henshaw, Forrest and Singh[4] estimate that 33.8 percent of the births to women below 100 percent of the federal poverty level were "mistimed," and another 24.8 percent were "unwanted[5]." A total of 58.6 percent of low income births are unintended.

While this is a likely group of possible candidates for family planning services, there are many other women in addition to these new mothers. To this end, Forrest and Singh estimated the

---

[1]The author would like to thank Dr. Joycelyn Elders, former Director of the Arkansas Department of Health, for proposing the project.

[2]A copy of the spreadsheet in Lotus 123 Release *3.1* format is available upon request.

[3]The data were weighted by age, sex, race, telephone noncoverage, and county size. As response rates vary according to chance and sociodemographic characteristics, weighted data more accurately reflect the true population composition.

[4]Forrest! Jacqueline Darroch and Susheela Singh, "The Sexual and Reproductive Behavior of American Women, 1982-1988," *Family Planning Perspectives,* 22,5:206, *1990.*

[5]Forrest and Singh state, "If the respondent reported that at the time a birth was conceived, she had not planned to have a baby, the birth was considered unintended. If the woman reported that she did not ever want to have another child or any child, the unintended birth was termed unwanted. If she did want to have a baby in the future, but at a later time, the unintended birth was termed mistimed. Since the responses are retrospective information gathered after the child was born, they may underestimate the extent of unintended fertility."

percentage distribution of women by, what they refer to as "exposure to the risk of unintended pregnancy." In essence, the table reports the current sexual activity (if any) and contracepting behavior (if any) of low income Arkansas women. (Please note that these data are for all women, as data specific to low income women were not available.)

As some women are not at risk, they are not in need of services. Approximately 11.5 percent of the female population 15 through 44 years of age have never had sexual intercourse. In Arkansas, that would yield 12,893 women who are in the 25 to 133 percent of poverty range, out of the previously estimated total of 112,891 women at that poverty level.

Another 29.7 percent of these women are sterile, 6.1 percent non-contraceptively and 23.6 percent contraceptively. Alternatively, 3.8 percent of the women were actively seeking to become pregnant.

Forrest and Singh" found that although most births among lower income women are unintended (mistimed or unwanted), some are indeed wanted. This latter group of women would presumably not be interested in availing themselves of family planning services, regardless of financial barriers.

The "candidates" would come from the remaining women: those 6.9 percent who have not had intercourse in the last three months, the 4.0 percent of women who are pregnant and 0.7 percent who are postpartum, the 36.7 percent who are using some reversible (as opposed to sterilization) contraceptive, and, most importantly, those 6.7 percent of the women who use no contraceptives at all.

At this point, the potential number and percent of mistimed and unwanted births is calculated. The number of births to women who have never had intercourse is, of course, zero. Likewise, those women who are seeking pregnancy or are sterile are not a consideration here. For those women who have had no intercourse in the past three months, an arbitrary figure of 25 percent was assigned. Applying the Henshaw and Forrest/Singh estimates of 58.6 percent mistimed and unwanted births to the pregnant and postpartum women yields an Arkansas estimate of 2,646 and 463 women, respectively. Those who use various reversible methods have a higher collective risk of pregnancy, based on utilization and efficacy.

Adding up this column results in an estimate of *32,764* women who needed contraceptive services, out of the original 112,891 Arkansas women at 25 to 133 percent of poverty.

Some 20,143 of the women in need are presently using a reversible method. Their utilization-efficacy rate was calculated as follows. Mosher found, for example, that 37 percent of the contracepting women below *150* percent of the poverty level relied upon female sterilization as compared with only *22* percent of those women with incomes more than 300 percent of the poverty level. Alternatively, 14 percent of the more affluent women relied upon male sterilization as the contraceptive method as opposed to only 4 percent of the low income women.[7]

The efficacy of various contraceptive methods is related to several factors including the intrinsic effectiveness of the product as well as the compliance rate; the latter is partly a function of sociodemographic characteristics of the women, such as education. Also using data from the National Survey of Family Growth, Jones and Forrest' estimated the percent of women below 200 percent of the poverty level who experienced a contraceptive failure during the first 12 months of use among women. Their estimates were standardized for age, race, and marital status. The pill, for example, had a 8.1 percent failure rate resulting in a 91.9 percent efficacy rate.

Grady et al.[9] in an earlier analysis of the same data obtained somewhat lower failure rates because they did not allow for method-specific differences in underreporting. Since Jones and Forrest did not publish failure rates for the IUD, Grady is the source of the IUD failure rates in this study. It should be noted that the use of the IUD as a contraceptive method has dropped precipitously since the early 1980s and represents a relatively small proportion of total contraception. As the literature on NORPLANT is limited, it is assumed, for the sake of this study, that utilization of NORPLANT in the immediate future is one percent and that it is 99 percent effective.

'Forrest, Jacqueline Darroch and Susheela Singh, "The Sexual and Reproductive Behavior of American Women, 1982-1988," Family *Planning Perspectives,* 22,5:206, *1990.*

[7]For the purposes of this study, male and female sterilization were combined.

'Jones, Elise F. and Jacqueline Darroch Forrest, "Contraceptive Failure in the United States: Revised Estimates from the 1982 National Survey of Family Growth," *Family Planning Perspectives,* **21**:103, *1989.*

'Grady, William R. et al. "Contraceptive Failure in the United States: Estimates from the 1982 National Survey of Family Growth," *Family Planning Perspectives,* **18**:200, 1986.

At this point, this distribution of birth control method utilization for women below 150 percent of the federal poverty level was multiplied by their corresponding efficacy rates. These utilization-efficacy rates were, in turn, multiplied times 112,891, resulting in the figure of 32,764 seen previously. This is the number of women in Arkansas aged 15 through 44 who are between 25 and 133 percent of poverty and are in need of family planning services.

At this point, estimates of the number of women aged 15-44 in Arkansas who are between 25 and 133 percent of the federal poverty level have been developed. The sexual and contraceptive status of these women has been identified. The number of these women who are at risk of unintended pregnancy has been estimated. Estimates of the number of these women who could have unintended births have been generated. And a figure which tells us how large a population must be served has been calculated. The next step is to estimate the actual number of births in the poverty population.

In 1991, there were a total of 35,456 births to Arkansas residents. Using data from the 1990 Arkansas Health Insurance Survey, the income distribution was determined. Using the results of a previous matching of birth certificates (for the period of October, 1986 through July, 1988) to Medicaid paid claims, the percent of those births were Medicaid births, at various federal poverty levels, was calculated.[10]

In Arkansas in 1991, an estimated 20.9 percent of the births were to Medicaid mothers who had incomes of less than 25 percent of the federal poverty levels. One-third (33.2 percent) of all births were to Medicaid mothers who had incomes of less than 133 percent of the federal poverty levels. The difference between the two levels is the additional number of women who delivered who would have eligible for family planning services if a higher level of Medicaid eligibility had been in place.

When this distribution was applied to the 1991 birth data, the result was a total of 4,343 women out of that year's birth cohort. When the contraceptive utilization-efficacy rates are applied, this figure is reduced to 2,872, the number of births that would not have occurred to these low income women had family planning services been available to them.

The 'final phase of this analysis involves the cost data. Based on actual Medicaid paid claims data[11] for the October, 1986 through July, 1988 period, the average birth cost $2,129.[12] This figure was updated by applying the medical component of the Consumer Price Index for the intervening years, bringing it up almost 50 percent to an estimated $3,190 in 1990. This figure is multiplied by the estimated 2,872 women who would not have delivered had they experienced the income-specific contraceptive utilization and efficacy rates discussed above. The total one year savings in direct Medicaid costs for deliveries alone would have been $9,162,222. The federal Medicaid share for Arkansas is more than 74 percent.

WIC and AFDC costs are also included in this analysis (although they would not be if the state were seeking a Medicaid waiver). The average monthly WIC food package in Arkansas is $26.46 while the administrative cost is $10.58. If the 2,872 low income births had been prevented, this would have resulted in an WIC savings of $6,383,540.

AFDC costs were reported to be about $2,306 per year. As with WIC, it is assumed that the average family will be receiving these services for five years. With about 1.98 children per family, the cost was allocated to a single child. The estimated AFDC savings for one child is $5,823. If the 2,872 low income births had been prevented, this would have resulted in an AFDC savings of $16,725,858.

Having estimated that the total costs of these unintended pregnancies comes to $32,271,621, the next step is to calculate the costs of providing family planning services to this population. Under current guidelines,[13] Medicaid will reimburse providers at a fixed rate for specific services and supplies. A typical package of services costs about $165 per year. While NORPLANT is more expensive initially, the five year costs are lower than the average of other contraceptive

---

[10]Senner, John, Arkansas Center for Health Statistics, unpublished data. These data are based on Arkansas Medicaid enrollment rates during the period of October, 1986 through July, 1988. During this time, eligibility for Medicaid maternity benefits increased from 37 percent of the federal poverty level to 75 percent and then to 100 percent. The data were smoothed using a logit transformation and then projected downward to 25 percent and upward to 133 percent and 185 percent.

[11]Senner, John, Arkansas Center for Health Statistics, personal communication, December 2, 1991.

[12]The medical component of the Consumer Price Index has gone up more than 25 percent since 1987. Therefore, this is a very conservative estimate of the costs.

[13]AGREEMENT for FAMILY PLANNING SERVICES AND SUPPLIES FOR INDIVIDUALS OF CHILDBEARING AGE WHO ARE ELIGIBLE UNDER THE ARKANSAS MEDICAL ASSISTANCE PROGRAM.

services. Consequently, this suggests that almost 20 women could be provided with Medicaid financed family planning services for an amount equal to the immediate cost of one delivery.

The total cost to Medicaid, WIC and AFDC for the 2,872 unwanted and mistimed births is estimated to be $32,271,621. If annual family planning services are estimated to total $163.67 per person and a total population of 35,283 women must be served in order to prevent the 2,872 births, the total cost of the services will be $5,774,933.

The benefit to cost ratio is calculated by taking the $32'271,621 in Medicaid, WIC, and AFDC costs and dividing by the $5,774,933 in family planning services. This results in a ratio of 5.6 to one; that is, $5.60 would be saved in Medicaid delivery, AFDC, and WIC costs within five years for every dollar spent on family planning services. Even if the savings were limited simply to the costs to Medicaid of the deliveries, the benefit to cost ratio would still be 1.6 to one.

It should be noted that this analysis is based on a birth cohort; thus, these savings are available with every new year. Moreover, the savings estimates are relatively conservative in that they do not include many other costs of raising the child such as other Medicaid expenses, education, etc. Indeed, in 1989 in the South, the cost of raising urban children from birth to age 18 at a moderate-cost level was estimated to be $114,483.[14]

After this study was completed, the author located another study by Forrest and Singh in which they concluded "for every government dollar spent on family planning services, from $2.90 to $6.20 (an average of $4.40) is saved as a result of averting [short-term] expenditures on medical services, welfare and nutritional services."[15] It is interesting to note that the methodology employed in this study is quite different from the Forrest and Singh analysis; moreover, the two studies were conducted completely independently of one another. Nonetheless, the results are quite similar with the 5.6 benefit to cost ratio falling well within the Forrest and Singh estimate of 2.9 to 6.2.

There are two conclusions to be drawn from this presentation. First, the analysis indicates that it would be cost-effective to expand Medicaid eligibility to 133 percent of the federal poverty level in order to provide family planning services to those women.

Second, computer spreadsheets are very useful for developing complex models such as this. Their flexibility allows the user to easily change a wide range of parameters, thereby quickly responding to suggestions and recommendations regarding the basic assumptions of the model.

---

[14] Edwards, Carolyn S. "USDA Estimates of the Cost of Raising a Child: A Guide to Their Use and Interpretation" USDA Miscellaneous Publication N. 1411, in "Updated Estimates of the cost of Raising a Child," Family Economics Review, Vol. 2, No. 4, pp.30-31.

[15] Forrest, Jacqueline Darroch and Susheela Singh, "Public-Sector Savings Resulting from Expenditures for Contraceptive Services," *Family Planning Perspectives,* 22,1:6, 1990.

Estimated Arkansas Medicaid/WIC/AFDC Savings
If Family Planning Services Were Expanded
From 25% to 133% of Federal Poverty Level

Number of Women Aged 15-44 in Arkansas:			519,852
Source: 1990 U.S. Census			
Est. No. Women 15–44, 25%–133% Federal Poverty Level:			1,311
Source: 1990 Arkansas Health Insurance Survey			
Total Women 15-44 in AHIS:			6,037
Source: 1990 Arkansas Health Insurance Survey			
% Women 15–44, 25–133% Federal Pov Level	1311 / 8037	=	21.7%
Source: 1990 Arkansas Health Insurance Survey			
Est. Arkansas Women 25%–133% FPL	21.7% X 519,852	=	112,891
Percent Mistimed Births Among < 100% Federal Poverty Level:			33.8%
Percent Unwanted Births Among < 100% Federal Poverty Level:			24.8%
Source: Forrest & Singh, FPP, 22:5, pg 272			
Percent of Low Income Population Using ANY Contraceptives:			57.8%
Source: Forrest & Singh, FPP, 22:5, pg 209			
Percent and Number At Risk of Unintended Pregnancy			
Never had intercourse (e.g., 11.5% X 112,891):	11.5%		12,983
No intercourse in past 3 months:	8.9%		7,790
Pregnant:	4.0%		4,518
Postpartum:	0.7%		790
Seeking pregnancy:	3.8%		4,290
Sterile noncontraceptively:	6.1%		6,886
Sterile contraceptively:	23.6%		26,642
Reversible method user:	36.7%		41,431
Nonuser:	6.7%		7,564
Source: Forrest & Singh, FPP, 22:5, pg 210			
Percent and Number of Mistimed / Unwanted Births			
Never had intercourse:	0%		0
No intercourse in past 3 months:	50%		3,895
Pregnant:	59%		2,646
Postpartum:	59%		483
Seeking pregnancy:	0%		0
Sterile noncontraceptively:	0%		0
Sterile contraceptively:	0%		0
Reversible method user:	50%		20,716
Nonuser:	100%		7,584
Total Est. 25%– 133% FPL Women Needing Contraception			**35,283**
Utilization of Various Contraceptive	Sterilization :		41.0%
Methods in Low Income Populations	Pill :		36.0%
Source: Mosher, FPP, 22:5, pg 201	IUD :		3.0%
	Diaphragm :		2.0%
	Condom :		13.0%
	Norplant :		1.0%
	Other :		4.0%
Efficacy of Contraceptive Methods	Sterilization :		100.0%
Source: Jones & Forrest, FPP, 21:3, pg 107	Pill :		91.9%
Grady et al., FPP, 18.5, pg 204	IUD :		93.0%
	Diaphragm :		76.6%
	Condom :		79.3%
	Norplant :		99.0%
	Other :		66.9%
1991 Arkansas Live Births:			35,456
Est Impact on Arkansas Medicaid Births With Various Fed Pov Levels			% Medicaid
At Various Federal Poverty Levels	25% :		20.9%
Source: Arkansas Center for Health	***37%*** :		23.5%
Statistics – Unpublished Data	***75%*** :		28.6%
	100% :		30.8%

	133% :		33.2%
	185%:		36.0%
Est Number & Percent Births in 25%-133% FPL Population		4,343	12.2%
Target Group: Est. Number and Percent 1991 Births That Would Not Have Occured In Expanded Eligibility Population If Contraceptive Utilization / Efficacy Rates Were In Effect		2,872	68.1%
Total Estimated 1990 Medicaid Cost per Birth:			$3,189.79
Source: Arkansas Center for Health Statistics and DHS Economic & Medical Services -- Unpublished Data			
Total 1990 Target Group Births Medicaid Costs	83,190 X 2,872	=	$9,162,222
Federal Medicaid Share	74%	=	$6,780,044
State Medicaid Share	26%	=	$2,382,178
WIC: Average Monthly Food Package	$26.46 X 12 =		$317.52
WIC: Average Monthly Administrative Cost	$10.58 X 12 =		$126.96
WIC: Costs per Person, For 5 Years	$444.48 x5=		$2,222.40
Source: ADH Division of Women, Infants and Children Unpublished Data for November, 1991			
AFDC: Costs per Family, 1987-88, For 5 Years	$2,305.92 x 5 =		$11,529.60
Source: Arkansas Statistical Abstract -- 1991: pg 275			
Average Number of Children per AFDC Case	1.98		
Est. Five Year AFDC Savings for One Child	$1,164.61 X5 =		$5,823.03
Total 1990 Target Group Births AFDC Costs	$5,823 X 2,872	=	$16,725,858
Federal AFDC Share	74%	=	$12,377,135
State AFDC Share	26%	=	$4,348,723
Estimated Short-term Medicaid, WIC, and AFDC Costs per Birth			$11 ,235.22
Estimated Total Short-term Medicaid, WIC, and AFDC Costs	$11,235 X 2,872	=	$32,271,621

Maximum Allowable Rates for Family Planning Services
In Arkansas Under Medicaid Effective July 1, 1991

	Procedure Code	Rate
Initial/Annual Visit	20847 :	$70.00
Periodic/Basic Visit	20848 :	$50.00
Urinalysis, routine, without microscopy	81002 :	$3.80
Gonadotropin, chorionic, qualitative	84703 :	$11.15
Blood count/hematocrit/hemoglobin/colorimetric	85014 :	83.52
Syphilis, precipitation or flocculation test, qualitative VDRL, RPR, ART	86592 :	$6.38
Culture, bacterial, any source, anaerobic	87075 :	$10.99
Cytopathology, smears, cervical or vaginal (Papanicolaou), screening and interpretation, up to three smears	88150 :	$9.00
Est. Total Annual Non-Norplant Family Planning Cost Per Person		$164.84
Norplant (Amortized over 5 years)		
Cost of the device		$345.00
Cost of Insertion and removal		$218.00
Cost of first visit		$48.38
Cost of five periodic visit		$96.15
Estimated Annual Cost for Norplant		$141.50
Estimated Annual Costs for All Forms of Contraceptive Service (5% usage of Norplant and 95% all other)		$163.67

Cost of 2,872 Unwanted/Mistimed Births, 1990	2,872 X $11,235	=	$32,271,621
Medicaid Costs for Family Planning	35,283 X $163.67	=	$5.774.933

Total Benefit to Cost Ratio: $32,271,621 / $5,774,933 :==========> 5.6

For Every Dollar Spent On Family Planning, This Many Dollars Are Saved:

Number of Clients To Whom Family Planning Services Could Be Provided With Savings	$32,271,621 / $163.67	=	197,171

Note: FPP = Family Planning Perspectives

THE IMPACT OF DISSEMINATION OF SMOKING CESSATION METHODS ON HEALTH CARE COSTS AMONG PREGNANT MEDICAID RECIPIENTS

Richard A. Windsor, Health Promotion Group
Neal Richard Boyd

INTRODUCTION

Multiple reports[1-6] in the last 30 years have confirmed smoking during pregnancy as a national priority. Active and passive cigarette smoke exposure causes maternal, fetal, and infant morbidity and mortality.[7] Although the risks of maternal smoking has been established, over one million women of the approximately 4.1 million women (25%) who delivered in recent years, smoked during pregnancy.[6] Of special concern is the 1,100,000 or more Medicaid patients.* Over 350,000 of these women (≥ 35%) smoked during pregnancy.'

Comprehensive meta-evaluations[10-11] of the literature have confirmed the efficacy[12-13] and cost-effectiveness[13-14] of health education methods for Black and White pregnant smokers enrolled in prenatal care. A consensus exists about what health education methods to use with pregnant smokers. The intervention should be routinely delivered by regular staff who have been trained to provide the following components: (1) a standardized cessation skills and risk counseling session of approximately 10 minutes which includes instruction on how to use a self-directed cessation guide; (2) clinic patient reinforcement methods at follow-up visits; and (3) social support methods.

Thus, the problem and one efficacious component of the solution is well described. The purpose of this paper is to examine the potential impact of nationwide dissemination and routine use of health education methods on the incidence of low birthweight and associated medical care costs among the Medicaid maternity cohort who smoke.

METHODS

The target population for this discussion is the annual Medicaid cohort: 1,112,000 in 1990.[8] The proportion of White Medicaid patients was approximately 50% (550,000), Black patients approximately 35% (385,000), and Hispanic patients about 12% (130,000).[8] Survey data confirm a 40%+ smoking prevalence rate among White patients and 20%+ prevalence among Black pregnant smokers.[9] Because the incidence of low birthweight (5%)[9] among Hispanic women is one of the lowest rates in the U. S. and smoking prevalence is also low (15%), they were excluded from this analysis. Also, no study has documented the efficacy of cessation methods for Hispanic maternity patients who smoke.

Estimated Behavioral Impact of Dissemination

The Medicaid maternity cohort typically receives only verbal advice to quit and information about the risks of smoking and the benefits of quitting.[10-11] While population studies of the smoking behavior and cessation rates during pregnancy of the Medicaid cohort are not available, public health evaluation studies[12-13] confirm a smoking cessation range of 2% to 4% from exposure to clinic advice and information.

Evaluation studies[12-13] conducted in public health settings confirm that among pregnant smokers who are provided efficacious health education methods tailored to their needs, the cessation rate can be increased from 4% to 18% to 20% among Blacks and from 2% to 10% to 14% among Whites.

Estimated Impact of Dissemination on Low Birthweight

Estimates of population attributable risk (PAR) were used to calculate the impact on low birthweight (LBW) incidence of routine use of "tested" health education methods. In our analysis we used a smoking population attributable risk (SPAR)[2] of 30% (White) and 25% (Black) to calculate the potential impact of smoking behavior change on LBW incidence among the Medicaid cohort. In 1990 the LBW rate for the Black Medicaid patients was approximately 15% and White Medicaid patients about 9%.[9]

Estimation of LBW Associated Medical Care Costs

The full impact of smoking, both qualitatively and quantitatively, cannot be characterized by one indicator. The emotional and financial impact of a LBW infant to its family is considerable. However, costs associated with the delivery of a LBW/premature infant represents one of the most prominent outcomes associated with smoking during pregnancy. In this analysis, we used the Office of Technology Assessment (OTA) estimate of the net incremental health care costs with discounting of a LBW birth in 1986: $9,000 -- low estimate and $23,000 -- high estimate.' These estimates included three components: 1) hospitalization and physician costs at birth; 2) rehospitalization costs in the first year of life (hospital costs only); 3) long-term health care costs of treating a LBW infant.[5] These estimates were adjusted to 1990 dollars using the Consumer Price Index -- Medical Care Component of the Department of Labor Statistics: 5.8% -- 1987, 6.9% -- 1988, 8.5% -- 1989, and 9.6% -- 1990. The

discounted, inflation adjusted estimates used in our calculations were $14,645 (low estimate) and $37,431 (high estimate).

RESULTS

Behavioral Impact -- Cessation

The potential behavioral impact of health education methods for the Medicaid cohort were derived using the following process. Quit rates from the literature[10-11] were examined to determine their validity and consistency. Because these rates were derived from studies in which specialists provided the methods, they represent **Efficacy rates.** Estimates of the potential impact of routine use were derived by reducing the efficacy rates by approximately 25%. The result represents an **Effectiveness rate:** quit rates likely to be observed in routine practice when the methods are delivered by normal staff. An additional **26,840** patients might be helped to quit smoking.

Impact of Dissemination on Low Birthweight Incidence

Data in Table 1 reflect estimates of impact on the incidence of LBW from the routine delivery of these methods to Medicaid pregnant smokers. Smoking attributable risks,[2] 25% for Black patients and 30% for White patients, were used to estimate the number of smoking attributable LBW infants. The potential rate was derived by multiplying the current LBW births for Blacks (50,050) and Whites **(43,500)** by the smoking attributable risks for each. These totals were multiplied by the current LBW rates[9] (Black -- 14.3% and White -- 8.7%) to obtain the annual of number of preventable LBW births. As shown, 1,789 Black and 1,135 White LBW births might be prevented by dissemination.

Estimated Economic Impact of Dissemination

If we were able to reduce the incidence of LBW from increased cessation, medical care costs incurred by the treatment of these LBW infants would be reduced (See Table). These estimates are attenuated so as to not over estimate the potential reduction in medical care costs. As noted, estimates of the economic benefit were based on the OTA discounted and inflation adjusted data: $14,645 (low estimate) and $37,431 (high estimate).[5] In order to produce a conservative estimate of the economic benefits, we attenuated our estimates by 20% and rounded the estimate to the nearest thousands. The low estimate of $14,645 was reduced to $12,000 and high estimate of $37,431 was reduced to $30,000.

Economic Benefit of Cessation

Using $12,000 (low) and $30,000 (high) per birthweight prevented, the estimated excess medical care costs of the 2,924 smoking attributable LBW births is $35,088,000 (2,924 x $12,000) and $87,720,000 (2,924 x $30,000). The Black low estimate is $21,468,000 (1,789 x $12,000) and high estimate is $53,670,000 (1,789 x $30,000). The White low and high estimates are $13,620,000 (1,135 x $12,000) and $34,050,000 (1,135 x $30,000).

TABLE 1
Estimated Impact of Dissemination of Tested Smoking Cessation Methods to the 1990 Medicaid Cohort

	Black		White		Total	
A. Behavior **Impact--Pregnant Smokers**						
1. Current Annual Cessation	4.0%	3,080	2.0%	4,400	3.0%	7,480
2. Potential Annual Cessation	16.0%	12,320	10.0%	22,000	12.0%	34,320
B. Obstetrical Impact--Low Birthweight Infants						
1. Number of LBW Infants	14.3%	50,050	8.1%	43,500	11.0%	93,550
2. Potential LBW's Preventable Smoking Attributable	13.8%	48,261	8.5%	42,365	10.7%	90,626

C. Economic **Impact--Low Birthweight Infants**
1. Excess Cost LBW Prevented **LOW $12,000** **High** $30,000

	Black	White	Total
2. Total Excess Cost LBW Prevented	Low $21,468,000 High $53,670,000	Low $13,620,000 High $34,050,000	Low $35,088,000 High $87,720,000
3. Total Intervention Cost	$539,000 (77,000 x $7)	$1,540,000 (220,000 x $7)	$2,079,000 (297,000 x $7)
4. Cost to Benefit Ratio	LOW 1 : $40 High 1 : $100	LOW 1 : $9 High1 : $22	LOW 1 : $17 High1 : $42

These cost benefit estimates of dissemination are calculated to determine the net cost difference -- economic benefit minus cost -- among all 297,000 patients who would receive the intervention rather than only patients who stopped smoking. Cost estimates to deliver the standardized intervention are based on expenditures of personnel time and materials. Because an agency perspective was used,[14,15,16] patient time, facilities cost, and intervention development costs were not utilized in our estimates. A nurse would be the usual patient educator. A salary (1990 dollars) of $30,000 plus a fringe benefit rate of 20% was used to estimate staff cost ($36,000 per year/2080 hours = $17.31 per hour). Personnel costs for routine delivery of a 10 minute intervention at the first prenatal visit and follow-up reinforcement (five minutes) would be about $4.33 per patient ($17.31 x 0.25). Each patient would receive and be taught how to use a cessation manual such as A Presnant Woman's Guide to Quit Smoking.[17] Although this cessation guide's cost is approximately $5.00, large volume printing (5,000 copies) would reduce the materials to about $2.00 per patient. Other materials, reproduction, and labor costs are about $0.40. The total cost per patient is about $7.00. The cost for the estimated 297,000 pregnant Medicaid smokers is approximately $2,079,000 per year.

Our Cost Benefit Ratio (CBR) analysis reveals a low estimate of $1 : $17 and high estimate of $1 : $42. The corresponding Black and White low and high estimates are: $1 : $40 (low) and $1 : $100 (high) and $1 : $9 (low) and $1 : $22 (high), respectively. The net benefit minus cost difference is $33,009,000 (low estimate) and $85,641,000 (high estimate) in favor of the intervention. An observed difference in net benefit minus cost difference was observed by race. The difference was more favorable for Black patients.

Sensitivity Analyses

Sensitivity analyses were performed of the cost benefit of dissemination. We varied intervention costs from $7.00 (low estimate) to $14.00 (high estimate). We varied the estimates of smoking attributable risk. Smoking attributable risk for Black patients was reduced from 0.25 to 0.20 and for White patients was reduced from 0.30 to 0.25. The estimated impact of dissemination on the number LBW infants was reduced from 2,924 to 2,377.

Costs for personnel and materials to provide the intervention will vary between states. If the intervention costs were increased by 100%, $7.00 per patient to $14.00 per patient, the total cost to deliver the intervention to the 297,000 pregnant smokers is $4,158,000. The overall CBR is low estimate = $1 : $7 and high estimate = $1 : $17. The net difference between benefit and cost still favors the intervention: $24,366,000 (low estimate) and $67,157,000 (high estimate). CBRs for Black patients are: low estimate -- $1 : $16 and high estimate -- $1 : $40. CBRs for White patients are: low estimate -- $1 : $4 and high estimate -- $1 : $9.

DISCUSSION

Studies[10-11] have demonstrated the efficacy of self-help smoking cessation methods in multiple settings. Cost-effectiveness[14] and cost benefit[13] of these methods have also been documented. Smoking prevalence among women of childbearing age, especially those who use public health maternity clinics (35%), has decreased very slowly, approximately one-half percent per year, during the last decade.[6,9] The LBW rate among the Medicaid population has remained stable for the last 10 years.' Accordingly, the dissemination of efficacious smoking cessation methods to all groups of pregnant women should be one essential step to help achieve state and national objectives: Year 2000 Objectives for smoking prevalence -- 10% and LBW -- 5%.[6] The Medicaid maternity cohort, annually one-fourth of the pregnant women, should represent a major target group for dissemination of tested methods for pregnant smokers.*

These analyses indicate that dissemination and routine use with the Medicaid cohort has the potential to reduce the smoking prevalence rate and produce an additional improvement in the LBW rate. An additional 26,840 pregnant smokers might be assisted in stopping smoking. These outcomes could be achieved at a low cost -- $7.00 to $14.00 per patient. The potential for cost savings from reductions in LBW appear to be substantial, even if the lower SPAR and CBR estimates were observed.

Results of sensitivity analyses do not affect the conclusions derived from these estimates. Variations in the smoking attributable risks, intervention costs, inflation, and discounting, do not substantially change the estimated cost benefit and cost savings of dissemination of these methods for the Medicaid cohort. If the intervention costs were increased by 100% and smoking attributable risks for Black and White patients were reduced to 0.15 and 0.20 respectively, the CBRs are higher than the $1 : $3.4 CBR estimates of the Institute of Medicine for prenatal care.'

These estimates should, however, be interpreted with some caution. The behavioral change rates used in this

analysis to estimate the potential impact of dissemination and routine use were derived from multiple clinical trials. While the efficacy (cessation) rates were attenuated prior to use in our cost calculations, other major factors, in addition to smoking, contribute to LBW. These include prenatal nutrition, genetics, demographics, socio-economic status, and substance abuse.[18] The impact on LBW of quitting smoking will be diminished when other factors such as these are present.

In order to achieve the quit rates used in our estimates, it also assumes that 90% or more will receive the methods and two-thirds will use the methods. While the etiology of LBW is complex and solutions are multi-dimensional, our analyses present strong support for dissemination and routine use of low cost smoking cessation methods to reduce LBW attributable medical care costs of the Medicaid cohort.

REFERENCES

1. U.S. Public Health Service. Smoking and Health. Report of the Advisory Committee to the Surgeon General of the Public Health Service. Department of Health, Education, and Welfare, Public Health Service, Center for Disease Control, DHEW Publication No. 1103, 1964, 387 pp.
2. U.S. Public Health Service. Reducing the Health Consequences of Smoking -- 25 Years of Progress. A Report of the Surgeon General. Department of Health and Human Services, Public Health Service, Centers for Disease Control, DHHS Publication No. 89-8411, 1989, 703 pp.
3. U.S. Public Health Service. The Health Benefits of Smoking Cessation: A Report of the Surgeon General. Department of Health and Human Services, Public Health Service, Centers for Disease Control, DHHS Publication No. 90-8416, 1990.
4. Institute of Medicine, Committee to Study the Prevention of Low Birthweight. Preventing Low Birthweight. Washington, D.C.: National Academy Press; 1985.
5. Office of Technology Assessment. Health Children: Investing in the future. Washington, D.C.: Office of Technology Assessment: 1988. OTA-H-345.
6. U.S. Public Health Service. Healthy People 2000: National Health Promotion and Disease Prevention Objectives. Washington, D.C.: Department of Health and Human Services; 1991. DHHS/PHS 91-50213.
7. Kleinman J, Mitchell B, Madan JH, et al. The effects of maternal smoking on fetal and infant mortality. Am J Epidemiol. 1988; 127 (2): 274-282.
8. Robert Pokras. National Center for Health Statistics, March, 1993, personal correspondence.
9. Fichtner RR, Sullivan KM, Zyrkowski CL, Trowbridge FL. Racial/ethnic differences in smoking, other risk factors, and low birth weight among low income pregnant women, 1978-1988. MMWR. 1990; 39: 13-21.
10. Windsor RA, Orleans CT. Guidelines and methodological standards for smoking cessation intervention research among pregnant women: improving the science and the art. Health Educ Q. 1986; 13: 131-161.
11. Windsor RA, Boyd NR, Orleans CT, et al. A meta-evaluation of smoking cessation intervention research among pregnant women: improving the science and art. Health Educ Q. In review.
12. Windsor RA, Cutter G, Morris J, et al. The effectiveness of smoking cessation methods for smokers in public health maternity clinics: a randomized trial. Am J Public Health. 1985; 75: 1389-1392.
13. Windsor RA, Lowe JB, Perkins LL, et al. Health education methods for pregnant smokers: its behavioral impact and cost benefit. Am J Public Health. 1993; 83: 201-206.
14. Windsor RA, Warner K, Cutter G. A cost-effectiveness analysis of self-help smoking cessation methods for pregnant smokers. Public Health Rep. 1988; 103: 83-89.
15. Warner KE, Luce BR. Cost-Benefit, Cost-Effectiveness Analysis in Health Care: Principles, Practices, and Potential. Ann Arbor, Michigan: Health Administration Press; 1982.
16. Russell LB. Evaluating Preventive Care: Report on a Workshop. Washington. D.C.: Brookings Institution; 1987.
17. Windsor RA, Smith DY. A Pregnant Woman's Guide to Quit Smoking. 4th Edition. Birmingham, Alabama: EBSCO Media; 1991.
18. Kramer M. Determinants of low birth weight: methodological assessment and meta-analysis. Bull WHO. 1987; 65 (5): 663-737.

THE OREGON MEDICAID PLAN: PRIORITIZING HEALTH SERVICES USING DATA AND VALUES

Barry F. Anderson, Portland State University

The 1989 Oregon Basic Health Services Plan (SB27, SB935, SB534) established a Health Services Commission (hereafter, the Commission) and charged it with reporting to the Governor and the Legislature "a list of health services ranked by priority, from the most important to the least important, representing the comparative benefits of each service to the entire population to be served". It also required the Commission to "actively solicit public involvement in a community meeting process to build a consensus on the values to be used to guide health resource allocation decisions."

The Commission began by dividing its task into fact and value components. Fact judgments regarding prediction of health states with and without treatment would be made by health care experts. Value judgments regarding preferences for health states would be made by representatives of the public. In this way, the Commission would be able to meet the legislative requirement to use public values to guide the allocation process and, at the same time, avoid the serious but all-too-common error of asking the public to evaluate alternatives directly.

Public Judgments of Value

At 47 community meetings held throughout the state, 1048 participants identified as important the following 13 values: Prevention, Quality of Life, Cost Effectiveness, Ability to Function, Equity, Effectiveness of Treatment, Benefits Many, Mental Health and Chemical Dependency, Personal Choice, Community Compassion, Impact on Society, Length of Life, and Personal Responsibility[1]. These were used qualitatively, as a checklist, to guide the structuring of the prioritization process.

In a random-sample telephone survey of 1001 persons, health-related quality of life was evaluated in terms of the Quality of Well-Being Scale[2,3]. The QWB, developed by the National Center for Health Statistics to operationalize the World Health Organization's definition of "health", measures health-relatedquality of life in terms of three function scales, Mobility, Physical Activity, and Social Activity, and 23 symptoms, e.g., "trouble learning, remembering, or thinking clearly". The QWB was selected by the Commission because it has been reasonably well validated[4,5,6], is relatively simple to administer, places morbidity and mortality on the same scale, and discriminates well over the full range of health states.

The results from the telephone survey were highly similar to those that Kaplan had obtained in face-to-face interviews, with the exception of three items. If these three items are excluded, the correlation between the Oregon weights and Kaplan weights is .92. On one function item, however, (confinement to a bed or wheelchair) the Oregon weight was higher than Kaplan's; and two symptom items (loss of consciousness and paralysis or fracture of the leg) the Oregon weights were lower. The Oregon weights for these three items were confirmed, and the Kaplan weights disconfirmed, by independent judgments obtained from Commission members. The difference between the Kaplan and Oregon methods of analysis may account for the differences on these items. Kaplan employed regression analysis on a data set with unknown multicollinearity and arbitrarily assigned any ambiguous weight to symptoms, rather than functions[7]. The Oregon method, on the other hand, avoided ambiguity in the weights by focusing on differences between matched scenarios. It can thus be argued that the results of the Oregon telephone survey are at least as valid as those obtained by Kaplan from one-on-one interviews.

Expert Judgments of Fact

The QWB served as the interface between the quantitative public judgments of value and expert judgments of fact. Health experts stated the probable consequences of treatment and non-treatment in terms of the same functions and symptoms to which representatives of the public had assigned QWB weights. Panels of licensed health care providers in over 54 specialty areas projected the probable consequences of designated treatments for specific conditions'. Experts expressed consequences in terms of up to five probabilistic scenarios, specifying each scenario in terms of QWB functions and symptoms, duration of health state, and scenario probability. The use of multiple scenarios made it possible to take into account the probability of return to previous health state (including spontaneous remission), the probability of death, and the probabilities of various levels of morbidity (including side effects of treatment).

The Benefit/Cost Formula

Ideally, the fact and value data would be combined by means of a formula.

A formula would make the basis for prioritization explicit, could be implemented by computer to handle large data sets and support sensitivity analyses, and would yield a single number to serve as the basis for ranking. In addition, a formula would decompose the problem into simpler problems, making it possible to assign judgment tasks to appropriate persons and enabling individual judges to take more complete and appropriate account of relevant information.[9,10,11].

The initial formula was a benefit/cost ratio, commonly used in cost-effectiveness analyses. Prioritization on the basis of benefit/cost ratios assures the maximum benefit for the dollars expended[12], provided all funds are expended and there are no significant synergisms. In the present case, all funds would be expended; and, to minimize synergisms, "treatment" was defined comprehensively to include diagnosis, drugs, and durable medical equipment, in addition to the primary treatment.

The version of the formula that the Commission initially sought to implement is as follows:

$$B/C = \frac{\Sigma pQALYS^{+} - \Sigma pQALYS^{o}}{costs^{+} - Costs^{o}}$$

" QALYS " refers to Quality-Adjusted Life Years, which is computed by summing QWB over remaining years of life. The Quality-Adjusted Life Year (QALY) is the measure of level and duration of health benefits that is in general use[12]. Loomes and McKenzie[13], even after criticizing it, acknowledge they know of no better measure, and Menzel[14], after weighing several competing arguments focusing on the duration component of QALYs, arrives at a similar conclusion. Ideally, QALYs would be computed by determining some measure of health status, such as the QWB, for each year and summing across years. As an approximation, QALYs were computed by multiplying "representative" QWB following treatment by duration in years, where "representative" QWB was taken to be that present five years after treatment.

Multiplying QWB by the number of years over which the condition continues assumes, in effect, that QWB is constant over time. Gafni & Torrance" have suggested two other possibilities: adaptation and sensitization. If people become adapted to a loss, the amount by which QWB diminishes is a decreasing function of time. If, on the other hand, people become increasingly incapacitated by the loss, the amount by which QWB diminishes is an increasing function of time. The fact that no non-trivial relationships were found in the survey between amount of experience with a condition and rated importance of that condition provides some support for the assumption that QWB does not change with time.

Expected QALYS is computed by multiplying the probability of each scenario by the QALYS for that scenario and then summing over scenarios. Expected utility is widely accepted as the rational way to take account of multiple probabilistic scenarios in decision making[16]. In the formula, "QALYS+" refers to QALYS with treatment, and "QALYS°" refers to QALYS without treatment. By subtracting QALYS° from QALYS', the formula focuses specifically on the marginal effects of treatment, with the result that a health service provided to a disabled person would have precisely the same marginal benefit as that same health service provided to any other person. The numerator is, thus, expected QALYS with treatment minus expected QALYS without treatment. The denominator is the cost of treatment minus the cost of non-treatment, though, ideally, it would be expected cost with treatments minus expected cost without treatment. In some cases, it may be more costly not to provide a treatment than to provide it.

The Categories

The Commission changed its basis for prioritization a number of times. Early on, the Commission shifted its attention from the 709 condition-treatment pairs to a grouping of these pairs into 17 categories, based on probability of death, expected improvement in QWB, and whether the condition was acute or chronic'. These categories were ranked on the basis of three attributes: Benefit to the Individual, Benefit to Society, and Basic Health Care. These three attributes were, in turn, based judgmentally on overlapping sets of the 13 values identified at the community meetings and were evaluated entirely by the members of the Commission, without distinguishing formally between facts and values. (There is no evidence that all these values actually influenced Commission judgments. On the contrary, there is abundant external evidence[17,18,19] to suggest that the human mind is incapable of processing wholistically as much information as evaluation of the categories would have required.) To rank within categories, a reduced version of the formula was employed.

Thus began a gradual decline in which duration of benefit, all measures of morbidity and eventually even the categories were dropped from consideration, and the combining rule deteriorated from the B/C ratio to a noncompensatory, lexicographic decision rule[20].

What led the Commission finally to abandon both the categories and the formula was pressure from the federal government, which directed that quality of well-being not be taken into account at all! But what led the Commission to distrust the formula initially were discrepancies between the ranking produced by the formula and that suggested by intuition. Three categories of discrepancies merit discussion: efficiency, benefit to society, and the QWB.

Efficiency. In evaluating health services, members of the Commission were unaccustomed to dividing benefits by cost, probably because such division is ordinarily inappropriate for clinical decisions. It was clear from a number of exercises in which the Commission participated that they tended not to take costs into account and, when thinking directly about costs, greatly underestimated them. At the extreme, Hadorn has argued vigorously for ignoring costs, altogether[21].

Yet Beauchamps & Childress[22]; Eddy[23]; Weinstein & Stason[12]; and many others have argued effectively for incorporating costs. For example, while a single heart transplant and a single appendectomy each save one life, the $300, 000 it costs to save one life by means of a heart transplant would save 37 lives if spent on appendectomies. Eddy[23] has made the point well: "If you want to check the results against your intuition, you should compare the volumes of different services that can be offered with a particular amount of resources" (p. 2138). Beauchamp and Childress[22] argue that cost-effectiveness analysis and its near relatives "may be morally required if they can help make reasoned assessments and judgments about trade-offs more perspicacious than they otherwise might be" (p. 228). Weinstein, Fineberg, et al.[24] state that "Above all, economic concerns bring into play the perspective of 'society,' or the aggregate of all patients present and future" (p. 229). In practical contexts, use of benefit/cost ratios, by improving the decision process, have been shown to increase enormously the actual benefit realized per dollar spent[25].

Cost-effectiveness analysis is considered by many to be inappropriate for decisions involving human lives because of the great weight placed on cost factors. However, cost-effectiveness analysis places no weight at all on cost vis-a-vis human life and health. All measures of benefit are in the numerator of the B/C ratio. The role of cost in the denominator is solely to ensure the greatest benefit, in terms of human values, given the available budgeted resources. At no point are cost factors traded off against non-cost factors (as they would be in benefit-cost analysis).

Benefit to society. An intuitive review of the formula-produced ranking suggested adding to the formula a judgmental component for effects on the treated individual's family and on the larger society. This component would have included control of contagious diseases, increases in productivity, decreases in crime, and control of population. Before this could be attempted, however, attention had turned to the categories.

The QWB ratinas. An intuitive review also suggested that too much weight was placed on trivial conditions, such as wearing eyeglasses. Subsequent research suggested that this was a consequence of employing a numerical rating method, rather than a tradeoff judgment method. For example, the mean rating for eyeglasses is .05 on a scale that extends from 0 for death to 1.00 for good health, yet not one person out of over 100 tested would trade one life for as few as 20 eyeglasses! In general, the tradeoff method places less weight on conditions ordinally far from death and greater weight on conditions close to death. Eddy[23] and Menzel[26] have suggested that the tradeoff method is to be preferred for public policy decision making, because it is more meaningfully related to the use to which the judgments will be put, allocating health benefits across people. In addition, it is subject to fewer judgmental biases[27] and seems to discriminate more appropriately among health states.

Recommendations

The Commission should revisit the formula and attempt to refine it. In particular, tradeoff methodology should be considered for measuring health status, whether or not the QWB is employed, and thought should be given to adding a judgmental component for benefit to society. Unlike the category system, the formula analyzes the problem in a manner that allows judgment tasks to be assigned to appropriate persons, enables individual judges to take more complete and appropriate account of relevant information, and opens the prioritization process more fully to public review, and it combines information in a manner that should result in greater health status benefits. These strengths are too great to be given up lightly.

References

1. Garland M, Hasmain R: Health care in common; Setting priorities in Oregon. Hastings Center Report 1990;20:16-18.
2. Patrick D L, Bush J.W, Chen M M: Toward an operational definition of health. Journal of Health and Social Behavior 1973;14:6-23.

3. Kaplan R M, Anderson J P: A general health policy model: Update and applications. Health Services Research 1988;23:203-235.
4. Kaplan R M: Human preference measurement for health decisions and the evaluation of long-term care. Lexington, MA: Lexington Books, 1982;157-188.
5. Kaplan R M, Bush J W, Berry C C: Health Status: Types of validity and the index of well-being. Health Services Research 1976;11:478-507.
6. Kaplan R M, Bush J W, Berry C C: Health Status Index: Category rating versus magnitude estimate for measuring levels of well-being. Medical Care 1979;17:501-525.
7. Kaplan R M: Personal Communication. 1989.
8. Oregon Health Services Commission: Prioritization of Health Services: A report the Governor and the Legislature. Portland, OR: Oregon Health Services Commission, 1991.
9. von Winterfeldt D, Edwards W: Decision Analysis and Behavioral Research. Cambridge: Cambridge U. Press, 1986;364.
10. Schum D A, Martin A W: Empirical studies of cascaded inference in jurisprudence: Methodological considerations (Tech. Rept. No. 80-01). Houston, TX: Rice U, Dept of Psychology, 1980.
11. Schum D A, Martin A W: Assessing the probative value of evidence in various inference structures (Tech. Rept. No. 81-02). Houston, TX: Rice U, Dept of Psychology, 1981.
12. Weinstein M C, Stason W B: (1977). Foundations of cost-effectiveness analysis for health and medical practices. The New England J Med 1977;296:716-721.
13. Loomes G, McKenzie L: The use of QALYs in health care decision making. Soc Sci Med 1989:28:299-308.
14. Menzel P T: Medical Costs, Moral Choices: A philosophy of health care economics in America. New Haven: Yale U Press, 1983;188-193.
15. Gafni A, Torrance G W: Risk attitude and time preference in health. Management Science 1984;30:440.
16. Raiffa H: Decision analysis. Reading, MA: Addison-Wesley, 1968.
17. Miller G A: The magical number seven plus or minus two: Some limits on our capacity for processing information. Psychological Review 1956;63:81-97.
18. Goldberg L R: Man versus model of man: A rationale, plus some evidence, for a method of improving on clinical inferences. Psychological Bulletin 1974;73:422-432.
19. Dawes R M, Corrigan B: Linear models in decision making. Psychological Bulletin 1974;81:95-106.
20. Oregon Health Services Commission: Prioritization of Health Services: A report to the Governor and the Legislature. Portland, OR: Oregon Health Services Commission, 1993.
21. Hadorn D C: Setting health care priorities in Oregon: Cost effectiveness meets the Rule of Rescue. JAMA 1991;265:2218-2225.
22. Beauchamp T L, Childress J F: Principles of Biomedical Ethics, 3rd ed. NY: Oxford U Press, 1989.
23. Eddy D M: Oregon's methods: Did cost-effectiveness analysis fail? JAMA 1991;266:2135-2141.
24. Weinstein M C, Fineberg H V, Elstein A S, Frazier H S, Newhauser D, Neutra R R, McNeil B J: Clinical Decision Analysis. Philadelphia: W B Saunders, 1980.
25. Anderson B F: Toward an Improved Human Resources Prioritizing System. Portland, OR: Institute for Oregon Policy Studies, Portland State U, 1982.
26. Menzel P T: Strong Medicine: The ethical rationing of health care. NY: Oxford U Press, 1990;84-86.
27. Poulton E C: Bias in quantifying judgments. Loudeu: Lawrence Erlbaum, 1989.

DISCUSSION: ETHICAL ASPECTS OF RATIONING BASED ON COST-BENEFIT ANALYSES

Paul T. Menzel, Pacific Lutheran University

Economic analyses of health care often raise thorny conceptual and ethical issues. It is crucial that in doing, publicizing and using them we are aware of both their economic and ethical limitations. I will comment on a number of such limitations in the papers in this session by David Hamburger, Douglas Murray, and Barry Anderson.

Hamburger concludes that Medicaid provision of child safety seats and seat belts would pay for itself in the long run. He appropriately notes the sensitivity of that conclusion to his assumptions about those devices' effectiveness and the degree to which they would be used. He also notes the sensitivity of his analysis to the distinction between direct and indirect net costs. If only direct net costs (direct costs minus direct savings) are considered, the maximum cost-effective price ("break-even cost") of a safety seat is only $17.58, hardly close to the $35 Medicaid would actually have pay to provide one. Once indirect net costs are included, however (e.g., the savings in later earnings not lost by the injury of a child), the break-even cost becomes $84.31 and the $35 actual price a bargain. So his conclusion is highly dependent on the correctness of his estimate of indirect costs.

Here two problems may afflict his calculations. First, the indirect savings in otherwise forgone earnings not lost should be time-discounted (he does not). That discounting will lower the estimated savings from preventing children's injuries. Second, the net savings from preventing a young child's death are lower than from preventing the death of a teenager: one has to invest more in supporting a younger child before she gets to the the earning stage.

But in any case, we undoubtedly ought to prevent these injuries even if the break-even cost is less than what we have to pay for safety seats. After all, the most important thing has yet to be considered: the value of the reduced pain and suffering and the preserved life that the safety seats accomplish. Hamburger, of course, knows this. My concern, however, is that we ought to write and talk about our cost-effectiveness analyses as if they were not just part of a cold, mechanical calculation. Cost-effectiveness analysis should serve only to clarify our choices that involve other less quantifiable factors. If society sees us as using cost-effectiveness as a method unto itself, it will get rightly suspicious that we are resorting to the easy crutch of a precise, numerical method to the detriment of the human values involved.

The general applicability that Hamburger alleges for his model is also more treacherous than he may think. Used on elderly or near-elderly populations, we run into another issue. While saving younger lives creates indirect savings (preserved later earnings), saving lives much later in life does almost the opposite: greater pension pay-out obligations, and greater health care expenses incurred in the extended years of life. This is why smokers, for example, cost the rest of the population very little money; the extra costs they run up in health care for smoking related diseases, in lost worktime from their illnesses, and in the lost earning years from smoking related mortality is almost completely counterbalanced by the savings in dying, say, at 70 instead of 80 (Manning et al., 1989). Or to put it the other way around, there are costs of quitting smoking, or of not ever smoking: longer likely years of retirement, requiring considerably more pension payouts, plus all the expenses for unrelated health problems that nonsmokers still run into in their 70's and 80's.

Murray's paper on the cost-effectiveness of expanding Medicaid elitigility for family planning services may raise fewer questions about the economic analysis itself, but it brings us to the edge of even larger moral guestions. What is left out in the economic analysis alone, of course, are questions like the following. In human terms, what is a child worth? Enoush to justify not planning its non-existence? And in human terms, what is the avoidance of an unwanted pregnancy and 'birth worth?

These are questions that should be answered by parents who either choose to have a child or avoid having one. Economic analysis need not deny this for a minute, but can in fact complement it. If giving prospective parents more choice by expanding the availability of family planning services does not compete for government funds with something else because it saves the state enough expense elsewhere, then of course they should be given the widened choice.

The one caveat here might be the possibility that the cultural and economic situation is one in which the society badly needs more children--for example, for its labor force. If it did, though, its first choice should be to encourage families to have children (by sharing more the costs of

childrearing, e.g.), not fail to give them a choice not to have children.

In any case, family planning services ought to be driven fundamentally by the value of parental choice, not economic expense or saving. Nothing, of course, that Murray says denies this, but it is nonetheless important for a cost-effectiveness analysis like his to be placed against this moral backdrop. It should not appear to use it as a decision vehicle by itself.

Anderson's analysis of the Oregon plan is completely on target in trying to separate more cleanly the roles of scientific medical data on the one hand and people's values about life and health on the other. The Oregon plan is probably indeed the mess that he describes, laudable as the state's attempt is to directly and publically take up the rationing issue. For one thing, the cost of services was considered at first, but then because the early commissioners could not stand its effects on the prioritization outcome, it got deleted. In the last and current version of their prioritization, cost get some role back, but only in a lexically secondary place: to break ties between treatments that .are equally likely to prevent death. Any number of examples can be constructed to show that none of us would choose such a principle if we thought about it. Would anyone prefer to be covered for one $1 million treatment that cured a rare lethal disease which we all had equal risk of contracting, rather than be covered for hundreds of very inexpensive treatments that each reduced the risk of death from their respective conditions by only 20 percent? The aggregate cost of one coverage set is equal to that for the other, but since the cost/treatment of the things in the second one is so much lower, each of us gains much more of a total reduction in our risk of death from it.

At an intermediate stage the commissioners subjectively rated the seventeen large categories of condition/treatment pairs by not only the conceptual benchmark of "benefit to the individual" but "essential to basic care." We should be using considerations of benefit to individual and society to form our judgments on what basic care is, not importing pre-existing judgments on that score at the start.

Also at an intermediate stage the Quality of Well-Being data was collected by asking everyone without targeting more intensively those who had some experience with and real knowledge of a condition. Questions ought also to have been used that made much clearer the trade-off uses to which the elicited data would be put. Considering quality of life is mandatory, even for the disabled--they surely do not want a treatment that greatly improves their quality of life to get no credit in the prioritization scheme [Menzel, 1992].

On all of that Anderson and I agree. My one significant point of disagreement with his analysis is on something quite different. He claims that the cost-effectiveness used at points in Oregon's plan (and the greater amount of it that he would have preferred) does not weigh dollars off against life or health. This reflects a standard point made by economists: cost-effectiveness analysis (CEA) only shows us how we can get more lifesaving or health improvement for our money; unlike cost-benefit analysis (CBA), it does not tell us when to stop spending on life or health, since unlike CBA it does not assign a surrogate dollar value to either.

The standard point is indeed correct, but it does not so cleanly get Anderson his claim about Oregon's plan. The context of the CEA influenced prioritization, is the need to draw a line on what Medicaid provides with a given amount of money appropriated by the legislature. When Oregon legislators and citizens see what services got in and what got left out, they are then left with the business of deciding whether the amount of money appropriated was sufficient or not. And that, given that they have the cut-off line that now results from a particular appropriation right in front of them to see, becomes a judgment about whether the amount of health and life bought with the services near the cut-off line is sufficient to justify increasing the appropriation or, on the other hand, so meager that they should consider reducing the total money available.

This is not an unusual context for CEA. While CEA itself does not make any value-of-life judgments, it is often directly part of the process in which we do have to make those kinds of decisions. Those who do and use CEA must own up to that or the society will see them as insensitive technocrats who pretend that some quantitative formula solves our problems. CEA helps, often immensely, to clarify what it is that we are really doing with our money, but it is no substitute for public discourse about what life and health are worth in terms of what assets we are prepared to invest in them.

References

Manning, W.G.; Keeler, E.G.; Newhouse, J.; Sloss, E.M.; and Wasserman, J. "The Taxes of Sin: Do Smokers and Drinkers Pay Their Way?" Journal of the American Medical Association 261: 11 (March 17, 1989): 1604-1609.

Menzel, P.T. "Oregon's Denial: Disabilities and Quality of Life." Hastings Center Report 22: 6 (November 1992): 21-25.

Session V

Comparable Methods and Data

Toward the year

2000

Refining the measures

MONITORING AND PRIORITIZATION OF CDC's CONSENSUS SET OF HEALTH STATUS INDICATORS BY STATE HEALTH DEPARTMENTS IN THE UNITED STATES.

Sharon L. Zucconi, PhD; Catherine A. Carson, PhD; and Donald R. Mattison, MD
University of Pittsburgh,

In July 1991, a consensus set of 18 health status indicators (Table 1) and a list of priority data needs to augment the indicators were released by the Centers for Disease Control and Prevention (CDC) to assist communities in assessing their general health status[1]. These indicators addressed the requirements of the Year 2000 Objectives Planning Act (Public Law, 1990) and Objective 22.1 of Healthy People 20002. The intent of these indicators was to provide a broad indication of the general health status of a community'.

Priority in selecting the indicators was given to measures for which data are readily available and that are commonly used in public health'. It is not known however whether these indicators are actually being monitored within each state. Particularly lacking in the literature is an assessment of the usefulness of the ***Consensus Set of Health Status Indicators*** for monitoring health status at the state and local level. Therefore, we surveyed health departments in the 50 states and the District of Columbia to determine how the CDC's ***Consensus Set of Health Status Indicators*** are being assessed in the United States.

I. METHODS

A four-page survey was mailed to directors of state health departments, as well as the District of Columbia (D.C.), to elicit information about efforts to monitor and prioritize CDC's ***Consensus Set of Health Status Indicators.*** Questions were also asked about sources and availability of data, as well as the adequacy of funds available in the health department to monitor the ***Consensus Set of Health Status Indicators.*** Health department directors were also queried about their provision of data (mortality, incidence and risk factor data) to local health departments and hospitals, and the associated usual fees and frequency of provision of these data.

Table 1

CDC's Consensus set of indicators* for assessing community health status and monitoring progress toward the year 2000 objectives -- United' States, July 1991

Indicators of health status outcome

1. Race / ethnicity-specific infant mortality, as measured by the rate (per 1000 live births) of deaths among infants < 1 year of age

Death rates (per 100,000 populations)** for:

2. Motor vehicle crashes
3. Work-related injury
4. Suicide
5. Lung cancer
6. Breast cancer
7. Cardiovascular disease
8. Homicide
9. All causes

Reported incidence (per 100,000 population) of:

10. Acquired immunodeficiency syndrome
11. Measles
12. Tuberculosis
13. Primary and secondary syphilis

Indicators of risk factors

14. Incidence of low birth weight, as measured by percentage of total number of live-born infants weighing < 2500 g at birth
15. Births to adolescents (females aged 10-17 years) as a percentage of total live births
16. Prenatal care, as measured by percentage of mothers delivering live infants who did not receive prenatal care during first trimester
17. Childhood poverty, as measured by the proportion of children < 15 years of age living in families at or below the poverty level
18. Proportion of persons living in counties exceeding U.S. Environmental Protection Agency standards for air quality during previous year

*Position or number of the indicator does not imply priority
**Age-adjusted to the 1940 standard population
Source: Centers for Disease Control and Prevention. ***Morbidity and Mortality Weekly Report.*** 1991; 40:449-451.

Tabulations and frequency distributions were used to describe state health departments' compliance with **CDC's *Consensus Set of Health Status Indicators.*** Prioritization was indicated by low, medium and high. A rank score was computed by assigning a numerical value to each category and summing the individual scores for each indicator; 1, 2, and 3 for low, medium and high, respectively. Analyses were completed with the use of **Statistix®** 4.0 for the PC (Analytical Software, St. Paul, MN; 1992).

II. RESULTS

Thirty eight surveys (74.5 percent) were returned as a result of the initial mailing in the summer of 1992. A second mailing in October 1992 yielded 8 additional responses, representing a cumulative response rate of 90.2 percent. A follow-up mailing to the 5 non-responding states provided a cumulative response rate of 100 percent.

Monitorins and Prioritization of Mortalitv Indicators

As shown in Figure 1, the majority of the ***Consensus Set of Health Status Indicators*** are being monitored in all states. In general, mortality indicators are often monitored. However, only 73 percent of the states monitor work-related injury deaths.

Monitorins and Prioritization of Incidence Indicators

Practically all health departments monitor indicators of AIDS, measles, tuberculosis, and syphilis incidence, which are diseases reportable to the CDC. The reported monitoring and prioritization of each of these indicators is depicted in Figure 2.

Monitorina and Prioritization of Risk Factor Indicators

As shown ***in*** Figure 3, the incidence of low birth weight, births to adolescents and prenatal care are being monitored in nearly all states; these indicators were considered the highest priority in most states. However, ***only 45*** percent are monitoring poor air quality; childhood poverty is being monitored in only 59 percent of the states.

Adequacy of Funds to Monitor Indicators

The survey elicited information about the adequacy of funds available in state health departments to monitor the ***Consensus Set of Health Status Indicators.*** Seventy-five percent of health departments reported that insufficient funds are available to adequately monitor these indicators.

Figure 1. Monitoring and Prioritization of CDC's Consensus Mortality Indicators

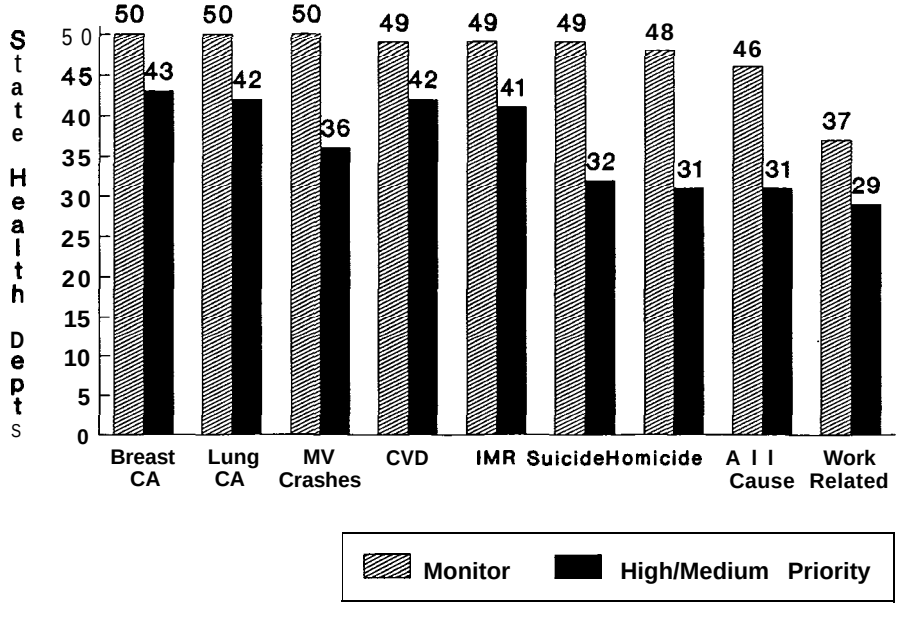


Figure 2. Monitoring and Prioritization of CDC's Consensus incidence Indicators

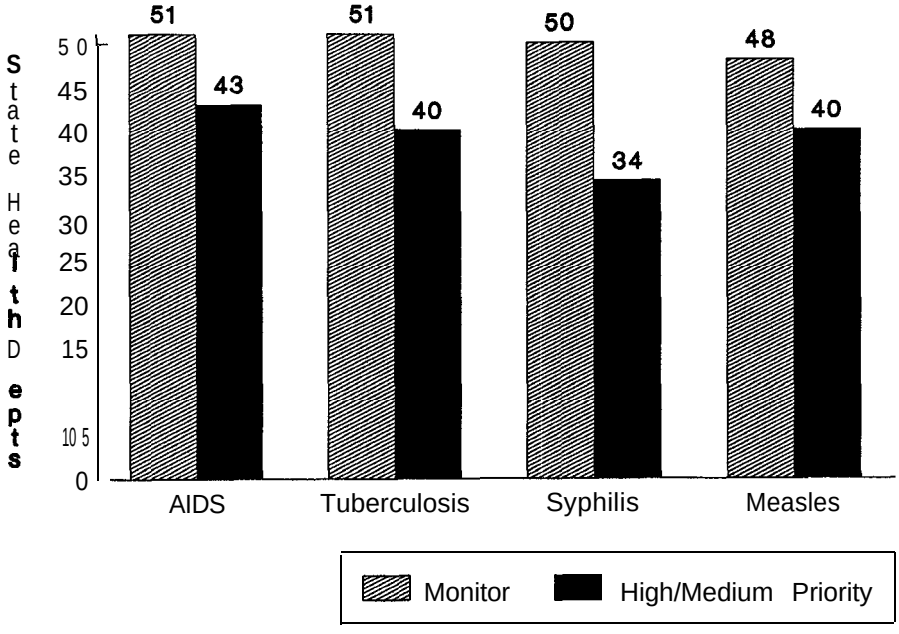
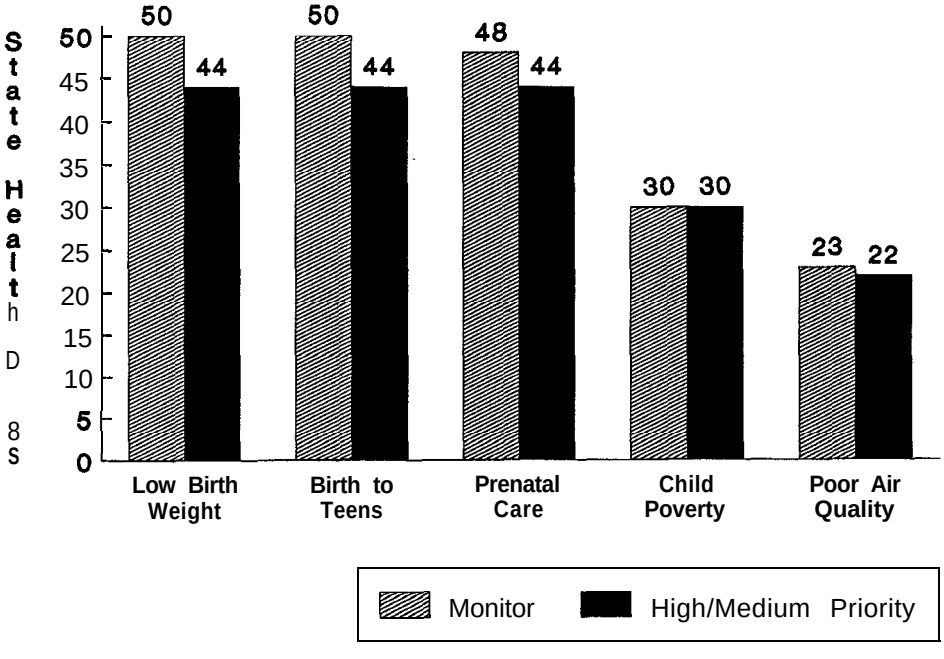


Figure 3. Monitoring and Prioritization of CDCs Consensus Risk Factor Indicators



Sources of Information Used to Monitor Indicators

The most common source of mortality indicator data was vital statistics. Two exceptions are: 1) Fatal Accident Reporting System (FARS) data and police reports for monitoring deaths from motor vehicle crashes; and 2) Worker's Compensation data from the Occupational Safety and Health Administration (OSHA).

Provision of Data to Local Health Departments and Hospitals

Data necessary for the monitoring of health status indicators at the local level are provided routinely by 36 states (71 percent) and on a request basis by 35 (69 percent). The frequency of provision of data depended on the type of data and the organization requesting these data. Data for the mortality indicators were mor'e likely to be provided than either incidence *or* risk factor data. Provision of mortality and incidence data to local health departments was reported by all states with local departments. However, risk factor data were often not available at the county or community level. Approximately one half of the states provided data to hospitals. The majority of health departments (n=34, 67 percent) do not charge a fee for provision of data. Several reasons were given for charges, e.g., computer time, special requests, photocopying charges, and personnel time.

III. DISCUSSION

This paper clearly demonstrates the widespread use of CDC's consensus set of 18 health status indicators by state health departments. However, it also highlights problems in data collection, maintenance and resource allocation to effectively monitor these indicators.

In July 1991, a consensus set of 18 health status indicators and list of priority data needs to augment the indicators were released by the Centers for Disease Control and Prevention (CDC), to address Objective 22.1 of *Healthy People* 2000. The *Consensus Set of Health Status Indicators* were identified, in part, because it was felt that data were readily available at national, state, and local levels. The results of this survey indicate that data for some indicators may actually not be readily available. And, the availability of data had a major influence on whether or not states placed a high priority on the need to monitor, and subsequent reported monitoring of indicators.

Our survey assessed efforts by states to monitor the indicators. It did not, however, address the difficulty that local health care reform initiatives may be having in assessing the health of the people in a given service area. As previously noted, states were less likely to monitor a given indicator if data were unavailable, perceived of lesser quality, or of lower priority. Health care institutions often lack the expertise to monitor epidemiologically-based indicators. Also, monitoring by health care institutions requires the data to be coded at the census tract or minor civil division level so the data may be aggregated for their specified service area. This is often considered a special request and data acquisition may not be timely. Mortality statistics are often viewed as poorer outcome measures because of the large time span between intervention and impact on the outcome, i.e, death.

Sources of information used to monitor each objective may vary for some indicators i.e, motor vehicle crash deaths, work-related injury deaths. The use of different data sources introduces the possibility of poor reproducibility of rates; differences in rates across communities may be due to different sources of data rather than actual discrepancies in events.

Differences in monitoring may be artifactual, i.e, events may be recorded but not considered an indicator unless health departments actually use the information for program planning and/or resource allocation. For example, all states have access to mortality data but they may not be used for health status indicators. The disparity between data collection and use of data for program planning and evaluation probably is greater at the local health department level. And, it is at the community level that most program planning and evaluation actually occurs.

States were often not monitoring two of the indicators in the *Consensus Set of Health Status Indicators:* childhood poverty; and the proportion of persons living in counties exceeding EPA standards for air quality ("poor air quality"). One of the reasons for not monitoring "poor air quality" may be due to a lack of available data for counties, i.e, not all counties have air quality stations. It was frequently mentioned that census data are available indicating the number of children under 18 living in poverty. However, the indicator is for children less than 15 years of age living in poverty. A solution to this problem would be to either have current population survey data compiled for children less than 15, or change the indicator to the proportion of children under 18 living in families at or below the poverty level.

Objective 22.1 states that the value of public health agencies in assessing the health status of the population is "enhanced when comparisons to some benchmarks, such as other similar populations or to national norms, can be made."[2(p. 554)] The lack of targets for some of the indicators may be a deterrent to

monitoring of the indicator, e.g., there are no targets for three of the 18 indicators -- all cause mortality, childhood poverty, and "poor air quality". Also, some of the available targets may not be appropriate for a state, e.g., a state with a predominately white population may not want a target for the total population, but would need a target for just the white Population. Conversely, many states would need the expansion of targets for race-ethnic specific indicators. Although definitions and national data have been recently published[4] for the *Consensus Set of Health Status Indicators,* technical support from the CDC is needed to assist communities in establishing local targets for the indicators.

We found differences in the perceived adequacy of funds available in each state health department to monitor the *Consensus Set of Health Status Indicators.* These findings suggest that in order to have consistent data monitoring at the state level, the Federal government should provide funds to enhance data collection abilities. Seven states have been given *Assessment Initiative Funds* by the Centers for Disease Control and Prevention' and grants from foundations are being used to enhance data collection efforts in several states. However, while several states may be receiving multiple sources of funding to enhance their capacity to use data for policy development and programmanagement, many other states are not receiving any funds.

Questions not addressed in the survey that need to be further examined include: usefulness of indicators in monitoring health status of a community to assess the effectiveness of health care reform initiatives at a local level; and evaluation of the form in which data are given to local health departments, i.e., level of computerization, ease of use, and how this affects the monitoring of indicators at the local level.

As stated in Objective 22.1 the indicators should include measures needed for area-wide health planning and allocating health resources among programs, areas, or regions. "The indicators should also provide measures needed to judge the results of health planning and to study changes in health over time"[2(p.554)]. This set of indicators should be evaluated periodically to serve as a broad indication of the general health status of a community. The key to evaluating progress and comparing results across geographic, political, and programmatic lines is comparability of data. CDC's *Consensus Set of Health Status Indicators* are widely recognized by the states and are in effect being monitored rather consistently, except for only a few indicators. This is a first step in achieving comparable data methods. However, the status of these efforts should be assessed often, as funding changes and indicators are revised.

The results of this survey are useful for assessing the current and future impact of the *Consensus Set* of Health *Status Indicators.* They also indicate a need for resource allocation by the Federal government to provide support to health officials at national, state, and local levels for assessment of these health status indicators in standard fashion for comparison across the nation, and for policy development and program planning for preventive and primary health services.

REFERENCES

1. Centers for Disease Control. Consensus set of health status indicators for the general assessment of community health status - United States. *MMWR;* 1991; 40(27):449-451.

2. Public Health Service. *Healthy People* 2000: *National Health Promotion and Disease Prevention Objectives.* Washington, D.C.: U.S. Department of Health and Human Services, Public Health Service, 1990; DHHS publication no. (PHS)91-50212.

3. Klein RJ, and Hawk SA. Health status indicators: Definitions and national data. *Healthy People 2000 Statistical Notes from the Centers for* Disease *Control/National Center for Health Statistics;* 1992; 1(3):1-8.

4. Centers for Disease Control and Prevention/NationalCenter for Health Statistics. CDC assessment initiative funds seven states. *Healthy People 2000 Statistics and Surveillance.* 1992; 1(4):1-6.

A NONLINEAR REGRESSION MODEL FOR THE PROJECTION OF EXPECTED VALUES FOR SELECTED HEALTH STATUS INDICATORS IN THE YEAR 2000

Chun-Lo K. Meng, Carol J. Garrett and Joseph D. Carney
Colorado Department of Health

Healthy People 2000 sets forth many different health status objectives for the nation. These objectives, ranging from teen pregnancy to heart disease deaths, from low birth weight to suicide, provide a series of well-defined and measurable targets for the nation to achieve by the Year 2000. As part of the state public health agency, we feel it is also very important to develop asetofhealthstatusobjectivesatthestatelevel. These statelevelobjectivescanthenbeusedtonkmitorprogress and assess the effects of state health planning and policy efforts.

The Year 2000 national objectives were established (as quoted from Healthy People 2000) "through the consensus of expert physicians, public health professionals, and epidemiologists" and "on the basis of trend evaluation and expert judgement." There are four main approaches to setting health status objectives at the state level: (1) adopt the national objectives without change, (2) adapt the nationalobjectivesto reflect state baseline data; (3) consult local experts; and (4) perform trend analysis. This papertookths last approach and used statistical modeling to predict the expected values of selected health status indicators for the Year 2000. These expected values were then compared to what the Colorado objectives would have been had they been adapted fran national objectives. Itishopsdthatthismathodcan inform the establishment of health status objectives appropriate for the state and provide an indication of the effort needed to achieve the objectives.

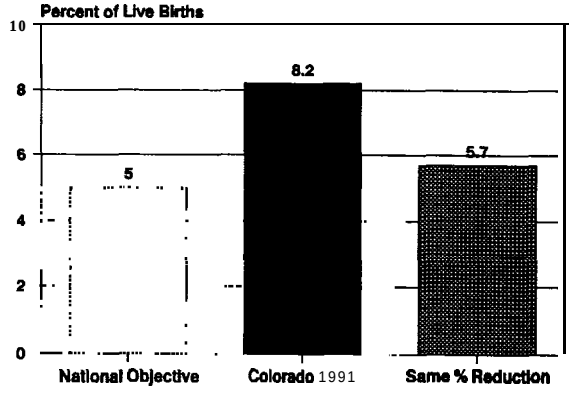
We first calculated yearly levels of 16 selected health status indicators for Coloradobyusingdata from Colorado birth and death certificate files. These indicators included incidence of low birth weight; percentage of pregnant women receiving first trimester prenatal care; fertilityrateamorqadolescentgirlsaged 15 through 17; infant mortality rate; neonatal mortality; and age-adjusted death rate for cancer, lung cancer, female breast cancer, coronary heart disease, stroke, chronic obstructive pulmonary disease, diabetes, unintentional injury, motor vehicle crashes, homicide and suicide. The definitions of these health status indicators were exactly the same as those used in **Healthy** People 2000 and/or various articles published by NCHS. U.S. 1940 population data were used as the standard for the calculationof all age-adjusted deathrates. Annual percentages or rates as appropriate were calculated for the period 1980 to 1991. Although we did examine the indicators for different race/ethnicity groups, only those for all races arepresentedhere.

During this 12-year period, there were seven indicators that showed no significant trend. They were: incidence of low birth weight; percentage of pregnant women receiving first trimester prenatal care; and age-adjustedrate for total cancer, breast cancer, diabetes, homicide, and suicide. Therefore, no trend analyses were done for these indicators. We did, however, compare the latest Colorado data to national objectives to see if national objectives would be appropriate for Colorado.

For sane of these seven indicators, the national objectivemaybedifficult for Colorado to achieve by the Year 2000. An example of this is the incidence of low birth weight (Figure 1). The percentage of births with low weight birth has remained very steady in Colorado since 1980. It started as 8.2 in1980 and ended as 8.2 in 1991. The national objective is to reduce low birth weight to an incidence of no more than 5 percent of total live births. This implies a 27.5 percent reduction from the baseline year of 1987 to the Year 2000. If we apply the same percentage reduction to Colorado's 1987 baseline rate, then the objective for Colorado would be 5.7 percent. This method of adapting national objectives to Colorado will be called "same percentage reduction" method and will be used throughout this paper. In **Figure 1**, the barlabeled "national objective" indicates the value of the Year 2000 national objective; the bar labeled with "same percent reduction" represents the Colorado objective when the percentage reduction for the nation from 1987 to 2000 is applied to Colorado data. Roth national and Colorado objectives for low birth weight incidence are much lower than what the Colorado rates have bsen since 1980.

Figure 1.

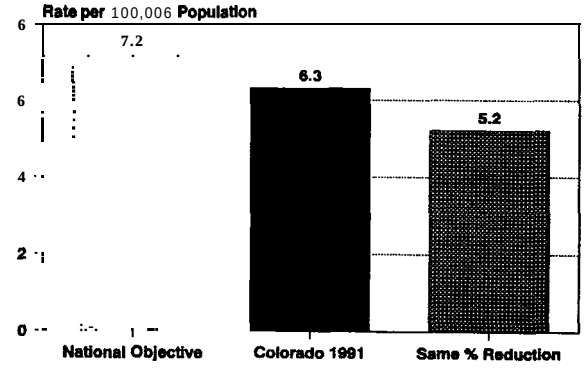
Percentage of Births with Low Birth Weight



For suns other indicators, Colorado had already achieved the Year 2000 national objective. An example of this is the age-adjusted hanicide rate. Colorado's age-adjusted death rate for hanicide was also fairly stable during the 12-year period with a rate of 6.3 in 1991 (Figure 2). The national objective is 7.2 for the Year 2000; the Colorado objective would be 5.2 if the same percentage reduction were applied.

Figure 2.

Age-Adjusted Homicide Rate



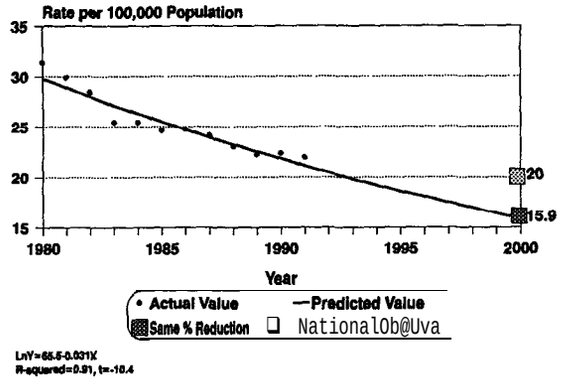
Thosehealthstatusindicators forwhichsignificant trends were identified during the 1980s and early 1990s include the fertility rate among adolescent girls aged 15 through 17; infant and neonatalmrtality rates; and age-adjusted deathratee forlungcancer, coronary heart disease, stroke, chronic obstructive pulmonary disease (COPD), unintentional injury, andmotorvehicle crashes. All these indicators show statistically significant upward or downward trends during the 12-year period. A simple regression time-series model was used for forecastingpurposes. Thehealth status indicatorswere the dependent variables. The year was used as the only independent variable. Linear as well as various nonlinear functional forms (loglinear, semilog and polynomial, etc.) were tried for each indicator and measures of goodness of fit were compared to determine whichoneto use for the projection. For most of the indicators, semilog specification had the best measures of goodness of fit, including R-squared and t-value, for Colorado data. Parameter estimates derived fran the regression' model then were used to calculate the predicted values of the health indicator for the Year 2000.

The comparison of thepredictedvalues for Colorado with the values adapted fmn the national objectives yielded three different results: (1)thepredictedvalue was very close to the objective adapted from national objective; (2)thepredicted value was not very close but was in the same direction as the objective; and (3) the predictedvalueandobjectivewereindifferentdirections. from the current point. Those indicators that fall into the first category include age-adjusted death rate for stroke, lung cancer, COPD, infant mortality rate, and neonatal mortality rate.

Colorado's age-adjusted stroke death rate has followed a smooth downward trend since the 1980s (Figure 3). Theregressionresultsusingthesmilog functional form and 1980 to 1991 data indicated an adjusted R-squared of 0.91 and t-value of -10.4. The predicted value for the Year 2000 is 15.9 deaths per 100,000 population. Healthy People 2000 targetsthenationalage-adjusted death rate from stroke at no more than 20 per 100,000 population. This represents a 34 percent reduction between 1987 and the Year 2000. If this 34 percent reduction-e appliedto the Colorado rate of 1987, the objective would also be 15.9 for the Year 2000.

Figure 3

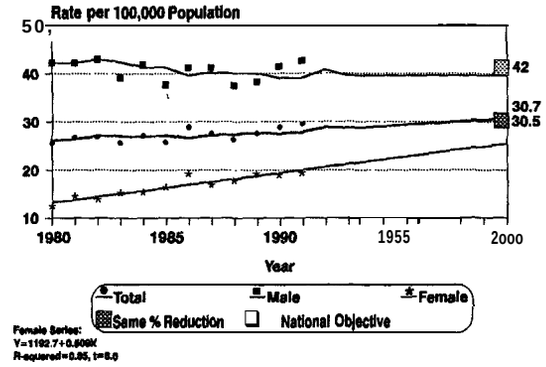
Age-Adjusted Stroke Death Rate: Colorado Residents, 1980-2000



The next example is the indicator of age-adjusted lung cancer death rate (Figure 4). Although the rate for male has been consistently higher than that for females, nosignificanttrendwasidentifiedforthemale rate during the 12-year period. The female age-adjusted rate, on the other hand, did show a significant upward trend. A linear regression model was applied to the female data (adjusted R^2=0.85 and t=8.0) and the predicted values for males were derived by using 3-year moving average method. The predictedvalues forthetotal age-adjusted lung cancer death ratewerethemighted-average of male and female predicted values. The Year 2000 predicted value would be 30.7 by using the method described above. The national objective is to slow the rise in lung cancer deaths to achieve a rate of no more than 42 per 100,000 people. The objectivewouldbe 30.5 for Colorado if the "same percentage reduction" method were applied.

Figure 4.

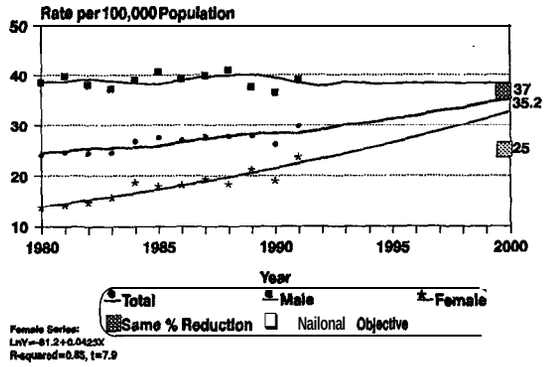
Age-Adjusted Lung Cancer Death Rate: Colorado Residents, 1980-2000



The trend for the Colorado age-adjusted rate of chronic obstructive pulmonary disease (COPD) deaths *(Figure 5)* was similar to that of lung cancer deaths. A semilog regression was applied to female data for the 1980 to 1991 period, and significant statistical results were obtained (adjusted R^2=0.85, t=7.9). The male Predicted values were franthemoving averagemethod. Again, the predicted values for the total were the weighted-averageofmaleandfemalepredictedvalues. The predicted value of COPD for Year 2000 is 35.2, and the Colorado objective would be 37.0 if the "same percentage reduction" method were applied. Roth figures are substantially higher than the national objective of slowingthe rise in deaths from COPD to a rate of no more than 25 per 100,000 people.

Figure 5.

Age-Adjusted COPD Death Rate: Colorado Residents, 1980-2000

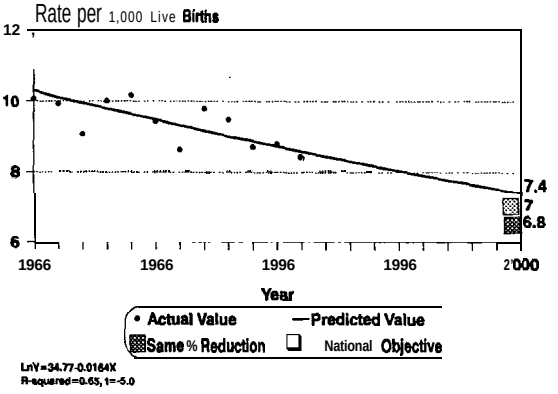


In the last decade and the early 19908, Colorado's infant mortality rate *had a* basically downward trend but thevariation from year to year was significant *(Figure* 6). Various nonlinear spscificationsweretriedto find the most appropriate one. The semilog functional form

was the one that had the best goodness of fit result, although the adjusted R-sguaredwas only 0.65 (t=-5.0). The predicted value for Year 2000 is 1.4 as the result of this model. Thenationalobjectiveis 7.0 and Colorado objective would be 6.8 if the "same percentage reduction" method were applied.

Figure 6.

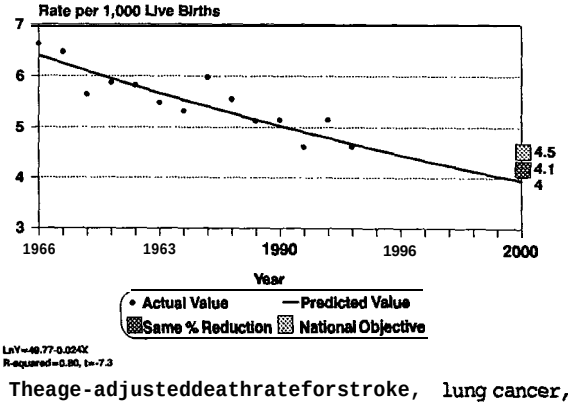
Infant Mortality Rate: Colorado Residents, 1980-2000



The Colorado resident neonatal mortality rate has followed a smoother downward trendthanthat of infant mortality (Figure 7). A semilcg regression model resulted in an adjusted R-sguaredof 0.80 and t-value of -7.3, both measures were higherthanthose for the infant mortality model. The predicted value of the neonatal mortality rate for the Year 2000 is 4.0 per 1,000 live births. Colorado objective would be 4.1 by applying the "same percentage reduction" method; the national objective is 4.5.

Figure 7.

Neonatal Mortality Rate: Colorado Residents, 1980-2000



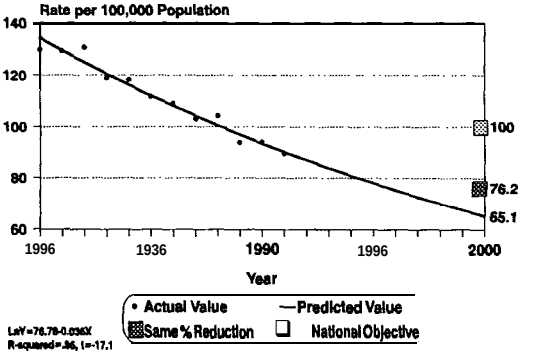
Theage-adjusteddeathrateforstroke, lung cancer, COPD, and the infant and neonatalmxtalityraterepresent five examples where the predicted values from simple time-seriesregressionmodelswereveryclosetotheobjectives adapted fran the national objectives by using "same percentage reduction" method. Host of them also showed verysignificantstatistical resultsas far as R-squared and t statistics were concerned.

There is another group of indicators where the statistical results (R^2 andt-values) frantheregression models were quite good, buttheepredictedvalueswerenot close to the objectives adapted from the national objectives. In at least two of the three, the data suggest a structural change in the trend after 1990. Three examples of this category are age-adjusted death rates for coronary heart disease, total unintentional injury and, motor vehicle crashes.

The first example is the indicator of age-adjusted coronary heart disease deathrate (Figure 8). This rate followed a very smooth downward trend duringthe 1980to 1991 period. Actually the statistical results fran the semilog model were the best among all the indicators that we tried. The adjusted R-squared is 0.96 andt-value is -17.1. When the model is applied for the Year 2000 projection, the predicted value is 65.1. The national objective istoreduce coronary heart diseases to no more than 100 per 100,000 people. Although Colorado's age-adjusteddeath rate franheart disease has alwaysbeen lower than the national rate, the objective, 76.2 calculated from "same percentage reduction" method, is still considerably higherthanthe predictedvalue fran the model.

Figure 8.

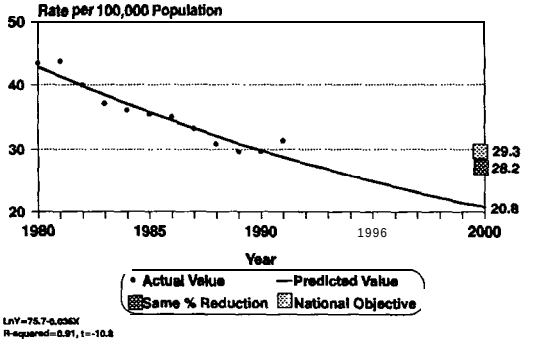
Age-Adjusted CHD Death Rate: Colorado Residents, 1980-2000



The age-adjusted death rate fmn unintentional injury has decreased steadily in Colorado over time (Figure 9), reachingthe lowest points in 1989 and1990 (both were 29.6). It is too early tote11 if the higher rate in 1991 (31.3) was a temporary variation or the beginning of anewtrend. A semilog model applied to the 12-year data resulted in an adjusted R-squared of 0.91, t-value of -10.8 and a predicted value for Year 2000 of 20.k. This predicted value is significantly lower than the national objective of 29.3 and Colorado objective of 28.2 calculated from "same percentage reduction" method. The large discrepancybetweenthepredictedvalue and the objective suggests that expert opinionhe reliedupmto set a realistic objective.

Figure 9.

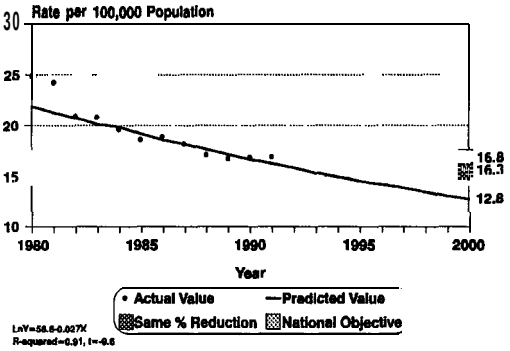
Age-Adjusted Unintentional Injury Death: Colorado Residents, 1980-2000



Rotor vehicle crashes accounted for more than half ofthetotalunintentionalinjury deathsoverthe 12-year period. Therefore,theage-adjusteddeathratefranmtor vehicle crashes followed the samepattern as that of the total unintentional injury deaths (Figure 10). Two outlier points of 1980 and 1981 were excluded fran the regressiontoimproveaccuracy. Althoughthestatistical results from the regression model were good (adjusted R^2=0.91 and t=-9.6), thepredictedvalueof 12.8 is again significantly lower than either of the objectives based on national data. The national objective is 16.8 per 100,000 people and the Colorado objective would be 16.3 if the "same percentage reduction" method were applied.

Figure 10.

Age-Adjusted Motor Vehicle Crash Death: Colorado Residents, 1980-2000



For the above three indicators, the canparisons between the predicted values andobjectives adapted fran the national objectives showsanedegreeof discrepancy. Thedirectionof change is the same, however. The next examplerepresentsacasewherethepredictedvalueand the objective are in opposite directions.

The national objective for teen pregnancies is to reduce pregnancies among adolescent girls aged 15 through 17to no more than 50 per 1,000 population. This implies a 29.7 percent decrease from 1987. Roth live births and induced terminations are included as pregnancies. Underreporting and data incunpletenese have always been major problems for the induced termination file in Colorado. We don't have enough confidence to use these data as part of this health indicator. Instead,teen fertility, the number of live births with age of mother as 15 to17 per 1,000 female population inthie agegroup, is used. Duetothe different definitions, the absolute value Of the national objective (50) is not appropriate here. The percentage of reduction (29.7%) to 20.9 is appropriateundertheassmptionthatinducedterminations represent a constautproportion of pregnancy outcomes overtime. Colorado's teen fertility rate, on the other hand, has followed an upward trend since 1980 arriving atarateof35.7 in 1991. Two nonlinear regression models were tried for the 12 years of data. Figure 11 shows that a semilog functional form yields an adjusted R-squared of 0.69, t-value of 5.0 and predicted value as 39.1.

When the model was changed to the polynomial functional form (Figure 12), both adjusted R-squared and t-values improved greatly (R^2=0.93 and t=-6.2 and 6.4, respectively) but the predicted value for Year 2000 jumped to 61.5. Foreithermdel, the wide gapbetweenpredicted valuefrantrendanalysieaudtheobjectiveindicateethat enormous efforts from family, school, community and public health agencies are needed to achievethetarget.

Figure 11.

Fertility Rate among Adolescent Girls Aged 15-17: Colorado Residents, 1980-2000

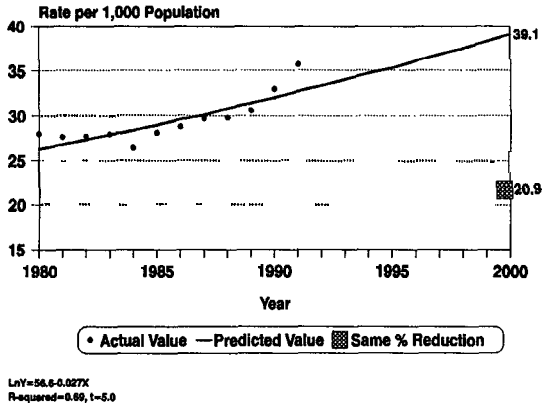
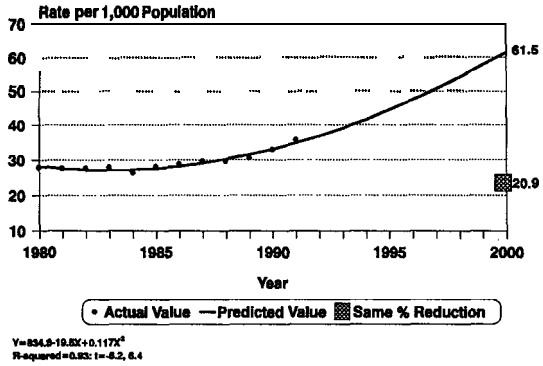


Figure 12.

Fertility Rate among Adolescent Girls Aged 15-17: Colorado Residents, 1980-2000



In summary, our application of simple time-series regression models to stateleveldatatoderiveYear2000 predicted values for various health status indicators suggests the following conclusions:

(1) Formostoftheindicators, theobjectivederived from the "same percentage reduction" method would be more appropriate for state use than the national objective itself.

(2) Formanyoftheindicatorsinwhichsimothtrends wereobserved inthepastdecade, thepredictedvalues frantheregressionmodelareveryclosetotheobjectives adapted fran national objectives. Good statistical results were also obtained franthemdel. Theseresults suggested that the adapted objective would be reached with noincreasedprogramorpolicyeffort. Theydonotsuggest that the "adapted" objective is the appropriate objective.

(3) For those indicatorewherethepredicted values derived from the projection model are not close to the adapted objectives (i.e., motor vehicle crashes), consultation with local experts to set realistic objectives is recommended.

(4) Those indicators where the objectives are in a direction opposite the value predicted from trend analysis (i.e., teenpregnancy), call for the greatest program, planning, and policy efforts. Statistical analysis will not be able to guide the determination of realistic objectives.

COMPUTERIZED DATABASE FOR STRENGTHENING CAPACITY FOR LOCAL AREA NEEDS ASSESSMENT, SURVEILLANCE, AND MONITORING OF YEAR 2000 OBJECTIVES: THE ILLINOIS EXPERIENCE

Merwyn R. Nelson, Illinois Department of Public Health
Paul D. Sarvela, Angela Oldfield

INTRODUCTION AND BACKGROUND

The Illinois Department of Public Health is taking a lead in redirecting the Illinois public health system by implementing the public health functions in local health jurisdictions. These functions were set forth in the Institute of Medicine Report, The Future of Public Health, and the Illinois Public Health Charter. They are: assessment, policy development, and assurance. State and local health departments must develop the capacity for necessary leadership to meet the health needs of populations they serve. In September of 1992, the Department committed funds from the Preventive Health and Health Services Block Grant for a series of planning activities to build capacity within local health jurisdictions resulting in the development of organizational assessments, community health needs assessment, and community health plans. Known as the Illinois Project for Local Assessment of Needs, or IPLAN, it is a major initiative in Illinois. The project started October 1, 1992 and involved a team of faculty from Southern Illinois University and staff from the Illinois Department of Public Health.

IPLAN is an outgrowth of Project Health. Project Health was a four-year effort by state and local public health agencies and public health academia to define and develop strategies for implementing an improved public health system in Illinois. One of the Project's major recommendations was that statewide and local needs assessments should be conducted to provide accurate, concise and defensible information to identify and describe public health needs. The participants in Project Health agreed that needs assessments in Illinois should be structured by a minimum set of standardized content elements to be used in each jurisdiction.

The Project Health participants structured community health indicators, or measurements, into six categories:

- Demographic and Socioeconomic Characteristics
- General Health and Access to Care Indicators
- Maternal and Child Health Indicators
- Chronic Disease Indicators
- Infectious Disease Indicators
- Environmental/Occupational Health and Injury Control Indicators

The health indicators within each of the six categories were developed through Project Health and approved by the Project's governing body. The community health indicators reflect many of the issues and problems of interest to public health agencies, including mortality, natality, morbidity, and prevalence of risk factors in the population. The final selection of indicators was based on what would be measurable at the county level and would be meaningful as a measure. Illinois community health indicators that are consistent with Healthy Peonle 2000 objectives are used for monitoring progress toward those objectives when appropriate. Other year 2000 objectives that are measurable at the state level but could not be measured at the county level were not selected as indicators by the Project Health members. Therefore, the developed system was based on indicators selected by a broad-based group of public health professionals in Illinois and are suited for local needs assessment and surveillance.

DATABASE DEVELOPMENT

To implement local area needs assessment with a uniform set of indicators available, the IPLAN team designed a database and assembled a set of indicators consistent with the efforts of Project Health. We will present the indicators here and a brief explanation of the selection process. The initial screen presents the selection of six categories plus a selection of sentinal events and a set of summary reports.

TIME: 10:30:15 am DATE: 07/18/93

IPLAN Indicator Selection Menu

COOK

1. Demographic and Socioeconomic Characteristics
2. General Health and Access lo Care Indicators
3. Maternal and Child Health Indicators
4. Chronic Disease Indicators
5. Infectious Disease Indicators
6. Environmental/Occupational/Injury Control Indicators
Sentinel Events
3: Summary Reports

Select corresponding number and press ENTER

Press ESC to return to previous menu

The screen for Demographic and Socioeconomic Characteristics includes 11 indicators selected by Project Health. These indicators represent important population characteristics which can have related health attributes. Several of the indicators are obtained from the 1990 Census of Population (Pop. distribution by race and Hispanic origin; Median age; % of non-graduates of High School; Poverty levels; % rural pop; and single parent households) and as such will only be available once for the decade. Dependency indicators (indicator # 1) are obtained from the US Census, but population estimates for total population and population age 65 and older will be produced throughout the decade and that part will be available each year. High School dropouts is available annually from the State Board of Education; Number of persons on food stamps will be available annually from the Illinois Department of Public Aid. The population who are unemployed for 1990 was obtained from the US Census but in subsequent years of the decade that indicator will be obtained from the Illinois Department of Employment Security. The number and percent of medicaid enrollees (and enrollees under age 21) will be obtained annually from the Illinois Department of Public Aid.

TIME: 10:30:15 am DATE: 07/18/93

IPLAN DATA SYSTEM

DEMOGRAPHIC AND SOCIOECONOMIC CHARACTERISTICS

COOK

1. Dependency Indicators
2. Race/Ethnicity Distribution
3. Median Age
4. Non-High School Graduates
5. High School Drop Outs
6. Poverty
7. Food Stamps
6. Rural Population
9. Unemployed
10. Medicaid Enrollees
11. Single-Parent Households

Select Indicator Number:

Press ESC to return to previous menu

The section for General Health and Access to Care indicators provides an overview of health status using general measures of mortality, years of potential life lost and life expectancy. The section also provides general measures of health care access which attempt to quantify the availability and use of basic health services and resources as well as the presence of financial barriers to access. Five indicators are based on vital statistics data (mortality rates; leading causes of mortality; life expectancy at birth; excess non-white deaths; and cause specific years of potential life lost-YPLL). For this section, age-adjusted rates were calculated for 1990 using the 1990 US Census counts for age groups. In 1990 and in subsequent years of the decade, crude rates and premature mortality based on deaths to persons age O-64 will be monitored since our demographic staff find severe accuracy problems with post-censal population estimates by 5-year or lo-year age groups for Illinois counties. For that reason, Illinois will depart from the Healthy People 2000 objectives for monitoring mortality rates. Age-adjusted death rates will be used only at the beginning and at the end of the decade for evaluating county mortality. Leading causes of mortality (and numbers of deaths) will be provided annually by race for counties as will the leading causes of YPLL. Life expectancy at birth is listed here on a tentative basis pending evaluation of a new estimation technique for life expectancy for larger counties. If successfully adapted to measuring Illinois mortality levels, this measure would be calculated annually for all but the small Illinois counties.

TIME: 10:30:15 am DATE: 07/18/93

IPLAN DATA SYSTEM

GENERAL HEALTH AND ACCESS TO CARE INDICATORS

COOK

1. Mortality Rates
2. Leading Causes of Mortality
3. Ufe Expectancy at Birth
4. Excess Non-white Deaths
5. Population Uninsured
6. Cause Spedfic Years of Potential Life Lost at Age 65
7. Percent Population - No Medical Physical in Past 2 Years
8. Medicaid Enrollees to Medicaid Physician Vendors Ratio
9. Advanced Life Support Emergency Care Vehkles
10. Population Residing in Primary Care Heakh Professional Shortage Areas (HPSA)
11. Population with Optimally Fluoridated Water Supplies

Select Indicator Number:

Press ESC to return to Previous Menu

A measure of excess non-white deaths also will be calculated annually for most Illinois counties where there is a significant non-white population. The other measures in this section relate to issues of access to care. Persons who are uninsured will be from data obtained from the US Government Accounting Office periodically through the decade. The data for population with no physical exam by a doctor in the past 2 years will come from the BRFS data and are estimated for combined counties that are rural or counties that are more urban, and counties that are in the Chicago metropolitan area, and for Chicago and for the suburban portion of Cook County. This measure will be updated annually. Medicaid recipients as a ratio to active medicaid physician vendors provides a measure of the availability of physicians working with medicaid recipients. These data are from the Illinois Department of Public Aid. The number of advanced life support emergency care vehicles in the county per 100,000 population and the number and % of population that reside in primary care health professional shortage areas will be obtained from data collected by IDPH.

The section on Maternal and Child Health provides an overview of the key components of maternal, infant, and child health and the risk factors that contribute to ill health and poor outcomes. In addition to the infant mortality rate measures, these indicators include important measures of the risks of death and disability such as incidence of low birthweight and receipt of prenatal care. The section contains 16 indicators, most of which will be available for counties each year of the decade. Because population estimates for age groups will not be prepared for Illinois counties, the percent of births that are to teens will be used as a measure that can be monitored annually while the teen birth rate can be measured for counties at the beginning and end of the decade.

TIME: 10:30:15 am DATE: 07/18/93

IPLAN DATA SYSTEM

MATERNAL AND CHILD HEALTH INDICATORS

COOK

1. Live Births
2. Infant Mortality
3. Low Birthweight
4. Mothers Who Smoke
5. Mothers Who Ddnk
6. Kessner Index
7. Mothers Begin Prenatal Care in First Trimester
9. Infant Urine Tests Positive for Cocaine
9. Leading Causes of Mortality (Children 1-4)
10. WIC - Low Weight For Height
11. Teen Birth Rate
12. Percent Births to Teens
13. Child Abuse / Neglect
14. Congenital Anomalies
15. Medicaid Deliveries
16. DPA Enrollees Receiving EPSDT

Select Indicator Number:

Press ESC to return to Previous Menu

The next section is on Chronic Disease Indicators which provides an overview of mortality, incidence, and hospitalization rates for selected chronic diseases that reflect the influence of lifestyle-related risks. The category also illustrates the prevalence of several risk factors controllable by each individual that can play an important role in the prevention and management of cardiovascular diseases, cancers, stroke, diabetes, and mental health problems. There are 14 indicators included in this section, nine are based on mortality for cause-specific cancers and other chronic diseases, while another is a group of cancer incidence for ten different types of cancers. Other measures provide county

TIME. 10:30:15 am DATE: 07/18/93

IPLAN DATA SYSTEM

CHRONIC DISEASE INDICATORS

COOK

Mortality Rate For:
1. Coronary Heart Disease
2. Cerebrovascular Disease
3. Cirrhosis of Liver

Mortality Rate For:
4. Breast Cancer
5. Lung Cancer
6. Colorectal Cancer
7. Cervical Cancer
8. Prostate Cancer
9. Childhood Cancer

Hospitalization Rate For:
10. Alcohol Dependence Syndrome
11. Total Psychoses
12. Diabetes

Percent of Population
13. Overweight
Smokers
Sedentary Lifestyles

Incidence Rate For:
14. Cancers

Select Indicator Number:

Press ESC to return to Previous Menu

level hospitalization rates for alcohol-dependence, total psychoses, and adult onset diabetes and the percent of population among similar counties which are overweight, which are smokers and which are not getting satisfactory exercise. The latter measures are from the BRFS which groups responses into 4 types of counties which are fairly similar.

The section on Infectious Disease Indicators presents an overview of available information on the incidence of reportable infectious diseases in Illinois and to report the immunization status of Illinois children. Even though the number of infectious disease cases is under-reported, surveillance systems are important in detecting both the causative agents and diseases and are essential components of modem prevention and control strategies. The eleven specific indicators included in this section are mainly reported at the county level and cover the major communicable diseases -- food-borne, sexually transmitted, vaccine preventable, etc.

TIME 10 30:15 am DATE: 07/18/93

IPLAN DATA SYSTEM

INFECTIOUS DISEASE INDICATORS

COOK

1. Syphilis
2. Gonorrhea
3. Chlamydia
4. AIDS
5. HIV Infection
6. Basic Series Vaccinations (Age 5)
7. Haemophilus Meningitis (Ages O-4 & O-24)
8. Infections by Foodborne Pathogens
9. Vaccine Preventable Diseases
Diphtheria
Pertussis
Tetanus
Measles
Mumps
Rubella
Polio
10. Hepatitis B
11. Tuberculosis

Select Indicator Number:

Press ESC to return to Previou Menu

The final section, Environmental, Occupational, and Injury Control Indicators, contains a diverse set of indicators related to health protection factors in the areas of environmental health, occupational safety and health, and injury control. Mortality from homicide, suicide and motor vehicle use are included here as are number of hospitalizations for non-fatal head and spinal cord injuries and for hip fractures. Occupational health indicators include sentinel occupational cancers, hospitalizations from occupational injuries, and occupational lung diseases. An indicator for numbers of children with high blood lead levels is included as are the numbers of reported assaults and numbers of alcohol-related motor vehicle crashes. While most measures are limited to county level, the hospitalizations and deaths are available at community level.

TIME: 10:30:15 am DATE: 07/18/93

IPLAN DATA SYSTEM

ENVIRONMENTAL, OCCUPATIONAL AND INJURY CONTROL INDICATORS

COOK

1. Environmental Indicators
Regulated Drinking Water / Private Wells
NPL Hazardous Waste Sites
Days Exceeding EPA Ambient Air Pollution Standards
2. Toxic Agents Released Into Air, Water, Soil
3. Mortality Due to Motor Vehicle Crashes
4. Mortality Due to Homicide
5. Mortality Due to Suidde
6. Hospitalization for Non-fatal Head/Spinal Cord Injuries
And for Hip Fractures
7. Alcohol-related Motor Vehicle Deaths
8. Occupational Diseases/Injuries
9. Blood Lead Levels in Children
10. Assaults

Select Indicator Number:

Press ESC to return to Previous Menu

Overlapping several of the above sections, is a set of sentinel indicators which cover health conditions that are considered preventable or controllable with regular primary care. The occurrence of sentinel events can be interpreted to indicate inadequate access to primary care. A separate screen selection presents the seven indicators: hospitalizations for dehydration in infants; among children age 1 to 17, the number of hospitalizations for rheumatic fever and asthma and the number of vaccine preventable diseases. Among adults, the number of cases of tuberculosis, and number of hospitalizations for uncontrolled hypertension are among the sentinel events while for adult females the number of cases of late stage cervical and breast cancer are included.

TIME 10 30 15 am DATE: 07/18/93

IPLAN DATA SYSTEM

SENTINEL EVENTS INDICATING
LACK OF ACCESS TO PRIMARY CARE

YEAR 1990	COOK	Illinois	U.S.	YEAR 2000
Infants (0-1)				
Hospitalization for Dehydration	1,904	3,511	N/A	N/A
Children (1-17)				
Hospitalization for Rheumatic Fever	24	40	N/A	N/A
Hospitalization for Asthma	5,304	9,596	N/A	N/A
Adults (>=18)				
Tuberculosis	783	1,035	N/A	N/A
Hospitalizations for Uncontrolled Hypertension	4,374	8,010	N/A	N/A
Adults (>=15)				
Late Breast Cancer	61.8 (1353.2)	56 4 (2767.0)	60.7	N/A
Black	44.2 (299.6)	60 2 (355.4)	61.0	N/A
White	44.3 (1023.8)	55 9 (2364 6)	61.8	N/A
				N/A
LateBlackrical Cancer	8.6 (0 59.6)	11.2 ((268.4)	12.0	N/A
White	4.0 (86.4)	4.6 (185.6)	4.6	N/A

Cancer: numberofcases is a 5-year average. PRINT? (Y/N)

As seen on the following sample screen of infant mortality, the database system shows county-level indicators with comparable state-level and national data plus the Healthy People 2000 objectives if applicable. This allows local users to make comparisons through use of consistent data elements. A separate screen is planned for development in the 2nd year of the project to show data for those of Hispanic origin.

TIME: 10:30:15 am DATE: 07/18/93

IPLAN DATA SYSTEM

INFANT MORTALITY
Rate par 1,000 live births

YEAR: 1990	COOK		Illinois		U.S.	YEAR 2000
Infant Mortality Rate	13.3	(1.266)	10.7	(2,087)	9.2	7.0
Asian/Pac.Is.	4.0	(13)	4.1	(20)	N/A	N/A
Black	22.6	(773)	22:o	(950)	16.0	11.0
White	8.4	(498)	7.6	(1.115)	7.6	N/A
Neonatal Mort. Rate	6.6	(856)	7.0	(1,365)	5.6	4.5
Asian/Pac.Is.	3.1	(10)	2.9	(14)	N/A	N/A
Black	13.9	(476)	13.4	(576)	11.6	7.0
White	6.2	(370)	5.3	(775)	4.6	N/A
Postneonatal Mort.	4.4	(430)	3.7	(722)	3.4	2.5
Asian/Pac.Is.	*.*	(3)	*.*	(6)	N/A	N/A
Slack	8.7	(297)	8.7	(374)	6.4	4.0
White	2.2	(128)	2.3	(340)	2.8	N/A

II < 10 deaths, no rate calculated PRINT? (Y/N)

TECHNICAL CHARACTERISTICS OF THE SYSTEM

The database system has been developed using FOXPRO and will operate on IBM-compatible AT PCS. This means that a computer that is based on a 286 or faster micro-processor will be able to function with this database. The system has been installed on Local Area Networks (LANs) and on stand-alone PCs. It will work in a Windows environment or with DOS.

The data assembled for the first year and the database software will require about 8 to 10 megabytes of space on a hard-drive. Because some indicators will not be monitored during the post-censal years, the space requirements for additional years will be somewhat smaller.

IPLAN INTEGRATES TRAINING WITH DATABASE USE

We believe that training is important not only from the staff development perspective but also from the quality assurance perspective. For those reasons, a comprehensive training program was designed to assist local and regional health department staff retrieve, analyze, and interpret their data sets. In addition, a variety of technical assistance activities took place during the project's time period.

TRAINING PROGRAM

The training program was learner-centered and focused on the following issues: benefits of a data-based needs assessment; introduction to the operation of the IPLAN data system; introduction and retrieval of the IPLAN data sets; analysis and interpretation of data; and prioritization of health problems.

Benefits of a data-based needs assessment. The benefits of conducting a needs assessment based on current epidemiologic data were discussed. Emphasis was placed on the advantages of setting program priorities based on objective data as opposed to emotion. Some benefits discussed related to conducting date-based needs assessments include the ability to: establish priorities based on needs and available resources, outline goals and objectives of a program, provide standards for evaluation of a program, and create an awareness of a health problem.

Introduction to the operation of the IPLAN data system. The local and regional health department staff also were provided with a brief orientation on the operation of the data system. They were trained on basic start-up procedures, navigation through the system, and powering down. Due to the "user friendliness" of the system, it took on average about five minutes for people to master the basic operations of the system. Computer requirements needed to run the IPLAN data system also were identified at this time.

Introduction and retrieval data in the IPLAN data system. Next, the data sets described earlier were introduced. Staff were instructed on the methods of identifying their county codes, community codes, and methods of grouping different counties together to compute averages over selected geographic regions. Then, methods of retrieving data from the various data groupings were introduced. The menu-driven nature of the system made the retrieval process easy to learn.

Analysis and interpretation of data. This was one of the most important elements of the training program. Through the IPLAN data system, local and regional health department personnel now had access to a number of important data sets. However, proper analysis and interpretation of the data now became a responsibility of the local health departments rather than the state program specialists.

The central theme of this aspect of the training was to help local health department staff determine if a health event was problematic or not. For example, is a county incidence rate of gonorrhea such as 348.8 per 100,000 a problem? In this example, the training involved the comparison of the county rate to state and national rates, and HEALTHY PEOPLE 2000 objectives. In addition, a discussion focusing on cautionary issues also took place. Such discussion covered ways that rates may be inflated or deflated as a result of special populations...for example, a college town with a high population of college students might be a factor in a high incidence of gonorrhea. This discussion illustrated the importance of considering age-adjusted rates, age- or race-specific rates, etc. In addition, the training emphasized that if a problem is detected, more than one year's worth of data should be evaluated to ensure that the problem is not a temporary one.

Prioritization of health problems. One of the major problems facing public health planners is the prioritization of health problems. The training program we used was based on a modified version of the APEXPH process which emphasizes community member involvement in setting priorities. In our training program we suggested that health department officials initially identify the major health problems based on the information from the IPLAN data system. Then, because the APEXPH program emphasizes "cause specific years potential life lost" we recommended that they list the top ten causes of YPLL for their area. Then, we recommended that the community group review the prioritized problems and YPLL identified by the health department and add to the list any problems the group feels should be added. Finally,, after a discussion of the data and community generated health problems, prioritization takes place using the nominal group process.

TECHNICAL ASSISTANCE

Several strategies were used to provide technical assistance to the local health department staff members: newsletters; interpretation guide, regional technical assistance meetings, and telephone consultation.

Newsletters. The grant activity included a quarterly newsletter. Columns were written by experts on a variety of topics. For example, the project biostatistician wrote an article on the computer system needed to run the IPLAN data system. The Chief of the Division of Epidemiologic Studies wrote a column emphasizing the strengths and weaknesses of the data set and "tips" on the interpretation of the epidemiologic data.

Documentation and interpretation guide. Health department personnel received two manuals along with the IPLAN data diskettes. The first manual was a documentation guide which included instructions on installing, operating, and powering down the system. Sources of the data and technical notes related to the data use were included in the documentation. The second manual, an interpretation guide, provided more detail related to each indicator and provided some examples of the use of the system. One example was the use of the system to identify indicators which would help health educators develop a comprehensive school health education program.

Regional technical assistance meetings. Regional technical assistance meetings were held throughout the state. At those meetings, IPLAN staff were available to help local health department personnel in any aspects of the IPLAN project.

Telephone consultation. IPLAN staff frequently provided telephone consultation to the local health department personnel. This consultation ranged from how to install the data system to help in forming the community groups.

APPLICATIONS OF THE COMPUTERIZED DATABASE SYSTEM

This system will be used for a number of applications by state and local health officials, as well as researchers.

The system will be used by the majority of local health departments in Illinois when preparing their community health needs assessment. This process has been started throughout the state. The system will be useful through the decade for surveillance of many health conditions in local communities. The system provides an ability to compare local health jurisdictions to state and U.S. conditions, including YRAR 2000 objectives. It also provides the ability to group counties or communities into meaningful comparison groups. In addition to needs assessment and surveillance, the dataset could provide information from the local health department to other data users : their community.

For example, a regional school official could obtain information to help them plan for effective school health programs. Indicators of interest to school officials could include: number of single parent households, poverty rates, number of Medicaid enrollees below the age of 21, years of potential life lost (YPLL), leading causes of death among children age l-4, percent of births to teens, number of child abuse cases, number of childhood cancers, number of children with high blood lead levels, numbers of hospitalizations among children for asthma, and number of vaccine-preventable diseases among children. Even the number of STD and HIV cases in the county will be of value to school officials since many of the cases are occurring among teenagers. While some of the indicators may be covering broader age groups than the school age population, school officials could still, be able to identify some of the most important health problems facing the school districts, and plan the health education/promotion and prevention programs accordingly.

SUMMARY

In summary, this system provides health indicators to local heaith department staff that will allow them to examine some health measures at the township and incorporated place level as well as county level geography. The system has been developed to meet a need for accurate, concise, and defensible information to identify and describe public health needs. These indicators will provide some of the information needed for their community needs assessments which, in turn, will be part of future certification processes.

A TELEPHONE SURVEY METHODOLOGY FOR LOCAL HEALTH DEPARTMENTS' COMMUNITY HEALTH STATUS AND RISK FACTOR ASSESSMENTS RELATED TO *HEALTHY PEOPLE* 2000

Bruce M. Brock, Information Transfer Systems, Inc.

In late 1990, *Healthy People* 2000 was released by DHHS to guide the public health apparatus through the decade from 1990 through the Year 2000. It sets 300 explicit and measurable objectives across 22 "priority areas," with 220 additional detailed objectives for specific population subgroups. Based on federal, state, and local input, *Healthy People 2000* is a consensus document that public health agencies at all governmental levels are working to support.

Healthy People 2000's 22 priority areas are broken down into four major groups. Health promotion, health protection, and preventive health services constitute the three main content areas, with the fourth area covering 'surveillance and data systems." Such systems are recognized as critical elements of the infrastructure needed to support the objective-setting, planning, policy-making, resource allocation, and evaluation process needed to successfully implement the specific objectives of the health promotion, health protection, and preventive health services areas. The surveillance and data systems priority area was judged to a priority based on a number of factors, including:

- The period from 1980 though 1990 was guided by a DHHS document similar to *Healthy People 2000,* titled *Objectives for the Nation,* which did not explicitly recognize the importance of surveillance and data systems for achieving the 1990 objectives. For some objectives there were no baseline data (from 1980) to provide the benchmark against which future data could be compared. Also, without baseline data, informed planning could not take place.

  When *A Midcourse Review* of the *Objectives for the Nation* was conducted by DHHS in 1986, there were numerous objectives for which no midcourse data were available to be used in combination with baseline data to determine whether there was evidence of progress toward reaching the objective. Without the required midcourse data, it was not possible to make midcourse corrections in areas empirically to be lacking.

  Once 1990 was reached, for many objectives, "outcome" data, required to determine whether the many specific and measurable *Objectives for the Nation* had been reached, were still unavailable.

- In 1988, the Institute of Medicine (IOM) published a landmark report, *The Future of Public Health.* In this broad-based review of public health practice at the state and local levels, the IOM Committee for the Study of the Future of Public Health strongly emphasized and recommended that "the core functions of public health agencies at all levels of government be assessment, policy development, and assurance."

  Concerning the "assessment" function, the IOM Committee went on to say: "In understanding of the determinants of health and of the nature and extent of community need is a fundamental prerequisite to sound decision-making about health. Accurate information serves the interests both of justice and the efficient use of available resources. Assessment is therefore a core governmental obligation in public health."

  The assessment function was believed to "facilitate good decisions in both the private and public sectors." The Committee therefore recommended that all public health agencies, at every level of government (including local), bear the responsibility for ensuring that they regularly and systematically collect, assemble, analyze, and make available information on the health of the community, including statistics on health status, community health needs, and epidemiologic and other studies of health problems. Among other sources of data, local community health-related surveys were recognized as an appropriate research strategy in this context. The IOM report emphasized that this basic "assessment" function of public health cannot be delegated. A fully developed assessment function was judged to be an essential part of the public health system, and it is one the IOM Committee believed to be "in large measure attainable."

Based on these and other factors, it was obvious that new surveillance and data systems had to be put into place if the Year 2000 objectives were to be met. Thus, as noted, *Healthy People 2000's* Priority Area 22, is dedicated to the institutionalization of such systems. It, like all Priority Areas, has a number of specific objectives.

Objective 22.1 forms the underpinnings for each of Priority Area 22's subsequent objectives. It addresses the need for surveillance and data systems to meet the needs of not only federal and state agencies, but also local health jurisdictions. It recognizes that, while there are a number of important surveillance and data systems that collect data on a national and state level, the data derived from these systems are rarely useful for informed planning and evaluation purposes at the "local" level. Objective 22.1 reads:

> **"Develop a set of health status indicators appropriate for federal, state and local health agencies and establish use of the set in at least 40 states."**

The associated narrative further states:

"A major function of public health agencies is to assess the health status of the population. The value of such an assessment is enhanced when comparisons to some benchmark, such as other similar populations or to national norms, can be made. These comparisons can be simplified and made more efficient if a limited set of health status indicators is monitored by jurisdictions nationwide.

A set of indicators will facilitate comparability of data on health status within and among States and local areas and would permit the valid comparison of local and State health data with national data... To be of optimum use, the indicators should include measures needed for area-wide health planning and allocating health resources among programs, areas, or regions. The indicators should also provide measures needed to judge the results of health planning and to study changes in health over time."

Objective 22.1 thus recognizes the need for local jurisdiction health departments to have access to "community" specific data to ensure that their planning and evaluation efforts are based on data which are accurate for their community.

In order to support the implementation of Objective 22.1, the Centers for Disease Prevention and Control (CDC) established Committee 22.1, which delivered its initial recommendations to the 1991 National Vital and Health Records Conference (and published them in *MMWR* - 7/91). Committee 22.1 is composed of members from key public health associations and organizations (including APHA, NACHO, ASTHO, PHF, and USCLHO), and is convened by Dr. M. Feinleib, M.D., Dr.P.H., Director of the National Center for Health Statistics (NCHS).

The Committee's recommendations included the identification of the minimum set of community health status indicators which, per Objective 22.1, were appropriate for federal, state and local health agencies. The Committee's primary products were two summary tables. TABLE 1, covered the "Consensus Set of Indicators for Assessing Community Health Status and Monitoring Progress Toward the Year 2000 Objectives," and TABLE 2, "Priority Data Needs to Augment the Consensus Set of Health Status Indicators."

Committee 22.1 recognized that, to evaluate a community's health, a multidimensional array of factors needs to be assessed. Some of these measures are retrievable from existing systems available to local and state health departments. These are primarily covered in TABLE 1, which unfortunately takes up too much space to display. Examples include birth and mortality data, as well as specific measures of morbidity. Much of the required data are mandated for collection by a variety of state and federal laws and regulations.

Other important elements to assess a community's health status identified by Committee 22.1, particularly those in TABLE 2 (Priority Data Needs) which is replicated below, are not routinely available at the local health jurisdiction level. This is true for the general class of health practices or behavioral risk factors (such as smoking, alcohol misuse, obesity, and others), and for service utilization-related factors (such as immunizations, Pap tests, mammograms, health insurance coverage, etc.).

Efforts to use state or national level survey data for local use in planning and evaluation for most of the topical areas covered by TABLE 2 have not fared particularly well, and there is good reason why. Data from existing state and national data collection systems (such as CDC's Behavioral Risk Factor Survey (BRFS), NCHS' National Health Interview Survey (NHIS) and National Health and Nutrition Examination Survey (NHANES), AHCPR's National Medical Expenditure Survey (NMES), and most of the "supplemental" health surveys conducted by the Bureau of the Census for various DHHS agencies), do not reveal, nor were their designs intended to reveal, the variability that unquestionably exists at the local level, masking real

COMMITTEE 22.1'9 TABLE 2. PRIORITY DATA NEEDS TO AUGMENT TEE CONSENSUS SET OF HEALTH STATUS. INDICATORS

Indicators of processes:
- Proportion of children 2 years of age who have been immunized with the basic series (as defined by the Immunization Practices Advisory Committee)
- Proportion of adults aged >= 65 years who have been immunized for pneumococcal pneumonia and influenza
- Proportion of assessed rivers, lakes, and estuaries that support beneficial uses (fishing and swimming approved)
- Proportion of women receiving a Papanicolaou smear at an interval appropriate for their age
- Proportion of women receiving a mammogram at an interval appropriate for their age
- Proportion of the population uninsured for medical care
- Proportion of the population without a regular source of primary care (including dental services)

Indicators of risk **factors (age-specific prevalence rates):**
- Cigarette smoking
- Alcohol misuse
- Obesity
- Hypertension
- Hypercholesterolemia
- Confirmed abuse and neglect of children

Indicators of health status outcomes:
- Percentage of children <5 years of age who are tested and have blood lead levels exceeding 15ug/dL
- Incidence of hepatitis B, per 100,000 population
- Proportion of children aged 6-8 and 15 years with one or more decayed primary or permanent teeth

differences within and between our communities. Even recent advances in creating estimates for small areas from large area data sources have not yielded data with sufficient accuracy for use in local planning, policy development, resource allocation, and evaluation purposes. Thus, they need to be collected via high quality, local community surveys.

Unless a standardized, survey data collection system is developed for local health jurisdictions, numerous data collection approaches and questionnaires will be implemented by local agencies across the country. These will suffer from lack of uniformity (contrary to the explicit intent of Objective 22.1), and thus will not be comparable enough to allow aggregation across jurisdictions for regional or state level estimates. Only a nationally coordinated effort to develop and test a core survey system will overcome this likely and counterproductive outcome.

In October 1992, Information Transfer Systems, Inc (ITS), of Ann Arbor, Michigan was awarded a DHHS contract (#213-92-0062) with the Agency for Health Care Policy and Research (AHCPR) to develop this local health jurisdiction survey system, including the "core questionnaire" and related survey and sampling methodology. To ensure that this effort was a nationally coordinated consensus-based effort, ITS was extremely fortunate to develop a working relationship with the Director of NCHS, Dr. Feinleib, and in particular, a newly created organization in NCHS which is focused on the Year 2000 objectives. Our primary contact in this group is Ms. Mary Anne Freedman, Special Assistant to Dr. Feinleib, who has overall responsibility for the coordination of the "surveillance and data systems" area. Ms. Freedman also participated in aspects of the Committee 22.1 process during 1991 as an NCHS staffer to the Committee. Additionally, a working relationship was developed with Ms. Freedman's key staff including Mr. Ken Keppel and Ms. Susan Hawk.

With the help of Ms. Freedman, her staff, and Dr. Feinleib, ITS coordinated a Committee 22.1 meeting during the APHA conference in November 1992. The result was an agreement to use Committee 22.1 as the coordinating and policy-making body which would work with ITS to review and comment on our approaches for survey instrumentation and research design development.

To guide ITS' work with Committee 22.1, a number of specific objectives were developed, including:

OBJECTIVE 1:

To identify those elements from Committee 22.1'9 Table 2, PRIORITY DATA NEEDS, which are amenable to a telephone survey approach for data collection.

OBJECTIVE 2:

To obtain extant survey instruments (such as BRFS, NHIS, NHANES, NMES, and CENSUS health supplements) which contain measures of the selected PRIORITY DATA NEEDS, and compare and contrast these survey items in a matrix format to allow similarities and differences in survey items covering the same dimensions to be easily identified.

OBJECTIVE 3:

To obtain methodologic documentation for each of the key survays to identify important similarities and differences in their sampling and data collection approaches, and to determine which of the design approaches are most cost-effective for the proposed local health jurisdiction survey(s).

OBJECTIVE 4:

To develop decision-rules for the selection of the specific survey items to be used in the Committee 22.1 recommended "core questionnaire", as well rules to aid in the selection of methodologic procedures to be used in the survey's administration by local health. jurisdictions.

OBJECTIVE 5:

To develop the final set of survey documents, including the draft questionnaire and survey method guidelines.

In servicing these objectives, ITS obtained and thoroughly researched all of the relevant survey instruments from CDC, NCHS, AHCPR and Census officials which measure any of the Priority Data Needs. The resulting matrix of relevant questionnaire items which compares and contrasts how these different instruments measure each topic of interest is available from ITS. Relevant documentation from CDC, NCHS, AHCPR and CENSUS officials representing each survey was also obtained and evaluated. Key.methodologic aspects of interest for each of the surveys included sampling approaches, sample size parameters and their basis, and interviewing methodology.

ITS staff then conducted face-to-face meetings with the leadership of most of the key surveys mentioned above and telephone conferences where meetings were not feasible due to scheduling conflicts. ITS also met with the leadership-of NCHS' Office of Methodology and its Questionnaire Development Research Laboratory.

Based on the various meetings and discussions, the following decision rules for the local health jurisdiction survey research design and the questionnaire(s) emerged:

- The entire research design must be as consistent as possible with Committee 22.1'9 "Priority Data Needs" recommendations, which aim to support *HEALTHY PEOPLE 2000* with a minimum core dataset.
- Scientifically validated procedures must be used throughout the research design.
- The total research design must focus on the "general population," including adults and children.
- The overall design must be as inexpensive as possible for localities to implement, while ensuring high quality data for planning and evaluation purposes.
- Random digit dialing (RDD) telephone sampling and computer assisted telephone interviewing (CATI) methods will be the most cost-effective overall strategy.
- The research design and questionnaire must be as comparable as possible with key federal and state surveys.

- The design must allow flexibility for oversampling or supplemental sampling of special population subgroups.
- The design must allow repeated local jurisdiction surveys to be compared, so that change in a community's health and risk factor status can be reliability detected over time.
- The questionnaire must focus on health-related "outcomes" rather than on knowledge or attitudes.
- The design must allow tailoring of the research design and questionnaire to meet local needs in a manner which does not impede the ability to compare the data collected across localities over time.
- The number of questionnaire items must be no greater than required, in order to minimize the questionnaire's length and associated respondent burden.
- A standardized "core" questionnaire will be devised containing questionnaire items which measure all of the Priority Data Needs which are amenable to a telephone approach of data collection.
- Its is expected that local jurisdictions will not typically modify the "core" (to ensure its comparability from community to community), but localities may add supplemental questions after the core is completed by respondents.
- Supplemental questionnaire modules will be offered covering a range of lower priority topics that were not included in Couunittee 22.1'9 Priority Data Needs list, but which will have value in certain communities.

In February 1993, ITS arranged a second meeting of Committee 22.1 and associated NCHS and CDC staff. ITS presented its recommendations concerning the overall research design parameters and the primary sources for questionnaire items which ITS believed best addressed Table 2's Priority Data Needs and the overall intentions of Committee 22.1. ITS was then charged by the Committee to draft the research design and questionnaire documents required by ITS' DHHS contract, in a manner consistent with the above decision rules. This does not, however, constitute a formal endorsement of ITS' work by Committee 22.1.

Outline 1, below, summaries the overall research design, and Outline 2 summarizes the content of the draft survey questionnaire(s) developed by ITS. Complete copies of all project documents are available though ITS.

OUTLINE 1

KEY RESEARCH DESIGN AND QUESTIONNAIRE PARAMETERS ARE SUMMARIZED IN THE FOLLOWING OUTLINE:

ACCEPTABLE RANDOM DIGIT DIALING SAMPLING DESIGNS:

- Mitofsky-Waxberg
- Brick-Waxberg
- State-of-the-art "list-assisted" (non-clustered) approach

RANDOM SELECTION OF AN ADULT (18 OR OLDER) IN EACH HOUSEHOLD BY THE "BIRTHDAY" METHOD:

- Reports on self
- Reports on other adults
- Reports on children unless other adult is more informed

SAMPLE SIZE PARAMETERS:

- Minimum of 500 households per community
- Should yield data on about 1,315 persons per community (2.63 adults and children per household based on the 1990 census)

OTHER DESIGN PARAMETERS:

- Relatively short data collection period during warm season (late spring/summer)
- CAT1 interviewing by professionals recommended

POTENTIAL DESIGN LIMITATIONS:

- Will not cover non-telephone households or individuals
- May not provide reliable data for all sub-populations groups of interest
- Sample sire may not be sufficient to reliably measure changes over time for important subgroups on all risk factors unless oversampling or supplemental sampling is part of the design
- Potentially inaccurate or missing data by proxy

BENEFITS:

- Is very consistent with NCHS Committee 22.1's Priority Data Needs
- Will be highly comparable to state BRFS (CDC) surveys
- Will be reasonably comparable with national level survey data
- Will allow comparisons of jurisdictions using basic design and questionnaire
- Will allow assessment of change overtime within communities, if design and questionnaire is repeated
- Will be largely consistent with APEXPH, MODEL STANDARDS, and other planning tools
- Allows tailoring after the "core" to meet special community needs
- Allows oversampling or supplemental sampling of special population subgroups

OUTLINE 2

QUESTIONNAIRE MODULES AVAILABLE FOR LOCAL HEALTH JURISDICTION SURVEYS

"CORE" QUESTIONS FOR ALL ADULTS IN THE HOUSEHOLD

- Overall (Subjective) Health Status
- Health Insurance And Sources Of Medical And Dental Care
- Hypertension
- Cholesterol
- Cigarette Smoking
- Alcohol Misuse
- Pregnancy, Pap Smears And Mammograms
- Adult Immunizations
- Obesity
- Abuse And Neglect Of Children
- Demographics

"CORE" QUESTIONNAIRE TOPICS FOR ALL CHILDREN IN THE HOUSEHOLD

- Overall Health Status
- Health Insurance And Utilization Of Medical And Dental Care
- Cigarette Smoking
- Alcohol Misuse
- Pregnancy Status
- Obesity
- Child Immunizations
- Child Demographics

SUPPLEMENTAL TOPICS DEVELOPED OR UNDER DEVELOPMENT

- Utilization Of Preventive And Non-Preventive Medical And Dental Care Services
- Prior Hospitalizations, Surgery, And Use Of Mental Health Services
- Detailed Alcohol Misuse
- Drug Misuse
- Smokeless Tobacco Use
- Colorectal Cancer Screening, Breast Self-Exam, Testicular Self-Exam
- Aids Knowledge And Testing
- Diet
- Exercise
- Activity Limitations
- Health Department Services: Knowledge, Use, Satisfaction, Attitudes
- Community Hospital(s): Knowledge, Use, Satisfaction, Attitudes
- Interest In Participating In Various Risk Factor Reduction Programs
- Mental Health And Nell-Being Assessment
- Social Support
- Prenatal Care
- Chronic Disease Checklist And Treatment/Control Status
- Symptom And Impairment Checklist
- Diabetes
- Angina And Respiratory System Assessment
- Injury Control
- Home Safety: Radon Testing, Storage Of Medicines, Toxic Chemicals Firearms, Smoke Detectors, Etc.
- Knowledge, Beliefs, And Attitudes Concerning Various Risk Factors

Session W

Medical Effectiveness Outcomes Research

Toward the year

2000

Refining the measures

INTEGRATING VALUES INTO CLINICAL CARE

David B. Pryor, Duke University Medical Center

(Not available for publication)

OUTCOMES RESEARCH FOR EFFECTIVENESS AND QUALITY: THE MAJOR ISSUES

Sheldon Greenfield, M.D., Tufts University

Summary of Discussion

It is now possible to use outcomes of care, not only mortality, but functional status and physiologic measures such as blood pressure, to measure the quality of care and to measure the effectiveness of treatments. The scientific and conceptual basis as well as the instruments have been advanced to the point where the barriers to assessing the outcomes are being overcome. These barriers: such as case mix, such as correct time interval for the assessment of care, such as determining the norms of outcomes, such as correctly measuring functional status, such as defining the factors that confound the interpretation of outcomes independent of quality or effectiveness of treatment, are being removed to make outcome data interpretable.

A new movement is afoot to include patient reported data, possibly combined with local administrative data.

EFFECTIVENESS OF THE NEW JERSEY HEALTHSTART INITIATIVE ON PRENATAL CARE AND INFANT OUTCOME, 1988 to 1991

Maryanne Florio, New Jersey State Department of Health
Susan Lenox Goldman, Leah Ziskin, A. I. Abdelrahman,
Lucille Wallington, Virginia M. Dato, Bruce Siegel

HealthStart is a major New Jersey (NJ) joint initiative of the Department of Health and the Department of Human Services to improve the accessibility, scope, and quality of maternal and child health services available to Medicaid-eligible pregnant women and children up to the age of two. It was initiated in February 1988 and has been operational statewide to the present time.

HealthStart has both a prenatal care and a pediatric component. The pediatric component provides preventive health care services, including immunizations, to Medicaid-eligible children under the age of two. Only the outcome of the maternity component of the HealthStart program will be addressed in this paper.

The HealthStart maternity care program provides a comprehensive package of medical and health support services to Medicaid-eligible pregnant women. HealthStart medical services include prenatal visits, intrapartum care and postpartum visits as recommended by the American College of Obstetricians and Gynecologists. Health support services include nutrition, social, and psychological assessment, guidance, and counseling: and health education, with an individualized care plan and case coordination of each of these services. Case coordination is considered the key component of the HealthStart program.

HealthStart services are provided to an increasingly larger population of poor women by expanding eligibility, to 100% of poverty in 1988, and 185% in 1991; and through presumptive eligibility which allows a woman to be determined eligible by the provider and receive services while waiting for her Medicaid eligibility to become approved. Services are provided in HealthStart-certified hospitals, community based agencies, and private practice settings, based on the standards and guidelines developed as part of the implementation plan. Increased Medicaid reimbursement is given to certified HealthStart providers.

The ability to do outcomes research in the HealthStart program was enhanced by the planning, legislation, and implementation of the evaluation prior to the implementation of the program. The evaluation assesses the following research questions:

1. How successful is the State in enrolling and providing services to the target population of low income pregnant women and children up to the age of two, who are at or below the Federal poverty level?
2. How successful are the providers of the program in delivering the services they have agreed to provide as part of the HealthStart Program?
3. What are the barriers to the delivery of services provided in the HealthStart package?
4. How effective is the HealthStart program in improving health outcomes of these women and their newborn babies?
5. How cost effective is the HealthStart program?

The answers to these questions are used to improve the quality of the HealthStart program, demonstrate the effectiveness and efficacy of the program, and justify the expenditures in a time of extensive budget constraints. The evaluation focuses on the entire program, not just the evaluation of one outcome, i.e., infant mortality.

In order to assess the effectiveness of the HealthStart package of prenatal care, information is needed to address each of the research questions. After much research, literature review, and discussion, three data collection sources were developed:

1. An in-depth summary of pregnancy information reported on each woman using the HealthStart Maternity Services Summary Data (MSSD) form:
2. A long-term Linked Data System (LDS) for comparison of the outcome of prenatal care of HealthStart women with all other groups of pregnant women in NJ;
3. A Tracking System implemented to provide information on the monitoring and quality assurance provided at each agency through site visits and follow-up, and renewal of certification every 18 months.

As with all data there are limitations to these data sets: 1) portions of the data are self-reported and therefore, some of the sensitive data, for example drug use, may be underreported: 2) there is a chance of error in completing forms, in data entry, or in data processing; 3) some of the data are incomplete: (The number of MSSD forms submitted is not equal to the number served. This appears to be random in all but one agency who submit their forms electronically, and experienced a delay because the electronic system was not implemented until 1991. Also, not all of the records were linked in the LDS. The percentage of non-linked records is higher in the Medicaid population, but this appears to be a random non-linkage of HealthStart and non-HealthStart Medicaid women.) 4) there is a potential selection bias for HealthStart women, i.e., the more motivated or healthier women may have chosen to participate in the program.

However, there are two ways in which this bias will be tested: first, a comparison of AFDC (less than 52% of poverty) with non-AFDC HealthStart and other Medicaid women will be conducted when there are sufficient non-AFDC women in the study for analysis: and secondly, a trend analysis will be conducted as more of the Medicaid population participants in HealthStart, to the point where, in 1992-1993, nearly the entire Medicaid population of pregnant women will have received HealthStart services.) The percentages of the data may not total 100% due to roundina. missina values, or multiple response items.

The MSSD form provides a comprehensive source of information on all aspects of the woman, her prenatal care and her the outcome of her pregnancy, including her demographics: risk factors; services provided; the outcome of the labor, delivery, and the pregnancy: and the health status of her baby through the postpartum visit. The MSSD data are reported for the 45,000 HealthStart women for whom an MSSD form was submitted between October 1988, the startup date for submission of the MSSD form, and December 1991.

The second major source of information, the HealthStart LDS, links the DOH Vital Statistics matched birth and death files, and the Hospital Uniform Billina (UB-82) data to the Medicaid eligibility and'claims files. The LDS is based on single live births to NJ residents with NJ hospital deliveries. It is used primarily to conduct a comparison of the prenatal care and birth outcomes of three groups of women who give birth in NJ: 1) Medicaid-eligible pregnant women in NJ who have received HealthStart services: 2) Medicaid-eligible pregnant women in NJ who have not received HealthStart services: and 3) all other pregnant women in NJ who did not receive Medicaid funds.

The linkage of the files for the baseline years 1985 through 1987, and the first two years of the implementation of the HealthStart program, 1988 and 1989. have been completed with a 90 to 92% link rate. There were approximately 90,000 single, live, in-hospital NJ births in 1985. This has increased each year, to 104,000 in 1989, with Medicaid coverina approximately 15% of the births during those years, and increasing to approximately 25% in 1991.

The HealthStart LDS for the first two years of implementation is included in the following analysis with this caveat. The HealthStart program was initiated in 1988, and therefore, data from the first two years of implementation portray a very early picture of the effectiveness of the program, and may not be indicative of its effectiveness after the program has been in progress for several years. In addition, the 1989 Linked Data are based on a preliminary file received just prior to this report, and have a lower linkage rate of Medicaid women than other women (85% compared to 92% respectively). This is being corrected by further manual work. Due to the delays, the file has not yet been scrutinized as carefully as those in prior years. However, the very early findings look promising, and if the data continue to show the improvement in the future, the HealthStart program will be considered a successful program in providing quality prenatal care and preventing adverse birth outcomes.

An analysis of the LDS shows that Medicaid women are significantly different from other women in age, race, Hispanic origin, marital status, and education. HealthStart women are approximately the same as other pregnant Medicaid women, and also different from other pregnant women in NJ in these characteristics. Approximately one-third of the HealthStart and other Medicaid mothers were teens, two-thirds in their twenties and early thirties, and 3% were over thirtv-four. In comparison, about 5% of all other women were teens, 85% were in their twenties' and early thirties, and 10% were older than thirty-four.

The MSSD data showed that the number of teens who received services decreased between 1988 and 1991. This is believed to be due at least in part to the number of teen pregnancy prevention and education programs provided in NJ, including the HealthStart teen prenatal care programs.

Based on the LDS, approximately 20% of HealthStart and other Medicaid women were married, whereas at least 85% of the non-Medicaid women were married. One-third of HealthStart and other Medicaid women were white and two-thirds black, whereas about 80% of the non-Medicaid women were white, 15% black, and 3% of other races. -Approximately 30% of HealthStart women were of Hispanic origin, compared to 25% of other Medicaid and 10% of non-Medicaid women.

When the MSSD data were used to calculate race with the inclusion of Hispanic origin, the percentage of HealthStart women who were white or black was reduced: however, the percentage of white decreased more than the black. Approximately 27% were white non-Hispanic, 40% black non-Hispanic, 30% Hispanic (black, white or other), and 3% of other races or ethnic origins, compared to 41% white, 49% black, and 29% of Hispanic origin when Hispanic origin is considered separately from race. The percentage of white and Hispanic increased in 1991. The explanation for this change most likely is the increase of the eligibility for Medicaid from 100% to 185% of poverty in that year.

There is an inverse relationship between education and participation in Medicaid or HealthStart, based on the LDS. As would be expected, those with the lowest level of education had the most participants in Medicaid, while those with a high school diploma were split evenly between Medicaid and non-Medicaid, and those with the highest

levels of education had the least Medicaid participation. This relationship remained relatively unchanged over the five years across these educational levels. The educational level of women receiving HealthStart was comparable to that of Medicaid.

Next, each of the research questions will be answered using the information collected to date. The data supporting the answers to questions four and five are preliminary data as described above, and further verification of the 1989 LDS combined with the LDS for 1990 and 1991 are needed in order to provide complete and conclusive answers to these questions.

RESEARCH QUESTION 1: How successful has the State been in enrolling and providing services to the target population of low income pregnant women and children?

Determining the number of women who participated in the HealthStart program each year was not as simple as it would seem. The Medicaid paid claims data include women who received services the previous year and are therefore not an accurate reflection of utilization in a given year. The utilization reports from each certified agency include women who transfer from one service site to another. Therefore, since women may be served by more than one provider, the numbers do not represent an unduplicated count of women served in the state. The forms submitted for each pregnant woman have a lag time of up to 12 months after delivery (six weeks to the postpartum visit, 60 days to submit the form, time for data entry, processing, and outputting the data) and therefore, are not complete or timely enough. The number of recipients is estimated based on the paid claims, the utilization reports and MSSDs submitted, and confirmed as the Medicaid paid claims data are linked to the birth and death files and hospital billing data.

Based on the number of Medicaid births in 1985 and the expansion in 1987 of Medicaid eligibility to 100% of the Federal poverty level, 15,000 pregnant women were originally projected as potential recipients of HealthStart services. Each year the actual number enrolled exceeded the goal set for the program. In 1991 the eligibility was expanded to 185% of the Federal Poverty Level, which increased the number of eligible women to approximately 30,000. It is estimated that by 1993 nearly the entire Medicaid population of pregnant women is being served by the HealthStart program.

Early entry into prenatal care is considered a major factor in assessing the effectiveness of a prenatal care program. Accurate measurement of onset of prenatal care is difficult because of the differing definitions used, and the use of self-reported versus staff-reported data.

The data on prenatal care reported on the LDS showed that 58% of the women receiving Medicaid began their prenatal care in the first trimester in 1985. This has decreased each year, to 50% by 1989. Approximately 85% of the non-Medicaid women in NJ began their prenatal care in the first trimester. This percentage showed some variability but no significant trend over time. Approximately 43% of HealthStart women began care in the first trimester in 1988. This percentage increased to 50% in 1989. This was significantly better than other Medicaid women. The MSSD data showed an increase in prenatal care for HealthStart women in the first trimester between 1988 and 1991, but lower percentages were reported than on the LDS. (The MSSD data are believed to be more accurate in this case because they are reported by prenatal care provider who keeps records of the visits, whereas the birth certificate data, from which the LDS is obtained, are self-reported after delivery and may not be as accurate in retrospect.)

Basic services are required at all HealthStart agencies. Data from the MSSD forms showed that basic services are provided at a rate of 95% to 100% by HealthStart providers statewide. Specialized services are those services provided beyond the basic HealthStart services. They include nutrition, health education, future family planning, childbirth education, assistance with substance abuse, social and psychological services, WIC, special laboratory work, and home visits.

Provision of specialized services increased between 1988 and 1991 in all areas except home visits. A possible reason for the decrease in home visits is the increase in patient volume at all provider sites, to the degree that staff are unable to provide services which extend to the clients' homes.

Provision of specialized services was analyzed by comparing them to the women's risk factors who received these services. Assessment of risk factors before and during pregnancy, and at delivery, based on the MSSD data, showed that the percentage of HealthStart women identified as at risk from any medical, obstetrical, social or psychological risk factor was 84% in 1991. This has been increasing since 1988.

Although the percentage of women who have medical or obstetrical risk factors during pregnancy has been rising, from 57% in 1988 to 62% in 1991, by the time of delivery these percentages decreased by about 40 percentage points. The percentage of women receiving HealthStart services who reported that they had social or psychological problems during pregnancy has also been increasing, from 28% in 1988 to 38% in 1991. However, the percentage of problems identified at delivery is also less than during pregnancy, 12% in 1988, and 17% in 1991.

The reported percentage in 1988 of HealthStart pregnant women who smoked (28%), drank (5%), or used illegal drugs (7%), has decreased less than two percentage points each over the four

years. However, by the time of delivery, smoking was reduced to 12%, drinkina to 1%. and use of drugs to 2%. It should be' noted that these figures are below state estimates of what would be the expected use of illegal substances. However, since these data are self-reported, the percentages are probably an underestimate of the actual substance use by this population.

The reductions in medical, social, psychological and substance abuse problems by the time of delivery is believed to be, in part, an indicator of resolution of some of the risk factors by provision of HealthStart specialized services during the course of prenatal care.

RESEARCH QUESTION 2: How successful have the providers of the program been in delivering the services they have agreed to provide?

Provider recruitment has been increasing each year. During the first year (1988), fifty comprehensive maternity care provider sites were certified to provide HealthStart services. BY 1990 there were seventy-four certified sites, and by the end of 1992 there were eighty-four comprehensive maternity care sites, recruited in all 21 counties in NJ, with the greatest concentration in the areas of greatest need.

Although HealthStart has been successful in recruiting a large number of providers, problems with expansion of the provider pool, coupled with increased utilization, especially in inner city and poorer areas, continue to be an issue. Fewer physicians are entering obstetrical care practice, resulting in a shortage of medical staff. Those who do enter obstetrical care often enter private practice and may not locate in the inner city and poor areas. In addition, physicians who are in private practice often limit or refuse to accept Medicaid clients as patients, citing low Medicaid reimbursement rates as one of the major reasons for their non-participation in the Medicaid program, even with the increased Medicaid reimbursement rates provided by the HealthStart program.

In order to assure that quality services are provided in compliance with the established HealthStart standards and guidelines, providers are monitored throughout their eighteen month certification period. One method of facilitate providers' ability to maintain their certification and remain eligible for renewal of their HealthStart certification.

In assessina the delivery of services by HealthStart providers, the data as described under research question one above showed that the agencies have provided basic and specialized services to a high degree of compliance with the HealthStart standards and guidelines.

RESEARCH QUESTION 3: What are the barriers to the delivery of services?

Each pregnant woman is asked to list and rank the barriers which prevented her from receiving early prenatal care services. An average of 64% of the HealthStart clients reported no barrier to their receiving prenatal services. Where barriers were identified, the woman's own inability, unwillingness or refusal to obtain services was ranked first by 14%. Language, transportation, cost, child care, and problems with scheduling were ranked highest by 5% or less. Only the language barrier had increased over the four years.

When HealthStart women who entered care late were asked to give their reasons on the MSSD form, 32% responded that the main reason was that prenatal care was not perceived as important or a priority. The next highest reason, given by 10% of the women, was a problem with money or insurance. Less than 5% said that problems with appointments, or other agency-related problems were the cause of delayed care.

RESEARCH QUESTION 4: How effective has the HealthStart program been in improving health outcomes of these women and their babies?

Based on the MSSD data, the percentage of HealthStart women with uncomplicated labor or delivery increased from 53% in 1988 to 65% in 1991. During the same time period, the percentage of HealthStart women giving birth by cesarean section has decreased from 20% to 18%.

Medicaid women had significantly more premature deliveries (19%) than other women (10%) between 1985 and 1989. HealthStart women had significantly fewer premature births than other Medicaid women in 1989. (Due to small numbers of prematurity (less than 37 weeks aestation), low birth weight (less than 2500 grams) and infant mortality

a woman receiving HealthStart services was nearly identical to that of other Medicaid women in 1988 and 1989.

The average infant birth costs also were higher for Medicaid ($988) than non-Medicaid ($893) in 1985, and increased steadily until 1987 for both groups of women, but then started to decline in 1988. The cost of the birth of the infant was lower for the HealthStart infant than that of other Medicaid infants, and more closely paralleled that of non-Medicaid infants in 1988 and 1989.

CONCLUSION

In answering the five major research questions, the preliminary LDS data comparisons and the use of the MSSD

and technical expertise in data collection, definition, and analysis in the past decade. The ability to produce good outcomes research has been and will be improving as our abilities to use good methodologies and technologies improves.

1. National Center for Health Statistics. Monthly vital statistics report: vol 38 no 13, provisional data. Hyattsville, Maryland: Public Health Service. 1990.

Session X

Health Data for Program Management

Toward the year

2000

Refining the measures

PUBLIC LAW 99-457 CHILD FIND AND TRACKING SYSTEM

Joyce M. Eatmon, Arkansas Center for Health Statistics

Introduction

In October of 1986, Congress enacted Public Law 99-457, which reauthorized the Education of the Handicapped Act (Public Law 94-142) and amended the act to include an early intervention services program. This act requires that states provide special services to developmentally disabled children under three years of age. Also required is a data management and child find system. Arkansas' approach in developing the Child Find and Tracking System is to draw upon data from existing data systems. Integration of data from multiple sources increases the likelihood that eligible clients will be identified, enhances the value of the individual data systems for administration, program management, policy analysis and research, and reduces the costs of implementing Public Law 99-457.

Child Find

There are two components to the Child Find and Tracking System. The first is the "child find" module. The second is concerned with "tracking" the child's progress. The "child find' component involves identifying those infants and young children who are developmentally delayed or at risk of being developmentally delayed. Once these infants and toddlers are identified, they may then receive specialized services to treat their problems and prevent the development 'of future disabilities. The "child find" component is accomplished by drawing upon data from at least two separate sources. They are Vital Records/Health Statistics, and High Risk Infant Case Management Program. Other possible data systems but not currently being used are Medicaid Paid Claim Files, Arkansas Reproductive Health Monitoring, Children's Medical Services, and Arkansas Department of Health Patient Management System.

The foundation of the child find system begins with Vital Records data. Birth certificate data provides a virtual wealth of data on every infant born in or to a resident of Arkansas. For example, socio-economic characteristics and medical risk factors of the mother, abnormal conditions of the newborn, congenital anomalies, gestational age, and birth weight. Since child find and early intervention is the identification of infants with serious medical needs there exists the, possibility that some of them may have died. For this reason, the child death file is linked to the birth file so that these children are not included with eligible child find children. Included in this paper are births occuring from January 1, 1989 through December 31, 1992. During this period, there were 148,417 births to Arkansas residents. 1,667 of these births were identified as either an infant or child death. After identifying certain selected conditions on the live birth file, 16,629 possible early intervention births were found.

The **second data system used for the child find system is the Arkansas Department of Health's High** Risk Infant Case Management Program. This four county pilot program involves targeted case management for low birth weight infants, infants with congenital anomalies, infants with teenage mothers and other socio-economic conditions such as substance abuse and an unstable home environment.

The reporting system for the High Risk Infant program includes two forms, the entry or identification and screening form and the tracking or encounter form. 883 high risk infant entry forms have been reported for infants and toddlers born from 1989 through 1992. 998 tracking forms tracking forms have been reported. Since there may be more than one tracking form per child, this file is aggregated to one record per child. Although the entry point for this system should be the entry form, tracking forms are often received for children we have not received entry forms for. To get an accurate count of the children in the High Risk data system, the two data bases are merged. After merging the entry form and tracking form data bases, 938 children were reported in the high risk data system.

To create the child find data base, the vital records data base and the high risk infant data base were merged. 853 high risk infants were matched to the live birth file. After merging the early intervention births and the high risk infant births, a total of 17,482 births were included in the child find data base. Only 225 of the High Risk infants were identified as possible early intervention children from the birth certificate.

There are definite differences between the two child find groups. High Risk Infant mothers have a greater percentage of mothers under 18 years of age at 28.8 percent while Vital Records mothers have 8.3 percent. When we look at education, we see that 27.4 percent of Vital Records mothers have less than 12 years of education while 60.9 percent of the High Risk mothers have less than 12 years of education. High Risk mothers are more likely to be unmarried at 71.8 percent. Vital Records mothers have 33.4 percent unmarried. High Risk mothers are also more likely to be Black with 72.9 percent. Vital Records mothers are 26.4 percent Black. High Risk mothers are more likely to receive late or no prenatal care with 18.5 percent, while 7.5 percent of Vital Records mothers received late or no prenatal care. High Risk mothers also are more likely to have had a previous birth. 21.0 percent had three or more previous live births in contrast to 9.7 percent Vital Records mothers. Weeks gestation also differs between the two sources. 21.7 percent of Vital Records births were less than 36 weeks gestation while 31.1 percent of High Risk births were less than 36 **weeks gestation. Percent of births with birth weight under 1500 grams were similar for both groups. However, the percent of births between 1500-2499 grams differed with 13.9 percent for Vital Records and 27.4 percent for High Risk.**

Tracking

The "tracking" of early intervention children is the second component of the child find and tracking system. The tracking component requires that the child be monitored to assure that appropriate services are being provided in a timely manner. This information is made available through the early intervention data collection system. Case Managers and other appropriate personnel report referrals made and the services provided for each child through the use of sixteen laptop computers available in the regional offices. The case managers may then update a child's record on a weekly basis.

The data collection system is composed of five data bases. Each one is tailored to a specific element of the Early Intervention System. They are Child Identification, Early Intervention Services Received, Early Intervention Services Needed, IFSP (Individualized Family Service Plan) Updates, and Transitions. All of these databases were designed with the purpose of providing ease in case management of eligible children, satisfying the federal reporting requirements, epidemiological research, linkages to other databases such as Vital Statistics and Medicaid, and evaluating the effectiveness of the Early Intervention System.

The first database, Child Identification, includes all children that have been referred for early intervention services. It has all the referral, demographic and identifying information on the child. Also included is information on parental consent, eligibility status and the developmental status of the child upon entry into the program.

Early Intervention Services Received contains information on the services that a child is receiving or has received. Included in this data base is information on the service location, provider, funding source, and dates of service. Related to the Early Intervention Services Received data base is the Early Intervention Services Needed data base. It contains information on the services that a child needs but is not receiving, including the barriers to the service.

Essential to case management is the IFSP. The IFSP Updates data base has basic information related to every IFSP written and assessments made on a child. Included in this data base is the type of IFSP and/or assessment that was performed on the child.

The transitions data base is provided for exiting the child from Early Intervention. Included in this data base is the developmental status of the child and what type of service the child will receive after exiting the program.

Conclusion

While the Child Find and Tracking System is primarily an administrative data base, there are certain benefits derived from its use. With the different data sources merged into one, further information becomes available that are not available when the sources remain separate. A new data base is created which enables case management, programmatic analyses, evaluation and research. With this large source of data, researchers may be able to identify those factors which are associated with the conditions that exist in early intervention children. With this knowledge, the early intervention system could be enhanced to provide better and more specialized services.

THE COMPETITIVE INITIATIVES PROGRAM--A NEW JERSEY RESPONSE TO THE ESCALATING COSTS OF HEALTH CARE

Marcia M. Sass, New Jersey Department of Health
C. Epstein, M.C. Fulcomer, R.M. Martin,
S.P. Ficara, and B. Siegel

When Governor Jim Florio took office in 1990, he stated his commitment to address the serious problem of escalating health care costs. Shortly thereafter, he established the Governor's Commission on Health Care Costs to study the sources of the problems and propose recommendations. This Commission published approximately 90 recommendations (Governor's, 1990) many of which were incorporated in Public Law 1991, Chapter 187, The Health Care Cost Reduction Act. This multi-faceted, omnibus health care reform act was aimed at reducing health care costs by introducing measures to assess, plan, build system infrastructure, and redirect patients to the most appropriate health care settings. Two key components of C.187, aimed at expanding nonhospital-based delivery systems for preventive and primary care, were among the recommendations from this Commission. A 0.53% assessment on all acute care hospital revenues in the State to establish the Health Care Cost Reduction Fund (HCCRF), serves as the funding source.

This paper will highlight uses of data in the development of programs and selection of sites and will address practical aspects of implementation and evaluation of a key component of C.187, New Jersey's Competitive Initiatives Program (CIP). This demonstration program is intended to strengthen relationships between urban hospitals with high uncompensated health care costs and community health centers in the delivery of primary care in an effort to reduce these health care costs. The second C.187 primary care demonstration initiative includes expanding services and hours of operation at Federally Qualified Health Centers (FQHCs). Because three of the CIP sites are connected with FQHCs, these will be mentioned throughout. Both of these demonstrations included mechanisms to reimburse the health centers either through the HCCRF or New Jersey Health Care Trust Fund monies (C.187, 1991). The use of key data sources will be emphasized throughout including the Current Population Survey; the Uniform Bill Patient Summary (UB-82) collected on approximately 1.3 million acute care hospital admissions each year; surveys of inpatient and emergency room (ER) utilization from 96 acute care hospitals: and Codman small area analysis data (Codman, 1993).

BACKGROUND INFORMATION THAT NEW JERSEY CONTINUES TO TRACK

Demographics of the Uninsured Population

In 1989 in New Jersey, there were more than 800,000 uninsured individuals under the age of 65. In terms of demographics, by race and ethnicity, the majority were white, although a disproportionate share were minorities. In particular, one in six African Americans and one in four Hispanics under the age of 65 had no health insurance coverage. For the population as a whole 11.7 % under age 65 lacked coverage. Even those with coverage often had limited benefits and lacked access to some services (NJDOH, 1992).

Of note is that not all uninsured persons are from the lowest income groups. According to a recent report on the uninsured in New Jersey (Kronick, 1990), 39% of all uninsured in New Jersey were in a family with incomes above 300% of the poverty level. About 428,000 uninsured persons, or 55% of New Jersey's uninsured population, were full time workers or dependents of full time workers. An additional 85,000 or 11% worked about 75% of full time or were dependents of such workers. Therefore, two-thirds of uninsured persons represent the working population. Many are self-employed or in small businesses with fewer than 100 persons, often in retail trade and service industries. Employees in these businesses usually lack the resources to obtain health insurance on their own.

In New Jersey, a UMDNJ Eagleton Poll (UMDNJ, 1992) indicated that 20% of those who had a health-related problem, and did not see a physician, failed to do so because of inadequate health insurance. Compared to New Jerseyans with health insurance, those without were five times less likely to have visited a physician even when they had a health problem. These individuals are often forced into crisis-oriented, episodic care provided in more costly settings such as the emergency room. New Jersey has made provisions through its hospital reimbursement system to provide care to anyone seeking services at any emergency room in the State. Through a surcharge on patients' bills, in effect until January, 1993, when new health care reform legislation took effect (C.160, 1992), it assured all residents access to hospital care by guaranteeing hospitals full compensation for

uncompensated care (charity care and bad debt) costs. For 1992, these costs were projected at $750 million (NJDOH, 1992a). As a companion to its reimbursement system, New Jersey has been collecting the UB-82, a rich database for both hospital inpatient, outpatient, and emergency room services. Information from this database converted into the Codman PANDORA database, as well as annual and quarterly surveys of inpatient utilization, have enabled the State to study the impacts of ambulatory care sensitive (ACS) conditions--those that generally could have been prevented or minimized with earlier preventive and primary care--and use of hospital outpatient services (Billings, Zeitel, Lukomnik, Carey, Blank & Newman, 1993; Millman, 1993).

Ambulatory Care Sensitive Conditions

One method of assessing the impact of barriers to preventive and primary care is to study hospitalizations for ACS conditions. Using this approach (Billings, Zeitel, Lukomnik, Carey, Blank & Newman, 1993; Millman, 1993), in 1991, there were about 183,481 such admissions--roughly one out of eight admissions--that might have been the result of a failure of New Jersey's primary care system. In 1990, the charges for these admissions amounted to over $945 million, and in 1991 (TABLE I), over $1.1 billion. The distribution of costs by payer closely parallels that for all inpatient admissions --meaning that early and adequate access may be a problem for many New Jerseyans. In 1991, this represented costs to the system exceeding $90,000,000 for those uninsured individuals with conditions that might have been preventable (Codman, 1993).

Princeton. The New Jersey adjusted rate was 23.7 per 1,000. Essex and Hudson Counties, which contain Newark and Jersey City respectively, among the areas with the highest rates on the list, have the highest percentages of minority residents in the State, 54.9% and 52.6% respectively (NJDOH, 1992).

Hospital Outpatient Departments

In 1989, New Jersey hospitals reported a total of about 10,000,000 outpatient visits. Of these, 21 percent were emergency room visits, and 18 percent were clinic visits. Of the 1.8 million clinic visits, 58% were for primary care. It is estimated that about one-third of the ER visits were for primary care as well. Of the total outpatient visits, 64% were made 'by Medicaid, self-pay (uninsured) and charity care (less than 250% below the poverty level) patients. At an average cost of $114 per OPD visit and $134 per ER visit, it is conservatively estimatedthatthe combined costs for the OPD visits for primary care in 1989 in these locations were at least $100 million. For primary care, the costs in this time period were roughly $200 million. Admissions to the hospital from the ER represent about 19% to 20% of all ER visits.

The New Jersev Medically Underserved Index

The development of this index was called for in C.187. It evaluates economic and health status indicators that point to relatively more underserved areas. Significant overlap is revealed in comparing high need areas identified by the ACS conditions and NJMUI rankings.

TABLE I
COSTS OF ACS INPATIENT ADMISSIONS TO NEW JERSEY HOSPITALS FOR 1991

Payer All	Cost of ACS Admissions*	Percent of Total Cost	Percent of Admissions
Medicaid	$ 114,604	10.2	10.8
Medicare	$ 530,633	47.8	47.8
Other/Federal	$ 328.828	4.5	4.1
Private Insurance	[illegible]	29.3	29.8
Self-Pay/Charity	$ 91,362	8.2	7.5
Total	$1,121,508	100.0	100.0

Source: (C) PANDORA 1991-1993, The Codman Research Group, Inc.
*Costs in 1000's

TABLE II rather dramatically illustrates the 1991 ACS condition age- and sex-adjusted admission rates for selected ZIP codes and counties in New Jersey. In rank order these range from 90.4 per 1000 admissions in areas in Newark to a low of 2.1 per 1,000 around

SELECTION OF SITES

A rigorous Request for Application (RFA), requiring hospital applicants to incorporate how their institutions compared on rates of ACS condition

admissions, percents of uncompensated care, emergency room use for nonemergent conditions, and use of outpatient clinics, in their proposals (NJDOH, 1991) was circulated in the fall of 1991. Selection of hospital sites was based on inclusion of these data in a comprehensive proposal to address the problems in the particular service area. Four large pilot projects, each receiving $1.3 million in each of two years, seven smaller "look-alike" sites receiving up to $200,000 per year, two planning grants, and eleven mixed grants up to $200,000 per year were recommended. Twenty-three have been funded and are currently in various stages of implementation.

but also of the sponsoring agency. It takes a sufficient staff with knowledge of the surrounding legislative/political environment, an understanding of and an ability to work with the original legislation and idiosyncrasies of the sponsoring agency, and the capacity of the sponsoring agency to provide clinical, reimbursement, and evaluation expertise to the sites. For example, groundbreaking policies to reimburse the health centers for charity care eligible patients from Chapter 83 Health Care Trust Fund monies and detailed guidelines for grant coverage of pharmaceuticals prescribed by the health centers for eligible patients had to be developed in addition to the

TABLE II
DRG DATABASE FOR 1991
SELECTED ZIP CODE AREAS, ALL AGES - ACS-TOTAL CONDITIONS
SORTED BY ADMISSION VALUE
ADJUSTED RATE CALCULATION

ZIP Code/County	ObsAdm	Adm/ARC	PtDays	LOS
Newark	938	90.43	983.23	10.9
Newark	1869	66.09	710.68	10.8
Paterson	1700	55.79	538.84	9.7
Jersey City	2581	50.09	488.14	9.7
Atlantic City	1960	45.70	390.70	8.5
Camden	1017	43.08	369.41	8.6
Trenton	662	42.09	344.22	8.2
Perth Amboy	1750	41.71	372.22	8.9
Paterson	679	41.69	369.17	8.9
Jersey City	1537	40.43	386.28	9.6
Salem	529	37.73	278.23	7.4
North Plainfield	1682	33.77	311.83	9.2
*CUMBERLAND COUNTY	4069	28.67	216.78	7.6
*ATLANTIC COUNTY	6792	28.55	233.53	8.2
**Total NJ	183481	23.74	210.42	8.9
Princeton	11	2.12	9.26	4.4

(C) PANDORA 1991-1993, The Codman Research Group, Inc.
* Total for county residents
** Total for all State residents

In addition, to the four main criteria, selection was based on the applicant's plan and timeframe for implementation, the ability and commitment of the applicant to develop nonhospital based comprehensive primary care services, and the likelihood that the program would impact health status indices, improve coordination of primary care delivery with other social services, and reduce ER use and hospital ACS condition admission rates. Further, the applicant had to be willing and able to participate in an evaluation.

REALITIES/PRACTICALITIES OF IMPLEMENTATION

Implementation of this type of program requires not only a partnership of the hospital and community based provider

evaluation strategy. In New Jersey, almost none of these had been considered when the original sites were selected. It has ranged from difficult to impossible to provide the necessarytechnicalassistance to sites for implementation of all aspects of the program. With the short two year timeframe given to these sites, it is unlikely that all aspects, such as the evaluation, will be accomplished as planned.

EVALUATION

For the larae sites. because of all of the confounding variables, including new health care reform legislation (C.160, 1992), general system data are insufficient to measure effects. Therefore, a retrospective/prospective design using a patient referral database, linked hospital ER and inpatient UB-82

data, and a health center prospective encounter database has been planned. Individuals enrolled in the CIP are to be tracked retrospectively for two years for hospital ER, ACS conditions, and primary care visits at the health center. Efforts in this area have been thwarted by layoffs, insufficient sponsoring agency staff, and lack of named UB-data.

Overall, there are some beginning indications that the FQHCs (among which three of the large CIP sites are connected), may be having some effect. For all the FQHCs, there were 65,678 primary care visits above the baseline at the end of the first year of expansion efforts. Though trends must be watched, total ER visits in the State appear to have dropped by more than 41,000 during this same time period. For the CIP sites which had later start dates, a couple of the smaller ones appear to be having some effect on ER use at the partner hospitals. For one that began expansion of its family health center in the third quarter of 1992, and succeeded in increasing primary care visits from about 50 to 300 per month, the partner hospital experienced a reduction of over 560 visits in the fourth quarter alone when compared with the previous year. At another active CIP site serving only charity care and self-pay patients, two local hospitals have experienced reductions of ER visits by at least 10% since the primary care site has become active (NJDOH, 1993).

SUMMARY

The CIP targeted urban hospitals with high emergency room utilization for primary care services, high uncompensated care costs (from charity care and bad debt patients), and high rates of hospitalizations. Through a rigorous request for application (RFA) process, these hospitals have been challenged to develop working relationships with lower cost community health centers to move patients into more appropriate care settings. These are very exciting programs. Use of data and practical aspects of implementation to enhance the likelihood of the success of such programs and their evaluation have been addressed. Two main conclusions can be drawn: 1) it is criticalto have sufficient sponsoring agency support staff in place when launching such programs; and 2) timeframes must be greater than two years when major infrastructure changes are planned to enable sites to respond to the challenge.

REFERENCES

Billings, J., Zeitel, L.,Lukomnik, J., Carey, T.S., blank, A.E., & Newman, L. (Spring, 1993). Impact of socio-economic status on hospital use in New York City. Health Affairs, 12, 162-173.

Codman Research Group, Inc. (1993). PANDORA 2.04a. Lebanon, New Hampshire: The Codman Research Group, Inc.

Governor's Commission on Health Care costs. (1990). CARE. Trenton, New Jersey: Governor's Office, State of New Jersey.

Kronick, R. (1990). Health insurance for the uninsured in New Jersey. La Jolla, CA: University of California, San Diego.

Millman, M., ed. (1993). Access to health care in America. Washington, D.C.: National Academy Press.

NJDOH. (1991). Competitive initiatives program request for applications. Trenton, New Jersey: New Jersey State Department of Health, Office of Health Policy and Research.

NJDOH. (1992). Preventive and primary care. New Jersey state health plan. Trenton, New Jersey: New Jersey State Department of Health, Office of Health Policy and Research.

NJDOH. (1992a). The New Jersey health care trust fund. A report to the governor and legislature. Trenton, New Jersey: New Jersey State Department of Health, Health Care for the Uninsured Program.

NJDOH. (1993). Unpublished B-2 utilization data. Trenton, New Jersey: New Jersey State Department of Health, Center for Health Statistics.

P.L. 1991, Chapter 187 (1991). The health care cost reduction act. Trenton, New Jersey: New Jersey State Legislature.

P.L. 1992, Chapter 160 (1992). The health care reform act of 1992. Trenton, New Jersey: New Jersey State Legislature.

UMDNJ-Eagleton Poll Results. (1992). Newark, New Jersey: University of Medicine and Dentistry of New Jersey News Service.

DEVELOPMENT OF A DATA SYSTEM FOR THE MANAGEMENT OF A COMPREHENSIVE CANCER SCREENING PROGRAM

Patricia M. Jamison
Centers for Disease Control and Prevention

(Not available for publication)

USING MEDICARE'S END STAGE RENAL DISEASE PROGRAM MANAGEMENT AND MEDICAL INFORMATION DATABASE FOR PROGRAM MONITORING AND POLICY DEVELOPMENT

Joel W. Greer, Health Care Financing Administration

INTRODUCTION

As the insurer for approximately 93 percent of chronic dialysis patients, the Medicare program has a responsibility to monitor and assess their medical care. The Health Care Financing Administration (HCFA), the Federal government agency administering Medicare, receives bills for most medical services provided to dialysis patients. The billing information and other data about Medicare end stage renal disease (ESRD) patients is compiled in a computerized database called the ESRD Program Management and Medical Information System. HCFA has begun using its exiting computerized information system to carry out its program oversight functions rather than undertake expensive and sometimes intrusive new data collection projects.

HCFA, like many private insurance companies, has designed its billing forms and procedures in an attempt to efficiently, quickly and accurately pay providers for covered and necessary medical services given to beneficiaries. Data derived from claims are not designed to be used for policy evaluation, quality assurance or biomedical research. However, health services researchers within and outside of HCFA have attempted to use such data for these purposes with varying degrees of success. On the surface, using claims data would appear to have some important advantages:

- The data are already computerized.
- There are large numbers of patients and services.
- The data cover a wide range of covered medical services and treatment settings--doctor visits, hospitalizations and outpatient clinics.
- The data contain charges and payments.

This paper uses the example of a newly approved pharmaceutical for chronic renal dialysis patients--human recombinant Erythropoietin (EPO)--to discuss the strengths and limitations of HCFA administrative data for

- monitoring patient treatment patterns and outcomes;
- evaluating HCFA policies and regulations; and,
- assessing and improving quality of care.

BACKGROUND

EPO is one of the first recombinant-DNA produced biologicals to receive Food and Drug Administration (FDA) approval for human use and to be widely marketed.' Because its first FDA approval was for the treatment of anemia associated with chronic renal failure, Medicare was the predominant purchaser. At the time of its approval, Amgen, Inc. was the sole seller in the United States. Amgen had invested heavily in developing EPO and getting it through the FDA review process. EPO was. safe and effective, but it was also expensive. It was estimated that EPO might cost $100 million in its first year with costs rising as its use spread to approximately a quarter billion dollars per year.[2] Given the unusual market structure with one seller and one major buyer and the problems of pricing a high-tech biological, it was apparent to the Federal government that special attention would have to be paid to monitoring the drug and assessing its medical and financial impacts.[3]

HCFA began coverage for EPO provided in outpatient dialysis facilities in June 1989 at an interim payment policy of $40 per administration with a $30 add-on for doses of 10,000 or more units.[4] Although the Government had expended much work in estimating an appropriate price and payment system, the policy was preliminary since much of the required information was unknown.[5] The Office of Technology Assessment (OTA) examined 9 payment systems and compared their strengths and weaknesses.[3] The Department of Health and Human Services office of the Inspector General (OIG) examined the costs of developing, manufacturing and distributing EPO that were relevant to setting a fair price for EPO.[3,5] However the OIG was given only limited access to the proprietary commercial and financial data necessary to estimate costs.[6] Finally, the dosing patterns and rate of diffusion into the dialysis population could only be crudely estimated from the results of the clinical trials.

HCFA believed the $40 per administration payment would

- cover dialysis facility costs for the pharmaceutical and its administration at an average dose of 3,500 units[7] and
- give Amgen a fair return on its investment in research and development.[2,3]

HCFA clearly stated that the initial rates were interim rates intended to allow timely access to EPO. A dialysis

facility would make a profit on some patients who responded to low doses and a loss on others who required high doses, but the payment covered overall costs at a clinically conservative average dose. It is not clear how the 3,500 unit average dose was derived. Based on the clinical trials, the OIG, OTA and HCFA expected doses to average 5,000 units.[3,5,8]

The dose actually administered might not be independent of the payment policy. The fixed payment per administration payment gave facilities a financial incentive (1) to limit EPO dosage and (2) to give EPO to many patients.[9] The lower the dose, the greater the profit per recipient; the greater the number of recipients, the greater the total profit. One of the important research questions investigated using billing data was the extent to which financial incentives influenced anemia treatment practices.

EPO BILLING PROCEDURES AND DATA

Billing data do not usually contain the detailed dosing and outcome measures needed for monitoring quality of care since billing procedures generally focus on efficient reimbursement. However for EPO, HCFA required providers to include on each bill the dose per administration, the number of administrations and the patient's most recent hematocrit. The clinical trials had used the hematocrit--a measure of the density of red blood cells in the blood--as the outcome measure.

Although the dosage and the hematocrit were collected primarily for reimbursement purposes, HCFA required more detailed information on EPO bills than was required for payment alone. The exact dose was collected although payment required knowing only whether the dose was greater than 10,000 units. HCFA wanted the hematocrit to determine whether EPO was medically necessary but it was collected to the nearest tenth of a percentage point. HCFA data often do not identify the dosage (only the number of administrations during a time period) and seldom contain a clinically appropriate outcome measure. The greater detail than are usually available from administrative data were one important element in the success of the EPO monitoring program.

Another critical element was timeliness. Due to delays in submission or processing, billing records may take many months to become complete.[10] HCFA's Bureau of Data Management and Strategy implemented a computerized fast-track data collection and monitoring system into which all EPO bills were copied as soon as they were paid. This allowed distribution within HCFA of a monthly worksheet that included the number of patients receiving EPO, charges, and the mean dose and hematocrit. Although these data were by month of bill processing rather than month of service, they were available about two months after the date of service.' EPO billing files created by the usual processing path take nine months from the date of service to become 95 percent complete. In addition to speeding up the data collection process, the EPO data files saved analysis time by performing much of the data gathering, formatting and cleaning steps. The 4-6 months saved by the expedited monitoring files were critical in making the data available to policy makers in a timely manner.

EARLY ANALYSES LED TO A CHANGE IN PAYMENT POLICY

Monitoring of EPO began soon after coverage of EPO. The monthly spreadsheets mentioned above were distributed to operations, policy and research sections within HCFA beginning in December 1989.' HCFA analyst Paul Eggers used the EPO data files to do person-level analyses. HCFA also awarded a contract to a research team lead by Neil Powe to study the use of EPO in the year following coverage of EPO. The Department of Health and Human Services' Office of the Inspector General (OIG) directly sampled dialysis facilities to create its own database of EPO billings.

There were preliminary results within 9 months of its first coverage by Medicare.[3,8,11]

- EPO was diffusing rapidly among the dialysis population--faster than had been expected.
- The mean dose was 2,600-2,700 units of EPO with about one-third of patients receiving the modal dose of 2,000 units.
- Hematocrits rose but remained below 30 for most patients. While there are no consensus guidelines on anemia therapy among dialysis patients, the clinical trials and the literature of the time suggested target ranges with minimum bounds of 32 or higher.
- There was no evidence of the dose-responsive rise in hematocrit found in the clinical trials.
- There was no evidence of physicians prescribing an initial high dose to raise hematocrits followed by a lower long-term maintenance dose.

While modified and expanded upon, all these findings were confirmed by later analyses.

In May 1990 the OIG began circulating a draft report on its investigation of EPO. The OIG concluded that the mean dose was 2,700 units and that at this dose facilities earned a gross profit of 42 percent. The OIG recommended a payment based on the number of units of EPO adminis-

tered to lower total Government expenditures. The OIG estimated Government payments to be $265 million per year (80 percent of EPO charges of $330 million) and increasing as EPO diffused through the dialysis population. The OIG recommended *@Eliminating the flat interim rate and establish[ing] a payment rate based on actual units of EPO administered up to a capped amount."[8] In its July 1990 comments to the OIG, HCFA expressed concern that payment based on dosage could create incentives to overperscribe EPO and said HCFA was studying the issue."

On June 14, 1990 the House Subcommittee on Health, Committee on Ways and Means held hearings on EPO. Representatives from the OIG, OTA and dialysis providers testified. The testimony focused on the large facility profits. There was little mention of the low hematocrits. All witnesses favored covering EPO when self-administered by home dialysis patients. Most witnesses supported, at least in principal, payments based on units of EPO dispensed, although dialysis facilities wanted no policy change at that time.[13]

In the Omnibus Reconciliation Act of 1990 Congress

- changed the payment policy changed to $11 per 1,000 units, effective January 1, 1991, and
- authorized HCFA to pay for EPO for self-administration to home dialysis patients.

Coverage of self-administered EPO required Congressional action since the Social Security Act generally prohibits Medicare from covering pharmaceuticals that are self-administered." Congressional intervention was not needed for the change in payment policy. HCFA was analyzing EPO utilization and outcome data and seriously considering alternatives to the interim payment policy. Congress acted quickly in part because it was Willing to use preliminary results as the basis for action and in part because it was willing to specify a dollar amount with only limited data. Nowhere in the published testimony or reports is there any discussion of an appropriate per unit payment rate. However by mandating a specific payment rate, Congress removed Administration flexibility. Future changes in payment rate will have to be made by Congress. current (August 1993) newspaper reports indicate that Congress Will change the rate to $10 per 1,000 units in OBRA 1993.

RESULTS OF THE POLICY CHANGE

Figure 1 shows the dose per administration and the mean hematocrit before and after the payment policy change. The change in payment policy which became effective January 1, 1991 appears to have had a large impact on physicians' prescription of EPO. In the first six months of EPO availability, average EPO doses were low. Providers begun increasing EPO doses at about the time the change in policy was announced. The rate of increase grew after the new policy went into effect. The current average dose per administration of about 4,100 units is consistent with dosing guidelines based on the clinical trials.

It is not possible to draw conclusions about causality from simple time trends such as presented in this graph. However, the low initial dose and its large increase following the change to a per unit payment is certainly consistent with the financial incentives of the payment policies. Unlike the old payment policy, the new one removed any financial incentive to limit EPO doses and may have created a small incentive to increase doses. Furthermore, more sophisticated analyses also suggest that the change in dosing patterns were caused by the change in policy.[15,16,17]

Under the new payment policy, Medicare EPO expenditures depend on the delivered dosage. Medicare payments for EPO decreased substantially after the new payment policy became effective. They dropped from an annualized rate of $319 million per year in the last three months of 1990 to $269 million per year in the first three months of 1991. Fueled by more EPO recipients and higher EPO dosing, annualized expenditures exceeded their highest level under the old policy in the second quarter of 1991 and have risen continuously since. Current estimates of annualized EPO expenditures exceed 500 million per year.

As shown in Figure 1, the hematocrits changed only a little despite the higher EPO doses. From fourth quarter of 1990 to fourth quarter of 1991 the mean hematocrit increased by only 1.0 percent while the mean EPO dose rose 26.5 percent. Over the same one year period, the percentage of patients receiving EPO for more than 3 months increased from 46.6 to 49.4 percent.' Analyses of person level data also show only a small dose-hematocrit relationship in clinical practice.[16] It is not known why hematocrits do not exhibit the dose-response rise seen in the clinical trials or remain below 30 despite the higher doses.

Dosing practices appear to be in line with those of the clinical trials, but the expected outcomes have not been achieved. Clinical trials establishing the efficacy and safety of new drugs take place in controlled practice settings and often on a selected sample of patients. For example, the phase III clinical trials of EPO excluded patients with systemic diseases.[18] More importantly,

efficacy is measured by prescribing exogenously determined dosages and measuring the resulting hematocrit. The efficiency of a drug in actual practice settings estimated from billing data measures something different. In clinical practice, the observed dose is likely to be based on the patient's response to previous EPO administrations. Physicians may increase the dose until the patient reaches a target hematocrit or until anemia symptoms moderate. As a result it is not necessarily surprising that effectiveness of EPO in actual practice is significantly lower than the efficacy of the trials. What may be of concern is the low average hematocrit and the surprising high number of persons with hematocrits below 30. Unfortunately, the billing data alone are not adequate to suggest the cause or causes of the lower outcomes."

THE EPO QUALITY ASSURANCE PROGRAM

The lower than expected hematocrits are one reason driving the choice of anemia treatment as the first ESRD Health Care Quality Improvement Program (HCQIP) topic. HCFA and the ESRD Networks are developing a patient data base to display the relations between EPO and hematocrit. Tables and graphs comparing each provider's treatment patterns and outcome measures to local and national averages are to be created using centralized data and computing capabilities. They will be given to local level Medical Review Boards and to the individual providers. The intention is to study at the local level how treatment and outcomes can be improved. It is hoped access to data on their own performance will stimulate providers to find ways to change practices in appropriate ways.[20]

Figures 2 and 3 are examples of the type of processed data that might be presented to a arbitrarily selected facility. Figure 2 displays the facility's average weekly dose of EPO--a measure of treatment pattern. The facility's performance is compared with the 95 percent confidence interval (CI) of its state. The CI is adjusted to reflect the number of patients in the facility in each quarter. As can be seen, the chosen facility was within the state range except for one quarter where it delivered higher doses.

Figure 3 shows the same facility's mean hematocrit and state CI. The facility was frequently below the state norm. The state as a whole was generally below the 30 percent target. This particular facility delivered doses of EPO similar to other facilities in its state, but experience lower outcomes. The most common factors believed to lower responsiveness to EPO are iron deficiency, inflammation and infection, and aluminum overload. Most of these are treatable, or at least can be ameliorated with additional medical intervention. There are also some relatively rare conditions which lower responsiveness to EPO and are not treatable. Non-responding patients can be tested and if necessary given large doses of EPO. With the advent of EPO, most patients can reach an acceptable hematocrit level. It is hoped that data on practice patterns and patient outcomes will give providers the data and motivation to refine their treatments to improve outcomes.

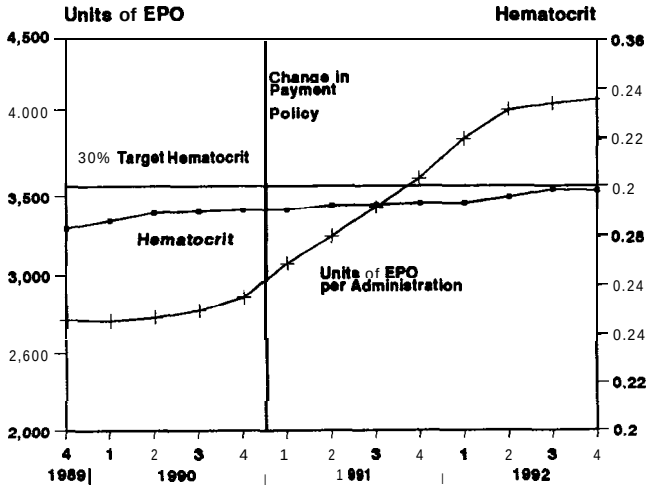
CONCLUSIONS

HCFA billing data for EPO provide a successful example of the use of data derived from billing and administrative forms. Data from the first period after EPO's approval showed rapid diffusion, but low doses, low outcomes and high facility profits. Analyses were begun soon after EPO was covered and were one reason for Congress legislating a new payment policy. Subsequent to the new reimbursement rules, EPO doses increased. Although the expected higher hematocrits have not materialized, this demonstrates that payment policies sometimes exist which provide incentives to improve patient outcomes. Due in part to the large variations in hematocrits, HCFA is now beginning to provide results from billing data to providers in expectation that patient outcomes can be further improved.

There are several reasons EPO billing data were so useful in changing policy. The first was the availability of data on dosing and outcomes from the bill itself. The second was timeliness. HCFA set up a data processing method before the first bill arrived that provided almost real time data. The OIG also started collecting data for its evaluation at an early date. The third was that the data were put into a usable form rapidly. This last step is the role of analysts: to put data into a form which will be meaningful to its audience. For Congress, the OIG emphasized the high profits being made by dialysis facilities. For clinicians, the Johns Hopkins analysts have published or in press seven articles in peer-reviewed journals. Finally, for providers, HCFA is designing data presentation formats based in part on Total Quality Management ideas.

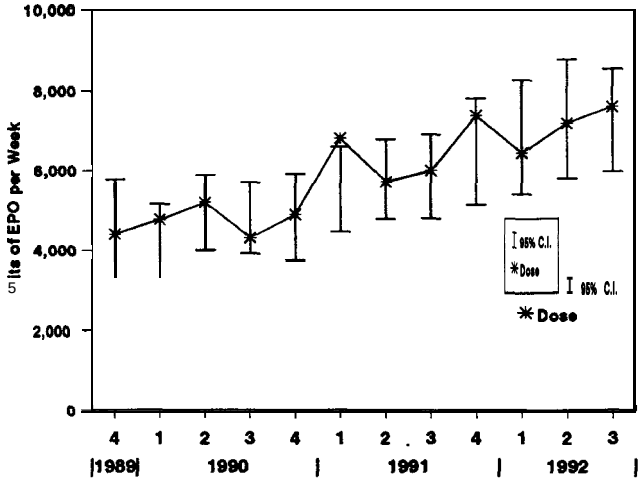
The views in this paper are the author's and do not reflect any position of HCFA.

Figure 1: Mean EPO Dose and Hematocrit



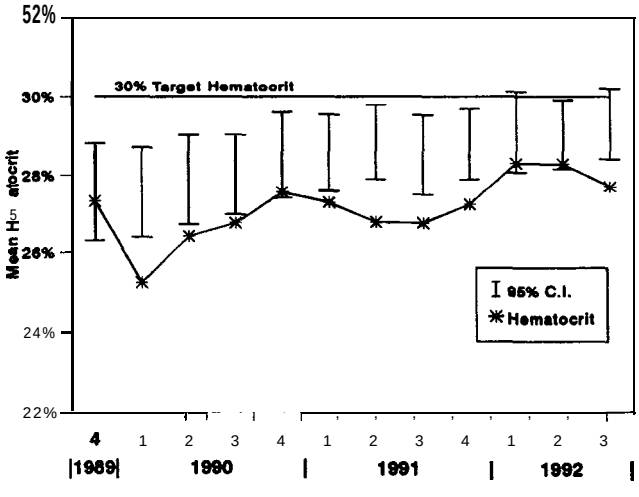
Source: HCFA, ESRD PMMIS. March 1993 update.

Figure 2: Mean Weekly Dose of EPO - Provider #1



Source: HCFA, ESRD PMMIS, March 1993 update

Figure 3: Mean Hematocrit - Provider #1



Source: HCFA, ESRD PMMIS. March 1993 update

REFERENCES

1. Powe NR, Griffiths RI, de Lissovoy G, et al. Access to Recombinant Erythropoietin by Medicare-Entitled Dialysis Patients in the First Year After FDA Approval. JAMA 1992;268(11):1434-40.

2. Mangano M. Testimony before the Subcommittee on Health, House Ways and Means Committee. June 14, 1990.

3. Office of Technology Assessment. Recombinant Erythropoietin: Payment Options for Medicare. Washington, DC: U.S. Government Printing Office, May 1990, OTA-H-451.

4. Health Care Financing Administration. Intermediary Manual--Part 3--Claims Process. Transmittal 1449.

5. Sullivan LW. Report to Congress: Payment for Erythropoietin. Washington, DC: Department of Heath and Human Services, July 1991.

6. Kramer GA. Testimony before the Subcommittee on Health, House Committee on Ways and Means, June 14, 1990.

7. Eggers PW. Recombinant Erythropoietin and Medicare Payment--Letter to the Editor. JAMA. 1991;266(19):2705-6.

8. Office of the Inspector General. The Effect of the Interim Payment Rate for the Drug Epogen on Medicare Expenditures and Dialysis Facility Operations--Draft. CIN: A-01-00512. May 4, 1991.

9. Greer JW, Eggers PW, Milam, R, Powe, NR. Medicare Payment for Recombinant Erythropoietin: Costs, Access and Outcomes. Presentation to the American Public Health Association Meetings, November 9, 1992.

10. United States Renal Data System. 1991 Annual Data Report. Bethesda, MD: National Institute of Diabetes and Digestive and Kidney Diseases, August 1991.

11. Eggers, PW. Personal Communications. December 1990, February 1991.

12. Wilensky, GR. Letter to the Inspector General. July 16, 1990. in Office of the Inspector General. The Effect of the Interim Payment Rate for the Drug Epogen on Medicare Expenditures and Dialysis Facility Operations. CIN: A-01-90-00512. September 4, 1990.

13. Hagen, MW. Testimony of the National Renal Administrators Association to the Subcommittee on Health, House Committee on Ways and Means, June 14, 1991.

14. Social Security Act. 42 CFR 410.29a.

15. Sisk JE, Gianfrancesco FD, Coster JM. Recombinant Erythropoietin and Medicare Payment. JAMA. 1991;266(2);247-52.

16. Powe NR, Griffiths RI, Greer JW, et al. Early Dosing Practices and Effectiveness of Recombinant Human Erythropoietin. Kidney International. 1993;43:1125-33.

17. Powe NR, Griffiths RI, de Lissovoy G, et al. Can Medicare Payment Policy Influence Erythropoietin Dosing? Journal of the American Society of Nephrology. 1992;3(3):430.

18. Eschbach JW, Abdulhadi MH, Browne JK, et al. Recombinant Human Erythropoietin in Anemic Patients with End-Stage Renal Disease: Results of a Phase III Multicenter Clinical Trial. Annals of Internal Medicine. 1989;111:992-1000.

19. Greer, JW. The End Stage Renal Disease Program--Letter. New England Journal of Medicine. 1993;329(2);141.

20. Jencks SF, Wilensky GR. The Health Care Quality Improvement Initiative: A New Approach to Quality Assurance in Medicare. JAMA. 1992;268:900-3

Session Y

Data Support for Medical Care Policy

Toward the year

2000

Refining the measures

DATA SOURCES USED IN PRIORITIZING HEALTH SERVICES UNDER THE OREGON HEALTH PLAN

D. D. Coffman, Oregon Health Services Commission
H. D. Klevit, P. R. Sipes-Meteler

ORIGIN OF THE OREGON HEALTH PLAN

In 1988 the Oregon State Legislature faced a budgetary shortfall in their Medicaid program. To alleviate this shortfall, they decided to discontinue the funding of an optional Medicaid service, organ transplants, so that an additional 1500 women could get necessary prenatal services, This decision was played up by local media when a 7 year old Medicaid eligible boy died of leukemia. And although the State never denied any treatment for the child, the public blamed his death on the new policy.

Health policy experts in Oregon then began to wonder why these two services should be pitted against one another. Perhaps there were still other services less beneficial than both of these that should not be funded in lean fiscal times.

A group of these people established the Oregon Priority Setting Project, which ordered broad categories of health care on a scale of l-10. While the results could not immediately be used in policy decisions, it was concluded that health services could be prioritized.

Dr. John Kitxhaber, then Senate President of the Oregon legislature and an emergency room physician, used this knowledge in crafting three pieces of legislation in 1989 that are now known as the Oregon Health Plan.

THE OREGON HEALTH PLAN LEGISLATION

The piece of 1989 legislation receiving the most attention is Senate Bill (SB) 27. To combat tighter and tighter restrictions for Medicaid eliaibilitv...which currently stands at about $500 a month for a family of three, SB 27 sets eligibility at a constant 100% of Federal Poverty Level. To compensate for the additional 120,000 people this would bring into the program, a fluctuating benefit package dependent on the legislature's funding down a prioritized list of health services would be incorporated. The 11-member Health Services Commission (Commission) was established to create this list and is made up of five physicians, four consumer representatives, one social worker and a public health nurse.

The second piece of legislation, SB 534, creates a high risk insurance pool for those with preexisting conditions who cannot obtain private insurance. Premiums are 150% of coxunnity ratings, with an assessment on insurance companies., Over 3,000 individuals are currently in this pool.

The final piece of the original 1989 legislation is SB 935, a "pay-or-play" mandate for employers with 25 or fewer employees. While on a voluntary basis initially, by July of 1995 these small businesses will have to pay a portion of the medical coverage for their employees and dependents, or pay into a state run pool. This could eventually affect another 300,000 uninsured Oregonians.

In 1991, three related bills were passed. Small market insurance reform is the goal of SB 1076, which ties the employer mandated benefit package to the standard benefit package under the Medicaid list. This will be the minimum package allowable and be guaranteed issue, guaranteed renewable, and rate banded. SB 1077 established the Health Resources Commission to explore cost control issues and technology assessment. And finally, SB 44 brought in the populations of the aged, blind and disabled who were initially excluded from the prioritieation process under 1989's SB 27.

THE HEALTH SERVICES COMMISSION

Originally, the Commission was allowed six short months to develop a prioritized list of health services, a task which had never been attempted anywhere. To complicate matters further, they were given very little guidance by the legislation, only "..to report to the Governor a list of health services ranked by priority, from the most important to the least important, representing the comparative benefits of each service to the entire population to be served." From this directive, the Commission envisioned a list of condition/treatment pairings ranked by a combination of both cost-effectiveness and public values.

Their first order of business was to choose a tool by which to measure benefits of treatment within the cost-effectiveness analysis. Three tools were considered: 1) the Sickness Impact Profile, 2) a category system developed by David Hadorn while working at RAND, and 3) the Quality of Well-Being (QWB) scale developed -by Dr. Robert Kaplan of UC San Diego, with the latter being chosen.

The QWB scale uses public values to quantify the measure of health states on a scale of 0, for death, and 1 being perfect health. To define a health state, one of 24 symptoms can be matched with one of two levels of health limitations in each of three areas: physical activity, social activity and mobility.

For example, while a symptom alone can represent minor vision problems correctable with glasses or contact lenses, matching the same symptom code with a limitation on mobility whereby you cannot

drive, limitations on physical activity due to the use of a cane or other assistance, and a limitation on social activity in not being able to perform certain major role activities, all can be put together to characterize blindness.

With a methodology base for their plan, the Commission created three subcommittees to delegate their work to. The first was the Health Outcomes Subcommittee. They would be responsible for collecting health outcomes from expert provider panels and the literature for each condition/treatment pair on the list. QWB scores could then be matched with the health outcomes information to determine the benefit of each treatment.

The second subcommittee created was the Social Values Subcommittee. They would oversee the 41 community meetings held throughout the state to solicit public values and a 1000 person random digit-dialed telephone survey to obtain the QWB scores for each symptom and health limitation.

Finally, a Subcommittee on Mental Health Care and Chemical Dependency was established to see whether or not these services could be prioritized in the same manner as physical health services.

DATA COLLECTION

In late 1989, data collection began with the Health Outcomes Subcommittee receiving testimony from 54 provider panels representing various medical specialties and sub-specialties. Each provider panel was asked what their practice consisted of, and what the effectiveness is for each of those treatments. International Classification of Disease version 9 (ICD-9) codes were matched for each condition, and Physicians ' Current Procedural Terminology version 4 (CPT-4) codes for each treatment were put into the database. They were asked what the duration of benefit of the treatment would be after the onset of the diagnosis. Then they were asked to give average expected outcomes of death, up to three possible residual effects, and an asymptomatic state of health for both with and without treatment scenarios. An example of the data collected for each condition/treatment pair is shown in Figure 1. These probabilities were asked to be given five years after onset of the diagnosis, so that the resulting net-benefit would act as a single Rehmann sum when multiplied by the duration of benefit.

As expected, with health outcomes research in its infancy, little was found in the literature. Only the Surveillance, Epidemiology, and End Results (SEER) Program, in the area of cancer treatments, gave outcomes over a spectrum of diseases in aggregate form acceptable for the Subcommittee's use.

Meanwhile, the Social Values Subcommittee was contracting with Oregon State

Figure 1. Example of Health Outcomes Information Collected Fran Expert Provider Groups

Condition: Stroke
Treatment: Medical Therapy
Duration of Benefit: 7 yrs.

	Without Treatment	With Treatment
Death	30%	10%
Residual Effect 1	50% MH,PB,SN,#5	40% MT,PW,SL,#5
Residual Effect 2	20% SN,#14	15% SL,#14
Residual Effect 3		30% #6
Asymptomatic	0%	5%

MH: In hospital
MT: Limitations in using transportation
PB: In bed or in wheelchair under other's control
PW: In wheelchair under own control or other limitation in walking
SN: Needs help with self-care activities
SL: Limited in role activities
#5: Missing, deformed, paralyzed or broken limbs
#6: Pain, stiffness, weakness, or numbness in chest, back, neck or limbs
#14: Trouble talking

University to conduct the telephone survey. Stratified sampling was used to ensure adequate representation from the different geographic regions of the State. Each respondent was asked to give a score from 0 (death) to 100 (perfect health) for each symptom and health limitation, which were then averaged and transformed to a scale of 0 to -1 so that each negative weight could be added using the QWB scale to take away from a perfect health score of one.

The results showed the highest weight of -0.560 given for being confined to a bed or a wheelchair under someone else's control and -0.253 for the symptom group including deformed or paralyzed extremities, The lowest symptom weights of -0.123 and -0.055 were given to having to take regular medications and having to wear eyeglasses, respectively.

THE MAY 2, 1990 DATA DUMP

The Health Services Commission was given an extra two months to complete their work and on May 2, 1990 a meeting was scheduled to hopeful 1 y finish the list. It was hoped that the cost-utility formula shown in Figure 2 would produce a reasonable list and then minor hand adjustments could be made to finalize it. As it turned out, cost data turned out to

Figure 2. Cost-Utility Ratio Fonda

$$R_n = \frac{c}{Y * \left[\sum_{i=1}^{5}(p_{i1} * QWB_{i1}) - \sum_{i=1}^{5}(p_{i2} * QWB_{i2}) \right]}$$

[With Treatment] [Without Treatment]

$$\text{with} \quad QWB_{ik} = I + \sum_{j=1}^{30} d_{ijk} w_j$$

where

R_n = the cost-utility ratio for the nth condition/treatment pair to be ranked. This value will be used in determining the actual rankings of health services from highest(O) to lowest(+∞).

c = cost with treatment, including all medications and ancillary services as well as the cost of the primary procedure.

Y = the years for which the treatment can be expected to benefit the patient with this condition. This may be the remain-&r of the patient's lifetime or some shorter amount of time.

p_{i1} = the probability that the ith outcome will occur five years hence with treatment.

d_{ijk} = an indicator variable denoting the presence (=1) or absence (=0) of the jth health limitation (mobility, physical or social activity) or chief complaint for the ith outcome either with treatment (for k=1) or without treatment (for k=2).

w_j = the weight given by Oregonian's to the jth health limitation or chief complaint ranging from 0 = no significant effect to -1 = death.

p_{i2} = the probability that the ith outcome will occur five years hence without treatment.

be the weakest link of the data chain. Many attempts were made to combine piecemeal billings in the State's computers and insurance company data to arrive at global cost of treatment. As the clock wound down towards the meeting, a last ditch effort at calculating cost was made just the night before. As in the case of the line for stroke, all lines that had medical therapy as a treatment were assigned a cost of $98.51. The resulting data dump for the 1743 condition/treatment pairs placed items in a counter-intuitive order. Anomalies such as appliances to correct malocclusion due to thumb sucking being placed above appendectomies were readily pointed out. The Commission found they could not easily come up with an acceptable product through hand adjustment.

THE WAY 1, 1991 PRIORITIZED LIST

The Commission was granted a one year extension to produce their first prioritized list. Following the disappointing results of the spring of 1990, a Subcommittee on Alternative Methodologies was formed to investigate other ways to quantifiably order the list, possibly incorporating the public values to some degree.

The Subcommittee determined that one of the major deficiencies of the cost-utility formula, even if reasonable data were used for costs, was that life saving treatments were generally not given enough weight, and many cosmetic treatments were given favorable ratios. This suggested a two-tiered system of ranking. Larger categories of care could be used for the first pass, taking into account broader issues such as whether or not the condition is life threatening, acute vs. chronic, or simply has no effective treatment at this time.

The categories of health care developed by the Oregon Priority Setting Project were modified by the Coxnnission to incorporate effectiveness of treatment and the possibility of fatality due to the condition. It was further decided that only the net-benefit portion of the cost-utility formula would be used to rank condition/treatment pairs within categories. While the categories helped to incorporate some of the community values conveyed through community meetings and public testimony, it was still felt that the computer alone could-not produce a list without the need for hand adjustment.

A total of seventeen categories of health care were identified. Each of the categories themselves were ranked by the Commission on the basis of three attributes: 1) value to society, 2) value to an individual, and 3) essential to a basic benefit package. These three attributes incorporated the thirteen public values expressed at the community meetings (eg. prevention, cost-effectiveness, benefits many). Each category was given a score of l-10 by each commissioner to indicate the degree to which they felt the categories addressed the values inherent to the three attributes. Each commissioner was then asked to assign a weight to each attribute (totalling 100%) to represent their overall importance in the category ranking process. The attribute weights and category scores were multiplied and summed for each commissioner as shown in Figure 3. The weighted scores for each category were then averaged over the eleven commissioners to arrive at a final number in the range of 1 to 10 by which to rank the seventeen categories. After some minor hand adjustments, the categories were ranked as they appear in Figure 4. The Commission further classified the categories as being essential to a basic benefits package, very important and therefore

Figure 3. Example Calculation of a Commissioner's Category Score

Attribute	Attribute Weight	Category Score	Total
Value to society	0.4	2	0.8
Value to an individual	0.2	6	1.2
Essential to a basic package	0.4	1	0.4
One commissioner's weighted score			2.4

should be included as covered benefits to the extent possible, and important to certain individuals but significantly less likely to be cost-effective or produce substantial long-term gain. The total number of condition/treatment pairs were collapsed down to 709 and each of these were assigned to one of the 17 ranked health care categories and ordered by net-benefit within category. The example of medical therapy for stroke was assigned to category 3 with a net-benefit of 0.3358, placing it about half way within the items of that category. After hand adjustment, the first prioritized list was submitted to the Oregon legislature on Way 1, 1991 at which time they approved the funding of 587 of the 709 lines.

Figure 4. Rankings of the 17 Categories of Health Care

Essential
1. Acute fatal, treatment prevents death with full recovery (eg. appendicitis)
2. Maternity care
3. Acute fatal, treatment prevents death without full recovery (eg. severe head in jury)
4. Preventive care for children
5. Chronic fatal, treatment improves life span and quality of life (eg. diabetes mellitus)
6. Reproductive services
7. Cunfort Care
8. Preventive dental care
9. Proven effective preventive care for adults (eg. mammograms)

Very Important
10. Acute nonfatal, treatment causes return to previous health state (eg. acute thyroiditis)
11. Chronic nonfatal, one-time treatment improves quality of life (eg. hip replacement)
12. Acute nonfatal, treatment without return to previous health state (eg. cornea1 laceration)
13. Chronic nonfatal, repetitive treatment improves quality of life (eg. chronic sinusitis)

Important to Certain Individuals
14. Acute nonfatal, treatment expedites recovery of self-limiting conditions (eg. common cold)
15. Infertility services
16. Less effective preventive care for adults (eg. sigmoidoscopy for persons less than 40 years of age)
17. Fatal or nonfatal, treatment causes minimal or no improvement in quality of life (eg. viral warts)

THE OCTOBER 30, 1992 PRIORITIZED LIST

It was the May 1991 list that accompanied the initial application to the Department of Health and Human Services (DHHS), requesting that they grant waivers of Medicaid law so that the Oregon Health Plan could be implemented. After reviewing the application for nearly a year, Secretary of Health Dr. Louis Sullivan notified Governor Barbara Roberts that the Medicaid waiver application was being denied on the belief that the methodology used in the process of prioritizing health services could possibly be in violation of the newly enacted Americans with Disabilities Act (ADA).

Specifically, DHHS conveyed concern with the telephone survey and some aspects of the seventeen categories of health care. In regards to the telephone survey, it was felt that the ADA might be violated when asking nondisabled persons about the effects of conditions (eg. vision problems) that could be considered disabling in certain cases. There were further concerns about the distinctions made in the categories between conditions whose treatment would result in a return to a person's state of health prior to onset versus a condition whose treatment leaves a person with residual effects in a majority of the cases. Objections were also voiced against the use of references to quality of life in defining a number of the categories. One of the categories, infertility services, was felt to constitute a disabled population in itself.

While the Commission disagreed with the conclusions of DHHS, they felt that the goal of providing health care access to the uninsured outweighed the methodology used in the process. So the Commission purged the database of the telephone survey results and removed the use of the categories from the prioritization process.

The new methodology focused on effectiveness of treatment as the primary means of ranking. A four step process was first used by the computer to rank services based on 1) the treatment's ability to prevent death. If more than one line tied in this capacity then, 2) the lines are ordered next on the ability of the treatment to return the patient to an asymptomatic state of health after saving the life. If more than one line was still tied then 3) the lines were ordered on the treatment's ability to take the patient from a symptomatic to an asymptomatic

state of health, and 4) any remaining pairs still tied were ordered on the average cost of treatment, with the higher cost pair ranked lower. The computer then listed lines tied on all four steps alphabetically by diagnosis.

Figure 5 shows this process applied to the example of stroke/medical therapy. The treatment prevents death in 30 - 10 = 20 cases out of 100. Furthermore, when death is prevented, 20 - (85 -70) = 5 out of these cases were taken to an asymptomatic state of health. Since only positive changes were measured, 0 cases out of 100 were said to have changed from a symptomatic to an asymptomatic state since 70 - 85 = -15. This information, plus an average cost of $32,500 placed the treatment of stroke at line 246 out of 688 lines. This compares to a ranking of 252 out of 709 on the original May 1, 1991 list. The new list dated October 30, 1992 with 568 of 688 lines covered was then resubmitted to DHHS.

THE APRIL 19, 1993 PRIORITIZED LIST

On March 19, 1993, Secretary of Health Donna Shalala granted the State of Oregon the necessary Medicaid waivers. This approval, however, was conditional on the Commission revisiting their methodology for prioritizing health services. They were asked to produce a new list within 60 days that did not use the distinction between treatments that resulted in an asymptomatic health state and those where symptoms were still present after treatment. DHHS felt that such a distinction could be construed by the courts as being discriminatory against those individuals with disabilities whose condition could never reach an asymptomatic state.

Again the Commission put the goal of health care access to an additional 120,000 Oregonians first in modifying their process. This meant deleting steps 2 and 3 of the October 30, 1992 methodology, leaving a two step computer ranking process of 1) ordering based on the ability of the treatment to prevent death, and if more than one line tied on this factor then 2) order by cost of treatment, with lower cost treatments ranking higher. Lines were again ordered alphabetically by diagnosis if they tied on all counts.

Figure 5. Use of Health Outcomes Data in October 1992 methodology

Condition: Stroke Treatment: Medical Therapy
Cost: $32,500

	Without Treatment	With Treatment	Positive Change
Death	30%	10%	20
Symptomatic	70%	85%	5
Asymptomatic	0%	5%	0

Because of the limitations imposed by such a pared down ordering based on the objective data, the Commission put more work into structuring the incorporation of subjective information. They identified a list of twelve criteria and hand moved lines up or down the list based on their adherence to these values. The twelve values used were A) general preventive services, B) comfort care, C) maternity care, D) family planning services, E) prevent a condition before treatment, F) medical ineffectiveness, G) prevent additional complications, H) prevent future costs, I) cosmetic services, J) self-limiting conditions, K) congruent conditions and L) public health risk. The only one of these values that is not self-explanatory is "congruent conditions." This judgement considered the fact that the organ system and/or etiology of the conditions are similar to that of another condition/treatment pair or that the outcomes of the condition/treatment pairs are congruent. For instance, most diseases of the eye which have effective treatments were placed in the 360-390 group of lines. Medical therapy for stroke, with avoidance of death in 20 out of 100 cases and an average cost of $32,500, ranked 266 on the new list, requiring no hand adjustment.

The resulting list of 688 lines which satisfied the conditions of the approved waiver was presented to the Oregon legislature on April 19, 1993 and had a correlation of r = 0.985 with the October 1992 list. This did not concern DHHS since at no point did they necessarily disapprove of the product, it was always the methodology used to produce the list that was at issue. The 1993 legislature agreed to fund 565 of the 688 lines on the list, with implementation of the Medicaid expansion to begin February 1, 1994. A list integrating mental health and chemical dependency services will go in effect January 1, 1995. The 1993 legislature also postponed implementation of the employer mandate portion of the plan, with employers of 26 or more employees required to offer health benefits on March 31, 1997 and employers with fewer than 26 employees given until January 1, 1998 to meet this requirement.

FUTURE IMPLICATIONS FOR USING DATA TO HELP MAKE MEDICAL CARE POLICY

The trials and tribulations of the State of Oregon in obtaining their Medicaid waiver raises some question as to the future use of data in making health policy decisions. While it appears that the collection and analysis of health outcomes probabilities is not in jeopardy, the matching of these probabilities with symp-

toms in order to measure benefits **requires caution.**

If this data is being used to compare different treatments for the same condition it is fairly straight forward since the same symptom complexes are involved and their corresponding probabilities can be compared directly. There seems to be obvious problems in comparing probabilities associated with different **symptom** complexes across conditions. One way to avoid possible ADA violations may be to use signs in place of symptoms. This **change would remove the patient values from consideration, which would not be an acceptable alternative for a publicly debated prioritized list.**

The Health Services Commission is exploring the possibility of holding a conference at the end of 1993 to further discuss the issues so that they may incorporate these new ideas into future iterations of the prioritized list.

Using HCFA's National Claims History File (NCH) for Health Policy Development

Stephen T. Parente, MS, MPH
Jonathan P. Weiner, DrPH

From
The Johns Hopkins Health Services Research
and Development Center

This work was supported by the Health Care Financing Administration, Health Standards and Qualify Bureau

Objectives of Presentation

- To explore the application of insurance claims data to the measurement of the quality and cost of medical care.
- To share several methodologic innovations used to construct NCH-based analytic files for quality of care analyses
- To outline the potential uses of the NCH database for health policy development and evaluation.

The Features of the National Claims History File (NCH)

- Ability to link Part B and Part A claims by patient or provider.
- 100% of Part A and Part B claims.
- Near on-line access to all claims records.
- All claims available regardless of carrier service areas.

The NCH is Not One Big File - It's a Family of Files

- Beneficiary file
- Physician and facility file (from carriers)
- Institutional Claims file
  - For Part B Services (e.g., MRIs)
  - For Part A Services (e.g., admissions)
- Non-institutional Part B file (e.g., Physician office services)

Potential Applications of the NCH for Health Policy Development

- Assessing basic concordance with clinical guidelines.
- QA-oriented profiles focusing on provider- and population-specific medical care utilization and quality.
- A risk-adjustment algorithm for prospective, population-based, reimbursement of health plans.

Advantages of Claims Data for Health Policy Analysis

- Extremely inexpensive when compared to the cost of primary data collection.
- Can be applied more unobtrusively than chart data or surveys.
- Easy to change the denominator: region, provider, patient
- Capture an astounding array of data. What they lack in depth, they make up for in breadth.

Disadvantages of Claims data for Health Policy Analysis

- The data are intended for paying bills, not monitoring episodes of care.
- The accuracy of a data item (e.g., CPT4) is a function of whether it is needed for payment.
- The technical & methodologic challenges are numerous.

Making Raw Claims Data Usable: Creating an Analytic File

- Clean data (e.g., delete duplicated and unneeded data).
- Merge key variables from other sources.
- Develop logical records (e.g., episodes, annual summaries) from claims.
- Aggregate multiple records by summarizing key variables (e.g., procedure, diagnosis) by a unit of analysis (e.g., provider).

Technical Feasibility Issues

- Size of NCH Database:

  30 millions records for one year of Part B Medicare data for Maryland, Iowa and Alabama patients.
- Variable record lengths.
- Multiple provider IDs (e.g., 3 group practice IDs and 1 solo practice equals 4 IDs).
- Multiple beneficiary IDs.

Validity Issues

- Accurate representation of procedure, diagnosis, DOS, patient and provider.

DEMPAQ **FY91** Results:

- Agreement on presence/absence of 6 profiled diagnoses ranged from 74% to 96% (kappa 0.27 - 0.67).
- For 6 profiled diagnoses in claim but not record, 18 of 20 cases diagnosis in record confirmed.

Profiled Diagnoses: Hypertension. Diabetes, IHD, COPD. CHF, Osteoarthritis

DEMPAQ Claims-based Products: Physician Quality of Care Profiles

- Condition-specific profiles:
  - Diabetes-
  - Ischemic Heart Disease
  - Chronic Obstructive Pulmonary Disease
  - Congestive Heart Failure
  - Hypertension
  - Ischemic Heart Disease
- Office Practice Profile
- Preventive Care Profile

Raw Claims to -a-Diabetes Profile: The DEMPAQ Steps

1. **Identify beneficiaries with 1 or more 'face to face' encounters with a primary diagnosis of diabetes.**
2. **Select all claims for these beneficiaries.**
3. **Aggregate these claims for each patient. Develop count fields of relevant services.**
4. **Assign a 'primary care source' (e.g., the primary care physician who saw a patient the most).**
5. **Summarize count fields by primary care source.**

Useful DEMPAQ Innovations for Health Policy Development

- Primary Care Source Algorithm
- Application of Johns Hopkins Ambulatory Care Groups (ACGs) to Medicare Data.
- Resource Value Units
- Group Practice Identification

Primary Care Source Algorithm
Methods. Results and Policy Implications

- METHOD: Identify the primary care physician to whom a patient had the most visits.
- DEMPAO RESULTS: Almost 75% of all patients can be assigned a primary care source.
- POLICY IMPLICATIONS: Medicare's transition to a managed care/gatekeeper organizational model could be realistic for roughly three-quarters of the Medicare population.

Johns Hopkins Ambulatory Care Groups
Methods, Results and Policy Implications

- METHOD: Using diagnosis information a case-mix category for every patient was assigned.
- DEMPAO RESULTS: All patients were assigned an ACG. Up to 45% of service utilization can be explained in a linear regression model.
- POLICY IMPLICATIONS: ACGs can be used to develop a risk-adjustment mechanism to predict individual beneficiary resource use.

Developing a Resource Value Table
Methods, Results and Policy Implications

- **METHOD: Average procedure costs (by DRG and HCPC)** and RBRVS weights were used to construct a table of standardized reimbursement.
- DEMPAO RESULTS: **All procedure cost variations were a function of quantity only, not price.**
- **POLICY IMPLICATIONS: Cost projections from a procedure fee schedule without specific geographic or facility price variations can be analyzed.**

Identifying Group Practices
Methods, Results and Policy Implications

- METHOD: Using employer identification ID from carrier-based provider files, group practices were identified.
- DEMPAO RESULTS: Provider billing activity was summarized on group and solo practices.
- POLICY IMPLICATIONS: Group Practices may have significantly different practice styles. This data would confirm or refute this hypothesis on a procedure or disease specific level.

Some- Suggestions to Make Claims Data Better for Health Policy Analysis

- An edit-checked Unique Physician Identification Code (UPIN) on every claim.
- Use of a unique non-changing patient identifier (e.g., SSN).
- Develop a procedure-specific list of outcome indications that must be entered to pay a claim.

ASSESSMEETS OF OUTCOMES FOR CORONARY ARTERY BYPASS GRAFT SURGERY IN NEW YORE STATE

Harold Rilburn, Jr., New York State Department of Health
Edward L. Hannan

The New York State Department of Health (NYSDOH) established the Cardiac Advisory Committee (CAC) in 1955. The role of the Committee is to provide advice and assistance in establishing and maintaining the highest possible quality of cardiac care. Composed of cardiac surgeons and cardiologists from across the State, the Committee's ad hoc status was formalized in statute in 1974. The Department currently relies on the Committee in setting professional standards, conducting site reviews and in the operation of the State's certificate of need process.

An important function of the CAC is the collection and review of aggregate statistics of hospital performance. This effort traditionally involved aggregate level information on the number of cardiac surgeries performed and the number of in-hospital patient deaths. Statistics were collected by individual hospitals and forwarded to the Department for compilation and dissemination to cardiac centers.

The shift in health care evaluation from an emphasis on structure and process to an emphasis on outcomes increased the Department's and the Committee's concern over the sizeable variation in the crude mortality rate among cardiac hospitals. Existing data was perceived as inadequate for addressing the key question of the extent to which crude rates reflected quality rather than case mix differences. CAC data provided only aggregate level information. The Department's universal discharge abstract data system suffers from the usual shortcomings of administrative data bases. These include inadequate clinical detail and the lack of information on the sequence of reported diagnoses.

The response to these shortcomings was the development of a cardiac specialty data system. This system, the Cardiac Surgery Reporting System (CSRS), was to provide data on individual patients, be clinically based, offer concurrent review opportunities and be closely monitored to ensure accurate, complete, and timely submission. The goals of the CSRS were 1) to help hospitals in their internal quality improvement efforts, 2) to help assess the advisability of performing surgery for specific patients by quantifying risk, 3) to provide outcome information to consumers and 4) to aid in prioritizing NYSDOH site visits. The CSRS data system was implemented in January 1989.

The CSRS data contains patient demographic information, pre-operative clinical factors, peri-operative complications of surgery, discharge status of the patient, and hospital and surgeon identifiers. The set of risk factors included in the 1991 version of the data appear in Figure 1. The data are submitted by each of the 31 cardiac centers on a quarterly basis.

The data entry and submission process uses NYSDOH's developed data entry software containing a number of checks for the appropriate range of values and internal consistency of the data entered. Data are downloaded to a disk at the end of each quarter using a transfer utility contained in the software. The information is due at the NYSDOH two months following the end of each calendar quarter. Data are examined by the CSRS Technical Support Unit at the end of each year. Data quality problems are identified and the hospital involved is contacted to confirm or correct the suspected error. Hospitals are also asked to confirm the number of cases submitted and the number of deaths occurring for the year.

Figure 1. Potential Risk Factors for Coronary Artery Bypass Graft Surgery

Age
Gender
Race
Surgical Priority (Elective, Urgent. Emergency)*
Height*
Weight*
Body Surface Area*
Body Mass Index*
Ejection Fraction
NYHA Functional Class
CCS Functional Class
Number of Vessels Diseased/Degree of Left Main Disease*
Previous Open Heart Operations
Previous Myocardial Infarction (<6 Hrs., 6-23 Hrs., l-7 Days, 8-14 Days, 15-21 Days, More Than 21 Days)
Stroke*
Carotid/Cerebrovascular Disease*
Aortoiliac Disease*
Femoral/Popliteal Disease*
Hemodynamic Instability*
Shock
Hypertension
IV NTG Within 24 Hrs. Before Operation
ECG Evidence of Left Ventricular Hypertrophy*
Congestive Heart Failure
Major Acute Structural Defect
Persistent Ventricular Arrhythmia*
Extensively Calcified Ascending Aorta*
Chronic Obstructive Pulmonary Disease
Diabetes Requiring Medication
Hepatic Failure*
Renal Failure
Immunosuppression Therapy*
Immunoincompetent Disease*
IABP Pre-Operation
Cardiac Catheterization Crash
PTCA Crash
Previous PTCA, This Admission*
PTCA Before This Admission*
Thrombolytic Therapy Within 7 Days*
Smoking History in Past 2 Weeks*
Moderate or Severe Myocardial Ischemia on Stress Test*

* Added in 1991

The NYSWH uses the corrected data to create a provider-specific quality of care rating. A statistical model of in-hospital patient mortality is developed using multiple logistic regression analysis. The model is used to identify the relative importance of each risk factor in predicting death and to generate a provider-specific expected mortality rate. This expected rate measures the average severity of illness of patients treated. Expected results are compared to actual results and a risk-adjusted mortality rate is calculated for each provider. This risk-adjusted mortality rate is calculated as the ratio of the actual to expected death rate scaled to the statewide mortality rate.

The final 1991 statistical model appears in Table 1. Column 1 contains the significant risk factors, column 2 contains the odds ratios, calculated as e^{b_i} where b_i is the unstandardized logistic regression coefficient for variable i. The final column reports the significance level of each risk factor. The model is an adequate fit as indicated by a C value of .8 and a

Table 1. Multivariate risk factor equation for hospital deaths in New York State in 1991

Patient Risk Factors for Hospital Deaths	Loaistic Coefficient	P-value
Demographic		
Age	0.0323	<.0001
Female Gender	0.4096	.0002
Coronary Disease		
Left Main Stenosis >50%	0.2425	.0349
3-System Stenoses 270%	0.2633	.0207
Myocardial Ischemia		
CCS Class IV	0.2435	.0418
Ventricular Function		
Previous MI <6 hours	0.6104	.0279
Ejection Fraction <.20	0.8503	.0025
Ejection Fraction .20-.39	0.4412	.0001
Secondary Conditions		
Congestive Heart Failure	0.3561	.0044
IABP Pre-op	0.4289	.0208
ECG Evidence, Left Ventricular Hypertrophy	0.3581	.0039
Persistent Ventricular Arrhythmia	0.4526	.0246
Severity of Atherosclerotic Process		
Coexisting Carotid Disease	0.4936	.0002
Aortoiliac Disease	0.3228	.0350
Calcified Aorta	0.6201	.0002
Secondary Conditions		
Hemodynamic Instability	1.0152	<.0001
Shock	1.9400	<.0001
Renal Failure	1.0564	.0001
Hepatic Failure	1.6951	.0261
Nonelective Surgery	0.3246	.0147
Previous Open Heart Operation (>1)	1.2324	<.0001

Hosmer-Lemeshow statistic of 3.03 on 8 degrees of freedom.

Table 2 shows the types of provider information made available to cardiac hospitals and the public at large. For each calendar year, this information includes data by hospital and by surgeon on the number of cases, number of deaths, actual mortality rate, expected mortality rate, the risk-adjusted mortality rate (RAMR) and a confidence interval for the RAMR. Providers for which the confidence interval does not contain the annual statewide rate has a statistically significant difference between the expected and actual mortality rate.

The quality assurance goals of the CSRS data system are met through information releases to both the hospitals and the public. The data submitted to the CSRS are analyzed annually. The analysis results including the provider rankings are submitted to the CAC for review and comment. Specific providers may be targeted for further review by the Committee. Once reviewed by the CAC, hospitals are provided with the ranking. This allows providers to compare their own program to others in terms of volume, case mix and in-hospital mortality with case mix differences controlled.

The NYSDOH supports the hospitals' internal quality assurance efforts by providing computer software that can be used by the hospitals to monitor outcomes. The software is integrated into the data entry program developed by the Department and provides two basic types of information. One program uses the most recent statistical model of patient death to calculate the predicted probability of death prior to surgery. A patient's risk factors are entered by the clinician and the predicted probability of in-hospital mortality rate is returned. This information is intended to complement rather than substitute for other information used in the care decision-making process. A second software product allows a hospital to examine the performance of cardiac surgeons operating at their hospital. The information appearing in Table 2 is produced for individual surgeons to reflect any time period of the hospital's choosing. Planned for the 1993 release of the data entry software is the capability to produce descriptive statistics for the patients treated. Frequency distributions for individual risk factors can be generated along with the mortality rate associated with the level of each factor. The software will also display this same information for the entire state as reported for the most recent calendar year.

The second pillar of the Department's quality initiative is the public release of provider-specific results for coronary artery bypass graft (CABG) surgery. The first release of hospital identified information for CABG operations occurred in late 1990. Timed to coincide with the release of the December 5, 1990 issue of the Journal of the American Medical Association, the NYSDOH held a press conference to announce the availability of hospital performance statistics. Information of this type was provided to all cardiac centers in the autumn of 1989 and again in the spring of 1990. The information released publicly in December 1990 was provided to hospitals one week in advance. A copy of the complete information was sent to both the Director of Cardiac Surgery and an appropriate executive officer at each hospital.

The release of surgeon-specific information was accepted in principle by the Department but was not a planned part of the December 1990 data release. In the Department's view, surgeon-specific outcome rates were not sufficiently stable to warrant release of a single year's data. A New York city newspaper requested the surgeon data under The Freedom of Information Act in early 1991. The NYSDOH refused to provide the information arguing that the relative frequency of low volume surgeons during a single year could result in errors in the judgements of the performance of individual providers. The New York State Supreme Court rejected the Department's argument in deciding the 1991 lawsuit brought by the newspaper and ordered the State to release the surgeon-specific data. The newspaper published surgeon-specific outcomes for 1989-1990 in December of 1991.

Table 2. Hospital Cases, Deaths, Actual, Predicted, and Risk-Adjusted Mortality Rates for Isolated Coronary Artery Bypass Graft Surgery, 1991

Hospital	Cases	Deaths	Act. MR	Exp. MR	RAMR	Conf. Int. For RAMR
1	831	26	3.13	2.85	3.38	(2.20, 4.95)
2	466	13	2.79	2.14	4.01	(2.13, 6.85)
3	59	4	6.78	2.81	7.42	(2.00, 19.00)
4	169	4	2.37	3.05	2.39	(0.64, 6.11)
5	325	8	2.46	3.35	2.26	(0.98, 4.46)
6	1151	26	2.26	2.52	2.76	(1.80, 4.04)
7	148	5	3.38	1.93	5.39	(1.74, 12.59)
8	507	18	3.55	3.17	3.45	(2.04, 5.45)
9	369	13	3.52	2.70	4.02	(2.14, 6.88)
10	608	22	3.62	3.63	3.07	(1.92, 4.64)
11	496	15	3.02	2.44	3.82	(2.14, 6.30)
12	305	5	1.64	2.82	1.79	(0.58, 4.18)
13	196	0	0.00	2.29	o.oo**	(0.00, 2.51)
14	497	15	3.02	2.84	3.28	(1.83, 5.40)
15	831	23	2.77	3.77	2.26	(1.43, 3.39)
16	465	16	3.44	3.25	3.25	(1.86, 5.28)
17	707	31	4.38	5.38	2.51	(1.70, 3.56)
18	275	10	3.64	2.95	3.80	(1.82, 6.99)
19	959	22	2.29	3.40	2.08	(1.30, 3.15)
20	1261	31	2.46	3.10	2.44	(1.66, 3.47)
21	521	11	2.11	2.56	2.53	(1.26, 4.54)
22	734	24	3.27	2.44	4.12	(2.64, 6.13)
23	437	20	4.58	2.12	6.64*	(4.06, 10.26)
24	481	22	4.57	2.13	6.61*	(4.14, 10.01)
25	331	14	4.23	3.64	3.57	(1.95, 5.99)
26	216	12	5.56	2.37	7.21*	(3.72, 12.59)
27	277	9	3.25	5.27	1.90	(0.87, 3.60)
28	266	12	4.51	2.71	5.12	(2.64, 8.95)
29	605	18	2.98	2.66	3.44	(2.04, 5.44)
30	451	11	2.44	4.24	1.77	(0.88, 3.17)
Total	1494460	3.08				

* Risk-Adjusted Rate is Significantly Higher Than Statewide Rate ($P<0.05$)
** Risk-Adjusted Rate is Significantly Lower Than Statewide Rate ($P<0.05$)
RAMR = Risk-Adjusted Mortality Rate

In December of 1992, the most current provider results were released by the Department. The information was based on all CABG surgeries performed during the period from 1989-1991. The policy adopted calls for the release of individual information only for those surgeons performing at least 200 CABG operations over a three-year period. This policy is endorsed by the CAC and annual releases are planned for the future.

Table 3 presents annual statewide information regarding the volume, actual mortality rate, expected mortality rate, and risk-adjusted mortality rate for CABG surgery in New York State between 1989 and 1991.

The volume of CABG operations grew from 12,269 to 14,944 (22.0%) during the three year period. The crude mortality rate during the period declined from 3.52% to 3.08% while the average patient severity increased from 2.68% to 3.66%. The risk-adjusted mortality rate fell from 4.25% in 1989 to 2.72% in 1991. Since the RAMR measures performance having controlled for differences among patients, these results show a major improvement for CABG surgery in New York State during the three year period. While this change undoubtedly reflects improvements in both surgical technique and training, the Department believes that the implementation of the CSRS data reporting and analysis system has also played an important role in achieving this outcome.

Critics of the policy of publicly releasing provider-specific information often charge that access will be reduced among the most difficult patients. Analysis of these data suggest that this is not a systematic occurrence. The number of operations and the expected mortality rate increased between 1989 and 1991. Both are inconsistent with access problems. An increase in the number and average level of severity of individual risk factors coded may account for the increase in the expected mortality rate. Data verification efforts for this period, however, did not find any systematic data reporting errors.

The Department of Health intends to continue to release annually hospital- and surgeon-specific information. As the Department's experience grows, both in the time periods covered and in the number of patients included, the analyses will become more sophisticated and specific. It is expected that these additional information will serve to increase the value and utility of these releases.

Table 3. Actual, Expected, and Risk-Adjusted Mortality Rate for Coronary Artery Bypass Graft Surgery in New York State, 1989-1991.

	1989	1990	1991
Actual	3.52%	3.14%	3.08%
Expected	2.68%	3.26%	3.66%
Risk-Adjusted	4.25%	3.11%	2.72%

Session **Z**

Prevention Effectiveness

Toward the year

2000

Refining the measures

COST EFFECTIVENESS OF WORKSITE CHOLESTEROL SCREENING

Robin D. Gorsky, University of New Hampshire
Tim E. Byers, Centers for Disease Control and Prevention

One of the three priorities of the Centers for Disease Control and Prevention (CDC), as stated by the director Dr. William Roper, is to determine the effectiveness of the prevention interventions supported by the CDC. Prevention effectiveness at the CDC is defined as: identifying efficacious and effective strategies to reduce morbidity and mortality and to improve the quality of life: determining the consequences of those strategies; determining the economic impact of those strategies: determining optimal methods for implementing those strategies; and evaluating the impact of prevention programs.1

As part of this initiative, programs within the CDC have begun to look at the costs of their prevention interventions and are considering innovative ways to improve the health status of communities. The Division of Nutrition, in the National Center for Chronic Disease Prevention and Health Promotion, was one of the first to participate in this effort.

Serum cholesterol is a major predictor of coronary heart disease.2 As the value of serum cholesterol increases, the risk of heart disease increases, particularly for those individuals with values over 200 mg%.

Coronary heart disease (CHD) treatment is expensive. One study reports the following costs associated with CHD, updated to 1990 dollars.3

Sudden death	$448
Non-sudden death	$6,263
Angina pectoris	$19,293
Myocardial infarction	$24,162
Coronary insufficiency	$25,599

The same study calculates the expected lifetime benefits from coronary heart disease averted, for persons less than 60 years old who lower their serum cholesterol by 5%.3 These figures represent the present value of health care dollars saved due to cases of CHD averted in the population lowering their cholesterol by 5%. Since most individuals do not suffer from CHD, the average cost saved over many people is quite low. Costs are updated to 1990 dollars.

Initial mg%	Present value of dollars saved
260	$12 - 93
300	$90 - 120
340	$118 - 161

These lifetime costs can be compared to the costs of a cholesterol-lowering intervention. Costs of a prevention intervention include

Provider time	X salary, fringe
Supplies, materials	X unit cost
Lab tests	X unit cost
Facilities cost, including utilities	
Equipment cost, including maintenance	
Administration and support of intervention	
Travel costs per provider	

In addition, there are participant costs and costs to the employer. Participant costs include: time taken for travel, waiting, and service; expenses for travel; and other expenses, such as for child or elder care. Employer costs include the lost productivity due to participant time away from the work station.

One way of reducing the costs of prevention interventions is to provide the intervention at the worksite. To determine the costs associated with lowering cholesterol at the worksite, costs were collected as part of a demonstration project entitled "The Four State Worksite Cholesterol Screening and Intervention Trial", with the State departments of health in Colorado, Minnesota, Missouri, and Washington, in collaboration with the Division of Nutrition, National Center for Chronic Disease Prevention and Health Promotion, at the Centers for Disease Control and Prevention. This was a randomized trial of worksite cholesterol screening plus a usual or special intervention offered to those with cholesterol levels ≥ 200 mg%.

The usual intervention was defined as the blood drawing plus five minutes of basic information about cholesterol and CHD. The special intervention included, in addition, a two-hour session on ways to reduce ones CHD risk. All participants with elevated cholesterol were followed up at six and twelve months to obtain an additional test for serum cholesterol.

Forty worksites were randomized to the usual or special intervention. There were 2084 workers screened, and 853 workers with cholesterol levels above 200 mg% were identified.

Cost data were collected by survey at each worksite in the randomized trial. Costs included supplies per person and per worksite, mileage from health department to/from worksite, mailing and telephone costs, and provider time for preparation, travel, services at the site, and follow-up of participants.

Total costs per individual screened were

usual intervention	$22.69 per person
Special intervention	$50.25 per person

Effects on serum cholesterol, obtained from the baseline and 12 month rescreening were

Usual intervention	1.9% reduction
Special intervention	5.3% reduction

The intervention cost was, per percent reduction in serum cholesterol

Usual intervention	$11.94
Special intervention	$9.48

It appears that lifetime benefits from a group of individuals who lower their cholesterol only 5% are greater than the worksite intervention costs associated with this benefit. The behavioral dietary intervention at the worksite, following usual cholesterol screening, can have a meaningful impact on long-term cholesterol levels and thus on the risk of CHD. Preventing CHD through cholesterol interventions at the worksite appears to be an effective prevention strategy.

References

1. Teutsch, SM. A framework for assessing the effectiveness of disease and injury prevention. *MMWR* 41(RR-3):i-iv, 1-12; 1992.

2. Stason WB, Costs and benefits of risk factor reduction for coronary heart disease: insights from screening and treatment of serum cholesterol, *American* Heart Journal 119(3, part 2):718-24; 1990.

3. Oster G, Epstein AM. Primary prevention and coronary heart disease: the economic benefits of lowering serum cholesterol. American Journal of Public Health 76(6):647-56; 1986.

AN ECONOMIC EVALUATION OF PREVENTING OR DELAYING THE ONSET OF THE COMPLICATIONS OF DIABETES

Erik Dasbach
Centers for Disease Control and Prevention

(Paper was not presented)

COST OF FOODBORNE ILLNESS AND PREVENTION INTERVENTIONS

Tanya Roberts, Economic Research Service'

Economists are starting to measure the cost of microbial foodborne illness in the United States. The traditional method of assessing costs is the cost-of-illness method (Rice, MacKenzie and Associates: Hodgson and Meiners) although new, willingness-to-pay methods have been developed by environmental economists (Cropper and Freeman: Mitchell and Carson).

This paper uses the cost-of-illness method and annual incidence-based disease estimates for selected foodborne pathogens and reports cost estimates at $5-6 billion annually. Possible prevention options are discussed along with methods of improving existing foodborne disease data bases to sharpen both the estimated disease burden and the costs of intervention strategies.

Theoretically, cost estimates should be as comprehensive as possible and include all the human illness costs, industry costs, and public health sector costs. Costs in Table 1 are costs associated with identifiable illnesses. The categories, to be more comprehensive, should be expanded to include costs of self-protection by consumers and industry and preventive costs by government--actions reducing the incidence of foodborne disease.

In practice, cost estimation is more limited. This analysis includes only medical costs, productivity losses, and special education/residential care because of some chronic conditions. The basis of the disease estimates is "best estimates" of the actual number of foodborne disease cases each year and is not limited to the much smaller number of outbreaks reported to the Centers for Disease Control and Prevention (CDC).

The present value of lifetime medical costs for those becoming ill in 1992 is estimated using nation-wide data bases such as the published Medicare reimbursement rates and per capita expenditures on physicians' services from the Health Care Financing Administration, the National Center for Health Statistic's National Hospital Discharge Survey, the American Hospital Association's Hospital Statistics, or Blue Cross/Blue Shield charges. The average cost to community hospitals per patient (Stat. Abstract) is used to compute hospitalization costs. In 1990, the hospitalization cost was estimated at $687 per patient. Updated to 1992 dollars (using the change in the hospital room CPI), this amounts to $817.43 per patient. Fees for laboratory tests, supplies, medications, and physician visits while hospitalized are assumed to be similar to the costs of hospitalization. Procedures such as hemodialysis and kidney transplant operations are computed as additional costs.

Medical costs also include nonhospital, ambulatory physician visits. The societal cost of a physician visit is estimated by dividing per capita annual national expenditures on physician services, $484 in 1990 updated using the physician services CPI to $545 (Stat. Abstract), by the average number of annual visits to physicians in 1989, 5.4 (Stat. Abstract), to arrive at $101 per visit.

Those with mild illnesses who do not seek a physician's care may also attempt to self-medicate: however, we do not have data on such expenses.

Table 1. Social Costs of Foodborne Illness

Costs to Individuals	Industry Costs
Medical costs	Product recall
Income or productivity loss	Plant closings and cleanup
Pain and suffering	Product liability costs*
Leisure time cost	Reduced product demand
Child care costs	Insurance administration
Risk aversion costs	
Travel costs	**Public Health Surveillance Costs**
Self-protection costs	Disease surveillance costs
Home modifications	Costs of investigating outbreak
Vocational/physical rehab.	Costs of cleanup
Residential care	

*In adding up costs, care must be taken to assure that product liability costs to firms are not already counted in the estimated pain and suffering cost to individuals. However, the legal and court expenses incurred by all parties are social costs.

Medical costs for chronic conditions are estimated by determining the time spent receiving a medical service and applying an average cost per service to compute total costs. For example, dialysis costs for those ill with hemolytic uremic syndrome (HUS) as a consequence of E. coli O157:H7 were computed by using the average number of days on dialysis required for HUS patients (12 days) (Martin) and multiplying it by the daily rate of Medicare reimbursement for dialysis ($123) (Eggers). Medicare billing records provide a major source of information on medical costs of specific diseases.

Productivity losses because workers were ill and missed work is approximated by the Average Weekly Earnings for nonsupervisory production workers in private nonagricultural jobs published by the Bureau of Labor Statistics (BLS) in the U.S. Department of Labor plus estimated fringe benefits. We use pre-tax wages and fringe benefits because we are using society's perspective and want to capture the worker's full contribution to society. The number of hours spent visiting a physician or days spent at the hospital are added to days spent at home for recuperation from an illness to arrive at the amount of time to be valued. We assume that for every day spent in the hospital, twice as many days will be spent at home recuperating. Time spent by parents caring for children, as well as paid caretakers, is included in the productivity loss estimates. In 1992 BLS reported that average weekly earnings were $364.30 (Handbook of Labor Statistics).

Also added is an estimate of health, vacation, and retirement benefits at 38% (Stat. Abstract), resulting in a daily wage of $100.55 or an hourly wage of $12.57 for a 40 hour week. Assuming a labor force participation rate of 84% for a typical workforce aged 25-44 years (Stat. Abstract), the average daily loss of productivity is $100.55 x .84, or $84. By estimating the amount of time lost from regular activities and multiplying it by the corresponding rate of earnings, a value for the time lost is derived. We assume that the value of an hour of time is the same for everyone, regardless of gender or race.

For those ill in subsequent years, a present value of the reduced stream of earnings is calculated. For those who die, Landefeld and Seskin's human capital/willingness-to-paymethodisused. It combines elements of both methods to generate the present value of expected lifetime after-tax income andhousekeeping services at a 3% real rate of return, adjusted for an annual one percent increase in labor productivity and a risk aversion premium that increases the estimates by 60%. We averaged Landefeld and Seskin's estimates of these values across gender within each age group.

Annual case and death estimates for foodborne bacterial pathogens in Table 2 come from various CDC sources (Roberts, Marks, and Lin). Sources for parasitic disease estimates are primarily the National Hospital Discharge Survey (Steahr and Roberts) and CDC sources (Roberts and Murrell).

Table 2. Estimated Partial Annual Medical costs and Productivity Losses, selected Foodborne Pathogens, 1992

Pathogen*	cases	Deaths	Costs
	#	#	$ million
Bacteria:			
Salmonella	1,920,000	960-1,920	1,188-1,588
Campylobacter jejuni or coli	2,100,000	120-360	907-1,016
Escherichia coli O157:H7	7,668-20,448	146-389	229610
Listeria monocytogenes	1,526-1,581	378-433	209-233
Parasites:			
Toxoplasma gondii	2,090	42	2,628
Trichinella spiralis	131	0	0.8
Taenia saginata	894	0	0.2
Taenia solium	210	0	0.1
Total Annual Medical and Productivity Costs = $5 to 6 billion/year			

*Analysis assumes 100% of human illnesses are foodborne for Campylobacter, Escherichia coli, Trichinella and the Taenias and assumes 96% of Salmonella cases, 85% of Listeria cases, and 50% of Toxoplasma cases are foodborne.

Incidence vs. Prevalence. These cost estimates are based on incidence, rather than prevalence, because the goal is to develop public health cost estimates to compare to the costs of various prevention programs. Incidence estimates are the annual increase in cases and associated disease costs. In contrast, prevalence estimates are all persons with acute foodborne disease or ongoing chronic complications. For example, it not only includes the persons with kidney failure because of *E. coli* O157:H7 from food eaten this year, but also all previous cases with kidney failure from this cause.

Prevalence estimates are greater than incidence estimates because the estimates include everyone sick today from the illness (regardless of the year they became ill). Whereas the incidence estimates only include those who came down with foodborne illness this year and their expected lifetime consequences as a result of this acute illness.

Incidence estimates are also useful because they indicate that the health problem is a current problem and not just a holdover illness (for example, polio contracted in 1950) for which an effective intervention has been found.

Intervention today which prevents future cases will eliminate all the economic costs of prevented cases--this is the benefit of disease prevention.

Intervention today will not reduce costs for prior cases unless a "cure" is discovered. For example, we can prevent future cases of polio with vaccination but we cannot cure those whose limbs were withered in an earlier- polio episode. These cases of withered limbs caused by polio would be included in the prevalence estimates of disease burden, whereas the incidence estimate only includes cases contracted in the current year. Thus, incidence estimates are generally the most appropriate for evaluation of the effectiveness of prevention programs.

Prevention Program Benefits and costs. Public health protection programs can take a variety of forms which may immediately improve health or have long run impacts that improve health. Table 3 contains illustrative examples of control options for foodborne pathogens as well as ideas from other health programs that might be adapted to food safety.

Short run improvements:

1. Reduction in acute illness costs--Consumers informed about food handling and consumption risks are more likely to act to reduce their risks (i.e. eating foods well-cooked) than uninformed consumers. These actions will reduce their exposure to foodborne pathogens and reduce acute foodborne illnesses.
2. Reduction in chronic illness costs--Food handlinu labels informincf consumers to cook their hamburgers well-done, may change the behavior of some cooks, which will reduce the cases of chronic kidney

Table 3. Potential Foodborne Disease Prevention Programs and Possible Response.

Prevention program	Response
Inform consumers about risks	High-risk consumers may change their behavior to protect themselves (Magat and Viscusi)
Create screening programs for high-risk consumers	Pregnant women screened for *Toxoplasma sondii* would be identified and may change their behavior (Roberts and Frenkel)
Create labels for foods tested for pathogens & enforce label to assure that products are "safer"	High-risk and risk-averse consumers may increase purchases (andperhaps paypricepremium) for safer foods and reduce consumption of high-risk foods, (e.g., rare hamburgers) (Magat and Viscusi)
Encourage and establish new control procedures for food firms	Witheconomicincentives, firmsdevelopnewcontrol production/handling procedures/equipment so disease incidence rates fall: HACCP and microbial testing are two examples of control procedures that farm/ slaughter/processor/retailer has taken
Expand foodborne disease data bases and include data on animal disease rates, carcass contamination rates, pathogens found on processed foods and at retail as well as human hospitalizations and deaths by ICD-9 codes for foodborne pathogens	New knowledge will lead to improved understanding of the severity of human foodborne illness, its causes, and where pathogens enter the food chain--leading in the long run to new prevention options

failure due to E. coli O157:H7 associated with rare hamburgers.

Long run improvements:
Programs sometimes have long run impacts, intended or serendipitous, which alter foodborne disease rates. These impacts may occur through increased societal awareness of the magnitude of the foodborne disease problem, in general, or increased awareness of specific pathogens stimulating research *on* control options unique to that pathogen.

1. Reduction in acute illness costs--Labels identifying low-risk foods could encourage food firms to find safer methods of producing foods so that their foods can qualify for the "safer" label.

2. Reduction in chronic illness costs--Creation of better data will permit foodborne disease linkages to be established that have not previously been established, turning private and public attention to finding ways to reduce foodborne pathogens from being transmitted through various points in the food chain from the farm to the table.

When evaluating the effectiveness of a prevention program in reducing the incidence of human illness, both the short and long run consequences need to be examined for both acute and chronic illnesses. (Mossel and Archer have identified many chronic illnesses associated with a variety of bacterial and parasitic foodborne pathogens.) Once the expected effectiveness of the intervention has been determined, then the public health protection benefits in dollars are the expected percentage reduction in medical costs, productivity losses, etc.

Prevention Program Costs. On the other side of the benefit/cost ledger are the expected public and private program costs associated with the prevention intervention. Is the prevention program a one-time intervention that prevents any reoccurrence of the disease? Or are various levels of intervention ongoing? What are the costs of this ongoing intervention and in what years are the costs incurred? Who bears the costs: industry, government, consumers? And how are the costs distributed within these groups?

Table 4. Data Needed for Estimating Microbial Health Risks

Data Need	Possible Solution	Likely Cost
Better recording of existing data	• Expand CDC's passive laboratory-based reporting for 9 foodborne microbial pathogens to the 30 with tests. (Expanding the use tests by state laboratories would increase costs more.) • Expand CDC's laboratory-based reporting to make surveillance active in selected counties ("sentinel counties").	Around $500,000/year Additional $1 million/year
Better demographic data on who becomes ill	Study various methods of making reporting active and estimate likely costs and benefits of increased identification of cases.	Around $75,000
Increased knowledge about risk factors	Study individual susceptibility as well as food handling und consumption practices for people becoming ill and control groups who do not become ill.	Around $500,000/pathogen
Increased knowledge of disease severities	Increase the number of pathogens investigated in depth with surveys of people becoming ill with specific foodborne pathogens.	Each survey may cost around $100,000.$200,000
Increased knowledge on deaths from foodborne sources	• Expand the "sentinel county" survey to Include all counties in the United States for a few selected pathogens to develop a solid baseline on foodborne deaths. • Study improving the identification of foodborne deaths in the National Death Index by improving entry of laboratory test data, by revising the death certificate to specifically ask for laboratory data, etc.	Around $1 million/pathogen for passive surveillance, much more for active surveillance Around $300,000
Increased knowledge about hospitalizations caused by foodborne pathogens	Study improving identification of foodborne pathogens in the National Hospital Discharge Survey (NIIDS) by I) examining a sample of the 150,000 hospitalizations under ICD-9 "008-unspecified intestinal infections" to see if medical records contain laboratory data, 2) examining septicemia/bacteremia records for 1 year to determine if foodborne pathogens are involved, 3) estimating the cost of revising the face sheet for entry of NIIDS data to include laboratory data and where sample taken (blood, stool, cerehrospinal fluid), and 4) comparing state systems.	Unknown, but likely to be less than $200,000/study
Quantification of chronic conditions caused by foodborne pathogens	Review the literature und estimate the likelihood thnt foodborne pathogens are important in certain chronic disease syndromes such as reactive arthritis, neurological disorders, cardiac dysfunction, food allergies, gastritis, kidney and liver dysfunction, etc.	Literature review around $100,000; original research more costly

Source: Roberts and Smallwood, Amer J. of Agricultural Economics, 1991.

Spending to improve estimates of foodborne illnesses and deaths appears to be cost-effective, considering the $5-6 billion cost for selected microbial foodborne pathogens in Table 2. Current inspection program costs are $700 million annually (GAO), significantly less than the estimated foodborne disease costs from microbial pathogens. While it may cost $8 to $24 million/year to build an adequate foodborne disease surveillance data base to identify new pathogens and estimate annual foodborne disease incidence rates, this is less than 5% of current inspection budgets. Roberts and Smallwood have identified several types of data collectionthatwould improve estimates of human foodborne illnesses (Table 4). Some new data collection efforts are underway: USDA's Food Safety and Inspection Service is starting to collect baseline data on the pathogens in slaughter animals: USDA's Animal and Plant Health Inspection Service is expanding its animal surveys to include human pathogens: CDC has initiated a pilot multi-center diarrhea1 diseases surveillance study; and Thomas Steahr, University of Connecticut, has started analysis of the National Hospital Discharge Survey by International Classification of Disease codes related to foodborne pathogens. What is needed now is integration and analysis of data from these various data bases to paint a more complete picture of the magnitude and severity of human foodborne illness.

Better identification of which pathogens are the most costly would pinpoint which pathogens should be targeted for benefit/cost analyses for control on farm, during processing, or during marketing. Better identification of who is at high-risk of foodborne disease would enable regulatory programs to be tailored to these high-risk groups. Better identification of high-risk foods would also inform consumers, some of whom may alter their food handling and food consumption practices.

1. The views expressed are those of the author and do not necessarily represent those of the Economic Research Service, United States Department of Agriculture. Review comments by Jordan Lin, Univ. of Florida, improved the quality of the paper.

References

Archer, D.L. 1984. Diarrhea1 Episodes and Diarrhea1 Disease: Acute Disease with Chronic Implications. Journal of Food Protection 47(4):321-328.

Bennett, J.V., S.D. Holmberq, M.F. Rogers, and S.L. Solomon. 1987. Infectious and Parasitic Diseases. In closing the Gap: The Burden of Unnecessary Illness, ed. R.W. Amler, and H.B. Dull, 102-114. New York, NY: oxford University Press.

Cropper, Maureen L. and A. Myrick Freeman III. 1991. Environmental Health Effects, ch. 6 in Measuring the Demand for Environmental Oualitv. John B. Braden and Charles D. Kolstad (Eds.), Amsterdam, The Netherlands: Elsevier Science publishers, pp. 165-211.

Eqqers, Paul. 1990. Medicare End Staae Renal Disease proaram Expenditures per Person bv Treatment Group, 1990. Health Care Finance Administration, Office of Research and Demonstration, Tables 48 and 49.

Fisher, A., L.G. Chestnut and D.H. Violette. 1989. The Value of Reducing Risks of Death: A Note on New Evidence. Journal of Policv Analysis and Management 8(1):88-100.

General Accounting Office. 1991. Federal Food Safety and Quality Programs. Washington, DC: United States Government Printing Office.

Griffin, Patricia and Robert Tauxe, 1991. The Epidemiology of Infections Caused by Escherichia coli O157:H7. Other Enterohemorrhaqic E. coli, and the Associated Hemolytic Uremic Syndrome, Epidemiologic Reviews 13:60-98.

Hodqson, Thomas A. and Mark R. Meiners. Cost-of-Illness Methodology: A Guide to Current Practices and Procedures. Health and Society 60(3):429-462.

Landefeld, J.S., and E.P. Seskin. 1982. The Economic Value of Life: Linking Theory to Practice. American Journal of Public Health-72(6):555-566.

Maqat, Wesley A. and Viscusi, Kip W. 1992. Informational Approaches to Regulation. Cambridge, Massachusetts, UK: The MIT Press.

Martin, D.L., K.L MacDonald, K.E. White, J.T. Soler and M.T. Osterholm. The Epidemiology and Clinical Aspects of the Hemolvtic Uremic Syndrome in Minnesota. New Enaland J. of Medicine 323(17):1161-1167.

Mitchell, R.C. and R.T. Carson. 1989. Using Surveys to Value Public Goods: The Contingent Valuation Method. Washington, D.C.: Resources for the Future.

Mossel, D.A.A. 1988. Impact of Foodborne Pathogens on Today's World, and Prospects for Management. Animal and Human Health 1(1):13-23.

National Oceanic and Atmospheric Administration, Dept of Commerce. 1993. Natural Resource Damage Assessments Under the Oil Pollution Act of 1990, Federal Resister 58(10-January 15):4601-4614.

Rice, D.P., E.J. MacKenzie, and Associates. 1989. Cost of Injury in the United States: A Report to Conaress. San Francisco, CA: Institute for Health and Aging, University of California and Injury Prevention Center. The Johns Hopkins University.

Roberts, T. and-J.K. Frenkel. 1990. "Estimating Income Losses and Other Preventable Costs Caused by Congenital Toxoplasmosis in People in the United States. J. of Amer. Vet. Med. Assoc. 196:249-256.

Roberts, T., and R. Pinner. 1990. Economic Impact of Disease Caused by Listeria monocytogenes, Chap. 22 in Foodborne Listeriosis. ed. A.J. Miller, J.L. Smith, and G.A. Somkuti. Amsterdam, The Netherlands: Elsevier Science, 137-149.

Roberts, T., and K.D. Murrell. 1993. "Economic Losses Caused by Foodborne Parasitic Diseases," in Loharanu, Paisan (Ed.) Cost-Benefit Aspects of Food Irradiation Processing, IAEA/WHO/FAO, forthcoming 1994.

Roberts, T., and D. Smallwood. 1991. Data Needs to Address Economic Issues in Food Safety. Amer. J. of Agricultural Economics 73:934-942.

Steahr, Thomas E. and Tanya Roberts. 1993. Microbial Foodborne Disease: Hospitalizations, Medical costs, and Potential Demand for Safer Food. Workino Paper; University of Connecticut.

U.S. Bureau-of the Census, Statistical Abstract of the United States. 1992 (112 edition), Washington, D.C. Tables No. 138, 162, 170, and 738.

Session AA

Barriers to Care

Toward the year

2000

Refining the measures

USING HOSPITAL RECORDS TO EVALUATE THE PERFORMANCE OF THE OUTPATIENT DELIVERY SYSTEM

John Billings
United Hospital Fund

(Paper was not presented)

SOCIOCULTURAL BARRIERS TO APPROPRIATE CARE'

Sarita L. Karon, University of Wisconsin – Madison
John A. Capitman

Introduction

Significant differences in use of lona-term care services exist by race/ethnicity and gender. Although there is much that we do know, there is even more that we do not know. This paper begins with a brief review of what is and is not known about racial/ethnic and gender differences in use of long-term care. After highlighting some of the reasons for our limited knowledge, it describes recent studies by the authors that address some of these limitations, and concludes with recommendations for future study.

A thorough review of the literature by Capitman et al. (1992) concluded that African Americans are less likely to use nursing home services than are whites. Women are more likely to use nursing home services than are men, but this finding is less consistent and less strong than the finding about race. Some studies suggest differences in use of home care services also exists, but it is more difficult to draw conclusions about these types of services. This is largely due to the differences in dependent variable definition. Home care encompass a broad range of services (e.g., nursing, home health aide, physical therapy), and also varies greatly by source of payment (e.g., Medicare, private pay, Medicaid). In general, review of the literature suggested the presence of an interaction by race and payment source. African Americans appear to receive more of the short-term, post-acute kinds of services paid for by Medicare than do whites; but for all other payment types, whites appear to be in an advantaged position. Women are typically found to use more of all types of home care services than do men, but these differences are attenuated when the availability of informal care services are taken into account.

In addition to differences in use of long-term care services, these populations differ in other significant ways. The "life chances", i.e., morbidity and mortality experiences, of racial groups differ greatly. Whites experience lower rates of disability at all ages up until age 75 or 80, at which point the sample sizes become so small as to make meaningful comparison difficult (Karon et al., 1992). Groups also differ in their experience of specific disease conditions. African Americans, Latinos, and Native Americans are all known to have greater rates of heart disease, stroke, and diabetes than do whites.

These groups also differ in the economic and personal resources available to cope with these issues. Each of these groups, as well as Pacific/Asians are more likely to be living in poverty or near poverty than are whites. They also are less likely to have enrolled in Medicare Part B or to have purchased private long-term care insurance. By contrast, whites are less likely to be Medicaid enrolled even after controlling for income (Davis et al., 1987). This does not mean that elders of color are in general more likely to have Medicaid coverage; only about one-half of poor elders where Medicaid enrollees in 1985 and so it is likely that a far greater share of elders of color are uninsured for long-term care.

In terms of personal resources, there are differences among racial/ethnic groups in the availability and type of family assistance that is available, and in family attitudes about caregiving. These attitudes may be changing in some communities. For instance, there is growing concern about deterioration of traditional family support structures among Latino elders (Maldonado, 1989; Sotomayor and Randolph, 1988). Little is known about how these differences in informal care availability and attitudes affect service use. It may be that strong cultural values on selflessness and familial responsibility can act as barriers to service use (Weibel-Orlando, 1989; Ewald, 1987).

Despite these findings, there remains much that we do not know. Our knowledge is limited in several ways. First, most of what we know is limited to differences between whites and African Americans. We know very little about the service use experiences of other racial groups (Pacific Islanders/Asians, Native Americans, Latinos), and even less about the experiences of ethnic subgroups (e.g., Koreans, Japanese as opposed to Pacific Islanders/ Asians). Second, we do not know why the differences described exist. For instance, to what extent do the observed differences result from personal preferences for services of particular types, or differences in resources available to access services, or differences in provider behavior toward members of different cultural subgroups? Third, we do not know whether or under what circumstances the differences are appropriate. Nor can we even agree on what is meant by appropriate care. Is care that meets a groups desire not to seek care from particular sources appropriate, if that desire is based upon years of discriminatory and inadequate care? Finally; we do not know the consequences of the differences.

There are several reasons for the limited extent of our knowledge. The first reason underlies all others, and that is that we simply have not asked the necessary questions. We have been complacent in letting the experiences of the dominant group stand for the whole, without question. Second, when we have thought to ask the questions, we have had inadequate conceptual frameworks. We have not given enough thought to the ways that we ask the questions and understand the answers. We have allowed our analytic models, which are constrained by available data, to stand for our conceptual frameworks, which can be broader and represent our assumptions and perspectives. Third, we have been limited by inadequate data. Small sample sizes have limited our comparisons to whites and African Americans, and have not allowed examination of many cultural subgroups. Finally, we have suffered from poor methodological approaches. An example of such a problem is the practice of combining all cultural subgroups into a single sample for analysis. In cases where the sample sizes are widely disparate, the experience of the larger group will mask the experience of the smaller. When groups are of approximately equal size, the effects of variables that operate in different directions for the different groups may cancel each other out, so that know effect is observed when in fact dramatic effects exist within each group.

The studies described here have begun to overcome some of these barriers. Each has been rooted in a broader conceptual framework, which has allowed us to ask deeper questions of the available data. Within the limitations of the available data, we have attempted to address some of the methodological limitations of earlier studies. The following sections briefly describe these studies and their findings. The report concludes with recommendations for future research.

Careaiver Gender and Access to Long Term Care

The first study addressed the effect of caregiver gender on access to formal long-term care services in a managed care system. The

caregiving literature is rife with evidence of women's greater likelihood to act as caregiver, and of differences in the kinds of tasks that women and men perform. However, the literature is sorely lacking in explanations for the observed phenomena. Attempts at explanation only serve to further describe differences in observed behavior, without addressing the underlying causes of these differences. This study (Karon, 1991) proposed one possible source of the difference, i.e., gender-role socialization.

The conceptual framework (Figure 1) emphasized that these socialized norms affect the actions of individuals as well as social groups and provider institutions. The conceptual framework also highlighted the inter-relatedness of people, groups, and institutions. Each acts to constrain as well as to impel the actions of the others. For instance, care managers must act within the constraints of the practices of the provider organization; but they can influence the provider organization by using discretion in ambiguous situations.

This study utilized data from three of the national Social/HMO demonstration sites. The Social/HMOs provide a full range of medical services, as well as a broad range of long-term care services to enrolled individuals. These include community based services such as home care; home health aides: physical, occupational, and speech therapies: and home delivered meals. Limited nursing home care, beyond that covered by Medicare, is also provided. Services are provided in a capitated, integrated system that is coordinated by care managers, and is funded by a combination of Medicare and private pay premiums. Access to the long-term care benefits is a function of need-based eligibility criteria that are comparable to each state's nursing home pre-admission screening criteria. Need is determined on the basis of a standardized comprehensive assessment, administered by care managers.

Data about the health and functional status of enrolled individuals, and their available caregiving networks were drawn from the Comprehensive Assessment Forms that are the basis for the care managers' determination of client eligibility for services and for care planning. Data about access to services were drawn from the actual care plans for the month following eligibility. The sample included 485 elderly women who were eligible for services and who had identified a primary caregiver who was a husband, child, or child-in-law. Nearly equal numbers identified male (240) and female (245) primary caregivers.

Tobit analysis was used to model access to services. Tobit was chosen because of the large number of individuals who had no services authorized. This could have occurred either because the care managers determined that their needs were being adequately met in other ways, and so they were not given access to care; or because the care managers offered services that the care recipients chose not to accept (no access desired, or the offered services were not the ones desired).

Results from the initial model, which considered only main effects of the variables, found no significant effect of caregiver gender on access to services. This is the point at which many studies would stop. The conceptual model used, however, suggested that gender role socialization would be an issue. We therefore ran a model that considered the interactions of caregiver gender with other variables that would be expected to be gender-typed. This iteration of the model resulted in very large differences in access by caregiver gender, where those with male caregivers were to have access to services than were those with female caregivers.

In order to gain more detailed understanding of the effects of caregiver gender, we divided he sample on the basis of caregiver gender and ran independent models. This was statistically equivalent to running a model with all terms interacted with caregiver gender. Looking at the data in this way allowed a clearer examination of the process by which members of each group gained access to services. Several differences were particularly striking. First, those whose caregivers were male had greater access to services when they had need for more "intimate" activities of daily living (ADLs), i.e., the ADLs that require full or partial nudity on the part of the care recipient. These included bathing, dressing, and toileting. Assistance with these types of activities can be considered to be "female" in nature. Additionally, there may be taboos about who can see the care receiving woman naked. Interestingly, this finding did not differ by relationship. It was as likely to be true when the caregiver was a husband as when he was a son. Some care managers, interviewed for another part of the study, reported that it was not unusual to have a client who said "my husband's never seen me naked!"

Another striking finding was the effect of needing assistance with heavy household chores, such as home maintenance, lawn care and snow removal, and vacuuming. This variable had a significant effect on access to service for both those with female and male caregivers. However, the direction of the effect was opposite for the two groups, in directions that corresponded to gender role stereotypes. When the caregiver was male, the client had less access to formal services. When the caregiver was female, the individual client had greater access to services. These findings are in keeping with gender stereotypes that suggest men may be better able to perform these physically demanding tasks. The findings held even though the male caregivers included husbands, who might themselves be expected to have some physical limitations, whereas the female caregivers were all daughters and daughters-in law who might be presumed to be younger and in stronger physical condition.

Access to formal services among those with a female caregiver generally responded to medical needs, such as need for assistance with medications. The more social kinds of needs that are stereotyped as female activities did not influence access to services for this sample.

These findings illustrate three key points. First, is that use of a stronger conceptual model can encourage lines of inquiry that might otherwise be ignored. Second, is that use of methodologies that consider each group's experience in its own right and can provide greater understanding of the unique experiences faced by each. Comparison of the experiences of the different groups can have important implications for the design of more equitable patterns of care. Third, use of data from the Social/HMOs emphasized the strength of these gender-based inequities, by controlling for a multitude of other sources of difference. Each of the sample members, by virtue of her membership in the Social/HMO, had access to a range of services, i.e., supply of services was not a limiting factor. Further, each Social/HMO member faced identical eligibility criteria for access; there were no differences in the infrastructure that would act to differentially limit access. Social/HMO members also experienced a level playing field, in terms of economic resources and the knowledge needed to access services. The finding of such strong gender-related differences in this sample suggests that the differences may be even more dramatic in other populations.

Race/Ethnicity and Gender and Access to Home Based Long Term Care

The second set of studies was concerned with the effects of race/ethnicity and gender on use of home based long-term care services. These studies were rooted in a complex conceptual framework

(Figure 2) that emphasizes the interplay of sociohistorical factors, individual needs and resources, individual practitioners and provider organizations in determining use of various types of services, defined by type and payment source. Race/ethnicity and gender are conceptualized as independent factors that are shaped by the historical, social, economic, and cultural milieu, and that in turn influence the actions of individuals and provider systems. An example of such a historical factor is the Tuskegee experiments, in which, unknown to them, 400 African American men were given only placebos as part of an experiment to determine the consequences of untreated syphilis. Such historical events as these can result in long-standing mistrust of service providers by African Americans or other groups with their own histories of abuse or mistreatment. Individual level determinants of service use include the life chances, informal resources, and other resources (economic, knowledge, etc.) reviewed earlier. Systems of care include the state and local system infrastructures, practices of provider organizations, and individual practitioners' beliefs and behaviors. Individuals' attempts to obtain service are filtered through these systems of care, and result in use of services defined by type and payment source.

Within this framework, three sets of studies were conducted using two different data bases. The first study used data from the 1984 National Long Term Care Survey (NLTCS) and the linked Medicare bill files. The dependent variable was the number of days of Medicare home health care provided in the year following completion of the NLTCS. The sample was restricted to individuals who were known to be impaired, i.e., who had passed the screening criteria and were included in the detailed survey. The second and third study used data from the 1984 wave of the Longitudinal Study on Aging (LSOA). Again, only impaired individuals were included in the sample, where impairment was defined as needing assistance with one or more ADLs or instrumental ADLs (IADLs). These studies differed in the dependent variable used. In one case, the dependent variable was defined as receipt of paid assistance with ADLs or IADLs. In the second case, the dependent variable indicated receipt of home care, including nursing and home health aide services, without regard to payment source.

Comparison of the findings of these three studies can be used to examine the importance of the dependent variable definition. Comparison of the findings for the dependent variable Medicare home care use with the dependent variable home care use begins to provide understanding of the difference payment source can make. Comparison of the findings for home care use and paid assistance with ADLs and IADLs highlights the importance of considering the nature of the services being provided.

Although the variables included in the equations and the variable definitions differed somewhat between the LSOA-based studies and the NLTCS-based study, each included consideration of several key concepts: race/ethnicity, gender, functional status, availability of informal supports, and living arrangement. Key findings in each of these areas are summarized in total and for several subsample comparisons (Tables l-5).

Table 1 shows the comparison when the analyses were conducted for the full samples (all races, both genders). In all cases, greater levels of functional impairment were associated with greater likelihood of receiving formal services; however, the level of impairment required to achieve this increased access varied by the type of service. Having greater amounts of informal support decreased the likelihood of receiving paid assistance with ADLs and IADLs. For both home care variables, however, the likelihood of service was increased with greater informal supports. This may suggest an advocacy role for family members. Living alone increased the likelihood of receiving Medicare home health and paid help with ADLs/IADLs, but had no impact of the receipt of other home care services. African Americans were less likely to receive paid assistance with ADLs/IADLs and with home care services, but there were no significant differences in the receipt of Medicare home health services. This may indicate that the nearly universal coverage of Medicare works to reduce inequities of access. Finally, women were found to be more likely to receive paid assistance with ADLs/IADLs, even after controlling for living arrangement and level of informal support, perhaps reflecting the kinds of gender role issues described previously whereby husbands are unable or unwilling to perform the stereotypically female kinds of care associated with ADLs and IADLs.

Table 2 shows the same comparisons as in the previous table, when the sample includes only whites. The findings are almost identical to those based on the full sample. This is an example of a case where the experience of the larger group dominates the findings for the whole. Without examining the groups independently, one might conclude that whites and African Americans face the same process of gaining access to care, except that in some cases African Americans receive less of it. Table 3 shows the results of the analyses when the sample includes only African Americans. Examination of this table shows that the potential conclusion described in the preceding paragraph is false, i.e., African Americans do not appear to experience the same process of accessing care as do whites. As is true for whites, African Americans have greater access to all types of services when they have greater levels of ADL impairment. The level of impairment required before this effect is observed, however, is much greater for African Americans. The most dramatic difference is in the level of impairment at which access to home care services is increased. For whites, this occurs with 2 or more ADL impairments, but for African Americans there is no increase in the likelihood of service use until the individual is impaired in 5 or more ADLs. For African Americans, living arrangement does not affect access to any of the three types of services. Among African Americans, women are marginally less likely than men to receive access to Medicare home care; this is in contrast with a positive effect of being female among whites.

Table 4 shows the results when only women are included in the sample. In almost all cases, the results are identical to those for the full sample. This is likely because, like the full sample, most of the women are white. Additionally, a majority of both full samples from each of the data bases were women. Despite the broad similarity, there is one interesting difference between women and the full sample. Women who live with a spouse have an increased likelihood of receiving paid assistance with ADLs and IADLs. There was no such living arrangement effect for the full sample. As observed in other analyses, this may be a result of the husbands' gender role socialization which leaves them ill prepared to provide certain kinds of assistance, and perhaps more inclined to purchase services when possible.

Table 5 shows the results when the sample includes men only. The effects of functional impairment are the same for men as for the full sample. Some interesting differences appear in the informal supports and living arrangements effects. There is no impact of informal supports on the use of Medicare home health; and there is no impact of living arrangement on receipt of paid assistance with ADLs/IADLs. The latter finding is another likely result of gender role socialization. Men may be able to call on a variety of resources for assistance with ADLs and

IADLs, whether the person who helps is a household member or not.

Implications for Future Study

The results presented in Tables 1-5 begin to illustrate the benefits to be gained by considering the experiences of population subgroups in a comparative manner, a? well as the importance of clearly specifying the type of service use being considered. The studies that are described here offer several lessons for future study.

First, we need to develop more thorough conceptual frameworks. The use of stronger conceptual frameworks can push us to reconsider the validity of analytic models that result in nonsignificant findings, leading us to develop more detailed models. Additionally, stronger conceptual frameworks can lead us to deeper understandings of significant results, and encourage us to ask questions that carry us beyond the seemingly obvious.

Second, we need better data, both in terms of who is measured and what is measured. In terms of who is measured, we need larger samples of racial subgroups; and, within racial groups, we need larger samples of ethnic subgroups. We must be able to differentiate racial groups by their countries of origin, but also by their date of immigration to this country, and degree acculturation.

In terms of what is measured, we are sorely lacking information in many areas. We need information on provider orientation. For instance, do individual providers hold racial stereotypes that affect their desire to provide care or abilities to provide appropriate care? Are the providers aware of the different types of health care issues facing different groups? Do provider organization make information available in the range of languages spoken by the people in their service areas?

We need information on consumer attitudes and knowledge. This information will help us understand to what extent differences in access to services is a result of individual preferences for services, as well as individual awareness of the services available and the methods used to access them. Such information has clear ramifications for the approaches used to remedy inequities in access.

We need information on the supply of services. Clearly, individuals cannot access services that do not exist in their geographic areas. Some of this information exists in sources such as the Area Resource- File, and can be utilized by linking multiple databases together.

We need information on the infrastructure, as well as the supply, of services. What are the institutions, organizations, and policies that facility or hinder access to care, and how do they differ by locale? Do they differ in the kinds of populations that live in their target areas?

We need better information on the social norms and family roles that exist in different cultural groups, keeping in mind that these may be especially fluid among groups with more recent histories of immigration. We must be careful to make sure that our knowledge of norms and roles is based on careful empirical evidence, and not on supposition, stereotype, or prejudice.

Clearly, it may be difficult to find all of this information in a single source. As already noted, the possibility exists for linking databases. And, as was illustrated by the studies described earlier, there is much to be gained by comparing findings across studies and databases.

Better understanding of the sociocultural barriers to appropriate care is not merely a matter of improved sources of data. It is also a matter of improved methods of using the data that we have. This includes more carefully defining our measures of utilization. We need to be specific in the types of services being measured, recognizing that the type of service may be as much a function of payment source as it is of the kind of care provided.

We can also improve our understanding by use of methods that consider the issues more thoroughly. One important methodological step is to consider the role of interactions of race/ethnicity and gender, with each other and with other variables in the equations. Our research must reflect our understanding that different groups experience the world differently, and may be treated differently by others in the world.

some of our understanding of these differences has come from qualitative research. We must not forget the value of these different types of research, and must work' to integrate them. Similarly, we must recognize the knowledge to be gained by considering the work of many disciplines. For instance, much of the approach underlying the studies described in this paper has been fostered by feminist research and theory. Use of multiple perspectives can only enrich our understanding of the world.

Our country is now engaged in a struggle to find new or modified health care paradigms, that will allow access to care for all members of our society. As we struggle to decide how to pay for such care, we must also struggle to understand what access to care means. Removal of financial barriers will not, by itself, create an equitable system. It is only by continuing to learn about each other, the ways that our histories and experiences shape who we serve and to whom we go for service, that we will begin to recognize and, finally, remove the many socioculturalbarriers to care.

---

1. The authors are grateful to the American Association of Retired Persons and to Work/Family Directions, Inc. for partial funding of portions of these studies.

REFERENCES

Capitman, J., B. Wethers, and E. Sadowsky (1992). Use of formal long term care services: evidence for the roles of race/ethnicity, gender, and informal care. Prepared for the American Association of Retired Persons. Institute for Health Policy, Brandeis University.

Davis, K., L. Lillie-Blanton, B. Lyons, F. Mullan, N. Powe and D. Rowland (1987). Health care for Black Americans: The public sector role in currents of health policy: Impact on Black Americans. The Milbank Quarterlv. Supplement 1(65), 213-247.

Ewald, P.D. (1987). Chronic health care needs of native village elders in interior Alaska. Final fellowship report, Division of Human Services, University of New England, Biddeford, Maine, l-82.

Karon, S.L. and J.A. Capitman (1992). The effects of race/ethnicity and gender on long-term care use: A comparative analysis. Report to the American Association of Retired Persons. Institute for Health Policy, Brandeis University.

Karon, Sarita L. The difference it makes: Caregiver gender and access to managed long-term care services in the social/hmo. Doctoral dissertation, Brandeis University, 1991.

Maldonado, D. (1989). Testimony presented to Congress. Hispanic and Indian Elderly: America's failure to care. Hearing before the Select Committee on Aging, House of Representatives, 101 Congress, August 7, 1989, Albuquerque, New Mexico. Comm. Pub. No. 101-730, Part I., U.S. Government Printing Office, Washington, D.C.

Table 1
COMPARISON OF KEY VARIABLES ACROSS ANALYSES: FULL SAMPLE

Variable Concept	Medicare Home Health	Paid Help with ADL/IADL	Use of Home Care Services
Functional Status	+ at all levels	+ at ≥ 3 ADLs	+ at ≥ 2 ADLs
Informal Supports	+ with greater numbers of supports	– with unpaid help from relatives	+ with unpaid help from distant relatives and non-relatives
Living Alone	+	+	ns
African American	ns	–	–
Female	ns	+	ns

Table 2
COMPARISON OF KEY VARIABLES ACROSS ANALYSES: WHITES

Variable	Medicare Home Health	Paid Help with ADL/IADL	Use of Home Care Services
Functional Status	+ for all levels	+ for ≥ 3 ADLs	+ for > 2 ADLs
Informal supports	+ for greater number of supports	with unpaid help from immediate family	+ with unpaid help from distant relatives
Living Alone	+	+	ns
Female			ns

Table 3
COMPARISON OF KEY VARIABLES ACROSS ANALYSES: AFRICAN AMERICANS

Variable	Medicare Home Health	Paid Help with ADL/IADL	Use of Home Care Services
Functional Status	+ at ≥ 1 ADLs	+ at 3 or 4 ADLs, marginally significant at 5 or more ADLs	+ at ≥ 5 ADLs
Informal Supports	ns	with unpaid help from immediate family	+ with unpaid help from nonrelatives
Living Alone	ns	ns	ns
Female	– (p < .10)	ns	ns

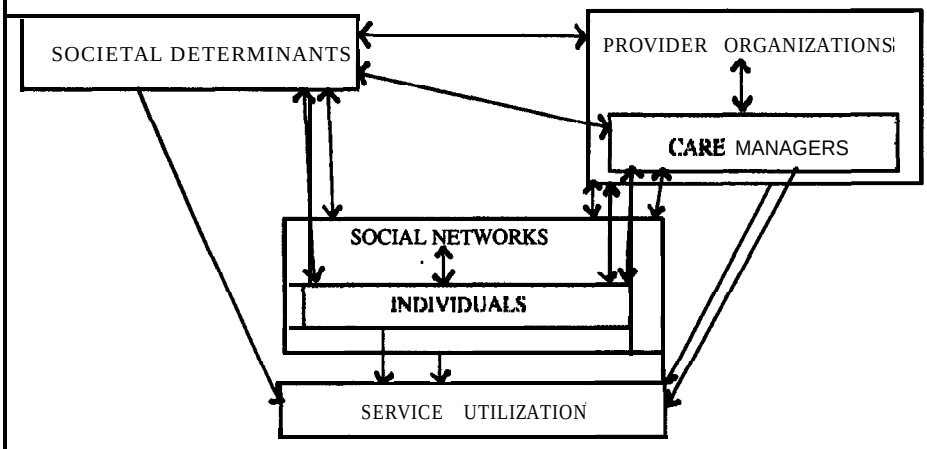
Table 4
COMPARISON OF KEY VARIABLES ACROSS ANALYSES: WOMEN

Variable	Medicare Home Health	Paid Help with ADL/IADL	Use of Home Care Services
Functional Status	+ at all levels	+ at ≥3 ADLs	+ at ≥ 3 ADLs (p<.08 at 2 ADLs)
Informal Supports	+ for greater numbers	– with unpaid help from immediate and distant family	+ with unpaid help from distant relatives and nonrelatives
Livina Alone	+	+	ns
African American	ns	(p < .08)	(p < .07)

Table 5
COMPARISON OF KEY VARIABLES ACROSS ANALYSES: MEN

Variable	Medicare Home Health	Paid Help with ADL/IADL	Use of Home Care Services
Functional Status	+ at all levels	+ at ≥ 3 ADLs	+ at ≥ 3 ADLs
Informal Supports	ns	unpaid help from immediate family	+ unpaid help from nonrelatives
Living Alone	+	ns	ns
African American	ns	+ for Hispanics ($p < .07$)	for African Americans ($p < .07$)

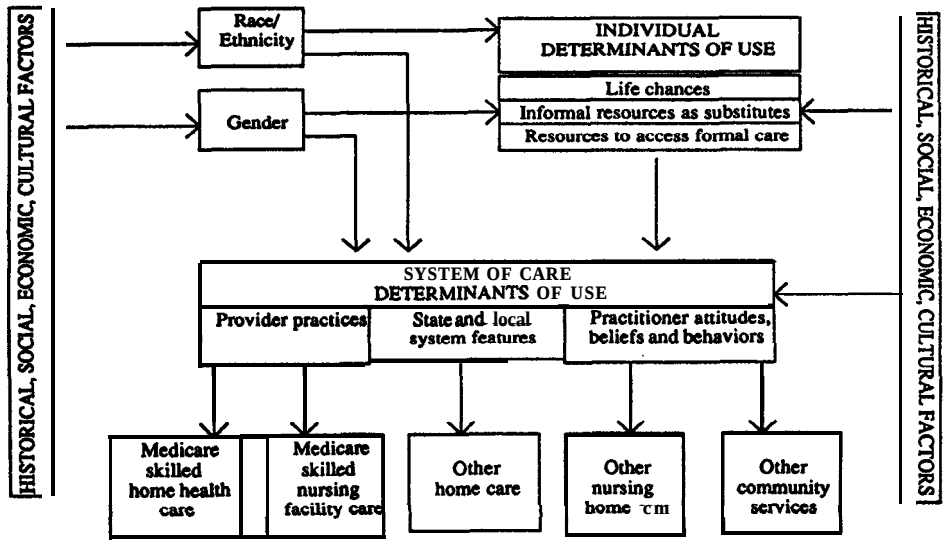
FIGURE 1



SOURCE: KARON 1991

FIGURE 2

A Framework for Understanding Race/Ethnicity and Gender Differences in Long-Term Care Use



Source: Capitman et al., 1992

EVALUATION OF BARRIERS TO PRENATAL CARE: A MODEL PROJECT

Mary E. Losch, University of Iowa
Herman A. Hein

I am here today to tell you about the Iowa barriers project: how it came about and how it runs on a day-to-day basis.

This project is designed to assess barriers experienced by Iowa women in obtaining prenatal care. It came into being almost two years ago as a result of Dr. Herman Hein's frustration at the continuing need to establish policy recommendations for reducing the infant mortality rate and having few data sources on which to base such recommendations. As a member of the national commission to prevent infant mortality and chairman of the sub-committee charged with assessing barriers to prenatal care, Dr. Hein began to seriously consider the basic components necessary to evaluate barriers to prenatal care at the state level. Primarily, he saw the need to 1) evaluate the accessibility of prenatal care to all Iowa mothers and 2) to provide a means of monitoring interventions that might be designed to reduce barriers to prenatal care.

Drawing on his considerable experience and professional contacts as the head of the State Perinatal Program, Dr. Hein brought together a number of individuals from the Iowa Department of Public Health, The University of Iowa Departments of Pediatrics, Obstetrics & Gynecology, and the Iowa Social Science Institute (which is my academic residence). The first order of business was to develop a survey instrument that could be used throughout the state. Drawing from several existing questionnaires on prenatal care including the 1988 national maternal health survey and adding a number of original questions, the instrument was pilot tested in three hospitals. These hospitals served differing populations of mothers both in terms of sociodemographic characteristics as well as the types of pregnancies (i.e., high-risk versus non-high risk).

In addition to variables affecting whether or not a pregnant woman sought prenatal care such as insurance coverage and transportation, additional items include satisfaction with care, open-ended queries about why it might have been difficult to get care, and any experiences with refusals to provide care. Sociodemographic questions about both the mother and father are also included.

Results of this first pilot test resulted in rewording of several questions and a second pilot test at The University of Iowa Hospital in order to determine whether or not the questionnaire could be filled out in a relatively short time and to make sure revised questions were indeed unambiguous and simple to answer.

Finding this to be the case, Dr. Hein then contacted The Robert Wood Johnson Foundation (RWJ) to determine their level of interest in possibly funding such a project. Following their positive disposition to the idea, Dr. Hein and I prepared a proposal to RWJ for a 3-year project to assess barriers to prenatal care experienced by Iowa women. As you might have guessed by now, the project was indeed funded and we are nearing the end of our second year.

To give you some idea of the costs involved, I'm going to show you a summary of the budget (See Table 1). At the same time, you can see the main personnel and equipment requirements. Please note that these figures are for direct costs only and do not include overhead or indirect costs. In addition, they do not reflect the many hours contributed by public health officials with whom we meet quarterly and hospital personnel who administer the questionnaire. Due in no small part to Dr. Hein's many years of good work with the state perinatal program, all of Iowa's maternity hospitals agreed to participate in the project and are generously providing their assistance free of charge.

Table 1
YEARLY AND OVERALL COSTS

	Year 1	Year 2	Year 3	Overall
Personnel*	44,000	31,300	39,200	120,500
Hardware				
Scanner	5,000	--	--	5,000
PC	3,200	--	--	3,200
Printer	400	--	--	400
Software				
Scanner	1,000	--	--	1,000
Word Proc.	200	--	--	200
Statistical	300	--	--	300
Printing	9,500	8,100	8,100	25,100
Postage	6,700	6,700	6,700	20,100
Comm. Costs	1,200	1,200	1,200	3,600
Travel Costs	3,000	3,000	3,000	9,000
TOTAL	$74,500	$56,300	$58,200	$189,000

*Includes Principal Investigator, Co-Principal Investigator, Project Coordinator, Project Consultant, Computer Consultant, Secretarial Support

Dr. Hein and I are budgeted at 5% each which probably does not reflect the actual time devoted to the project. Ten percent would probably be more accurate. A second member of the state perinatal program acts as liaison between the project coordinator and the hospitals. Ten percent has been more than adequate for this role. The project coordinator provides day-to-day oversight of the project; she is responsible for mailing questionnaires to the hospitals, tracking returns from hospitals, entering and analyzing data. This position was originally budgeted at 50% time and we have since expanded it to 60%. Even at that, expanded requests for specialized analysis from various public health officials and local agencies have resulted in the need to hire additional personnel to assist with these day-to-day responsibilities.

In the first year of the project, funding for

computer consultation was included to assist with the installation and programming of an NCS optical scanner. Because the design was a census rather than a sample of new mothers, we decided that the initial investment in a scanner would more than pay for itself when considering the data entry costs associated with tens of thousands of questionnaires that are returned to us each year. The only other equipment necessary are a microcomputer with large storage capacity and the power to manipulate large data sets and a printer.

When we began the project, Iowa's large hospitals were contacted first and provided with information about the study and instructions about the questionnaire's administration. Bather than prescribe a standard set of instructions on who should approach the mothers, staff on the maternal units were encouraged to make their own determinations about who should be responsible for the administration of the questionnaires. Some hospitals prefer to have the questionnaire administered at the same time the birth certificate information is provided, others separate the two.

After successfully bringing the large hospitals on board for the project, the many other smaller hospitals were contacted and all are now participating in the project. Returns are made monthly by the large hospitals and those hospitals with very few births per year make returns quarterly. Based on tracking sheets completed by hospital staff, the response rate for the project ranges from 50 to >95% with an average of approximately 80%. We are now waiting for the actual birth data for 1992 in order to calculate our exact response rate.

I would now like to show you a few examples of the types of data we have drawn from the barriers project and how they can be used. First I'll show you a city comparison (See Table 3). Using this approach you can also analyze data within a city, for example by zip codes (See Tables 4-5). And although these data are extremely valuable for assessing barriers and characteristics of these mothers and hence planning interventions, an additional value is the ability for this design to monitor such interventions and evaluate their efficacy.

Table 3
METHOD FOR FEEDING BABY BY CITY

	ALL	DM	DV	SC	WL	CR
Now						
Breast Milk Only	41	49	49	22	39	45
Mostly Breast Milk	12	10	91	8	8	9
Breast Milk = Formula	3	2	2	8	2	2
Mostly Formula	2	2		2		312
Formula Only	42	37	38	49	50	42
In 6 Weeks						
Breast Milk Only	24	28	31	17	22	25
Mostly Breast Milk	18	19	17	16	14	17
Breast Milk = Formula	8	8	7	10	7	7
Mostly Formula	6	6	5	8	6	6
Formula Only	44	39	40	49	51	45

All entries reflect percentages

Based on infant mortality data and these barriers data, it is clear that one metropolitan area in Iowa stands out as having a number of negative indices when it comes to prenatal care and perinatal outcomes. We are now beginning work on a demonstration project designed to positively impact the delivery of prenatal care in this area. When this new project is implemented, the barriers project will be used to evaluate the effectiveness of the intervention. In short, if the project is working, our numbers from this should look better.

Finally, in order to further enhance the value of these data, we have recently worked out a system whereby the questionnaire number from the barriers survey will be placed in the birth certificate database for each mother completing the barriers instrument. We will then be able to directly assess the possible links between access to prenatal care and perinatal outcomes.

In sum, the barriers project is indeed serving as an important data source for public health officials across the state of Iowa and we hope will continue to provide valuable information as a tool for monitoring the important reforms which will take place in the near future.

Table 4
PERCENT OF BIRTH CONTROL USED BY DESIRABILITY OF PG BY AREA

Desirability of Pregnancy	50314	50315	50313	50265
No, But in the Future	30	32	43	49
No, Nor in the Future	38	47	50	55

All entries reflect percentages

Table 5
DESIRABILITY OF PREGNANCY BY AREA

Desirability of Pregnancy	50314	50315	50313	50265
Yes, & Earlier	22	32	27	32
Yes, at That Time	17	31	32	45
No, But in the Future	52	33	35	20
No, Nor in the Future	9	4	6	3

All entries reflect percentages

BARRIERS TO CARE

Discussant:

P. Ellen Parsons
National Center for Health Statistics

(Not available for publication)

Session BB

CDC's Recommendations on Public Health Surveillance Systems

Toward the year

2000

Refining the measures

NATIONAL PUBLIC HEALTH SURVEILLANCE: WHERE DO WE GO FROM HERE?

Elaine D. Eaker
Center for Disease Control and Prevention

(Not available for publication)

THE FUTURE OF MICROCOMPUTERS IN EPIDEMIOLOGY

Andrew G. Dean
Centers for Disease Control and Prevention

(Not available for publication)

THE ELECTRONIC INFORMATION HIGHWAY: A PATHWAY FOR PUBLIC HEALTH INFORMATION

David A. Ross and Edward L. Baker, CDC

Among the most significant challenges in public health, is the challenge to improve the knowledge base for preventive services policy making through better and more timely information. Increasing and improving data access to build this knowledge base is clearly one of our major challenges. Reform of the health system will underscore the need for improved information systems.(l)

There are significant opportunities corresponding to these challenges. Primary among these opportunities is the potential to interconnect local, state and federal public health officials by using available advanced electronic network communications technologies and by connecting these organizations to national telecommunications networks.(2)

This paper addresses the impact these technologies are having on how we at CDC conceive our information mission and how these technologies will be used in the future.

Why speak about a network for public health officials at a session on surveillance? Simply put, the presence of large-scale inter networking taken in conjunction with expanding automation at the point of service deliver -- be it acute care, primary care, or preventive care -- is changing the landscape of possibilities for population assessments of public health status.

Health reform policy needs are already driving information collection. The need for data characterizing access to and demand for service, the cost effectiveness profile of the care rendered, and the outcomes resulting from that care are coming under intense scrutiny by all members of the health care system. Preventive health and population-based services will be forced to a higher standard of proof of effectiveness. Assessment, policy development and assurance as described in the Institute of Medicine's Future of Public Health (3) will be the cornerstone for public health practice. The assessment functions offer important opportunities for promoting public health leadership and will be enhanced through the fusion of computing and communications technologies.

We can no longer discuss assessment of disease trends, population disease phenomena, nor the effectiveness of prevention programs in the absence of a discussion of the technologies that enable the public health assessment function. Clearly, electronic communications between and among distributed sources of data for purposes of making population-based assessments is becoming more feasible and more essential to policy makers.(4)

What is CDC INPHO? The CDC Information Network for Public Health Officials (CDC INPHO) is an initiative whereby CDC:

1. is assisting state and local health departments in building their own electronic communications networks and in linking these local networks to the national information highways;

2. is providing the public health personnel a common access to CDC information resources; and

3. is setting the stage for a dynamic yet consistent interchange of information that can be relevant for local community health prevention programs, state prevention efforts, and for federal programs.

Therefore, CDC INPHO is an umbrella reference describing activities that will bring about electronic linkaae between and among public health partners; describing the implementation of CDC WONDER/PC as the vehicle for providing uniform access to CDC information resources; and describing the pathways upon which surveillance data and other information will be able to move. CDC INPHO is not a computer program that unifies or integrates CDC surveillance systems.

How will the aoals of CDC INPHO be implemented? The effort involves CDC working in partnership with state and local health leaders. At present CDC INPHO is guided through the collaboration of three CDC components: the Public Health Practice Program Office, the Information Resources Management Office and the Epidemiology Program Office.

To establish electronic linkage the CDC team will develop model or prototype state communications networks; link key public health organizations using the electronic mail facilities of CDC WONDER/PC; convene a national workshop to share information and assure that our activities support state and local health needs; provide information, technical standards, and technical consultation when requested; and assist public health agencies in developing information system capacity.

To offer uniform access to public health practitioners CDC will continue to enhance and expand the databases accessible through CDC WONDER/PC while also expanding its functionality. Coming releases of CDC WONDER/PC will include access to:

- the Prevention Guidelines Database, a comprehensive compilation of all CDC disease prevention and health promotion guidelines and recommendations,
- the Training Resources Directory, a database of all training and educational opportunities available through CDC's developing Public Health Training Network (CDC PHTN), and
- the MMWR made full-text searchable and including charts and graphs.

Finallv, CDC INPHO is collaborating with the Georgia Division of Public Health and the 19 District Health Officers to develop and test the utility of meeting local health needs for community health profile information usable for local programmatic decision making while also meeting state and federal disease surveillance needs through unified electronic capture, data transmission, results analysis and widespread distribution and notification.

The Georaia INPHO Project. This project is a partnership among public health, a private foundation, the Emory University -School of Public Health, the Medical College of Georgia, and the Georgia Center for Advanced Telecommunications Technologies located at Georgia Tech. This innovative 3-year project is producing a communications environment within Georgia that connects all levels of public health to one another and to CDC.

The project will position the Georgia Division of Public Health as national leaders in their use of advanced computing and communications technologies applied to solving public health emergencies, to bringing timely prevention information to community public health officials, and to empowering local community health planning efforts through information and automation.

CDC and The Woodruff Foundation served as catalysts to bring the academic centers and state and local governments together as partners for a common purpose. The partnership between GCATT technologists, Georgia State government departments (DHR and DOAS), and CDC has stimulated a vision for the state government's high-speed wide-area communications network. The collaboration among GDPH, Emory's School of Public Health, MCG, and CDC is providing a test bed for evaluating the impact of improved information and sophisticated analytical tools upon local health planning, service coordination, distance-based training and primary prevention efforts. None of these accomplishments would have been possible without the full collaboration of each of the partners.

Our objectives are three-fold. First, we will test the effectiveness and utility new information products have in assist public health practitioners and primary care physicians. Second, we will implement and evaluate the impact that a comprehensive electronic mail system has upon the interactions within the state public health leadership. And third, we will automate those health departments that presently lack the necessary hardware and software to participate in the electronic frontier.

CDC INPHO in Georgia has become an example for demonstrating GCATT's vision for Georgia as the world's leader in advanced telecommunications. CDC INPHO is building an organizational tie between MCG's Telemedicine project and the State Health Department that will serve to improve access to and coordination of preventive and primary care services in rural Georgia communities. Finally, the Georgia implementation of CDC INPHO is bringing the public health education capacities of Emory University to the worksite in each District Health Office thereby improving the competencies and capabilities of community public health department and their public health practitioners .

Among the many activities that comprise the Georgia project, the community surveillance database or community health profile will guide much of the system development in Georgia. The phenomenon of clinic and service-based data systems developed using client-server technology is beginning to drive how we engineer surveillance collection systems. In the distributed client-server world data reside locally, the information is contributed from each component organization, data systems will aggregate data hierarchically, and systems are being developed to measure and monitor prevention program effectiveness.

The challenge presented to us in Georgia is to combine federal surveillance needs with the State desire to monitor programs, to develop a singular surveillance data capture software that eliminates the multiple dissimilar CDC categorical disease surveillance systems and that interfaces with the clinic-based patient tracking system called HOST (for Health Outcomes and Services Tracking), and that is capable of producing a dynamic profile of the health patterns and problems confronting Georgia.

Discussions have just begun and a complete solution is probably several years away. It is most important to recognize that the communications capacity we are building in Georgia is changing the way in which local public health authorities envision meeting their prevention mission.

Most probably health reform legislation will transform public health. To be maximally effective public health leaders must shape public policy through timely data. We believe that CDC INPHO will open the door for rapid transmission and effective communication of the data that will be required to shape policy debates.

References:

1. Roper, William L., "Public Health in the New American Health System", Discussion **paper,** February 1993.

2. Scientific American, Special Issue, September 1991.
Gore, Al, "Infrastructure for the Global Village", pgs. 108-111.
Ray, Alan c., "Computers, Networks and Education", pgs. 100-107.
Tesler, Lawrence G., "Networked Computing in the 1990s", pgs. 54-61.

3. The Future of Public Health, Institute of Medicine, National Academy Press, Washington, D.C., 1988

4. "Results of the Survey on Data and Information Systems Issues for White House Task Force on Health Care Reform", 1993, pgs. l-7.

Appendix

Toward the year

2000

Refining the measures

PROGRAM SESSION ORGANIZERS

Jack R. Anderson** (Opening Ceremony)
Acting Deputy Director
Office of the Center Director

W. Edward Bacon, Ph.D.** (Y)
Director
Division of Health Care Statistics

Ronette R. Briefel, Dr.P.H., R.D.** (F)
Coordinator for Nutrition Monitoring and Related Research
Division of Health Examination Statistics

Lester R. Curtin, Ph.D.** (M)
Chief, Statistical Methods Staff
Office of Research and Methodology

Marcie Cynamon, M.A.** (K)
Special Assistant to the Director
Division of Health Interview Statistics

Pennifer Erickson** (A)
Chief, Clearinghouse on Health Indexes
Division of Epidemiology and Health Promotion

Trena M. Ezzati-Rice, M.S.** (T)
Biomedical Mathematical Statistician
Office of Research and Methodology

Lois Fingerhut, M.A.** (G)
Special Assistant for Injury Epidemiology
Office of Analysis and Epidemiology

J. Michael Fitzmaurice, Ph.D. (W)
Director, Office of Science and Data Development
Agency for Health Care Policy and Research
2101 East Jefferson Street, Room 601
Rockville, Maryland 20852-4908

Mary Anne Freedman, M.A.** (V)
Special Assistant to the Director
Healthy People 2000 Staff
Office of the Center Director

Patricia M. Golden, B.A., M.P.H.** (P)
Special Assistant to the Director
Division of Epidemiology and Health Promotion

Marjorie Greenberg, M.A.** (R)
Evaluation Officer, Office of Planning and Extramural Programs

Robert L. Heuser** (S)
Chief, Natality, Marriage, and Divorce Statistics Branch
Division of Vital Statistics

Peter L. Hurley** (First Plenary)
Associate Director, Office of Vital and Health Statistics Systems

Robert A. Israel** (N)
Associate Director
Office of International Statistics

Laura Kann, Ph.D.** (c)
Chief, Surveillance Research Section
Division of Adolescent and School Health
Center for Chronic Disease Prevention and Health Promotion
Centers for Disease Control and Prevention
1600 Clifton Road, NE., MS K-33
Atlanta, Georgia 30333

John L. Kiely, Ph.D.** (B)
Visiting Scientist, Division of Analysis
Office of Analysis and Epidemiology

Felicia LeClere, Ph.D.** (E)
Health Statistician
Division of Health Interview Statistics

Elizabeth Lowe (X)
(formerly with NCHS)
Foundation for Health Services Research
1350 Connecticut Avenue, NW., Suite 1100
Washington, D.C. 20036

Ronald W. Manderscheid, Ph.D. (D)
Acting Director, Division of State and Community Systems Development
Center for Mental Health Services
5600 Fishers Lane, Rockwall 2, Suite 501
Rockville, Maryland 20857

Mary Moien, M.S.** (0)
Assistant to the Director
Office of the Center Director

Gregory Pappas, M.D., Ph.D.** (Q)
Special Assistant to the Director
Office of Planning and Extramural Programs

Paul J. Placek, Ph.D.** (U)
Statistician, Office of Vital and Health Statistics Systems

Harry M. Rosenberg, Ph.D.** (J)
Chief, Mortality Statistics Branch
Division of Vital Statistics

Trish N. Royston (AA)
Deputy Director, Division of Information and Analysis
Health Resources and Services Administration
5600 Fishers Lane, Room 14-33
Rockville, Maryland 20857

Sandra S. Smith, M.P.H.** (I and Third Plenary)
Public Affairs Officer
Office of Data Processing and Services

( ) Indicates session

Donna F. Stroup, Ph.D., M.Sc. (BB)
Director, Division of Surveillance
and Epidemiology
Epidemiology Program Office
Centers for Disease Control and Prevention
1600 Clifton Road, NE., MS C-08
Atlanta, Georgia 30333

Steven M. Teutsch, M.D., M.P.H. (Z)
Director, Division of Surveillance and
Epidemiologic Studies
Epidemiology Program Office
Centers for Disease Control and Prevention
1600 Clifton Road, NE., MS C-08
Atlanta, Georgia 30333

Joan F. Van Nostrand, M.P.A.** (H)
Coordinator of Data on Aging
Office of Vital and Health Statistics
Systems

James A. Weed, Ph.D.** (L)
Deputy Director
Division of Vital Statistics

Ronald W. Wilson** (Second Plenary)
Director, Division of Epidemiology and
Health Promotion

**National Center for Health Statistics
6525 Belcrest Road
Hyattsville, Maryland 20782

PROGRAM SESSION CHAIRPERSONS

Ronette R. Briefel, Dr.P.H., R.D.** (F)
Coordinator for Nutrition Monitoring
and Related Research
Division of Health Examination Statistics

Trena M. Ezzati-Rice, M.S.** (T)
Biomedical Mathematical Statistician
Office of Research and Methodology

Jacob Feldman, Ph.D.** (A)
Associate Director
Office of Analysis and Epidemiology

Lois Fingerhut, M.A.** (G)
Special Assistant for Injury Epidemiology
Office of Analysis and Epidemiology

J. Michael Fitzmaurice, Ph.D. (W)
Director, Office of Science and Data
Development
Agency for Health Care Policy and Research
2101 East Jefferson Street, Room 601
Rockville, Maryland 20852-4908

Mary Anne Freedman, M.A.** (BB)
Special Assistant to the Director
Healthy People 2000 Staff
Office of the Center Director

Anne Haddix, Ph.D. (Z)
Economist
2537 Circlewood Road
Atlanta, Georgia 30345

Roderick Harrison (J)
Chief, Racial Statistics Branch
Population Division, Room 2318-3
U.S. Bureau of the Census
Suitland, Maryland 20233

Dorothy S. Harshbarger, M.S. (L)
State Registrar and Director
Alabama center for Health Statistics
P.O. Box 5625
Montgomery, Alabama 36103-5625

Gerry E. Hendershot, Ph.D.** (K-Substitute)
Chief, Illness Disability Statistics
Branch
Division of Health Interview Statistics

Edward L. Hunter** (X)
Chief, Planning, Budget, and Legislation
Staff
Office of Planning and Extramural Programs

Robert A. Israel** (N)
Associate Director
Office of International Statistics

Jared B. Jobe, Ph.D.** (M-Substitute)
Psychologist, Office of Research and
Methodology

John L. Kiely, Ph.D.** (B)
Visiting Scientist, Division of Analysis
Office of Analysis and Epidemilogy

Michael R. Lavoie, M.A. (S)
Chief, Health Statistics and
Measurement Center
Georgia Department of Human Resources
47 Trinity Avenue, SW., Room 217-H
Atlanta, Georgia 30334

Felicia LeClere, Ph.D.** (E)
Health Statistician
Division of Health Interview Statistics

Ronald W. Manderscheid, Ph.D. (D)
Acting Director, Division of State and
Community Systems Development
Center for Mental Health Services
5600 Fishers Lane, Rockwall 2, Suite 501
Rockville, Maryland 20857

J. Michael McGinnis, M.D. (First Plenary)
Deputy Assistant Secretary for Health
Office of Disease Prevention and Health
Promotion
Department of Health and Human Services
330 C Street, SW., Room 2132
Washington, D.C. 20201

Michael L. Millman, Ph.D. (AA)
Senior Staff Fellow, Office of Planning,
Evaluation, and Legislation
Health Resources and Services
Administration
5600 Fishers Lane, Room 14-33
Rockville, Maryland 20857

Mary Moien, M.S.** (0)
Assistant to the Director
Office of the Center Director

Paul J. Placek, Ph.D.** (U)
Statistician, Office of Vital and
Health Statistics Systems

James G. Ross, M.S. (C)
Vice President
Macro International, Inc.
8630 Fenton Street, Suite 406
Silver Spring, Maryland 20910

Lisa Miller Schalick** (Q)
Office of Planning and Extramural
Programs

Sandra S. Smith, M.P.H.** (I)
Public Affairs Officer
Office of Data Processing and Services

Elliot M. Stone, M.V.C. (Y)
Executive Director, Massachusetts
Health Data Consortium
400-l Totten Pond Road
Waltham, Massachusetts 02154

( ) Indicates session

Joan F. Van Nostrand, M.P.A.** (H)
Coordinator of Data on Aging
Office of Vital and Health
Statistics Systems

Signe Wetrogan (P)
Statistician
Population Division
U.S. Bureau of the Census
Washington, D.C. 20233

David R. Williams, Ph.D., M.P.H. (R)
Associate Research Scientist
Institute for Social Research
P.O. Box 1248, University of Michigan
Ann Arbor, Michigan 48106

P. Douglas Williams, M.S.** (V)
Chief, State and Academic Research
Applications
Office of Planning and Extramural
Programs

Ronald W. Wilson** (Second Plenary)
Director, Division of Epidemiology
and Health Promotion

Nicholas Zill, Ph.D. (Third Plenary)
Vice President
Westat, Inc.
1650 Research Boulevard
Rockville, Maryland 20850-3129

**National Center for Health Statistics
6525 Belcrest Road
Hyattsville, Maryland 20782

Lu Ann Aday, Ph.D. (Third Plenary)
Professor of Behavioral Sciences
School of Public Health
University of Texas
P.O. Box 20186, RAS East 311
Houston, Texas 77225

Mohammad N. Akhter, M.D., M.P.H.
(Second Plenary)
Commissioner of Health, D.C. Department
of Health and Human Services
1660 L Street, NW., Suite 1200
Washington, D.C. 20036

Barbara M. Altman, Ph.D. (K)
Research Fellow, Agency for Health
Care Policy and Research
2102 East Jefferson Street, Suite 500
Rockville, Maryland 20852

Hortensia Amaro, Ph.D. (R)
Associate Professor, Boston University
School of Public Health
85 East Newton Street, M-840
Boston, Massachusetts 02118

Laurie M. Anderson (T)
Epidemiologist, Cancer Prevention
Center for Chronic Disease Prevention
and Health Promotion
Centers for Disease Control and Prevention
1600 Clifton Road, K-55
Atlanta, Georgia 30333

Barry F. Anderson, Ph.D. (U)
Professor of Psychology
Portland State University
P.O. Box 751
Portland, Oregon 97207-0751

Jacqueline L. Angel, Ph.D. (E)
Sociology Department
University of Texas at Austin
336 Burdine Hall
Austin, Texas 78712

Andrew Asher, M.P.P. (0)
Research Analyst
Mathematica Policy Research, Inc.
600 Maryland Avenue, SW., Suite 550
Washington, D.C. 20024

Nancy D. Barker, M.S. (T)
Mathematical Statistician, Epidemiology
Program Office
Centers for Disease Control and Prevention
1600 Clifton Road, MS C-08
Atlanta, Georgia 30333

Trude Bennett, Dr.P.H., M.S.W. (L)
Assistant Professor, Department of
Maternal and Child Health
School of Public Health
University of North Carolina
CB 7400, Rosenau Hall
Chapel Hill, North Carolina 27599-7400

Karl E. Bergmann, M.D. (N)
Head, Division of Public Health
Statistics, Federal Health Office, D1
P.O. Box 330013, Werner-Voss-Damm 62
D-1000 Berlin 42, Germany

Neal Richard Boyd, Jr., Ed.D. (U-Presented
for Windsor)
Cancer Prevention Research Fellow
National Cancer Institute
Executive Plaza North, Room 320
Bethesda, Maryland 20892

Coleen A. Boyle, Ph.D. (K)
Division of Birth Defects and
Developmental Disabilities
Centers for Disease Control and Prevention
4770 Buford Highway, MS F-15
Atlanta, Georgia 30341

Bruce M. Brock, Ph.D., M.P.H. (V)
President
Information Transfer Systems, Inc.
209 East Washington Street, Suite 200
Ann Arbor, Michigan 48104-2007

Scott Campbell Brown, Ph.D. (K)
Education Research Specialist
National Institute for Disability and
Rehabilitation Research
400 Maryland Avenue, SW., Room 3520
Washington, D.C. 20202-2641

Jay S. Buechner, Ph.D. (T)
Chief, Office of Health Statistics
Rhode Island Department of Health
3 Capitol Hill, Cannon Building
Providence, Rhode Island 02908-5097

Carol A. Burnett, M.S. (L)
Epidemiologist, Surveillance Branch
National Institute for Occupational
Safety and Health
4676 Columbia Parkway, MS R-18
Cincinnati, Ohio 45226

Lorraine L. Cameron, Ph.D. (T)
Epidemiologist, Division of Surveillance
Hazards Evaluation and Field Studies
National Institute for Occupational
Safety and Health
4676 Columbia Parkway, MS R-18
Cincinnati, Ohio 45226

Mary Chamie, Ph.D., M.P.H. (K)
Statistician, Demographic and Social
Statistics Office
United Nations
Plaza 2, Room 1586
New York, New York 10017

Darren D. Coffman, B.S., M.S. (Y)
Research Manager
Oregon Health Services Commission
800 NE. Oregon Street 17, Suite 550
Portland, Oregon 97232

( ) Indicates session

Victor Cohn (I)
Senior Writer and Columnist
The Health Section, Washington Post
1150 15th Street, NW.
Washington, D.C. 20071

Larry S. Corder, Ph.D. (H)
Associate Research Professor
Center for Demographic Studies
Duke University
2117 Campus Drive
Durham, North Carolina 27706

Lester R. Curtin, Ph.D.** (P)
Chief, Statistical Methods Staff
Office of Research and Methodology

Peter D'Ascoli, M.D., M.P.H. (E)
Faculty Member
School of Public Health
University of Minnesota
420 Delaware Street, SE., Box 97
Minneapolis, Minnesota 55455

Arthur J. Davidson, M.D., M.S.P.H. (L)
Director, Data Management
Denver Public Health Department
605 Bannock Street
Denver, Colorado 80204

Sam Davis (S)
Statistician
Population Division
U.S. Bureau of the Census
Washington, D.C. 20233

Andrew G. Dean, M.D., M.P.H. (BB)
Chief, Systems Development Branch
Division of Surveillance and Epidemiology
Centers for Disease Control and Prevention
1600 Clifton Road, NE.
Atlanta, Georgia 30333

Paula Diehr, Ph.D. (A)
Professor, Biostatistics and Health Services
Department of Biostatistics, SC-32
University of Washington
Seattle, Washington 98195

Walter R. Dowdle, Ph.D. (First Plenary)
Deputy Director, Centers for Disease Control and Prevention
1600 Clifton Road, NE., MS D-14
Atlanta, Georgia 30333

Elaine D. Eaker, Sc.D. (BB)
Assistant Director for Science
Division of Surveillance and Epidemiology
Epidemiology Program Office
Centers for Disease Control and Prevention
1600 Clifton Road, NE., MS c-08
Atlanta, Georgia 30333

Joyce M. Eatmon, M.P.A. (X)
Medical Economist
Center for Health Statistics
Department of Health
4815 West Markham, Slot 19
Little Rock, Arkansas 72205-3867

Cecile H. Edwards, Ph.D. (B)
Graduate Professor, Nutritional Sciences
College of Allied Health Sciences
2400 Sixth Street, NW.
Washington, D.C. 20059

Pennifer Erickson** (A)
Chief, Clearinghouse on Health Indexes
Division of Epidemiology and Health Promotion

Manning Feinleib, M.D., Dr.P.H.** (First Plenary)
Director

Maryanne J. Florio, M.Ed. (W)
Research Scientist, Maternal and Child Health Epidemiology and Program Planning
Department of Health
363 West State Street, CN 364
Trenton, New Jersey 08625

Blanche Frank, Ph.D. (C)
Chief, Bureau of Applied Studies
State Office of Alcoholism and Substance Abuse
55 West 125th Street
New York, New York 10027

Emma L. Frazier, Ph.D. (P)
Chief, Data Management Section
Center for Chronic Disease Prevention and Health Promotion
Centers for Disease Control and Prevention
1600 Clifton Road, NE., MS K30
Atlanta, Georgia 30333

Brian K. Gallagher, M.S. (0)
Program Research Specialist, Perinatal Health Unit, NYS Department of Health
Corning Tower Building, Room 780
Albany, New York 12237

Clifton R. Gaus, M.H.A., Sc.D. (First Plenary)
Senior Advisor, OASH
Department of Health and Human Services
200 Independence Avenue, SW
Washington DC 20201

Jane F. Gentleman, Ph.D. (M)
Chief, Health Status Section
Canadian Centre for Health Information
Statistics Canada, RHC-18A
Ottawa, Ontario, Canada K1A 0T6

Robert S. Gold, Ph.D. (C)
Senior Science Advisor
Macro International, Inc.
8630 Fenton Street, Suite 406
Silver Spring, Maryland 20910

Robin D. Gorsky, Ph.D. (Z)
Associate Professor, Health Management and Policy, University of New Hampshire
R.R.l, Box 161D, Harmony Grove Road
Northwood, New Hampshire 03261

Sheldon Greenfield, M.D. (W)
Senior Scientist, Health Institute
Department of Medicine, Tufts University
750 Washington Street, 345
Boston, Massachusetts 02111

Joel W. Greer, Ph.D. (X)
Economist, Research and Demonstrations
Health Care Financing Administration
6325 Security Boulevard, OM 2504
Baltimore, Maryland 21207

Corazon M. Halasan, Ph.D. (J)
Vital Records and Health Statistics
Public Health Division
1190 St. Francis Drive
Santa Fe, New Mexico 87502-6110

David P. Hamburger, M.D. (U)
Injury Control Fellow
Children's National Medical Center
12809 Circle Drive
Rockville, Maryland 20850

Mary J. Harris, R.N., Dr.P.H. (G)
Assistant Professor, Community Health and Gerontology
University of Texas School of Nursing
1100 Mechanic
Galveston TX 77555-1029

Robert Hartford, Ph.D.** (N)
Deputy Director
Office of International Statistics

Marilyn J. Henderson, M.P.A. (D)
Assistant Chief, Survey and Analysis Branch
Center for Mental Health Services
5600 Fishers Lane, Room 18C-07
Rockville, Maryland 20857

Alan B. Humphrey, Ph.D. (Q)
Associate Professor, College of Business Administration
University of Rhode Island
301 Ballentine Hall
Kingston, Rhode Island 02881-0802

Patricia M. Jamison, M.P.H. (X)
Computer Specialist, Division of Cancer Prevention and Control
Centers for Disease Control and Prevention
4770 Buford Highway, NE., K-55
Atlanta, Georgia 30341-3724

Jared B. Jobe, Ph.D.** (L)
Psychologist, Office of Research and Methodology

Timothy P. Johnson, Ph.D. (R)
Assistant Director, Survey Research Laboratory
University of Illinois at Chicago
910 West Van Buren, Suite 500
Chicago, Illinois 60607

Clifford Johnson, M.S.P.H.** (F)
Special Assistant for Analysis
Division of Health Examination Statistics

Sarita L. Karon, Ph.D. (AA)
Assistant Scientist, Center for Health Systems Research and Analysis
University of Wisconsin
610 Walnut Street, Room 1158
Madison, Wisconsin 53705

Gerald Kaufman, Ph.D. (M-Presented for Logrillo)
Principal Research Scientist
AIDS Institute
New York State Department of Health
Empire State Plaza
Albany, New York 12237

Harold Xilburn, M.A. (Y)
Program Research Specialist
Health Policy and Management
State University of New York
2019 Tower Building, Empire State Plaza
Albany, New York 12237

Russell S. Kirby, Ph.D. (B)
Department of Pediatrics and CARE
University of Arkansas Medical Sciences
4301 West Markham Street, Slot 512
Little Rock, Arkansas 72205

Susan M. Xrebs-Smith, Ph.D. (F)
Nutritionist, Applied Research Branch
National Cancer Institute, DC PC
9000 Rockville Pike
Bethesda, Maryland 20892

James S. Larson, Ph.D. (D)
Professor, Institute of Government
Univ of Arkansas, Public Administration
2801 South University
Little Rock, Arkansas 72204-1099

Meg Leavy (C-Presented for Gold)
Project Director
Macro International, Inc.
8630 Fenton Street, Suite 300
Silver Spring, Maryland 20910

Andrew C. Leon, Ph.D. (D)
Assistant Professor of Biostatistics
Department of Psychiatry, Cornell Univ
525 East 68th Street
New York, New York 10021

Vito M. Logrillo, M.P.H. (M)
Director, Information Resources Management
AIDS Institute, NYS Department of Health
Empire State Plaza, Tower Building
Albany, New York 12237-0020

Mary E. Losch, Ph.D. (AA)
Program Director, Social Science Institute, The University of Iowa
345 Schaeffer Hall
Iowa City, Iowa 52242-1409

James T. Massey, Ph.D.** (P)
Chief, Survey Design Staff
Office of Research and Methodology

Patricia McKinney, M.S., R.D. (F)
Program Analyst, Analysis and Evaluation
Food and Nutrition Service
U.S. Department of Agriculture
3101 Park Center Drive, 2nd Floor
Alexandria, Virginia 22302

David Mechanic, Ph.D. (Second Plenary)
Institute for Health, Health Care Policy, and Aging Research, Rutgers University
30 College Avenue
New Brunswick, New Jersey 08903

Chun-Lo Meng, Ph.D. (V)
Director, Statistical Analysis Unit
Colorado Department of Health
4300 Cherry Creek Drive South, HSVRD-HS-A5
Denver, Colorado 80222-1530

Paul T. Menzel, Ph.D. (U)
Professor, Department of Philosophy and Dean of Humanities
Pacific Lutheran University
Tacoma, Washington 98447

Ele Meux, Ph.D. (0)
Section Chief, Patient Discharge Data
California Office of State-wide Health Planning and Development
818 K Street, Room 100
Sacramento, California 95814

Keith J. Mueller, Ph.D. (R)
Director, Center for Rural Health Research
University of Nebraska Medical Center
600 South 42nd Street
Omaha, Nebraska 68198-4350

Andreas Muller, Ph.D. (Q)
Associate Professor, Health Services
University of Arkansas at Little Rock
2801 University Avenue, Room 207
Little Rock, Arkansas 72204

Douglas R. Murray, M.S. (U)
Director, Center for Health Statistics
Arkansas Department of Health
4815 West Markham Sreet, Slot 19
Little Rock, Arkansas 72205-3867

Raymond D. Nashold, Ph.D. (J)
Director, Center for Health Statistics
State Registrar, Division of Health
1 West Wilson Street, Room 172
Madison, Wisconsin 53702

Merwyn R. Nelson, Ph.D. (V)
Chief, Center for Health Statistics
Illinois Department of Public Health
535 West Jefferson
Springfield, Illinois 62761

Edward Ng, Ph.D., M.A. (N)
Analyst, Occupational and Environmental Health Research, Statistics Canada
Coats Building, Tunney's Pasture
Ottawa, Ontario, Canada K1A OT6

Alan Otten (I)
Columnist
Washington Bureau, Wall Street Journal
1025 Connecticut Avenue, NW.
Washington, D.C. 20036

Mary D. Overpeck, Dr.P.H. (C)
Epidemiologist, Epidemiology Branch
National Institute of Child Health
6100 Executive Boulevard, Room 7B03
Bethesda, Maryland 20892

Victoria Vespe Ozonoff, Ph.D. (G)
Director, Health Resources Statistics
Bureau of Health Statistics, Research and Evaluation
Department of Public Health
150 Tremont Street
Boston, Massachusetts 02111

Elsie Pamuk (Q)
Division of Nutrition
Centers for Disease Control and Prevention
4770 Buford Highway, NE.
Atlanta, Georgia 30341-3724

Stephen T. Parente, M.P.H., M.S. (Y)
Project Coordinator, Johns Hopkins Health Services Research and Development Center
624 North Broadway, Room 600
Baltimore, Maryland 21205

P. Ellen Parsons, Ph.D.** (AA)
Statistician, Division of Health Interview Statistics

Donald L. Patrick, Ph.D., M.S.P.H. (First Plenary)
Director, Social and Behavioral Sciences
University of Washington
Department of Health Services, SC-37
Seattle, Washington 98195

David B. Pryor, M.D. (W)
Associate Professor of Medicine
Duke University Medical Center
Box 3531
Durham, North Carolina 27710

Roger Roberge, M.A. (A)
Analyst, Health Analysis and Modeling Group
Statistics Canada
Tunney's Pasture, RHC Building, 24th
Ottawa, Ontario, Canada K1A OT6

Tanya Roberts, Ph.D. (Z)
Senior Economist, Economic Research
U.S. Department of Agriculture
1301 New York Avenue, NW.
Washington, D.C. 20005-4788

Rhonda Roland (I)
Senior Producer
Cable News Network
1 CNN Center, Box 105366
Atlanta, Georgia 30348-5366

Wilfred Rosenbaum, M.Math. (M)
Department of Computing Science
Computational Epidemiology Laboratory
Simon Fraser University
Burnaby, British Columbia, Canada V5A 1S6

David A. Ross, Sc.D. (BB)
Assistant Director for Information and Communication Services
Public Health Practice Program Office
Centers for Disease Control and Prevention
1600 Clifton Road, NE.
Atlanta, Georgia 30333

Marcel Edward Salive, M.D., M.P.H. (H)
Epidemiologist, Epidemiology, Demography, and Biometry Program
National Institute on Aging
National Institutes of Health
7201 Wisconsin Avenue, Suite 3C309
Bethesda, Maryland 20892

Marcia M. Sass, Sc.D. (X)
Assistant Director, Health Initiatives Evaluation
New Jersey Department of Health
CN 360, Health and Agriculture Building
Trenton, New Jersey 08625-0360

Harry L. Savitt, Ph.D., M.P.H. (G)
Deputy Director, Division of Beneficiary Studies
Office of Research and Demonstrations
Health Care Financing Administration
6325 Security Boulevard, Room 2504 OM
Baltimore, Maryland 21207

William Scanlon (H)
Codirector, Center for Health Policy Studies, Georgetown University
2233 Wisconsin Avenue, NW., Suite 525
Washington, D.C. 20007

Rachel Schwartz, M.P.H. (B)
Associate Director
National Perinatal Information Center
One State Street, Suite 102
Providence RI 02908-5035

Ravi X. Sharma, Ph.D. (B)
Assistant Professor, Department of Health Services Administration
University of Pittsburgh
130 Desoto Street, 228 Parran Hall
Pittsburgh, Pennsylvania 15261

T. Joseph Sheehan, Ph.D. (J)
Professor, Epidemiology and Biostatistics
Department of Community Medicine
University of Connecticut School of Medicine
'Farmington, Connecticut 06030-1910

Monroe Sirken, Ph.D.** (T)
Associate Director
Office of Research and Methodology

Jane E. Sisk, Ph.D. (First Plenary)
Professor, Division of Health Policy and Management
Columbia University School of Public Health
600 West 168th Street
New York, New York 10032

Sharman Stephens (Third Plenary-Presented for; Thorpe)
Policy Analyst, Office of the Assistant Secretary for Planning and Evaluation
Office of Health Policy
Department of Health and Human Services
200 Independence Avenue, SW., Room 446F
Washington, D.C. 20201

Jonathan Reuel Sugarman, M.D., M.P.H. (R)
Medical Epidemiologist, Division of Research, Evaluation and Epidemiology
Portland Area Indian Health Service
2201 Sixth Avenue, Room 300
Seattle, Washington 98121

Barbara Sugland, Sc.D., M.P.H. (E)
Research Associate
Child Trends, Inc.
2100 M Street, NW., Suite 610
Washington, D.C. 20037

Dong Suh, M.P.P.(R)
Research Associate
Asian American Health Forum
116 New Montgomery, Suite 531
San Francisco, California 94105

David A. Swanson, Ph.D., M.A., B.Sc. (S)
Senior Demographic Specialist
Arkansas Institute for Economic Advancement
University of Arkansas
2801 South University Avenue
Little Rock, Arkansas 72204-1099

Karen Smith Thiel, Ph.D. (E)
Director, Healthy Start Evaluation
Health Resources and Services Administration
5600 Fishers Lane, Room 14-36
Rockvilla, Maryland 20857

Kenneth E. Thorpe, Ph.D. (Third Plenary)
Deputy Assistant Secretary for Health Policy
Department of Health and Human Services
200 Independence Avenue, SW., Room 4423
Washington, D.C. 20201

David Thurman, M.D., M.P.H. (G)
Medical Epidemiologist, National Center for Injury Prevention and Control
Centers for Disease control and Prevention
4770 Buford Highway, NE., MS F-41
Atlanta, Georgia 30341

Clarissa Tom, M.P.H. (P)
Manager
Asian American Health Forum
116 New Montgomery, Suite 531
San Francisco, California 94105

Joan Turek-Brezina, Ph.D. (0)
Director, Technical and Computer Support Division
Office of the Assistant Secretary for Planning and Evaluation
200 Independence Avenue, SW.
Washington, D.C. 20201

Dennis Utter (M-Presented for Walsh)
Chief of Design, Quality Control and State Data Branch
National Center for Statistics and Analysis
National Highway Traffic and Safety Administration
400 7th Street, SW., Room 6125
Washington, D.C. 20590

Therese van Houten, D.S.W. (C)
Evaluation Director
Macro International, Inc.
8630 Fenton Street, Suite 406
Silver Spring, Maryland 20910

Lois M. Verbrugge, Ph.D., M.P.H. (H)
Distinguished Research Scientist
Institute of Gerontology
University of Michigan
300 North Ingalls
Ann Arbor, Michigan 48109-2007

William H. Walsh, Jr. (M)
Director, National Center for Statistics and Analysis, New Hampshire TSA
400 7th Street, SW., Room 6125
Washington, D.C. 20590

Signe Wetrogan (S)
Statistician
Population Division
U.S. Bureau of the Census
Washington, D.C. 20233

Amanda White, B.A. (F)
Principal Social Survey Officer
Office of Population Census and Surveys
St. Catherine's House, 10 Kingsway
London WC2B 6JP Great Britain

Gale G. Whiteneck, Ph.D. (K)
Director of Research
Craig Hospital
3425 South Clarkson
Englewood, Colorado 80110

Russell Wilkins, M.Urb. (H)
Senior Analyst
Canadian Centre for Health Information
Statistics Canada, RHC-18N
Ottawa, Ontario, Canada K1A 0T6

Michelle A. Williams, Sc.D. (B)
Assistant Professor of Epidemiology
School of Public Health and Community Medicine
1124 Columbia Street, MS MP381
Seattle, Washington 98104

P. Douglas Williams, M.S.** (R)
Chief, State and Academic Research Applications
Office of Planning and Extramural Programs

Richard A. Windsor, Ph.D., M.P.H. (U)
Special Assistant to the Director
National Heart, Lung, and Blood Institute
NIH, Building 31, Room 5A48
Bethesda, Maryland 20892

Ruth Zambrana, Ph.D. (R)
Consultant, Agency for Health Care Policy and Research
2101 East Jefferson Street
Rockville, Maryland 20852

Alvan 0. Zarate, Ph.D.** (N)
Special Assistant to the Director
Office of the Center Director

Sharon L. Zucconi, Ph.D. (v)
Assistant Professor and Policy Analyst
University of Pittsburgh
School of Public Health
130 DeSoto Street, A655
Pittsburgh, Pennsylvania 15261

**National Center for Health Statistics
6525 Belcrest Road
Hyattsville, Maryland 20782

EXHIBITORS

Centers for Disease Control and Prevention
Center for Mental Health Services
Indian Health Service
National Center for Health Statistics
National Clearinghouse for Alcohol and Drug Information
National Institute on Aging
Public Health Reports

DEPARTMENT OF
HEALTH & HUMAN SERVICES
Public Health Service
Centers for Disease Control and Prevention
National Center for Health Statistics
6525 Belcrest Road, Room 1100
Hyattsville, Maryland 20782

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

SPECIAL
FOURTH-CLASS MAIL
POSTAGE AND FEES PAID
PHS/NCHS
PERMIT No. G-281

ADDRESS CORRECTION REQUESTED1

DHHS Publication No. (PHS) 94-1214