September 2010

An Assessment of the State of the Art for Measuring the Burden of Illness

Final Literature Review

Prepared for

Ansalan Stewart
U.S. Department of Health and Human Services
Washington, DC 20201

Prepared by

Amanda Honeycutt
Linda Brown
Steven Couper
Kevin Smith
Thomas Hoerger
Alex Hardee
Jeffrey Matthews
RTI International
3040 Cornwallis Road
Research Triangle Park, NC 27709

RTI Project Number 0212050.005.001

RTI Project Number
0212050.005.001

An Assessment of the State of the Art for Measuring the Burden of Illness

Final Literature Review

September 2010

Prepared for

Ansalan Stewart
U.S. Department of Health and Human Services
Washington, DC 20201

Prepared by

Amanda Honeycutt
Linda Brown
Steven Couper
Kevin Smith
Thomas Hoerger
Alex Hardee
Jeffrey Matthews
RTI International
3040 Cornwallis Road
Research Triangle Park, NC 27709

Contents

Section			Page
Executive Summary			**ES-1**
1.	**Introduction**		**1-1**
2.	**Literature Review Methodology**		**2-1**
3.	**Epidemiologic Burden of Illness**		**3-1**
	3.1	Epidemiology	3-1
	3.2	Descriptive Epidemiology	3-2
		3.2.1 Measures of Incidence and Prevalence	3-2
		3.2.2 Mortality	3-5
		3.2.3 Causes of Death	3-8
		3.2.4 Survival	3-8
		3.2.5 Life Expectancy and Life Tables	3-9
		3.2.6 Years of Life Lost	3-11
		3.2.7 Age Adjustment/Standardization	3-13
	3.3	Analytical Epidemiology	3-14
		3.3.1 Absolute Measures	3-15
		3.3.2 Relative Measures	3-16
		3.3.3 Tests of Statistical Significance	3-20
4.	**Economic Burden of Illness**		**4-1**
	4.1	Direct Medical Spending	4-3
	4.2	Direct Nonmedical Spending	4-7
	4.3	Indirect Costs Resulting from Excess Morbidity	4-7
	4.4	Indirect Costs Resulting from Early Mortality	4-9
	4.5	Willingness to Pay	4-9
	4.6	Additional Economic Burden Estimation Issues	4-11
	4.7	Emerging Areas in Economic Burden Measurement	4-11
5.	**Quality of Life Burden of Illness Measures**		**5-1**
	5.1	Function Measures	5-1
	5.2	Overall Health Status Measures	5-4

5.3 Utility 5-4
5.4 Generic Health Indexes 5-6
5.5 Disease-Targeted Measures of HRQoL 5-8
5.6 Quality-Adjusted Life Years 5-8
5.7 Summary 5-11
5.8 New Initiatives 5-11

6. Literature Review Summary **6-1**

References **R-1**

Attachments

1 Bibliography with Abstracts

2 National Health Expenditures Aggregate, per Capita Amounts, Percent Distribution, and Average Annual Percent Growth, by Source of Funds: Selected Calendar Years 1960–2008

3 National Health Expenditures, by Source of Funds and Type of Expenditure: Calendar years 2003–2008

4 Cost-of-Illness Summaries for Selected Conditions, January 2006

5 Examples of Health-Related Quality of Life Questionaires

6 Ranking of the Clinical Preventable Burden (CPB) and Cost-Effectiveness (CE) for Various Services

Figures

Number **Page**

3-1. Relationship between Epidemiology Studies..... 3-2

5-1. Example of a Standard Gamble Tree 5-5

Tables

Number **Page**

1-1. Diseases with the Greatest Burden of Illness and Six Measures of Burden of Illness 1-2

3-1. Comparison of Incidence and Prevalence 3-3
3-2. Example of Incidence: Physician-Diagnosed Hypertension among Adults Aged 20 or Older in the Province of Ontario from 1997 to 2004..... 3-4
3-3. Example of Prevalence: Diagnosed Hypertension among Adults Aged 20 or Older in the Province of Ontario from 1995 to 2005..... 3-5
3-4. Factors Influencing Prevalence 3-5
3-5. U.S. Mortality Rates (Per 100,000 Persons) by Race and Sex, 2000–2006 3-6
3-6. Example of Mortality: Maternal Mortality Ratio (Uncertainty Interval) per 100,000 Live Births by Region..... 3-7
3-7. Leading Causes of Death in the United States, 2006..... 3-8
3-8. Esophageal Cancer 5-Year Relative Survival Rates, 1975–2004, by Diagnosis Year, Sex, and Race..... 3-9
3-9. Estimated Life Expectancy at Birth and Selected Ages for Men and Women in Selected Countries.....3-10
3-10. Life for the Total Population: United States, 2005.....3-11
3-11. Deaths, Deaths before Age 75, and YLLs per 100,000 Population Younger than Age 75, United States, 2006 3-12
3-12. Crude and Age-Standardized Death Rates (per 100,000) for Ischemic Heart Disease in Three Selected Countries, 2002.....3-14
3-13. Symbolic Representation of Subjects With and Without a Disease or Outcome by Exposure Status.....3-16
3-14. Relative Risk Example: Risk of Death in Two Cities with High and Low Levels of Pollution 3-17
3-15. Symbolic Representation of Cases and Controls With and Without an Exposure ...3-18
3-16. Odds Ratio Example: Oral Cancer Risk Associated with Consumption of Liquor in Men, Puerto Rico, 1992–1995 3-18
3-17. Incidence Rates, Relative Risks, Excess Risks, Population AttribuRisk, and Corresponding 95% Confidence Intervals of Adenocarcinoma of the Lung within the Iowa Women's Health Study, 1986–1998 3-20
3-18. Relation of Tubal Infertility to History of Sexually Transmitted Diseases 3-22

4-1. Economic Burden of Illness Measures 4-4

5-1. Health Status/Quality of Life Burden of Illness Measures 5-2

5-2. Characteristics of Five Commonly Used Generic Health Indexes..... 5-7
5-3. Potential Increase in Quality-Adjusted Life Years (QALYs) from Fuller Utilization of Select Services 5-9
5-4. Leading Global Diseases in DALYs in 2001.....5-10

EXECUTIVE SUMMARY

Burden of illness measures provide information about the wide-ranging impacts of disease or specific diseases or risk factors on society, government, and the individuals affected by disease. Some burden of illness measures capture the number of people affected by a given disease or risk factor, whereas others capture the impact of disease on longevity, costs, and quality of life.

Because so many different burden of illness measures are used to describe the impact of disease or specific diseases on a population, it may be challenging for policy makers to select the best measure for addressing the policy problem at hand. However, different burden measures capture different aspects of disease effects, and consequently policy makers should consider multiple measures of disease burden when evaluating and establishing priorities for health care spending and research. For example, prostate cancer is the most burdensome type of cancer when measured in terms of new diagnoses per year, but it ranks seventh behind lung and other types of cancers in terms of impact on years of life lost. By considering multiple measures of burden, policy makers can make more informed decisions about which conditions contribute to the greatest overall disease burden. Only when goals are specific to a burden measure should a single burden measure be used. For example, when policy makers are interested in reducing the length of inpatient stays, a measure of number of days of inpatient stay may be adequate for informing decisions.

In this report, we describe our approach to reviewing the literature on the burden of illness, discuss approaches for generating burden measures, and provide examples of burden measures and new approaches from the literature for three types of measures: epidemiologic, economic, and quality of life. Epidemiologic burden of illness measures describe the extent of illness in terms of incidence (the number of new cases of disease), prevalence (the total number of people with a disease or history of a disease), and the impact on deaths. These burden measures are also used as the foundation for generating economic and quality of life burden of disease estimates. Economic burden of illness measures capture the full economic costs of disease, including health care costs, nonmedical spending, transportation costs for treatment, and other opportunity costs of illness, such as productivity losses or marital dissolution resulting from illness. Opportunity costs are the value of health and non-health outcomes that patients and their families and friends are unable to enjoy as a result of disease. They differ from accounting, or financial, costs in that opportunity costs value even those costs or losses for which no monetary cost is incurred, such as productivity losses. Health status and quality of life measures of the burden of disease use patient-reported assessments of health status to characterize the impact of disease across multiple domains (e.g., physical functioning, cognitive functioning, social functioning) using a single measure of health status or quality of life impact. A feature of health status and quality of life measures is that they can capture multiple impacts of

disease, including those impacts that are difficult to value in monetary terms, using a common metric to facilitate comparisons of disease burden across countries, across treatment interventions, or across diseases.

Different approaches to valuing disease costs or quality of life impacts can lead to vastly different estimates of burden, and different values for the same burden measure may create confusion among policy makers as they try to select the best burden estimates for a given disease. Efforts are underway to improve consistency in burden measurement for health care costs and for measuring patient-reported health outcomes. Improved estimates of disease burden may be of use to health officials and policy makers in making decisions about priorities for future research and about which diseases and intervention methods to target with public health resources.

Given the many different types of burden of disease estimates published in government statistics or in peer-reviewed research, policy makers may find it challenging to make use of the burden of illness information with which they are presented. Because different burden measures capture different aspects of how a disease affects the population, policy makers may find that the best approach is to consider multiple measures of disease burden when evaluating and establishing priorities for health care spending and research. For example, diseases that are most burdensome in terms of the number of people affected may not be the most burdensome in terms of impacts on longevity. However, in cases where policy decisions focus on a specific burden of illness measure, such as limiting hospitalization costs, that single burden measure may be useful for guiding policy decisions.

1. INTRODUCTION

Policy makers need information on the broad impacts of diseases or disease risk factors to inform public health decision making. For example, information about how a population is affected by disease can be used to establish a baseline and goals for disease prevention and health promotion efforts and can inform priority setting for disease prevention and management. Many different burden of illness measures have been used historically to describe the wide-ranging impacts of disease or of specific diseases or risk factors on the population. Burden of illness measures provide information about the impact of disease or specific diseases on society, government, and the individuals affected by disease. Some burden of illness measures capture the number of people affected by a given disease or risk factor, whereas others capture the impact of disease on longevity, costs, and quality of life.

Because so many different burden of illness measures are used to describe the impact of disease or specific diseases on a population, it may be challenging for policy makers to select the best measure for addressing the policy problem at hand. However, different burden measures capture different aspects of disease effects, and therefore policy makers should consider multiple measures of disease burden when evaluating and establishing priorities for health care spending and research. For example, as Brown, Lipscomb, and Snyder (2001) show, using cancer as an example, diseases that are most burdensome in terms of the number of people affected may not be the most burdensome in terms of impacts on longevity. They report that prostate cancer ranks highest among cancers in terms of new diagnoses per year, but the total years of life lost from prostate cancer is far lower than for lung cancer and, among all cancers, ranks seventh in terms of impact on years of life lost (Brown et al., 2001). Similarly, among cancers, lung cancer has by far the greatest impact on reducing quality of life across established market economies,[1] followed by breast, colorectal, and prostate cancers (Brown et al., 2001).

Thacker et al. (2008) provide a broader example of how relying on a single measure of burden of illness could lead policy makers to overlook the impact of diseases with large burdens when evaluated using other burden of illness measures. Thacker et al. examined the leading causes of public health burden in the United States using eight burden of illness measures. Their findings on the five diseases with the highest levels of burden vary considerably across the six measures shown in Table 1-1, which capture disease impacts on years of life lost, quality of life, underlying causes of death, days hospitalized, and costs. Heart disease, cancer, and injuries/trauma rank among the most burdensome illnesses across measures of years of life lost, quality of life (measured using disability-adjusted life

[1]The specific countries included are Western Europe, the United States, Canada, Japan, Australia, and New Zealand. The quality of life impact is captured using the disability-adjusted life year.

Table 1-1. Diseases with the Greatest Burden of Illness and Six Measures of Burden of Illness

Measure	Years of Life Lost (YLL) Before 75	Quality of Life (Disability-Adjusted Life Years [DALYs])	Underlying Causes of Death	Hospital Days	Quality of Life (Disability)	Costly Conditions
1	Malignant neoplasms	Ischemic heart disease	Tobacco	Perinatal conditions	Arthritis	Heart disease
2	Diseases of heart	Cerebrovascular disease	Poor diet/ physical inactivity	Septicemia	Back problems	Trauma
3	Unintentional injuries	Motor vehicle crashes	Alcohol	Psychoses	Heart trouble	Cancer
4	Suicide	Depression	Microbial agents	Malignant neoplasms	Respiratory problems	Pulmonary conditions
5	Homicide	Lung cancer	Toxic agents	Pneumonia	Hearing problems	Mental disorders

Source: Thacker et al. (2008).

years [DALYs]), and costs. The underlying causes of death are consistent with the burden measures for years of life lost, quality of life, and costs. For example, tobacco, poor diet, and physical inactivity are leading risk factors for heart disease and cancer, and alcohol contributes to certain injuries and trauma. However, when considering hospital days and disability, Thacker et al. (2008) show that conditions other than heart disease, cancer, and injuries are most burdensome.

By considering multiple measures of burden, policy makers can make more informed decisions about which conditions contribute to the greatest overall burden when general information is required about the impact of a particular disease or diseases. When goals are specific to a burden measure—for example, if policy makers are interested in determining what conditions to focus on to reduce inpatient lengths of stay—that burden measure should provide policy makers with adequate information to address the problem.

In this report, we describe our approach to reviewing the literature on the burden of illness, discuss approaches for generating burden measures, and present examples from the literature for three types of burden of illness measures: epidemiologic, economic, and quality of life. We then discuss the policy implications of our findings on the various burden of illness measures and the various ways of measuring each component of disease burden.

2. LITERATURE REVIEW METHODOLOGY

The goal of this literature review is to summarize the published literature on burden of illness methods, measures, and new and emerging approaches. We have identified three broad categories of burden of illness measures for which we find key articles and provide examples of recent burden measures: epidemiologic, economic, and health status measurement/quality of life measures. A fourth category includes articles that are more generally related to burden of illness measurement or that deal with more than one of the other three burden of illness categories.

The first step for the literature review was to compile references from reviews of economic and quality of life burden of illness measures. For example, we included many of the articles cited in *Cost-of-illness studies—A primer (*Segel, 2006); *Valuation of morbidity losses: Meta-analysis of willingness-to-pay and health status measures* (Van Houtven et al., 2003); *A review of the use of health status measures in economic evaluation* (Brazier, Deverill, & Green, 1999); and the July 2009 supplement to *Medical Care*.

We then searched the PubMed database using combinations of the terms "epidemiology," "burden," "cost of illness," "quality of life," "quality adjusted life years," "economic cost," and "methods." For example, one search used "economic cost" AND "burden" AND "methods." All of the PubMed searches excluded articles not published in English, not relating to humans, and published before 2000. Due to the vast literature available, our searches typically produced a few hundred results, many of which were applications of specific burden of illness methods or examples of measures for specific diseases rather than more general descriptions of the methods. To further limit the number of articles for review, we eliminated any articles with titles that were not related to burden of illness or that focused on reporting burden of illness estimates for a particular disease or illness. We reviewed the abstracts of the remaining articles to further eliminate articles that were not clearly related to burden of illness measurement. We saved the full electronic articles for all remaining peer-reviewed articles with a primary focus on burden of illness estimation or measurement. We provide an annotated bibliography of these articles in Attachment 1. These articles were reviewed with the intent of providing a background document that describes various types of burden of illness measures and approaches for measurement and provides some examples of how the different approaches have been applied. As we collected articles using Google Scholar, we occasionally identified other closely related articles that have been included in the literature review. We also shared the list with RTI project team members who are experts in, respectively, epidemiologic, economic, and quality of life burden of illness measures to obtain their feedback on additional articles to include.

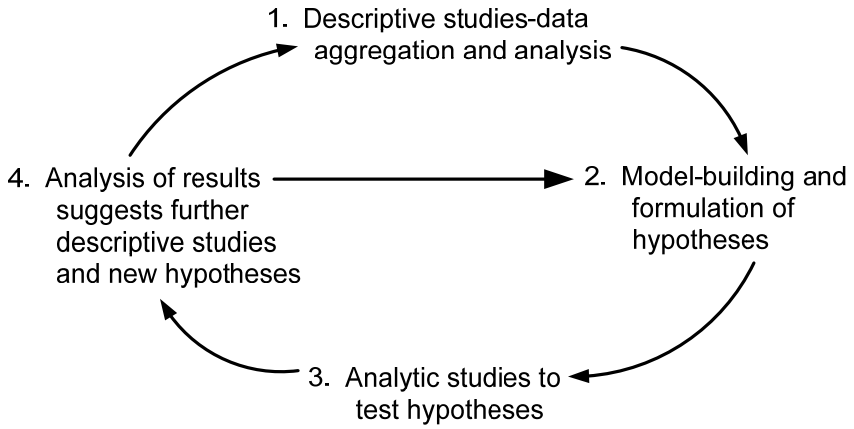
3. EPIDEMIOLOGIC BURDEN OF ILLNESS

Epidemiology is considered to be the basic science of public health and thus is a natural starting point for estimating the burden of disease or illness. It also provides the foundation for quantifying economic and health status measures of disease burden (Spasoff, 1999). The first published burden of disease study was written in 1662 by Englishman John Graunt (1620–1674) and used weekly church funeral logs to identify disease outbreaks (McKenna & Zohrabian, 2009). Since then, the field of epidemiology has evolved, and sophisticated studies of disease burden elucidate individual and environmental influences on health. The Framingham heart study is an example of a longitudinal study that tracks mortality and the causes of chronic heart disease across generations of participants (Rothman, Greenland, & Lash, 2008). In this section, we define epidemiology, descriptive epidemiology, and analytical epidemiology; and we describe key burden measures and tools from an epidemiological perspective.

3.1 Epidemiology

The word "epidemiology" comes from the Greek words *epi*, meaning on or upon; *demos*, meaning people; and *logos*, meaning the study of. Epidemiology is the study of the occurrence and distribution of diseases, causes of death, and behaviors and their determinants in populations as well as the application of the knowledge obtained to control health problems (Porta, 2008). Epidemiology is a data-driven quantitative scientific discipline that relies heavily on an objective and systematic approach to the collection and analysis of data (Porta, 2008). In addition, it is integrative in that it cuts across several fields, including biology, economics, and social/behavioral sciences to develop hypotheses (Lilienfeld & Lilienfeld, 1980). There are two main types of epidemiological studies—descriptive and analytic—that collect, analyze, and interpret information on the distribution and determinants of disease, respectively (Dicker et al., 2009). Figure 3-1 illustrates the relationship between these two types of epidemiologic studies.

Figure 3-1. Relationship between Epidemiology Studies



Source: Reproduced from Mausner and Kramer (1985).

3.2 Descriptive Epidemiology

Descriptive epidemiology studies are observational in nature and are designed to describe existing health outcomes according to three categories of variables: person, place, and time. Person variables include demographic characteristics, such as age, sex, race, and socioeconomic status; place variables include the location where a health event occurred; and time variables refer to the duration, age, birth cohort, or time trend in which a health event occurred. One essential role of descriptive studies is hypothesis generation. A well-known historical example of a hypothesis-generating descriptive study is the exploratory analysis of stomach and colon cancer mortality rates among ethnic Japanese living in Japan and California. By comparing the cancer mortality rates among individuals living in these locales and between first- and second-generation Japanese immigrants in California and Hawaii, epidemiologists were able to develop a hypothesis that environmental factors such as diet and lifestyle were more important risk factors for stomach and colon cancer than genetic factors (Boslaugh, 2008). Important descriptive measures, tools, and techniques that we describe below include incidence, prevalence, mortality, survival, life expectancy, and age adjustment/standardization.

3.2.1 Measures of Incidence and Prevalence

Incidence is the rate of occurrence of new cases of disease arising in a given period of time in a specified population, whereas prevalence is the proportion of individuals with a disease or condition in a specified population at a specific point in time (Table 3-1). Together, these measures form the basis of measuring disease occurrence and enable epidemiologists to gauge/estimate the overall magnitude (prevalence) of a health problem or determine the short-term trends (incidence) in a population. For example, there may be low incidence and high prevalence for a chronic disease, such as diabetes; or high incidence and low prevalence for a disease that lasts only a short time, such as the common cold.

Table 3-1. Comparison of Incidence and Prevalence

	Incidence	Prevalence
Numerator	Number of **new** cases of disease during a specified period of time	Number of **existing** cases of disease at a given point in time
Denominator	Population at risk	Population at risk
Focus	Whether the event is a new case; time of onset of the disease	Presence or absence of a disease; time period is arbitrary; rather a snapshot in time
Uses	Expresses the risk of becoming ill; the main measure of acute diseases or conditions, but also used for chronic diseases; more useful for studies of causation	Estimates the probability of the population being ill at the period of time being studied; useful in the study of the burden of chronic diseases and implication for health services

Incidence Further Explained

Incidence of a disease is calculated as follows:

$$\text{Incidence rate} = \frac{\text{Number of new cases of disease over a defined time period}}{\text{Person-time experience of the population}}$$

where person-time experience is the number of persons multiplied by the period over which they were monitored; this is often called person-years. For each individual in the population, the time of observation is the period that the person remains disease-free. The denominator used for the calculation of incidence is therefore the sum of all of the disease-free person-time periods during the period of observation of the population at risk. Because it may not be possible to measure disease-free periods precisely, the denominator is often calculated approximately by multiplying the average size of the study population by the length of the study period. This is reasonably accurate if the population is large and stable and incidence is low.

In practice, the incidence rate is typically used to describe the number of new cases that develop in a year in a specified population (Brownson et al., 1993). The numerator refers only to new disease events. The incidence rate is always expressed per unit of time (e.g., per 1,000 population per year for a fairly common disease such as hypertension or per 100,000 population per year for a fairly rare disease such as cancer).

In Table 3-2, for example, the age- and sex-adjusted incidence rate per 1,000 population of physician-diagnosed hypertension among adults aged 20 or older in Ontario, Canada, was compared for two time periods (1997 and 2004) and two age groups (20–49 years and ≥50 years). The age- and sex-adjusted incidence of hypertension increased from 25.5 per 1,000 adults in 1997 to 32.1 per 1,000 adults in 2004, a relative increase of 25.7%.

Table 3-2. Example of Incidence: Physician-Diagnosed Hypertension among Adults Aged 20 or Older in the Province of Ontario from 1997 to 2004

	1997		2004	
Age Group	No. with Hypertension	Rate per 1,000[a]	No. with Hypertension	Rate per 1,000[a]
≥20 year	140,137	25.5	171,338	32.1
20–49 year	43,576	9.4	60,147	12.0
≥50 year	96,561	51.4	111,191	64.3

[a]Rates are adjusted for age and sex using 2001 Canadian census data.

Source: Tu et al. (2008).

Prevalence Further Explained

Prevalence of a disease is calculated as follows:

$$\text{Prevalence} = \frac{\text{Number of existing cases of disease at a specified time}}{\text{Number of persons in the population at that specified time}}$$

The numerator refers to existing cases of disease. Data on the population at risk (i.e., those currently free of the disease who could get the disease) are not always available for use in the denominator; thus, in many studies, the total population in the study area is used as an approximation. The occurrence of disease is measured at a point (point prevalence) or period (period prevalence) in time rather than over an interval. Prevalence is often expressed as cases per 100 (percentage) or per 1,000 population and has to be multiplied by the appropriate factor: 10^n.

Table 3-3 looks at changes in the prevalence of diagnosed hypertension in Ontario for similar age groups and time periods as those shown in Table 3-2. The number of adults with existing hypertension more than doubled from 1995 (153.1 per 1,000 adults) to 2005 (244.8 per 1,000 adults), a relative increase of 60.0%. Also, the number of adults with existing hypertension in 2005 is much greater than the number of newly diagnosed cases in 2004 because most people diagnosed with hypertension have a long period of survival. This large and growing increase has the potential to overwhelm the health care system and to have financial implications for provincial drug plans (Tu et al., 2008).

Table 3-3. Example of Prevalence: Diagnosed Hypertension among Adults Aged 20 or Older in the Province of Ontario from 1995 to 2005

	1995		2005	
Age Group	No. with Hypertension	Rate per 1,000[a]	No. with Hypertension	Rate per 1,000[a]
≥20 year	1,139,478	153.1	2,311,042	244.8
20–49 year	238,462	50.2	460,246	82.1
≥50 year	901,016	318.6	1,850,796	506.7

[a]Rates are adjusted for age and sex using 2001 Canadian census data.

Source: Tu et al. (2008).

Because prevalence can be influenced by many factors unrelated to the cause of the disease, incidence studies usually provide stronger evidence of causality. However, measures of prevalence are helpful in assessing the need for preventive action, health care, and the planning of health services. Table 3-4 presents examples of the many factors that either increase or decrease prevalence.

Table 3-4. Factors Influencing Prevalence

Increased by	Decreased by
Longer duration of the disease	Shorter duration of the disease
Prolongation of life of patients without cure	High case-fatality rate from disease
Increase in new cases (increase in incidence)	Decrease in new cases (decrease in incidence)
In-migration of cases	In-migration of healthy people
Out-migration of healthy people	Out-migration of cases
In-migration of susceptible people	Improved cure rate of cases

3.2.2 Mortality

Mortality, most commonly expressed as a mortality rate, is the total number of deaths in a population or deaths due to a specific health cause, scaled to the size of that population and per a unit of time. Mortality rates can focus on many different segments of the population and can be expressed in several forms, including crude mortality rate, cause-specific mortality rate, age-specific mortality rate, maternal mortality rate, and infant mortality rate. Crude mortality rate is generally expressed as the total number of deaths per 100,000 population per year, whereas infant mortality rate is usually expressed per 1,000 live births per year. Mortality rate is calculated as follows:

$$\text{Mortality rate} = \frac{\text{Number of deaths during a specified period}}{\text{Number of persons at risk for dying during same period}}$$

Table 3-5 shows an example of recent U.S. total mortality rates by race and sex. Striking differences are apparent with rates for black males, the group with the highest rates, almost twice that of white females, the group with the lowest rates.

Table 3-5. U.S. Mortality Rates (Per 100,000 Persons) by Race and Sex, 2000–2006

	All Races			White			Black		
Year	Both Sexes	Male	Female	Both Sexes	Male	Female	Both Sexes	Male	Female
2006	776.5	924.8	657.8	764.4	908.2	648.2	982.0	1,215.6	813.0
2005	798.8	951.1	677.6	785.3	933.2	666.5	1,016.5	1,252.9	845.7
2004	800.8	955.7	679.2	786.3	936.9	666.9	1,027.3	1,269.4	855.3
2003	832.7	994.3	706.2	817.0	973.9	693.1	1,065.9	1,319.1	901.8
2002	845.3	1,013.7	715.2	829.0	992.9	701.3	1,083.3	1,341.4	901.8
2001	754.5	1,029.1	721.8	836.5	1,006.1	706.7	1,101.2	1,375.0	912.5
2000	869.0	1,053.8	731.4	849.8	1,029.4	715.3	1,121.4	1,403.5	927.6

Source: http://www.cdc.gov/NCHS/data/nvsr/nvsr57/nvsr57_14.pdf.

A specific measure of mortality, the maternal mortality rate (MMR) is calculated by dividing the total maternal mortality rate by the general fertility rate for the period and is expressed per 100,000 births by multiplying the product by 1,000. Table 3-6 summarizes maternal mortality by region from 1980 to 2008 for 181 countries. In most regions, maternal deaths have declined substantially. However, the analysis draws attention to the important adverse effect of the HIV epidemic on the MMR, particularly in east, southern, and west Sub-Saharan Africa where the high MMR is stable or continues to rise. Also striking are the rising MMRs observed for the United States, especially the MMR of 17 observed in 2008—a ratio higher than Asia-Pacific, high income; Australasia; and central and western Europe.

Table 3-6. Example of Mortality: Maternal Mortality Ratio (Uncertainty Interval) per 100,000 Live Births by Region

Maternal Mortality Ratio (uncertainty interval) per 100,000 Live Births				
	1980	**1990**	**2000**	**2008**
Asia-Pacific, high income	28 (26–31)	14 (13–15)	10 (9–11)	8 (8–9)
Asia, central	105 (96–115)	72 (68–77)	60 (56–64)	48 (45–52)
Asia, east	162 (142–183)	86 (76–98)	55 (48–62)	40 (35–46)
Asia, south	788 (568–1099)	560 (391–794)	402 (293–555)	323 (232–444)
Asia, southeast	438 (337–573)	248 (187–337)	212 (155–293)	152 (112–212)
Australasia	9 (8–11)	7 (6–8)	6 (5–7)	6 (5–7)
Caribbean	426 (293–613)	348 (234–518)	323 (218–483)	254 (168–372)
Europe, central	47 (43–51)	34 (31–37)	18 (17–20)	13 (12–14)
Europe, eastern	54 (49–60)	43 (39–48)	41 (37–45)	32 (29–35)
Europe, western	16 (15–17)	10 (10–11)	8 (8–9)	7 (7–8)
Latin America, Andean	326 (248–426)	229 (176–295)	156 (116–205)	103 (77–134)
Latin America, central	125 (114–137)	85 (77–94)	70 (64–78)	57 (51–63)
Latin America, southern	76 (69–84)	54 (49–60)	44 (39–49)	41 (36–45)
Latin America, tropical	150 (87–240)	113 (66–184)	71 (47–107)	57 (37–87)
North Africa/Middle East	299 (250–355)	183 (154–218)	111 (92–135)	76 (61–94)
North America, high income	12 (10–13)	11 (10–12)	13 (11–15)	16 (14–18)
Canada	7 (6–9)	6 (5–7)	6 (5–7)	7 (5–8)
USA	12 (10–14)	12 (10–13)	13 (12–15)	17 (15–19)
Oceania	517 (334–784)	416 (252–649)	329 (202–518)	279 (174–434)
Sub-Saharan Africa, central	711 (487–1072)	732 (488–1101)	770 (535–1,108)	586 (392–839)
Sub-Saharan Africa, east	707 (586–854)	690 (574–842)	776 (639–948)	508 (430–610)
Sub-Saharan Africa, southern	242 (184–319)	171 (132–222)	373 (280–499)	381 (288–496)
Sub-Saharan Africa, west	683 (577–818)	582 (485–709)	742 (608–915)	629 (508–787)

Source: Hogan et al. (2010).

3.2.3 Causes of Death

Cause of death, demographic, and other descriptive data obtained from the U.S. National Vital Statistics System are used to present the characteristics of those dying in the United States, to determine life expectancy (see Section 3.2.5), and to compare mortality trends with other countries (see Table 3-6). Cause-specific deaths in a population can be assessed by looking at a variety of measures, such as the leading causes of death, the leading causes of cancer death, and the types of accidental deaths. Table 3-7 presents the 10 leading causes of death in the United States in 2006. Heart disease and cancer, the number one and two causes of death, respectively, were responsible for almost 50% of all U.S. deaths. In the United States, cause of death varies greatly by race and sex. For example, in 2006, males were more than twice as likely to die from accidents as females (male to female ratio = 2.2) and blacks were more than twice as likely as whites to die from diabetes mellitus (black to white ratio = 2.1).

Table 3-7. Leading Causes of Death in the United States, 2006

Rank	Cause of Death	2006 Deaths	Percentage of Total Deaths (%)	Ratio: Male to Female	Ratio: Black to White
1.	Diseases of heart	631,636	26.0	1.4	1.3
2.	Malignant neoplasms	559,888	23.1	1.5	1.3
3.	Cerebrovascular diseases	137,119	5.7	1.4	1.2
4.	Chronic lower respiratory diseases	124,583	5.1	1.3	0.7
5.	Accidents (unintentional injuries)	121,559	5.0	2.2	0.9
6.	Diabetes mellitus	72,449	3.0	1.4	2.1
7.	Alzheimer's disease	72,432	3.0	1.4	2.1
8.	Influenza and pneumonia	56,236	2.3	1.4	1.1
9.	Nephritis, nephritic syndrome and nephrosis	45,344	1.9	1.4	2.3
10.	Septicemia	34,234	1.4	1.2	2.1

Source: http://www.cdc.gov/NCHS/data/nvsr/nvsr57/nvsr57_14.pdf.

3.2.4 Survival

The survival rate is the proportion of people in a specified group alive at the beginning of a specified time interval who survive to the end of the time interval. Survival rates are calculated to address the survivability of diseases for set periods of time, typically 1, 5, and 10 years after diagnosis. Five-year relative survival rates for cancer compare people with a

particular cancer to similar people (e.g., for a comparable race-, sex-, and time period-specific cohort) in the general population and are expressed as percentages. Table 3-8 presents an example of 5-year relative survival rates (based on follow-up through 2005) following a diagnosis of esophageal cancer. Survival rates increased consistently for all race-sex groups over the 40-year period. However, the survival rates for black men were substantially lower than the rates for the other groups during each time period.

Table 3-8. Esophageal Cancer 5-Year Relative Survival Rates, 1975–2004, by Diagnosis Year, Sex, and Race

Esophageal Cancer 5-Year Relative Survival Rates (%)[a]				
Year of Diagnosis	**White Male**	**White Female**	**Black Male**	**Black Female**
1975–1979	5.1%	6.2%	2.2%	6.8%
1985–1989	10.8%	10.5%	6.7%	10.3%
1996–2004	17.6%	19.7%	9.0%	15.6%

[a]Based on data from nine population-based Surveillance, Epidemiology, and End Results (SEER) registries.

Source: Brown & Devesa (2009).

3.2.5 Life Expectancy and Life Tables

Life expectancy is defined as the average number of years an individual at a given age is expected to live if current mortality rates continue. It is most commonly used as an indicator of overall population health. For example, in Table 3-9, women in Japan have the highest life expectancy at <1 year (85.9 years) and at representative ages, whereas women in Zimbabwe have the lowest life expectancy at <1 year (42.7 years) and ages 30 to 34 (20.6 years). The lowest life expectancy rates for ages 50 to 54 and 70 to 74 are observed for men in Zimbabwe (17.4 years) and men in the Congo (9.0 years), respectively. In countries with high infant and child mortality rates the life expectancy at birth is highly sensitive to the rate of death in the first few years of life. In addition, high rates of maternal mortality affect life expectancy at birth and throughout the reproductive years. For example, the under-5 mortality rate in 2006 per 1,000 live births for both sexes combined was 85 in Zimbabwe, but only 3 in Japan (World Health Organization, 2006). The MMR in 2005 per 100,000 live births in Zimbabwe and Japan were 880 and 6, respectively (World Health Organization, 2007).

Table 3-9. Estimated Life Expectancy at Birth and Selected Ages for Men and Women in Selected Countries

Life Expectancy at Birth and Selected Ages (Years), 2006								
	Men				Women			
	Age Range				Age Range			
Country	<1	30–34	50-54	70-74	<1	30-34	50-54	70-74
Japan	79.2	50.0	31.2	14.9	85.9	56.5	37.2	19.2
Canada	78.3	49.4	30.6	14.4	82.9	53.6	34.4	17.2
USA	75.5	47.2	29.0	13.8	80.4	51.5	32.7	16.2
China	71.7	44.3	25.9	11.1	75.2	47.9	29.0	12.6
Egypt	66.1	39.9	22.3	9.5	70.5	43.7	25.1	10.4
India	61.8	38.7	22.1	10.1	63.8	41.7	24.2	10.4
Congo	53.1	34.1	20.4	9.0	55.2	36.3	22.9	10.0
Zimbabwe	43.9	21.2	17.4	9.6	42.7	20.6	22.2	10.9

Source: World Health Organization (2008).

Life tables are a statistical/actuarial tool used to describe the pattern of mortality and survival in a population. They are typically used to portray expectation of life at various ages and also can provide information on numbers of individuals who survive to various ages, median age at death, age-specific death rates, and the probability of dying at certain ages. Life tables are typically constructed separately for men and women because of their substantially different mortality rates, and they often include various health risks (e.g., smoking) to illustrate the resulting effect on one's likelihood of dying. Table 3-10 is an excerpt from a U.S. life table that calculates for selected ages the number of additional years of life a person of age X is expected to live. It is interesting to note that life expectancy increases with the number of years lived. For example, life expectancy is 77.4 years at birth (0–1) but 88.5 years at age 80 (80–81). The analysis of life tables to determine mortality and its derivative life expectancy is an important tool that can be incorporated into studies of burden of disease to provide empirical information on the health status of the population for evidence-based public health policy making.

Table 3-10. Life Table for the Total Population: United States, 2005

Age	Probability of Dying between Ages x to x+1	Number Surviving to Age x	Number Dying between Ages x to x+1	Person-Years Lived between Ages x to x+1	Total # of Person-Years Lived Above Age x	Expectation of Life at Age x
0–1	0.006879	100,000	688	99,398	7,744,259	77.4
1–2	0.000463	99,312	46	99,290	7,649,471	77.0
20–21	0.000894	98,713	88	98,668	5,762,196	58.4
40–41	0.001943	96,488	187	96,394	3,808,021	39.5
60–61	0.009473	87,966	833	87,550	1,943,926	22.1
80–81	0.058457	53,338	3,118	51,779	455,623	8.5

Source: U.S. Department of Health and Human Services (2010).

3.2.6 Years of Life Lost

Years of life lost (YLLs), sometimes referred to as years of potential life lost, takes into account both the number of deaths attributable to a cause of death in a year and the number of years that a person would have lived if he or she had not died from that cause in the year. This concept captures the notion that a death at age 20 creates a larger burden than a death at age 85 because the person dying at 20 loses more years than a person dying at 85.

A general measure of YLL is given by the equation

$$YLL = \sum_i d_i w_i$$

where d_i is the number of deaths at age i, and w_i is the weight given to a death at this age. As emphasized by Gardner and Sanborn (1990), different authors use different weights to calculate YLLs, and the weighting scheme used will affect the relative ranking of YLLs from specific causes. One of the simplest weighting schemes is to set the weight equal to life expectancy at birth minus the age of death. For example, if 10 deaths occur at age 15 in a country with a life expectancy at birth of 75 years, deaths at that age would contribute 600 YLLs (= 10* (75[15]). This measure can be calculated using data on deaths by age and a single value of life expectancy at birth. A slightly more complicated weighting scheme sets the weight at age i equal to the life expectancy at age i. In this case, the weight for a death at age 0 might be set equal to 77.9 years (the life expectancy at birth in the United States), while the weight for a death at age 75 would be set equal to 11.7 years (the life expectancy conditional on living to age 75) (Xu et al., 2010). More complicated weighting schemes can also be applied, usually based on specific value judgments. For example, only working years

(such as ages 20 to 65) might receive values in the weights. A measure based on this weighting scheme could be interpreted as a measure of productive years of life lost.

To create standardized estimates of YLLs for different diseases, the World Health Organization (WHO) applies the same life expectancy for deaths in all regions of the world. In addition, it discounts life years by 3% annually from the date of death and uses age weights that are lower for years lived at younger and older ages (World Health Organization, 2010). The WHO YLL estimates are notable because they form one of the two components in calculating disability-adjusted life years (DALYs).

YLLs provide more information about the burden of disease than simple death counts because they also account for YLLs at each age. As a result, causes of death that rank highly in the number of deaths may or may not be highly ranked in YLLs. For example, in U.S. estimates of potential YLLs under age 75 (where w_i = 75 – age at death), heart disease ranks first in total deaths, second in deaths under age 75, tenth in the average years lost per death, and third in YLLs under age 75 (Table 3-11). In contrast, causes of death that typically occur at earlier ages, such as accidents, suicides, and homicides, rank much higher in YLLs (second, fourth, and fifth, respectively) than they rank in total deaths (fifth, eleventh, and seventeenth, respectively). Like other purely death-related measures, YLLs do not provide information on morbidity-related burden.

Table 3-11. Deaths, Deaths before Age 75, and YLLs per 100,000 Population Younger than Age 75, United States, 2006

Cause of Death	Deaths		Deaths before age 75		Average Years Lost per Death		YLL per 100,000 Population under Age 75	
	Total	Rank	Total	Rank	Average	Rank	YLLs	Rank
Diseases of heart	631,636	1	213,846	2	14.96	7	1,138.0	3
Malignant neoplasms	559,888	2	309,927	1	14.38	10	1,585.7	1
Cerebrovascular diseases	137,119	3	38,332	5	14.61	9	199.3	6
Chronic lower respiratory diseases	124,583	4	45,974	4	11.09	11	181.4	8
Accidents	121,599	5	93,330	3	35.09	2	1,165.4	2
Diabetes mellitus	72,449	6	35,581	6	14.74	8	186.6	7
Alzheimer's disease	72,432	7	4,584	21	NR	NR	Not in top 11	Not in top 11
Influenza and pneumonia	56,326	8	13,041	13	16.98	6	78.8	11
Nephritis, nephrotic syndrome and nephrosis	45,344	9	15,420	10	NR	NR	Not in top 11	Not in top 11

(continued)

Table 3-11. Deaths, Deaths before Age 75, and YLLs per 100,000 Population Younger than Age 75, United States, 2006 (continued)

	Deaths		Deaths before age 75		Average Years Lost per Death		YLL per 100,000 Population under Age 75	
Cause of Death	Total	Rank	Total	Rank	Average	Rank	YLLs	Rank
Septicemia	34,234	10	14,097	12	NR	NR	Not in top 11	Not in top 11
Chronic liver disease and cirrhosis	27,555	13	22,751	8	19.44	5	157.4	9
HIV	12,113	21	11,995	14	29.17	4	124.5	10
Suicide	33,300	11	30,385	7	32.30	3	349.2	4
Homicide	18,573	17	18,192	9	43.41	1	281.0	5
All causes	2,426,264		1,056,934		19.79		7,442.3	

Sources: Deaths and Deaths before age 75: Table 10 from Deaths: Final Data for 2006. NVSR Volume 57, Number 14; YLL per 1000,000 population under age 75 from National Center for Health Statistics. Health, United States, 2009: With Special Feature on Medical Technology. Hyattsville, MD. 2010; Average years lost per death calculated by authors based on estimate of 2006 population under age 75 = 281,054,976 from http://www.census.gov/popest/national/asrh/NC-EST2006/NC-EST2006-01.xls.

3.2.7 Age Adjustment/Standardization

Age adjustment or age standardization is a technique commonly used by demographers and epidemiologists to compare the incidence or mortality of different populations when the age profiles of the populations are different. It allows epidemiologists to assess the health of various populations based on the incidence or mortality rates they would have had if they had similar demographic structures. Age-adjusted data are typically standardized to the age structure of a given population and a given reference year. In Table 3-12, the crude death rate for ischemic heart, a disease more common in older adults, is much higher for Finland than Botswana and Brazil. However, after age adjustment to the WHO Global Standard population, the death rate for ischemic heart disease is very similar for all countries: 119–120/100,000 population. Finland's much higher crude rate was the result of a much larger elderly population than that of the other two countries, resulting in higher amounts of ischemic heart disease for the population as a whole.

Table 3-12. Crude and Age-Standardized Death Rates (per 100,000) for Ischemic Heart Disease in Three Selected Countries, 2002

Crude and Age-Standardized Death Rates (per 100,000) for Heart Disease, 2002		
Country	**Crude Death Rate**	**Age-Standardized Death Rate[a]**
Botswana	39	119
Brazil	79	119
Finland	240	120

Source: World Health Organization (2004).

[a]Rates are age-adjusted to the WHO Global Standard Population.

3.3 Analytical Epidemiology

Analytical epidemiology studies search for causes and effects. Epidemiologists conduct analytical studies to quantify the association between health exposures and outcomes and to test the hypotheses of causal relationships often developed through descriptive studies. Although epidemiology by itself can never prove that a particular exposure was the exact cause of a certain health outcome, it often provides sufficient evidence for public health officials to take the appropriate control and prevention measures.

Epidemiologists test hypotheses through the use of three study designs: *cohort*, which tracks a group or groups of subjects who either have or have not been exposed to an etiologic agent to see if they develop the disease or health outcome of interest; *case-control*, which involves the enrollment of a group of subjects with a particular disease or health-related complication and a similar sized group without the disease to compare past exposures; and *cross-sectional*, in which a sample of persons from a population is enrolled and their exposures and health outcomes are measured simultaneously, ignoring any time variables and focusing on the prevalence of the health outcome (Dicker et al., 2009).

Once an analytical study has been designed to test the hypothesis and the data have been gathered, the data need to be analyzed to determine whether there is an association between the exposure and the outcome being studied. Some of the most important measures of association used in these analyses are absolute measures of comparison (e.g., rate or risk difference/excess risk, attributable proportion/attributable fraction) that are based on the differences between two measures of disease frequency and relative measures of comparison (e.g., relative risk, odds ratio, attributable risk) that are based on the ratio of two measures of disease frequency.

Both absolute and relative measures contribute to the understanding of the effect of an exposure on disease occurrence. The advantage of absolute measures obtained by

subtracting one risk or rate from another is that they provide direct information about the public health impact of a particular exposure. A large difference generally indicates an important problem, regardless of the size of the baseline rate (Brownson et al., 1993). The advantage of relative measures obtained by comparing/dividing one risk or rate by another is that they generally give information about the strength of the association between exposure and disease (e.g., smokers are 10 times more likely than nonsmokers to develop lung cancer) and are most useful for etiologic research (Aschengrau & Seage, 2008). However, a doubling of the incidence or mortality rate may not indicate an important public health problem if the baseline rate is very low.

3.3.1 Absolute Measures

Rate or Risk Difference/Excess Risk

The rate or risk difference, also called the excess risk, is an absolute measure of comparison based on the difference between two measures of disease frequency. Typically, absolute calculations are applied to either exposed individuals or total populations. For example, health professionals would use the population risk difference to understand the overall health impact of a certain exposure on the health of a specific population (Aschengrau & Seage, 2008).

The rate or risk difference is expressed as follows:

$$RD = R_e - R_{u,}$$

where **RD** is the rate or risk difference, $\mathbf{R_e}$ is the rate or risk in the exposed population, and $\mathbf{R_u}$ is the rate or risk in the unexposed population (Aschengrau & Seage, 2008).

Excess risks in the form of rate differences for adenocarcinoma of the lung among women associated with different levels of pack-years of smoking are shown in the "Relative Risk, Excess Risk, and Attributable Risk Example" section (see Table 3-17 below). For example, the table shows that, compared with those who never smoked, 20 to 29 pack-years of smoking was associated with an excess risk of 123 cases of lung cancer per 100,000 population per year.

Attributable Proportion or Attributable Fraction

The attributable proportion in the total population (AP_t) describes the proportion of disease among the total population that would be eliminated if the exposure were eliminated. It is used when an exposure is considered a cause of the disease. The AP_t is very useful for determining priorities for disease prevention and general public health action. Policy makers are more likely to address health issues that have a high attributable proportion, because the impact on health and disease would be greater for potentially the same amount of resources.

The attributable proportion (**AP_t**) is expressed as follows:

$$AP_t = [(R_t - R_u)/R_t] \times 100$$

where **R_t** is the incidence rate or prevalence in the total population, and **R_u** is the incidence rate or prevalence in the unexposed population. For example, the proportion of disease (adenocarcinoma of the lung) among the total population (women in Iowa) that would be eliminated if the exposure (smoking) were eliminated can be estimated from Table 3-17 (presented below in the "Relative Risk, Excess Risk, and Attributable Risk Example" section) by calculating the incidence rate of disease in the total population, 124/100,000, as the weighted average of the incidence rate in each smoking category (never and 1–19, 20–39, 40–59, and ≥60 pack-years of smoking). The attributable proportion, calculated using the incidence rate in the total population (124/1000) and the incidence rate in the unexposed never smokers (19/100,000), would be 124–19/124 x 100 or 84.7. Based on these data, 84.7% of the adenocarcinomas of the lung among Iowa women would be eliminated if smoking were eliminated.

3.3.2 Relative Measures

Relative Risk

The relative risk (RR) is the measure of association used for cohort studies and cross-sectional studies, because it compares the health outcome among different exposure groups. It can be expressed as the risk of disease in the exposed divided by the risk of disease in the unexposed, a usage synonymous with the risk ratio; or as the ratio of the incidence or mortality rate in the exposed group divided by the incidence or mortality rate in the exposed group divided by the incidence or mortality rate in the unexposed group, a usage synonymous with the rate ratio. Relative risk is calculated as follows:

$$RR = \frac{a/a+b}{c/c+d}$$

where **RR** is the relative risk, **a/a+b** is the rate or risk among the exposed, and **c/c+d** is the rate or risk among the unexposed as depicted in Table 3-13.

Table 3-13. Symbolic Representation of Subjects With and Without a Disease or Outcome by Exposure Status

	Disease or Outcome		
Exposure	**Yes**	**No**	**Total**
Yes	a	b	a+b
No	c	d	c+d

The example in Table 3-14 uses data from a study of air pollution and mortality in six U.S. cities (Dockery et al., 1993); the exposed group is residents who lived in a city with high levels of air pollution from particulates and sulfur dioxide (Steubenville, OH), and the unexposed group is residents who lived in a city with low levels of these pollutants (Topeka, KA). The outcome measure is death. The risk of death in the exposed group is 291/1,351 = 0.215 (or 215.4/1000); the risk in the unexposed group is 156/1,239 = 0.126 (or 125.9/1000). The resulting crude RR is 0.215/0.125 = 1.72. Thus, there is a 1.72 increased risk of death among Steubenville residents compared with Topeka residents.

Table 3-14. Relative Risk Example: Risk of Death in Two Cities with High and Low Levels of Pollution

	Dead		
Exposure	**Yes**	**No**	**Total**
Lived in city with high pollution (Steubenville, Ohio)	291	1,060	1,351
Lived in city with low pollution (Topeka, Kansas)	156	1,083	1,239
Total	447	2,143	2,590

Source: Dockery et al. (1993).

Odds Ratio

In a case-control study, the risk of disease cannot be directly calculated because the population at risk is not known. Instead, epidemiologists calculate a number called an odds ratio that functions as a rate or risk in a case-control study. In a case-control study, we can calculate either the disease odds ratio (the ratio of the odds of being a case among the exposed [**a / b**] divided by the odds of being a case among the unexposed [**c / d**]) or the exposure odds ratio (the ratio of the odds of being exposed among the cases [**a / c**] divided by the odds of being exposed among the controls [**b / d**]). These odds ratios are algebraically equivalent and reduce to the *cross-product ratio* **ad / bc**. Thus, the odds ratio can be calculated as follows:

$$OR = \frac{ad}{bc}$$

where **OR** is the odd ratio, and **ad / bc** is the cross-product ratio as depicted in Table 3-15.

Table 3-15. Symbolic Representation of Cases and Controls With and Without an Exposure

	Disease Status	
Exposure	**Cases**	**Controls**
Yes	A	B
No	C	D

In the example in Table 3-16, from a population-based, case-control study of oral cancer among Puerto Rican men (Huang et al., 2003), the exposure is liquor use and the health outcome is oral cancer. The odds ratio is calculated as 258 x 147 / 22 x 257 = 6.7. Thus, men who drank liquor had more than 6 times the risks of having oral cancer compared with men who did not drink liquor.

Table 3-16. Odds Ratio Example: Oral Cancer Risk Associated with Consumption of Liquor in Men, Puerto Rico, 1992–1995

	Oral Cancer		
Liquor Use	**Yes (Case)**	**No (Control)**	**Odds Ratio**
Yes	258	257	6.7
No	22	147	1.0

Source: Huang et al. (2003).

Interpretation of the Numeric Value of the Risk Ratio or Odds Ratio

- If the numeric value = 1.0, there is no association between the exposure and the health outcome; people who were exposed are no more or less likely to develop the health outcome than the unexposed.
- If the numeric value is >1.0, there is a positive association between the exposure and the health outcome (i.e., the exposure increases the risk of disease). For example, if the RR/OR = 5, people who were exposed are 5 times more likely to develop the health outcome than those who are unexposed.
- If the numeric value is <1.0, there is a negative association between the exposure and the health outcome (i.e., the exposure decreases the risk of disease or has a protective effect). For example, if the RR/OR = 0.5, people who were exposed are half as likely to develop the health outcome as those who are unexposed.

Attributable Risk

The attributable risk is the maximum proportion of a disease that can be attributed to an etiologic factor. It is used when the etiologic factor is thought to be causally related to the disease. For example, in the study of adenocarcinoma of the lung within the Iowa Women's

Health Study, 60% of the cancers are attributed to smoking based on the relative risk for ever smokers compared with never smokers. The attributable risk is expressed as follows:

$$AR = \frac{b(r-1)}{b(r-1) + 1}$$

where **AR** is the attributable risk, **r** is the relative risk of developing a disease in an exposed population versus an unexposed population, and **b** is the proportion of the total population classified as exposed.

When the frequency of an etiologic factor in a population is low and the relative risk for resulting disease is also low (e.g., alcohol use in women and breast cancer), only a small proportion of the disease can be attributed to the factor. Conversely, with a high relative risk and high proportion of the population exposed to the factor (e.g., smoking and the risk of lung cancer in men), a much larger percentage of the disease can be attributed to the factor (Aschengrau & Seage, 2008).

Relative Risk, Excess Risk, and Attributable Risk Example

Table 3-17 is an example from the Iowa Woman's Health Study that illustrates the relationship between incidence rates, relative risks (incidence risk ratios), excess risks (incidence risk differences), and population attributable risks in a population of white postmenopausal women. The excess risk or risk difference provides an absolute measure of exposure effect on lung cancer risk among smokers compared with never smokers and shows that cigarette smoking is a very important risk factor for adenocarcinoma in postmenopausal white women. For example, the excess risk among women with 20 to 39 pack-years of smoking equals the incidence rate for 20 to 39 pack-years minus the incidence rate for never smokers (142 − 19 = 123). The relative risk and population attributable risk are relative measures and are more strongly driven by the background risk (i.e., the incidence rate among never smokers) than the excess risk. For example, the relative risk of lung cancer among women with 20 to 39 pack-years of smoking was 7.5 times greater than among nonsmokers (142/19). The population attributable risk estimates the potential public health significance (i.e., the percentage of decrease in the incidence of adenocarcinoma of the lung in this population [60%] that would result if active cigarette smoking were completely eliminated).

Table 3-17. Incidence Rates, Relative Risks, Excess Risks, Population Attributable Risk, and Corresponding 95% Confidence Intervals of Adenocarcinoma of the Lung within the Iowa Women's Health Study, 1986–1998

	Pack-Years of Smoking				
Risk Measure	**Never**	**1–19**	**20–39**	**40–59**	**≥60**
	Value (95% CI)				
Incidence rate[a]	19(15–25)	46 (31–67)	142 (112–181)	181 (136–242)	293 (202–424)
Relative risk[b]	1.0	2.4 (1.5–3.8)	7.5 (5.3–10.7)	9.4 (6.4–13.8)	15.5 (9.8–24.3)
Excess risk[b]	ref	27 (9–45)	123 (89–157)	162 (109–215)	273 (165–381)
Population attributable risk percent[c]			60 (43–68)		

[a]Age adjusted incidence rates (per 100,00).

[b]Age adjusted relative risks.

[c]Multivariate adjusted population attributable risk per 100,000 person-years.

Source: Yang, Cerhan, & Vierkant (2002).

3.3.3 Tests of Statistical Significance

The purpose of an analytic epidemiology study is to estimate the true relationship, or association, between an exposure and a health outcome. The ability to obtain a correct estimate of the association between exposure and outcome may be lessened by certain threats to internal validity, including confounding bias, other types of bias, and the role of chance. Confounding bias is a confusion of effects, in which an apparent association between an exposure and an outcome is actually due to a third variable that is associated with both the exposure and the outcome. The threat to internal validity due to known confounding variables can be avoided to some extent in the design of an epidemiology study, and the effect of known confounders can be adjusted for in the analysis of the study. The design, conduct, and analysis of epidemiology studies to minimize confounding and other types of bias is a major goal in epidemiology methods and is beyond the scope of this report.

The role of chance in the estimation of the association between an exposure and an outcome is addressed by testing for statistical significance. A test of significance is an indication of the reliability of the association between exposure and the observed health outcome, or the likelihood that the observed result might have occurred by chance, in the absence of a true association. There are two main ways of expressing the results of tests of statistical significance: P values and confidence intervals.

P Value

The P value is defined as the probability of obtaining the observed result and more extreme results by chance alone, given that the null hypothesis is true (e.g., that there is no relationship between exposure and disease; the RR/OR = 1.0 or the rate difference = 0). The compatibility of the study data with the null hypothesis is evaluated using a statistical test, such as the chi-square test. The P value is a continuous statistic ranging from 0.0 to 1.0, but scientists commonly use a cutoff point of 0.05 to determine whether to reject the null hypothesis. Results are considered statistically significant when the P value is ≤0.05. A major criticism of reliance on P values is that they are sometimes considered in isolation from the point estimate of the magnitude of the association, such as the relative risk or odds ratio. The P value alone provides no information about the strength of the association; it indicates only the role of chance, given the variability of the association. The variability of an association is generally lower when a sample size is large. From the P value alone, one cannot determine whether a P value is small because there is little variability in the association or because the sample size is very large (Aschengrau & Seage, 2008).

Confidence Interval

Most epidemiologists use confidence intervals rather than P values to describe variability and the role of change. A 95% confidence interval (CI) is the range of values that has a 95% chance of containing the true point estimate of the association (Boslaugh, 2008), a value that would be obtained if correct and complete information was obtained on all members of the population being studied, including the complete elimination of all confounding bias and other bias. Interpreting the 95% CI allows epidemiologists to evaluate the point estimate of the strength of the association in the context of the variability in the estimate. The 95% CI can also be used to test for significance and determine if the null hypothesis is rejected. The method for testing significance with the 95% CI for a relative measure such as a relative risk or an odds ratio is as follows:

- To have a statistically significant association between exposure and outcome, the 95% CI should not include 1.0.
- A 95% CI range below 1.0 suggests a statistically significantly lower risk of the outcome in the exposed population.
- A 95% CI range above 1.0 suggests a statistically significantly higher risk of the outcome in the exposed population.

Many epidemiologists prefer confidence intervals to P values because they communicate both the magnitude and the variability of the estimated association (Aschengrau & Seage, 2008).

Example of 95% CI and P Value

The study in Table 3-18 examined the association between exposures to various sexually transmitted diseases and the health outcome of tubal infertility. The ORs for gonorrhea and

Table 3-18. Relation of Tubal Infertility to History of Sexually Transmitted Diseases

Disease	Odds Ratio	95% CI
Gonnorrhea	2.4*	1.3–4.4
Trichomonas	1.9*	1.3–2.8
Yeast	1.3	1.0–1.7
Other vaginitis	1.7	1.0–2.7
Herpes	0.9	0.5–1.8
Genital warts	0.4	0.2–1.0

* $P < 0.05$.

Source: Grodstein, Goldman, & Cramer (1993).

trichomonas each were significantly elevated, the ORs for yeast and other vaginitis were elevated but of borderline significance (the CI included 1.0), the OR for herpes was nonsignificantly reduced, and the OR for genital warts was reduced but of borderline significance. P values indicating significance at $p \leq 0.05$ from a hypothetical chi-square test have been added to the table for illustrative purposes. Findings with significant P values are designated by the symbol "*".

4. ECONOMIC BURDEN OF ILLNESS

An important component of the economic burden of illness, and one that policy makers often consider when highlighting the need for disease prevention services, is health care spending. The Centers for Medicare & Medicaid Services (CMS) provides annual calculations of national health expenditures by payer, by type of expenditure, and relative to gross domestic product (GDP). Per capita national spending on health care rose from $148 in 1960 (current dollars) to almost $7,700 in 2008 (current dollars) (CMS, 2010).[2] The government's share of spending on health care has increased from 24.5% in 1960 to 47.3% in 2008, and although health care spending represented a mere 5.2% of the U.S. GDP in 1960, by 2008 health care spending accounted for 16.2% of GDP. Examples of the national health expenditure statistics published annually by CMS are provided in Attachments 2 and 3.

Although health care spending represents a large and growing share of the full economic burden of illness, other important components of the economic burden of illness are nonmedical spending, work loss resulting from increased morbidity or early mortality, and the impact on family members' employment or patients' psychological well-being. Estimates of the economic burden of illness attempt to capture the full economic costs of a disease, where the economic costs reflect the "opportunity costs" of having a disease. Opportunity costs are the value of health and non-health outcomes that patients and their families and friends are unable to enjoy as a result of the disease. For example, an opportunity cost of multiple sclerosis (MS) is the value of lost productivity for the patient and family members who switch jobs or quit working as a result of the illness. Opportunity costs differ from accounting, or financial, costs in that opportunity costs value even those costs or losses for which no monetary cost is incurred, such as productivity losses.

Most economic burden of illness measures are valued by applying cost-of-illness methods. In their 1982 paper, Hodgson and Meiners discuss recommendations of a 1978 Public Health Services task force on how to perform cost-of-illness studies to ensure that future studies use consistent approaches, thereby improving the comparability of studies. Hodgson and Meiners (1982) note that "the approach most frequently used by analysts to establish values for illness, disease, and health care services and programs is to identify the cost-generating components and to attribute a monetary value to them" (p. 431). They recommend including both direct and indirect costs in cost-of-illness analyses. Direct costs consist of medical and nonmedical spending to diagnose, treat, manage, and live with an illness (e.g., doctor visits, transportation costs, family spending for household help); and indirect costs include productivity losses that arise when people are unable to work because of increased morbidity or early death and psychosocial costs, such as the costs of financial strain or uncertainty over a person's future health and well-being. However, after describing

[2]These current dollar figures do not account for the impact of inflation on health care spending.

all possible costs of illness, Hodgson and Meiners acknowledge the limitations of attempting to include all costs associated with an illness and provide recommendations for cost-of-illness studies that use the human capital approach. Key among these recommendations are that cost-of-illness researchers (1) specify the costs to be included in a study; (2) clearly describe the methods and data; (3) use a range of discount rates from 2.6% to 10% to discount foregone future benefits (e.g., mortality-related productivity losses); (4) avoid double-counting of costs by excluding transfer payments, for example; (5) include nonmedical and psychosocial costs whenever relevant; and (6) conduct sensitivity analyses to examine the impact on disease cost estimates of uncertainty in key parameter values.

Since the Hodgson and Meiners cost-of-illness recommendations were published in 1982, hundreds of cost-of-illness analyses have been conducted to characterize the full or partial economic burden of specific diseases. The National Institutes of Health (NIH) has compiled a comprehensive set of estimates of direct and indirect costs of almost all of the diseases for which NIH conducts and supports research (Kirschstein, 2000). The first set of these estimates was provided in September 1995, in response to a request from the Senate Committee on Appropriations, and Committee members were especially interested in costs for the top 15 causes of mortality as identified by the Centers for Disease Control and Prevention (CDC). In response to the initial and subsequent requests, NIH has developed tables and reports that show disease-specific estimates of the direct and indirect costs of illness alongside the level of NIH support for each disease. The 2000 table of diseases and costs was recently updated by RTI to include additional estimates of disease and risk factor costs that have been published since 2000. This table (Attachment 4) provides direct and indirect cost estimates for almost 75 diseases or disease risk factors.

In the past 15 years, a number of texts have been published that offer limited guidance for estimating the economic burden of illness (see, e.g., Gold et al., 1996; Haddix, Teutsch, & Corso, 2003; Drummond & McGuire, 2001). These texts primarily provide approaches for performing full economic evaluations of health care or preventive health interventions and not just assessments of the cost of an illness or disease. For example, Gold et al. (1996) provide recommendations and specific methods to perform cost-effectiveness analyses that conform to the recommendations of the U.S. Panel on Cost-Effectiveness in Health and Medicine. Haddix, Teutsch, and Corso (2003) provide similar advice and examples for estimating the cost-effectiveness or cost-benefit of public health policies and practices. Yet, because evaluating the cost-effectiveness or cost-benefit of a health care or public health intervention necessarily involves valuing the impact of the intervention on health outcomes, these texts describe approaches for identifying which cost categories to include in disease cost analyses and how to estimate those costs.

Since 2000, the need to provide guidance on a consistent set of methods for economic burden of illness estimation has once again emerged. Several studies have shown that cost-of-illness estimates for a given disease vary widely, even when the same data are used

(Akobundu et al., 2006; Honeycutt et al., 2009). Moreover, estimated medical costs for any given disease are not linked to national health care spending and may exceed aggregate health spending when summed across diseases (Kirschstein, 2000; Trogdon et al., 2008). To help establish guidelines for estimating health care costs of disease and to identify specific areas for future research on disease costing approaches, disease costing researchers met in December 2007 to discuss approaches for standardizing disease and intervention cost estimation approaches. Their recommendations for standardizing disease cost estimation and future research needs are provided in the July 2009 supplement to *Medical Care*, "Health Care Costing: Data, Methods, Future Directions."

In this section, we briefly describe the literature on five economic burden of illness measures: direct medical spending, direct nonmedical spending, indirect costs resulting from excess morbidity, indirect costs resulting from early mortality, and broader disease burden estimates (i.e., willingness-to-pay) that capture the psychosocial costs of illness in addition to direct costs and productivity losses. Table 4-1 summarizes economic burden of illness measures, estimation approaches, and key references related to economic burden measurement.

4.1 Direct Medical Spending

The direct medical costs attributable to a disease or risk factor include all of the costs to prevent, diagnose, treat, and manage a disease, including inpatient care; outpatient care; physical and occupational therapy; emergency department services; prescription and non-prescription drugs; and medical supplies and devices, such as hearing aids and syringes (Haddix, Corso, & Gorsky, 2003). Methods for estimating the medical expenditures attributable to a disease vary widely and are often selected based on the data available for analysis (Akobundu et al., 2006). The approaches also vary in the ways they allocate disease costs to individuals. Some studies use simple algorithms to allocate medical spending across primary and secondary diagnoses using health insurance data. Others use survey data on health care utilization and allocate costs based on survey responses indicating which costs were for a specific disease or compare costs for people with and without the disease, controlling for other observable characteristics (i.e., bottom-up approaches). Others use aggregate health care spending and allocate spending to any particular disease using epidemiologic information on risks of developing the disease (i.e., top-down approaches). In this subsection, we briefly describe each cost estimation approach and provide examples from the literature.

Table 4-1. Economic Burden of Illness Measures

Economic Burden Construct	Measure Description	Estimation Approach	Challenges	Key References
Direct medical spending	Value of medical goods and services to prevent, treat, diagnose, or manage disease	▪ Health care cost modeling: compare costs for people with and without the disease; econometric models account for high proportion with no health care spending and a few with high spending ▪ Typically uses secondary data sources ▪ Estimate spending on a disease from various perspectives: – payer – patient – government – societal	▪ Spending attributable to a given disease can vary widely depending on methods used ▪ Sum of spending across diseases may lead to implausibly high aggregate cost estimates (i.e., double counting of disease health care costs) ▪ Data limitations: costs often based on charges or reimbursements in health care claims	Haddix, Corso, & Gorsky (2003) Luce et al. (1996) Brouwer, Rutten, & Koopmanschap (2001) Manning (1998) Manning & Mullahy (2001) Buntin & Zaslavsky (2004) Gileskie & Mroz (2004) Diehr et al. (1999)
Direct nonmedical spending	Value of nonmedical goods and services to treat or manage disease	▪ Calculate costs for patient transportation, home and auto modifications, special education, etc. ▪ Include value of informal caregiving provided by family or friends ▪ Typically uses secondary data sources	▪ Requires data on nonmedical impacts of disease from patients and/or family members	Haddix, Corso, & Gorsky (2003) Luce et al. (1996) Brouwer, Rutten, & Koopmanschap (2001)
Indirect costs resulting from excess morbidity	Economic losses resulting from living with illness or disease	▪ Human capital or friction cost methods to estimate productivity losses attributable to disease. ▪ Measured as costs of – absenteeism, – presenteeism, – early retirement, – changing jobs	▪ Human capital approaches value potential economic losses; do not assign a value to morbidity-related losses in children or older adults ▪ Willingness-to-pay approaches are difficult to apply when both mortality and morbidity effects are present	Sculpher (2001) Luce et al. (1996) Koopmanschap et al. (1995) Johnson et al. (1997)

(continued)

Table 4-1. Economic Burden of Illness Measures (continued)

Economic Burden Construct	Measure Description	Estimation Approach	Challenges	Key References
Indirect costs resulting from excess morbidity (continued)		▪ Typically uses secondary data sources ▪ Friction cost method attempts to capture productivity losses only until a job replacement is found ▪ Willingness to pay estimation of value of preventing short-term health conditions		
Indirect costs resulting from early mortality	Value of losses resulting from early death caused by a disease	▪ Value of a statistical life (VSL); captures non-pecuniary losses in addition to economic impacts of death ▪ Human capital or friction cost approach to calculate present value of the stream of future productivity losses resulting from early death ▪ Typically uses published data or secondary data sources	▪ VSL estimates vary widely; age, health status, and income-specific measures of VSL are not widely accepted ▪ Human capital approach assigns no value to deaths in older adults and low value to deaths in young children	Sculpher (2001) Luce et al. (1996) Freeman (2003) Koopmanschap et al. (1995) Viscusi (2003)
Comprehensive monetary estimate of disease burden	Valuation of health benefits of preventing or reducing the incidence/ prevalence of disease	▪ Willingness to pay to avoid a case of disease (for those without the disease) or to cure an illness (for those with the disease) ▪ Primary data collection usually required	▪ Often requires extensive new data collection to capture people's valuations of different disease states	Freeman (1993) Johnson et al. (1997) Portney (1994)

Studies that use data from health insurance claims make use of available information, which typically includes data on charges for individuals' health care services, diagnoses, and health care procedures, but very limited information on patient demographics or the presence of other health conditions (e.g., those conditions that are present but not coded as the diagnosis for which a visit was made). In some cases, actual reimbursement amounts

may also be available, which are likely to differ considerably from charges, given negotiated rates for health care services (Segel, 2006). Studies that use health care claims data may assign all health care charges to the primary diagnosis listed for a procedure or may assign costs to all listed diagnoses based on some algorithm (e.g., 60% to primary diagnosis and the remaining 40% split proportionately across secondary diagnoses [Ward et al., 2000]). Ward et al. (2000) showed that cost estimates for chronic obstructive pulmonary disease (COPD) that used only primary diagnoses of COPD were one-third the value of cost estimates that used primary and secondary diagnoses of COPD. One example is the study by Van Houtven et al. (2008) that used Medicare claims data to estimate the costs of six health conditions in older adults with environmental risk factors by including medical costs based on the primary diagnosis listed in hospital, skilled nursing, and other health care services claims reimbursable by Medicare fee-for-service health insurance.

When nationally representative survey data are available that include information on individuals' demographic characteristics and their medical spending, such as in the National Hospital Discharge Survey, the National Ambulatory Care Survey, or the Medical Expenditure Panel Survey (MEPS), total cost estimates may be estimated by multiplying the utilization of a medical good or service by unit cost (Bloom et al., 2001). For example, Rice and Miller (1998) used this type of bottom-up cost estimation approach to estimate the cost of mental disorders in the United States. Another example used patient surveys from people with MS to estimate average health care utilization among people with MS, then multiplied by unit costs for each of the health care services used to estimate the average per-patient cost of each service (Kobelt et al., 2006). The total medical costs of MS in the United States were calculated by summing across all health care services typically used by MS patients and then multiplying by MS prevalence in the United States (Kobelt et al., 2006).

Regression approaches are also frequently used to estimate the level of health care spending attributable to a disease by comparing spending for people with the disease to spending for people with similar demographic conditions who do not have the disease (Lipscomb et al., 2009). An advantage of using econometric models for health care spending is that they attempt to eliminate potential bias in estimates of disease costs by directly modeling features, such as a large number of people with no annual health care spending and a small number with very high health care spending, that may bias mean cost estimates for people with and without the disease (Manning & Mullahy, 2001).

A challenge of applying these regression-based approaches is that no single best model exists that can be applied in all cases. Rather, analysts must first examine features of the health care spending data to determine which modeling approach would provide the best fit (Manning & Mullahy, 2001; Buntin & Zaslavsky, 2004). Another challenge is that these models may lead to wide confidence intervals, suggesting a broad range for estimated health care costs attributable to any given disease (Manning & Mullahy, 2001). Yelin et al. (2001) used a regression-based approach and MEPS data to estimate the costs attributable

to musculoskeletal conditions and found that, although average medical spending was $3,578 per year among people with musculoskeletal conditions, the medical costs attributable to these conditions were $364 to $723, depending on econometric modeling approach.

Another approach that is sometimes used to estimate medical spending attributable to a disease is the top-down approach, which uses aggregate data on health care spending and allocates costs to specific diseases based on the likelihood of developing that disease in the population. Liu et al. (2002) used a top-down approach to estimate the annual costs of coronary heart disease in the United Kingdom; Hodgson and Cohen (1999) combined the top-down approach with a regression-based approach to estimate the cost of diabetes and its complications in the United States.

4.2 Direct Nonmedical Spending

The direct nonmedical costs attributable to a disease or risk factor include transportation costs to obtain medical services, the cost of home and automobile modifications, expenses for professional caregiving or housekeeping services, and the value of informal caregiving (i.e., care provided by family and friends at no charge to the patient) (Haddix, Corso, & Gorsky, 2003; Hodgson & Meiners, 1982). Informal care costs may represent a very large component of the overall costs of diseases such as Alzheimer's disease or Parkinson's disease, which disproportionately affect older adults and for which family members often provide care. However, a challenge of estimating informal care is that typically no payment is made for the care provided by family members or friends. McDaid (2001) describes the challenges of estimating informal care costs for people with Alzheimer's disease.

Estimating nonmedical costs generally requires the availability of survey data collected from patients on their utilization of nonmedical services. Using the survey data, researchers typically calculate the mean utilization of each nonmedical good or service and then apply a unit cost to each. Because nonmedical goods and services are often paid for by the patients themselves or their families, these cost estimates tend to reflect out-of-pocket spending to manage the disease. Examples of nonmedical costs included in a recent analysis of the annual costs of MS in the United States are wheelchairs, scooters, housing ramps, auto lifts for wheelchairs, and informal caregiving provided by family and friends (Kobelt et al., 2006).

4.3 Indirect Costs Resulting from Excess Morbidity

Indirect costs resulting from excess morbidity consist of the resources lost when individuals are sick and cannot perform their usual day-to-day activities or when the illness leads to additional costs to society, such as the impact of substance abuse on crime. Morbidity-related indirect costs often include productivity losses due to absenteeism (missing work because of illness), presenteeism (reduced on-the-job productivity because of illness), early

retirement, lost household productivity, and/or changing to a lower-paying job because of illness (Hodgson & Meiners, 1982; Haddix, Corso, & Gorsky, 2003). Two main approaches exist for valuing these losses: the human capital approach and the friction cost approach. Willingness to pay (WTP) for reduced morbidity may also be used to measure these indirect costs, but because the WTP estimation approaches are the same, regardless of whether the interest is in morbidity-related or mortality-related costs, we describe WTP estimation in Section 4.5.

The human capital approach values the lost production of a patient (or caregiver) by applying earnings estimates to time lost from productive activity. Absenteeism, or days lost from work because of an illness, is generally valued by applying estimated mean earnings by age group and gender (Sculpher, 2001). Presenteeism may be difficult to value because of the difficulty of observing and measuring reduced output in employees with an illness. Pauly et al. (2008) surveyed managers to estimate lost output due to presenteeism and showed that lost output varies considerably across occupations by job characteristics, from 12.5% to 75% of wages. Early retirement cost estimates require information from patients on the percentage who retired early because of an illness. These estimates are then applied to mean earnings estimates to calculate the present value of all future earnings lost as a result of early retirement (Kobelt et al., 2006). Similarly, the impact of job changes may be estimated using data from patients on the likelihood of shifting to a lower-paying job as the result of illness and the subsequent reduction in pay. These estimates may then be used to calculate the present value of earnings losses resulting from job changes. To estimate productivity losses for people who work outside the labor market (e.g., homemakers), household productivity cost estimates should be used (Grosse, 2009). These cost estimates are based on time-diary studies of the U.S. population.

The human capital approach has been criticized because it assigns zero value to morbidity-related losses in people who are retired and very low values to infants and young children because the productivity losses occur far into the future and are discounted. An annual discount rate of 3% to 5% is typically applied.

The friction cost method differs from the human capital approach in that it estimates productivity losses only for the friction period—the period during which a replacement employee must be identified, hired, and trained to perform the job (Koopmanschap et al., 1995; Koopmanschap & van Ineveld, 1992). According to this approach, once the friction period has passed, no additional productivity losses are incurred. However, a challenge of applying the friction cost method is that the "right" friction period may be difficult to estimate, because it varies over time and across geographic areas and depends on the unemployment rate. Morbidity-related indirect cost estimates vary widely, depending on whether the human capital or friction cost approach is used. For example, Goeree et al. (1999) found that the productivity losses of schizophrenia using the human capital approach were almost 70 times higher than estimates that used the friction cost method.

4.4 Indirect Costs Resulting from Early Mortality

Indirect costs resulting from excess mortality capture the resources lost when individuals lose expected life-years as the result of disease(s). Cost-of-illness analyses typically use a human capital approach to include the productivity losses associated with life-years lost (Hodgson & Meiners, 1982). The human capital approach estimates the productivity losses associated with life-years lost as the present value of the future stream of earnings lost as a result of death from a disease before the age of average life expectancy is reached, where annual discount rates from zero to 10% are applied to all future years of earnings losses. Friction cost methods may also be used to estimate productivity losses associated with the years of life lost from a disease. Some have criticized the use of the human capital and friction cost approaches on grounds that human life should not be valued based on productivity losses (Landefeld & Seskin, 1982).

An alternative measure to valuing loss of life is the WTP approach, which values small changes in the probability of death or survival based on individual preferences (Freeman, 1993). WTP has most often been estimated from labor market data that compare wages in occupations with a higher risk of death to wages in occupations with a low risk of death (Freeman, 1993). The resulting estimates of the value of a statistical life (VSL) vary widely, from less than $100,000 to more than $25 million (Mrozek & Taylor, 2002; U.S. EPA, 2000). However, the median VSL is about $7 million based on studies that compare wages to risks of death across occupations (Viscusi, 2008). Mrozek & Taylor (2002) have estimated a lower range of VSL in a meta-analysis that used 33 labor-market VSL studies published between 1974 and 1999. They found a VSL mid-range of $1.5 to $2.5 million when assuming the use of best practices in VSL models (Mrozek & Taylor, 2002). Aldy and Viscusi (2007) have also shown that VSL differs by age, with VSL rising at first and then falling with age.

Rice and Hodgson (1982) point out that human capital and WTP methods "are conceptually different, serve different purposes, and measure different aspects of threats to health" (p. 536). The human capital approach provides a consistent approach for valuing future earnings losses due to early death, but it does not capture the value of pain and suffering or loss of leisure time, which are included in WTP estimates (Rice & Hodgson, 1982). However, WTP approaches can be difficult to implement in practice, often requiring the collection of data on individuals' preferences (stated or observed) to estimate. In addition, WTP captures the value of all benefits of policy or intervention; it is therefore difficult to separate the value of reduced probability of death from other benefits, such as improved health outcomes (Rice & Hodgson, 1982).

4.5 Willingness to Pay

The WTP approach is an alternative to the cost-of-illness approach that values the prevention of illness and death from disease as the sum of what people are willing to pay to reduce their own risks plus the sum of the additional amounts that people are willing to pay

to prevent illness and death in others (Freeman, 1993). The main approach for estimating WTP uses wage differentials between high- and low-risk occupations (Freeman, 1993). This approach is often referred to as a "revealed preference" approach because it uses behavioral observations and assumes that people's occupational choices reflect their preferences over risk and pay. Another approach for estimating WTP is to ask people directly about how much they value specific reductions in illness and death risks (Freeman, 1993). Such survey methods are known as the "stated preference" approach.

A key advantage of the WTP approach is that it can capture in a single measure all of the benefits of disease prevention, including the value of productivity losses, pain and suffering, and even out-of-pocket medical spending. In addition, if the disease impacts are limited to short-term impacts that do not include death, valuations of those impacts can also be performed using WTP (Johnson et al., 1997).

However, from a practical standpoint, it may be challenging to implement WTP approaches. For example, it is often difficult to identify economic situations, such as occupational choice, that easily lend themselves to the estimation of reduced injury and death to provide relevant estimates for the disease of interest. Furthermore, survey approaches, although they may be tailored to address the specific features of the disease of interest, are expensive to conduct, and problems with over- and under-estimation of WTP values have been widely noted (Portney, 1994). For example, Hirth et al. (2000) examined how WTP for a quality-adjusted life year (QALY) differed depending on the approach used to collect information about people's willingness to pay to extend or improve life. Using a revealed preference approach, comparing people who choose to use safety measures such as seatbelts and smoke detectors to those who do not led to an estimated median QALY value of $93,402 (Hirth et al., 2000). Using a contingent valuation approach, where people are asked directly about their willingness to pay to accept increased health or mortality risks, the estimated median QALY estimate was $161,305 (Hirth et al., 2000). The highest median value for a QALY of $428,286 was found using revealed preferences from higher wages for higher risk occupations (Hirth et al., 2000). All three of these WTP estimates exceeded the estimate of $24,777 per QALY found using a human capital approach (Hirth et al., 2000).

In general, despite the advantages of WTP approaches and their ability to capture the full range of costs of a disease, disease cost estimates are usually generated using a cost-of-illness approach that estimates direct costs using survey or health insurance data on health care utilization and values indirect costs using a human capital approach. The human capital approach is viewed as providing a lower bound for what people would be willing to pay to entirely avoid a case of a disease (Hirth et al., 2000).

4.6 Additional Economic Burden Estimation Issues

Other issues that need to be considered when using cost-of-illness approaches are what analysis perspective to use, what discount rate to apply to future costs and benefits, how to identify people with the disease, and whether to estimate costs for the prevalent or incident population (Segel, 2006). For example, costs estimated from the perspective of the health care system should include all medical costs but do not capture productivity losses or nonmedical costs (Luce et al., 1996). Regarding discounting issues, any costs or benefits that occur in the future should be discounted to reflect individual preferences for income today rather than in the future; the recommended discount rate for use in cost-effectiveness analyses is 3% (Gold et al., 1996). Different ways of defining an illness, such as using narrowly defined diagnosis codes versus broader diagnosis codes or including costs for diagnoses other than the primary diagnosis, may have substantial impacts on estimated medical costs of disease costs (Javitz et al., 2004; Ward et al., 2000). Finally, cost-of-illness studies are typically conducted for the prevalent population over some period of time, such as 1 year (Segel, 2006). If the interest is in lifetime costs for individuals with a disease, as is often the case when valuing the potential cost savings of disease prevention, an incidence-based approach, which requires information on disease impacts and survival rates over the lifetime, should be used (Hodgson, 1988).

4.7 Emerging Areas in Economic Burden Measurement

Recent efforts to improve the measurement of economic burden of illness have focused on improved approaches for estimating health care costs attributable to disease. In July 2009, *Medical Care* devoted an entire issue to disease costing methods and future directions. Lipscomb et al. (2009) summarize some of the innovative work that is leading to improved disease costing approaches and offer recommendations for future research. They describe the need for disease-specific National Health Expenditure Accounts (NHEAs) to enable comparisons of spending on any given disease to future health status (Rosen & Cutler, 2009). Current annual NHEAs show health care spending by payer and by service (Heffler, Nuccio, & Freeland, 2009). A more productivity-oriented view of health care spending would also provide estimates of spending by disease and by payer and/or service for a set of predefined diseases.

Because disease-based expenditures are sensitive to the method of allocation, the Bureau of Economic Analysis is currently exploring possible methods to provide more than one set of measures of spending by disease (Aizcorbe et al., 2008). Researchers have experimented with computer algorithms to sift through health claims data to allocate spending to more than 500 types of diseases (Aizcorbe & Nestoriak, 2007). This would allow researchers without medical expertise to easily apply the algorithm and obtain measures of expenditure. Rosen and Cutler (2007) conducted a study to compare how existing approaches allocate spending across diseases to infer the value of medical care at the disease level. These

approaches might provide a consistent linkage between the micro-cost estimates for specific diseases and the macro-cost estimates for aggregate national health care. They would also help to eliminate "adding up" problems of disease-specific cost estimates by ensuring that the sum of disease-specific health care spending does not exceed total annual health expenditures (as discussed in Trogdon et al., 2008).

Although Lipscomb et al. (2009) focus on methods for estimating the health care costs of disease or the costs of interventions, they also argue for the need to conduct studies to compare human capital and WTP approaches for valuing morbidity- and mortality-related losses. Of particular interest is the extent to which both human capital and WTP estimates differ across population subgroups, such as individuals with lower labor market participation rates. Because human capital estimates are based on labor market and household productivity estimates, individuals who do not work outside the home are necessarily assigned lower values. Similarly, WTP estimates are bounded by income because individuals cannot report that they would be willing to pay more than they are actually able to pay for a small reduction in risk. The extent to which income affects WTP estimates has not been fully explored. In addition, although a great deal of recent work has focused on developing statistical approaches for estimating the health care costs attributable to a disease or risk factor, particularly the estimation of generalized linear models, Basu and Manning (2009) recommend several additional areas of research on modeling health care cost modeling and the development of cost predictions.

5. QUALITY OF LIFE BURDEN OF ILLNESS MEASURES

Burden of illness measures are defined by the way in which burden is conceptualized and operationalized. Table 5-1 summarizes the most commonly used health status and quality of life measures. A key feature of these measures is that they involve some patient-reported assessment of health status. However, the measures differ in that they focus on different hypothetical burden-related health constructs. In this section, we examine the primary constructs evaluated and the ways in which they have been measured in the research literature.

5.1 Function Measures

One class of measures focuses on the construct of functional status. These function measures, which are also called generic health status profiles, provide a coordinated summary of health of an individual for several domains of health. This allows an overall picture of health to be obtained. The most popular function measure is the SF-36 and the briefer SF-12 (Ware, 2010). Both versions examine eight attributes: physical functioning, role limitations due to physical health problems, bodily pain, general health, vitality, social functioning, role limitations due to emotional problems, and mental health (Glasziou et al., 2007). Multi-item scales measure each domain. This multidimensional framework enables investigators to create a profile for individual patients. A profile may reveal that a patient scores highly on some domains, but lower on others.

Profiles, however, can often display a confusing array of conflicting scores. As a result, responses are frequently aggregated into the primary underlying domains of physical and mental function. In the case of the SF-12 or SF-36, responses are typically combined to produce a normed Physical Component Score (PCS) and a Mental Component Score (MCS). Functional status questionnaires are frequently classified as either generic or disease-specific. Generic measures, like the SF-36, contain broadly worded questions that are relevant to many different diseases and health states. Disease-specific instruments, on the other hand, introduce domains and item content that frequently pertain only to the identified disease. Examples of disease-specific questionnaires include the Paediatric Asthma Quality of Life Questionnaire for children with asthma (Osman & Silverman, 1996), the Minnesota Living with Heart Failure Questionnaire (Rector et al., 1987), and the McGill Quality of Life Questionnaire (MQOL) for people with life-threatening illness (Lua et al., 2005). It is generally believed that disease-specific instruments are more sensitive to functional effects than generic questionnaires. Disease-specific measures allow for increased focus on disease complications and thus can be responsive to a patient's condition (Drummond et al., 2005).

Table 5-1. Health Status/Quality of Life Burden of Illness Measures

Health Status/ Quality of Life Burden Construct	Measure(s)	Metric/ Scoring Algorithm	Time Frame	Disease-Specific?	Instruments	Literature Review Reference(s)	Example in Literature
Utility	Utility; preference for selected health state	0=death, 1="full" health Items weighted by societal preferences	Current health state	No	EQ-5D, SF-6D, Health Utilities Index II and III, Quality of Well-Being Scale (QWB), HALex, TTO, VAS	Craig et al. (2009), Ferreira et al. (2008), Fryback et al. (2009), Glasziou et al. (2007), Janssen et al. (2008), Knies et al. (2009), Krabbe et al. (1997), Luo et al. (2009), Rowen et al. (2009), Paz et al. (2009), Seymour et al. (2009), Whynes (2009)	The EQ-5D utility score is 0.597 for patients who have moderate problems on all five dimensions (Shaw et al., 2005).
Functional status	Physical, mental, or social function level	Continuous scores with higher levels indicating better function; often normed to general population with mean=50, SD=10	Usually status in past week, past 2 weeks, or past month	No	SF-36, SF-12, WHO-QOL, EORTC-QOL, FACT; EQ-5D, Health Assessment Questionnaire	Cheak-Zamora et al. (2009), Chuang et al. (2009)	The mean functional status of patients with three chronic conditions was 36.4 on the Physical Component Score (PCS) and 47.2 on the Mental Component Score (MCS) using the SF-12v2 (Cheak-Zamora et al., 2009).
Adjusted life years	Quality-adjusted life years (QALYs)	Years of full health; accumulated utility values over time	Interval between baseline and follow-up measure-ments	No	Derived from utility measures	Knapp et al. (2007), Raisch (2000)	Dividing the increased costs by the increase in utility, a program for screening smokers for lung cancer cost $42,500 per QALY gained (Mahadevia et al., 2003).
Time lost to disability	Disability-adjusted life years (DALYs)	Years lost to condition-related disability and premature death	Optimal life expec-tancy	Yes	None	Murray (1994), Hollinghurst et al. (2000)	The total global burden of road traffic accidents, measured in DALYs, is 35.06 million years (Mathers & Loncar, 2006).

(continued)

Table 5-1. Health Status/Quality of Life Burden of Illness Measures (continued)

Health Status/ Quality of Life Burden Construct	Measure(s)	Metric/ Scoring Algorithm	Time Frame	Disease-Specific?	Instruments	Literature Review Reference(s)	Example in Literature
Adjusted life expectancy	Health-adjusted life expectancy (HALE)	Years of life expectancy in full health	Lifetime	No	None	Gold et al. (2002), Zhang et al. (2008)	For U.S. men, the life expectancy from age 20 is 63.1 years but the HALE for smoking men is 55.4 years and 58.5 for obese men (van Baal et al., 2006).
Disutility	Decrease in utility due to specific disease	Same as utility (0=death, 1=full health); difference in mean utility between those with and without disease	Current status	Yes	Derived from utility measures	Franks et al. (2006), McKenzie & van der Pol (2008)	The disutility of diabetes mellitus without complications is −0.035 (Sullivan et al., 2005).

A related measure attempts to characterize the amount of time people are unhealthy. Developed by the Centers for Disease Control and Prevention (CDC) for use in the ongoing Behavioral Risk Factor Surveillance System, this measure consists of one item eliciting the number of days in the past month that a person's physical health was "not good" and a second item for mental health. Responses to the two function items are summed to provide an overall estimate of unhealthy days.

Function scales have been measured in many different ways by various instruments. A major federal initiative, the Patient-Reported Outcome Measurement Information System (PROMIS), is underway in an effort to standardize the process. PROMIS is developing an extensive catalog of health-related domains and a bank of items for each domain. Using Item Response Theory, the items are being calibrated for each domain so that domain scores can be measured on the same metric even when different sets of items have been administered in studies.

5.2 Overall Health Status Measures

Another general approach is to ask respondents to make a single global assessment of their health status. The most widely used of all health status measures is patient-rated health status: "In general, would you say your health is excellent, very good, good, fair, or poor?" Mean scores for this simple question (sometimes known as the EVGGFP) can be used to contrast patients or subgroups within the general population. Another method of eliciting global assessments is the visual analog scale (VAS). VASs are presented in the form of a bar ranging from the worst imaginable health state at one end of the bar to the best imaginable health state on the other end. Subjects place a mark on the bar to indicate their status, which is then converted into a numerical score based on the distance from the two endpoints.

Rather than simply cataloging chronic conditions, respondents may also be asked to rate the impact of their comorbidities. One method of measuring severity is to ask about the extent to which a condition interferes with usual activities of daily living. This interference of usual activities includes conditions that hamper the ability to perform routine tasks at home and work; the ability to perform normal self-care, such as eating, washing, and dressing; and the ability to participate in leisure activities (Bleichrodt & Johannesson, 1997). Another approach is to combine comorbidities into a global measure, such as mortality risk. This is exemplified by the Charlson Comorbidity Index (Charlson et al., 1994), which estimates the risk of dying in the next year as a function of multiple chronic conditions.

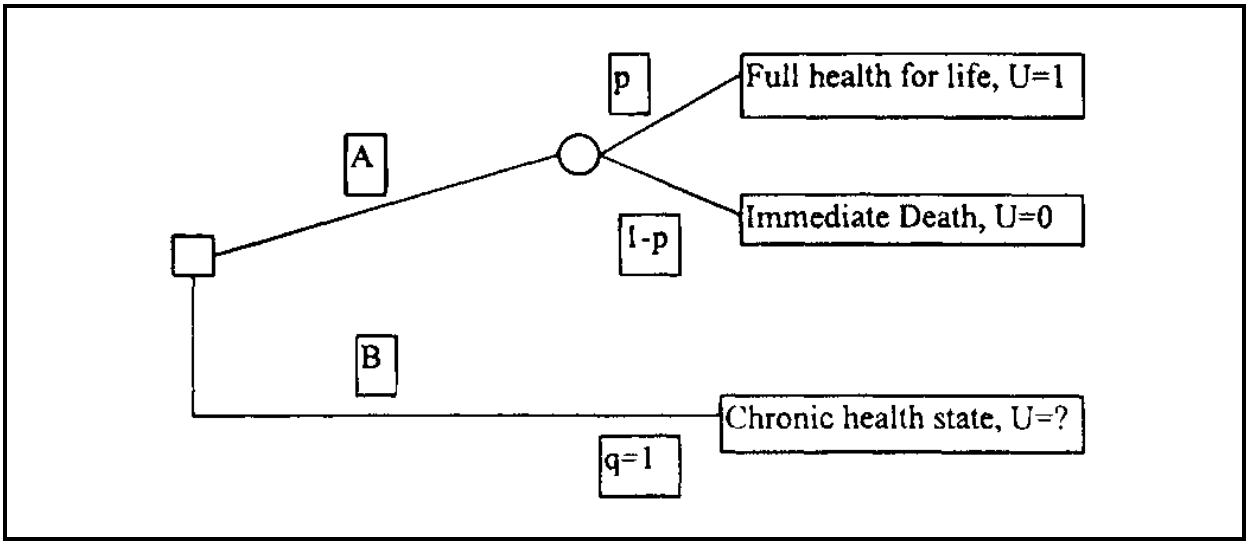
5.3 Utility

Utility is the value associated with a particular health state. Developed from the axioms of expected utility theory, all utility measures are based on a continuum ranging from 0 (representing death) to 1 (representing optimal health). However, in extreme cases such as

a permanent vegetative state, it is possible for utility to be less than 0, indicating a state worse than death. Individual health states are arrayed along this continuum based on comparative preferences. The health states refer to varying combinations of symptoms and functioning problems, such as limited mobility, severe pain, or moderate depression. In contrast to the multidimensional function measures, utility consolidates these disparate elements into a single value.

Utility values have been estimated in two stages. The first stage involves direct assessments made by a small group of subjects. The two primary valuation methodologies are the standard gamble and the time tradeoff approaches. In the standard gamble, subjects must choose between remaining in a designated health state and accepting a treatment that offers the possibility of a return to optimal health but also involves a risk of dying prematurely. Figure 5-1 provides an example of a standard gamble tree.

Figure 5-1. Example of a Standard Gamble Tree



Source: Law, Pathak, & McCord (1997)

In this basic standard gamble tree, a patient must choose between one of two choices: A or B. Option A includes two possibilities: (1) returning to full health (utility of 1) with probability p or (2) experiencing immediate death (utility of 0) with probability 1-p. Option B results in the status quo: namely a patient stays at his or her current condition. The goal is to find the individual's current utility. By changing the probability p, the probability of being returned to full health, which also impacts the probability of immediate death, one can solve for and discover the individual's current utility. This is done by solving the following equation: $p*1 + (1 - p)*0 = 1*(\text{current utility})$.

In time tradeoff, subjects are confronted with a choice between varying lengths of time in a healthy state as opposed to more time in a less healthy state. With either approach, the point of indifference between risk or time in competing states is used to determine the utility value for that particular health state. The resulting utility values are nearly always based on social preferences (those of the general population) because utilities reported by patients diagnosed with chronic conditions have been found to be systematically higher. This discrepancy between the reports of patients and the general public is known in the literature as the "disability paradox" (Albrecht & Devlieger, 1999).

In the second stage of the process, the directly assessed utility values are correlated with a small number of health state descriptors. An algorithm is then developed that predicts utility as a function of the descriptors, where the predicted utilities are often referred to as indirect utility measures. The algorithm is then applied to self-reported status for the descriptors to derive a utility value.

The flip side of utility is a concept known as "disutility," which refers to the negative consequences of having a particular health condition. The disutility for a specific condition is the difference in mean utilities between those who have the condition and those who do not.

5.4 Generic Health Indexes

Generic health status profiles, such as the SF-36, provide a comprehensive means for establishing an individual's health status and judging the effectiveness of health care interventions. However, they cannot be used in cost-utility analysis because they are not based on preferences. They describe, but do not value, health. Generic health indexes, which combine health status measures with utility valuations, were thus developed for use in valuing disease impacts and in economic valuation. They estimate a preference-based single index measure for health based on general population values. Links to examples of several common generic health status questionnaires are provided in Attachment 5.

The development of health-related quality of life (HRQoL) measures involves first characterizing a health state or disease condition and then valuing the state or condition based on general population preferences for each health state relative to others.[3]

The most commonly used generic indexes for estimating HRQoL are the SF-6D, the QWB, the HUI–2 and 3, and the EQ-5D. The acronyms stand for, respectively, short form, quality of well-being, health utility index, and EuroQol. Each of these indexes characterizes health status according to functioning in multiple health domains. Table 5-2 summarizes information on indexes currently in use (Fryback, 2010; IOM, 2006).

[3]Utility scores are generated using preference weights from a sample of the general population in a specific area. This population may have different preferences than others, so generally an index's preference scores are valid only when applied to a similar population. For example, because the SF-6D's preferences were established in the UK, its valuations may not apply to non-UK populations.

Table 5-2. Characteristics of Five Commonly Used Generic Health Indexes

	SF-6D	QWB[a]	HUI-2	HUI-3	EQ-5D
Dimensions	6	4	7	8	5
Description[b]	8, 17, 18, 19, 20, 21	12, 13, 14, 19	6, 7, 8, 9 10, 11, 13	1, 2, 3, 4, 5, 6, 7, 8	4, 8, 10, 15, 16
Health States	18,000	1,548	24,000	972,000	243
Derived From	SF-36	Index of Well-Being	HUI-1	HUI-1	N/A (original)
Preference Weight Method	standard gamble	visual analog scale	VAS/SG	VAS/SG	time trade-off
Population Norm	Sheffield, UK (n=611)	San Diego, CA (n=866)	Hamilton, ON (n=293)	Hamilton, ON (n=504)	16 countries (nbar = 2100)

[a]The QWB-SA (self-administered) is a less burdensome update to the original QWB instrument, which took 12 to 15 minutes to administer by a trained interviewer.

[b]Categories: 1 = vision, 2 = hearing, 3 = speech, 4 = ambulation, 5 = dexterity, 6 = emotion, 7 = cognition, 8 = pain, 9 = sensation, 10 = self-care, 11 = fertility, 12 = acute and chronic symptoms, 13 = mobility, 14 = physical activity, 15 = usual activity, 16 = anxiety/depression, 17 = mental health, 18 = physical functioning, 19 = social functioning, 20 = role limitations, 21 = vitality.

To apply preference weights to the indexes' numerous health states, researchers have elicited social utilities from samples of people. These preference weights are established using several different valuation tasks, which are described in Sections 5.2 and 5.3. Here we give a specific example of utility valuation for the EQ-5D, the most commonly used generic index worldwide. Respondents were asked to describe their own health using the EuroQol descriptive system, then to rank a set of 13 health states in comparison to the near-perfect health ("1 1 1 1 1") state. They were instructed to select a length of time in the near-perfect health state that they considered equivalent to 10 years in the state in question; the shorter the time given, the worse the target state. The utility of the state in question is thus the years responded divided by 10. Summary statistics were compiled from the sample subpopulations of each of the 16 countries to which EuroQol preference weights apply (Dolan et al., 1996).

After preference weights from a representative subpopulation are established for an index, those weights may be applied to other subpopulations. For example, health status information may be collected from a target population (e.g., clinical trial participants) and then valued according to the preference weights that have been established for a given index. These weights are then combined into a single measure using a scoring algorithm, which varies by index. For example, the HUI's scoring algorithm is multiplicative, which allows for preference-weighted scores in one domain to be related to scores in another domain. In contrast, the QWB scoring algorithm is additive. Researchers who use any of these generic HRQoL estimates to characterize disease burden must be aware of differences

across indexes in the elicitation of preferences and the scoring algorithm assumptions (Fryback, 2010).

It is important to note that the five quality of life indexes use different surveys, valuation techniques, and scoring algorithms; they also scale health states differently. Fryback et al. (2010) demonstrated that their scores are somewhat comparable by developing two-part linear crosswalks among indexes, but there is considerable error in the transformations for above-average health states. Despite the indexes' differences, all provide a similar summary measure for respondents: a valuation of health status between 0 and 1. Debate continues about which one is "best," and, because the best index for one situation or study may not be best for another, a global standard has not emerged.

5.5 Disease-Targeted Measures of HRQoL

Specific HRQoL measures assess aspects of HRQoL that are particularly relevant to people with the characteristic of interest, such as a particular age group or disease. In comparison to generic quality of life measures, disease-targeted measures are generally more sensitive to smaller differences and changes over time because they are selected especially for a given condition. For example, in a study of HRQoL in men treated for prostate cancer, there were no differences in SF-36 scores for those treated with surgery, radiation, watchful waiting, or a control group (Litwin et al., 1995). Disease-targeted measures, however—which targeted three organ systems (sexual, urinary, and bowel function)—found worse HRQoL among groups that had undergone surgery and radiation. Other recently developed disease-targeted measures are the National Eye Institute 25-Item Visual Function Questionnaire (NEI VFQ-25) (Mangione et al., 1998, 2001) and the Scleroderma Gastrointestinal Tract 1.0 survey (SSC-GIT 1.0) (Khanna et al., 2007).

The choice between generic or disease-targeted HRQoL measures depends on study aims, methodological concerns, and practical constraints (Cherepanov & Hays, in press). In some cases, a combination of measures may be used; for example, the Kidney Disease Quality of Life (KFQOL) was designed with the SF-36 as its generic core and with 11 kidney-disease targeted domains added to assess the specific effects of kidney disease (Hays et al., 1994).

5.6 Quality-Adjusted Life Years

Utility is a point-in-time measure of quality. Quantity of life can be introduced into the equation by incorporating a time dimension to create additional quality-adjusted measures. The average utility experienced during the course of a year is a quality-adjusted life year (QALY). These values may be summed over time. For example, someone with QALYs of 0.80, 0.73, and 0.65 over 3 consecutive years has accumulated 2.18 QALYs during that time. QALYs gained are the denominator in cost-utility ratios used to evaluate the cost-effectiveness of health interventions.

Table 5-3 provides an example of the use of QALYs from the literature. In the study, Masiocek et al. (2006) considered the likely increase in QALYs from increasing the utilization of clinical prevention and screening services from their current utilization levels up to 90% utilization. They found that tobacco cessation services would have the greatest impact in terms of QALYs saved, with aspirin use having the second largest impact.

Table 5-3. Potential Increase in Quality-Adjusted Life Years (QALYs) from Fuller Utilization of Select Services

Services	Current % Receiving Services Nationally	Additional QALYs Saved if Current % Receiving Services Increased to 90%
Tobacco-use screening and brief intervention	35%	1,300,000
Colorectal cancer screening	35%	310,000
Influenza vaccine—adults	36% among adults aged 50 to 64 years; 65% among adults aged ≥65 years	110,000
Breast cancer screening	68%	91,000
Cervical cancer screening	79%	29,000
Chlamydia screening	40%	19,000
Pneumococcal vaccine—adults	56%	16,000
Cholesterol screening	87%	12,000
Based on Limited Available Data, Utilization Rates of 50% Were Assigned to the Following Services:		
Aspirin chemoprophylaxis	50%	590,000
Problem drinking screening and brief counseling	50%	71,000
Vision screening—adults	50%	31,000

Source: Maciosek et al. (2006).

Another influential measure, developed by the Global Burden of Disease Project, is the disability-adjusted life year (DALY). DALYs are the sum of years of life lost to premature death (YLL) and years of health life lost due to injury or illness (YLD) (Mathers & Loncar, 2006). YLL measures deaths and years lost for males younger than age 80 and females younger than age 82.5. YLDs are calculated by looking at the incidence of a disease and multiplying this by the average duration of the disease and a scale factor for the severity of the disease (Mathers & Loncar, 2006). DALYs are becoming an increasingly popular measure on the global scale because of their ability to differentiate between disease burdens (e.g., infant mortality) that affect primarily third-world countries and chronic diseases that

disproportionately harm more developed regions. DALYs are estimated at the level of a disease, whereas QALYs are estimated at the person level. Although DALYs offer many advantages for comparing disease impacts across countries, they are not sensitive enough to be used in clinical research to evaluate the effectiveness or cost-effectiveness of a specific intervention on improving health outcomes.

To quantify the worldwide burden of illness, the Global Burden of Disease Study measures all health losses due to injury, illness, and premature death in terms of DALYs, broken into YLL and YLD. Table 5-4 shows the primary diseases contributing to DALYs and the fraction of the total global burden of illness contributed by each disease.

Table 5-4. Leading Global Diseases in DALYs in 2001

	Primary Cause	DALYs (millions of years)	% of DALYs
1	Perinatal condition	90.48	5.9
2	Lower respiratory infection	85.92	5.6
3	Ischemic heart disease	84.27	5.5
4	Cerebrovascular disease	72.02	4.7
5	HIV/AIDS	71.46	4.7
6	Diarrheal disease	59.14	3.9
7	Unipolar depressive disorder	51.84	3.4
8	Malaria	39.97	2.6
9	Chronic obstructive pulmonary disease	38.74	2.5
10	Tuberculosis	36.09	2.3

Source: Mathers & Loncar (2006).

The total worldwide burden of illness was estimated at 1.48 billion DALYs in 2002. The average global burden of disease was 250 DALYs per 1,000 people (Mathers & Loncar, 2006). The three diseases contributing primarily to this burden are perinatal conditions (90.48 million), lower respiratory infections (85.92 million), and ischemic heart disease (84.27 billion). On a global scale, primary health problems affect countries differently depending on income level. For example, in low- to middle-income regions, total DALYs are driven primarily by the YLL resulting from conditions such as perinatal disease. Of the 56 million people who died in 2001, 10.5 million were infants and children younger than age 5, and of these, 99% of the deaths occurred in low- and middle-income countries. In contrast, the primary burden in high-income regions is YLD, primarily among those aged 60 or older (Mathers & Loncar, 2006). Future projections of the global burden of disease suggest that despite a projected population increase of 27% through 2030, disease burden is projected to increase by only 3%, to 1.54 billion DALYs in 2030 (Mathers & Loncar, 2006).

5.7 Summary

In this section, four major types of health status and quality of life measurements have been identified (see Table 5-1). There are multiple measures within each category, and any of these measures can be used to portray burden of illness.

These health status and quality of life measures differ in two important ways from the other burden of illness measures described in this report. First, with only two exceptions, time lost to disability and disutility, these measures are not disease-specific. As a result, they can be applied to all members of the general population to monitor population health regardless of whether respondents have any chronic health conditions. Moreover, they are also suitable for patients diagnosed with multiple conditions. Second, with the exception of health-adjusted life expectancy and DALYs, each measure requires patient-reported assessments of personal health. An emphasis on patient-reported outcomes is one of the defining features of quality of life measurement.

The primary importance of these various epidemiologic, economic and quality of life measures lies in their ability to influence policy change. Policy makers and researchers seek to devote resources and time to interventions that are cost-effective and reduce clinically preventable burden (CPB) (Maciosek et al., 2006). CPB looks at the reduction in burden if available interventions were applied as recommended to almost all individuals at risk within a given population. Maciosek et al. (2006) rank 25 interventions by their cost-effectiveness in implementation and the amount of CPB that would be achieved by increasing their use (Attachment 6). Their findings serve as "a tool to help decision makers at multiple levels choose where to improve utilization rates by indicating which services are most consequential and cost effective for the population or individuals" (Maciosek et al., 2006, p. 55). By providing comparative information on economic cost and quality of life for multiple diseases or potential preventive interventions, policy makers can more easily use the information in making health care policy decisions.

5.8 New Initiatives

One of the largest new initiatives in the status measurement arena is the PROMIS program. This federally funded program is a 10-year study designed to develop, validate, and standardize item banks to measure patient-reported outcomes (Cella et al., 2007). Part of this effort will include the creation of a domain framework for physical function, pain, emotional distress, social health, and global health (PROMIS, 2010). The project aims to establish a technologically advanced resource to be used on a national scale that will allow for "accurate and efficient measurement of patient-reported symptoms and other health outcomes in clinical practice" (PROMIS, 2010). The ultimate goal of this project is standardized domain scoring, allowing researchers to compute scores on the same metric while using different sets of items from a calibrated data bank. The recent literature calls for standardization initiatives, as there is an urgent need to better quantify and measure

symptoms and outcomes. Consequently, further research on the reliability, validity, and interpretation of health status and quality of life measures is expected to increase in the near future.

6. LITERATURE REVIEW SUMMARY

Burden of illness research is receiving a great deal of attention from policy makers and researchers alike. Burden of illness estimates of all types are useful for informing policy, but the wide variation in methods and estimates for the same burden of illness measures suggests the need for a set of uniform, standardized methodologies for estimating and reporting burden of illness (Corso et al., 2006; Hofstetter & Hammitt, 2002; Lopez, 2005; Murray & Frenk, 2008). Some new efforts focus on standardizing approaches to collect data and estimate disease burden across different regions of the world, particularly in developing countries (Corso et al., 2006), and on consolidating burden measures across diseases and injuries (Franks et al., 2006; Lopez, 2005). Improved estimates of disease burden may be especially useful to health officials and policy makers for making decisions about priorities for future research and identifying specific diseases and intervention methods to target with available public health resources (Lyons et al., 2007; Thacker et al., 2006). A strong correlation has been observed between National Institutes of Health (NIH) spending and the burden of disease, as measured by disability-adjusted life years (DALYs) (McKenna & Zohrabian, 2009), suggesting that information about disease burden may be guiding, to some extent, U.S. decisions about research priorities.

In this literature review, we have described approaches for estimating three types of burden of disease measures: epidemiologic, economic, and quality of life. Each of these types of burden measures captures different elements of the impacts of disease on a population, and all three types of burden measures should be considered when making policy decisions that affect the allocation of prevention and treatment resources across diseases. Epidemiologic measures typically represent the extent of illness (prevalence and incidence) and the impact of illness or risk factors on years of life lost. Economic measures focus on valuing the opportunity costs that arise when individuals have a particular disease or risk factor. These costs capture both the monetary outlays required to diagnose, treat, or manage disease (health care spending and non-health spending) and the nonmonetary impacts of disease, such as productivity losses for patients and their caregivers. Quality of life measures capture the burden of disease in terms of the impact on functional status and utility preferences. An advantage of some quality of life measures is that they combine mortality- and morbidity-related impacts of disease into a single measure.

Faced with many different measures of the burden of a specific disease or risk factor, it may be challenging for policy makers to decide how to make use of all the burden of illness information with which they are presented. Because different burden measures capture different aspects of how a disease affects the population, policy makers may find that the best approach is to consider multiple measures of disease burden when evaluating and establishing priorities for health care spending and research. For example, diseases that are most burdensome in terms of the number of people affected may not be the most

burdensome in terms of impacts on longevity. Using estimates from the published literature, Thacker et al. (2008) showed that although cancer was responsible for the greatest number of years of preventable life lost, heart disease is the largest contributor to DALYs and costs.

In making broad decisions about how to allocate public health resources, it is important for policy makers to understand the burden of specific illnesses or risk factors in terms of their impacts on life expectancy, costs, health status, and quality of life. Only by understanding how an illness or risk factor affects morbidity, mortality, costs, and quality of life can policy makers make informed decisions. However, in cases where policy decisions focus on a specific burden of illness measure, such as the goal of limiting hospitalization costs, that specific burden measure (i.e., hospital costs) may be used to guide policy decisions.

Another important finding that has emerged from our literature review on burden of illness is that standardized approaches are needed to measure disease burden. Different approaches to valuing disease costs or quality of life impacts can lead to vastly different estimates of burden, and different values for the same burden measure may create confusion among policy makers as they try to select the best burden estimates for a given disease. Efforts are underway to improve consistency in burden measurement for health care costs and for measuring patient-reported health outcomes. These efforts are likely to lead to more consistency in future estimates of disease burden, which will ultimately benefit policy makers and the population as a whole, as better estimates of disease burden may contribute to well-informed decisions about public health resource allocation.

REFERENCES

2009 Atlas of CKD and ESRD. (2009). *United States Renal Data System*. National Institute of Diabetes and Digestive and Kidney Diseases. http://www.usrds.org/atlas.htm.

Adami, H. O., Hunter, D., & Trichopoulos, D. (2008). *Textbook of cancer epidemiology*. 2nd ed. New York, NY: Oxford University Press.

Aizcorbe, A. M., Retus, B. A., & Smith, S. (2008). *Toward a health care satellite account. BEA briefing*. Retrieved August 5, 2009, from http://www.bea.gov/national/health_care_satellite_account.htm.

Aizcorbe, A. M., & Nestoriak, N. (2007, December). *Changes in treatment intensity, treatment substitution, and price indexes for health care services*. Paper presented at the National Bureau of Economic Research Productivity Workshop, Cambridge, MA.

Albrecht, G. L., & Devlieger, P. J. (1999). The disability paradox: High quality of life against all odds. *Social Science and Medicine*, *48*, 977–988.

Akobundu, E., Ju, J., Blatt, L., & Mullins, C. D. (2006). Cost-of-illness studies: A review of current methods. *Pharmacoeconomics, 24*(9), 869–890.

Aldy, J. E., & Viscusi, W. K. (2007). Age differences in the value of statistical life: Revealed preference evidence. *Resources for the Future Discussion Paper, 07–05*.

Aschengrau, A., & Seage, G. R., III (2008). *Essentials of epidemiology in public health*, 2nd ed. Sudbury, MA: Jones and Bartlett Publishers, LLC.

Basu, A., & Manning, D. G. (2009). *Issues for the next generation of health care cost analyses*. *Medical Care, 47*(7 Suppl 1), S109–S114.

Beck, J. R., Kassirer, J. P., & Pauker, S. G. (1982). A convenient approximation of life expectancy (the "DEALE"). I. Validation of the method. *American Journal of Medicine, 73*(6), 883–888.

Bleichrodt, H., & Johanneson, M. (1997). Standard gamble, time trade-off and rating scale: experimental results on the ranking properties of QALYs. *Journal of Health Economics, 16*, 155–175.

Bloom, B. S., Bruno, D. J., Maman, D. Y, & Jayadevappa, R. (2001). Usefulness of U.S. cost-of-illness studies in healthcare decision making. *Pharmacoeconomics, 19*(2), 207–213.

Boslaugh, S. (2008). *Encyclopedia of epidemiology*, Volume 1. Thousand Oaks, CA: SAGE Publications, Inc.

Brazier, J., Deverill, M., & Green, C. (1999). A review of the use of health status measures in economic evaluation. *Journal of Health Services Research & Policy, 4*(3), 174–184.

Brouwer, W., Rutten, F., & Koopmanschap, M. (2001). Costing in economic evaluations. In M. F. Drummond & A. McGuire (Eds.), *Economic evaluation in health care: Merging theory with practice* (Chapter 4). New York: Oxford University Press.

Brown, L. M., & Devesa, S. S. (2009). Epidemiology and risk of esophageal cancer: Clinical. In B. A. Jobe, C. R. Thomas, & J. G. Hunter (Eds.), *Esophageal cancer principles and practice* (pp. 103–113). New York: Demos Medical Publishing, LLC.

Brown, M. L., Lipscomb, J., & Snyder, C. (2001). The burden of illness of cancer: Economic cost and quality of life. *Annual Review of Public Health*, *22*, 91–113.

Brownson, R. C., Remington, P. L., & Davis, J. R. (1993). *Chronic disease epidemiology and control*. Washington, D. C.: American Public Health Association.

Buntin, M. B., & Zaslavsky, A. M. (2004). Too much ado about two-part models and transformation? Comparing methods of modeling Medicare expenditures. *Journal of Health Economics, 23*(3)*,* 525–542.

Cella, D., Yount, S., Rothrock, N., Gershon, R., Cook, K., Reeve, B., et al. (2007). The patient-reported outcomes measurement information system (PROMIS): Progress of an NIH roadmap cooperative group during its first two years. *Medical Care, 45*(5 Suppl 1), S3–S11.

Cheak-Zamora, N. C., Wyrwich, K. W., & McBride, T. D. (2009). Reliability and validity of the SF-12v2 in the medical expenditure panel survey. *Quality of Life Research, 18*(6), 727–735.

Cherepanov, D., & Hays, R. D. (in press). Health and quality of life outcomes: The role of patient-reported measures. In J. L. Magnabosco & R.W. Manderscheid (Eds.), *Outcomes measurement in the human services: Cross-cutting issues and methods* (2nd edition). Washington, DC: National Association of Social Workers Press.

Chuang, L. H., & Kind, P. (2009). Converting the SF-12 into the EQ-5D: An empirical comparison of methodologies. *Pharmacoeconomics, 27*(6), 491–505.

Corso, P., Finkelstein, E., Miller, T., Fiebelkorn, I., & Zaloshnja, E. (2006). Incidence and lifetime costs of injuries in the United States. *Injury Prevention, 12*(4), 212–218.

Craig, B. M., Busschbach, J. J., & Salomon, J. A. (2009). Modeling ranking, time trade-off, and visual analog scale values for EQ-5D health states: A review and comparison of methods. *Medical Care, 47*(6), 634–641.

CureResearch.com. Prevalence and incidence of common cold. symptoms, diseases and diagnosi*s.* Retrieved February 2, 2010, from http://www.cureresearch.com/c/cold/prevalence.htm.

Dicker, R., Coronado, F., Koo, D., & Parrish, R. G. (2006). *Principles of epidemiology in public health practice*, 3rd ed. [SS-1000]. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, Office of Workforce and Career Development.

Diehr, Yanez, P., Ash, A. Hornbrook, M., & Lin, D. Y. (1999). Methods for analyzing health care utilization and costs. *Annual Review of Public Health*, *20*, 125–144.

Dockery, D. W., Pope, C. A., Xu, X, Spengler, J. D., Ware, J. H., Fay, M. E., et al. (1993). An association between air pollution and mortality in six U.S. cities. *New England Journal of Medicine*, 329, 1753–1759.

Drummond, M., & McGuire, A. (Eds.). (2001). *Economic evaluation in health care: Merging theory with practice*. New York, NY: Oxford University Press.

Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., & Stoddart, G. L. (2005) *Methods for the economic evaluation of health care programmes*. 3rd ed. New York: Oxford University Press.

Encyclopedia Britannica. (2010). Mortality. Retrieved February 1, 2010, from http://www.britannica.com/EBchecked/topic/393100/mortality.

Ferreira, P. L., Ferreira, L. N., & Pereira, L. N. (2008). How consistent are health utility values? *Quality of Life Research, 17*(7), 1031–1042.

Flegal, K. M., Graubard, B. I., Williamson, D. F., & Gail, M. H. (2005). Excess deaths associated with underweight, overweight, and obesity. *Journal of the American Medical Association, 293*(15), 1861–1867.

Franks, P., Hanmer, J., & Fryback, D. G. (2006). Relative disutilities of 47 risk factors and conditions assessed with seven preference-based health status measures in a national U.S. sample: Toward consistency in cost-effectiveness analyses. *Medical Care, 44*(5), 478–485.

Freeman, A. M. (1993). *The measurement of environmental and resource values: theory and methods*. 1st ed. Washington, DC: Resources for the Future.

Freeman, A. M. (2003). *The measurement of environmental and resource values: theory and methods*. 2nd ed. Washington, DC: Resources for the Future.

Fryback, D. G., Palta, M., Cherepanov, D., Bolt, D. & Kim, J. S. (2010). Comparison of 5 health-related quality-of-life indexes using item response theory analysis. *Medical Decision Making*, *30*, 5–15.

Gilleskie, D. B., & Mroz, T. A. (2004). A flexible approach for estimating the effects of covariates on health expenditures. *Journal of Health Economics, 23*(2), 391–418.

Glasziou, P., Alexander, J., Beller, E., & Clarke, P. (2007). Which health-related quality of life score? A comparison of alternative utility measures in patients with Type 2 diabetes in the ADVANCE trial. *Health and Quality of Life Outcomes, 5*, 21.

Goeree, R., O'Brien, B. J., Blackhouse, G., Agro, K., & Goering, P. (1999). The valuation of productivity costs due to premature mortality: A comparison of the human-capital and friction-cost methods for schizophrenia. *Canadian Journal of Psychiatry*, *44*(5), 455–463.

Gold, M. R., Siegel, J. E., Russell, L. B., & Weinstein, M. C. (Eds.). (1996). *Cost-effectiveness in health and medicine*. New York: Oxford University Press.

Gold, M. R., Stevenson, D. & Fryback, D. G. (2002). HALYS and QALYS and DALYS, Oh my: Similarities and differences in summary measures of population health. *Annual Review of Public Health, 23,* 115–134.

Gregg, E. W., Cadwell, B. L., Cheng, Y. J., Cowie, C. C., Williams, D. E., Geiss, L., et al. (2004). Trends in the prevalence and ratio of diagnosed to undiagnosed diabetes according to obesity levels in the U.S. *Diabetes Care, 27*(12), 2806–2812.

Grodstein, F., Goldman, M. B., & Cramer, D. W. (1993). Relation of tubal infertility to history of sexually transmitted diseases. *American Journal of Epidemiology*, 137(5), 577–584.

Grosse, S. D., Krueger, K. V., & Mvundura, M. (2009). Economic productivity by age and sex: 2007 estimates for the United States. *Medical Care*, 47(7 Suppl 1), S94–S103.

Haddix, A., Teutsch, S., & Corso, P. (2003). Chapter 4. Costs. In A. Haddix, S. Teutsch, & P. Corso (Eds.), *Prevention effectiveness: A guide to decision analysis and economic evaluation*. 2nd ed. New York: Oxford University Press.

Haddix, A., Teutsch, S., & Corso, P. (2003). *Prevention effectiveness: A guide to decision analysis and economic evaluation*. New York: Oxford University Press.

Hajjar, I., Kotchen, J. M., & Kotchen, T. A. (2006). Hypertension: Trends in prevalence, incidence, and control. *Annual Review of Public Health, 27*, 465–490.

Hays, R. D., Kallich, J. D., Mapes, D. L., Coons, S. J., & Carter, W. B. (1994). Development of the Kidney Disease Quality of Life (KDQOL) Instrument. *Quality of Life Research*, *3*, 329–338.

Heffler, S., Nuccio, O., & Freeland, M. (2009). An overview of the NHEA with implications for cost analysis researchers. *Medical Care, 47*(7 Suppl 1), S37–S43.

Hilditch, J. R., Lewis, J., Peter, A., van Maris, B., Ross, A., Franssen, E., Guyatt, G. H., et al. (2008). A menopause-specific quality of life questionnaire: Development and psychometric properties. *Maturitas, 61*(1-2), 107–121.

Hirth, R. A., Chernew, M. E., Miller, E. & Fendrick, A. M. (2000). Willingness to pay for a quality-adjusted life year: In search of a standard. *Medical Decision Making, 20*(3), 332–342.

Hodgson, T. A. (1988). Annual costs of illness versus lifetime costs of illness and implications of structural change. *Drug Information Journal, 22*, 323–341.

Hodgson, T. A., & Meiners, M. R. (1982). Cost-of-illness methodology: A guide to current practices and procedures. *Milbank Memorial Fund Quarterly Health and Society, 60*(3), 429–462.

Hodgson, T. A., & Cohen, A. J. (1999). Medical care expenditures for diabetes, its chronic complications, and its comorbidities. *Preventive Medicine, 29,* 173–186.

Hoerger, T. J. (2006). Controversies in obesity mortality: A tale of two studies. *Health Promotion Economics Issue Brief, 1*(1), 1–4.

Hofstetter, P., & Hammitt, J. K. (2002). Selecting human health metrics for environmental decision-support tools. *Risk Analysis, 22*(5), 965–983.

Hogan, M. C., Foreman, K. J., Naghavi, M., Ahn, S. Y., Wang, M., Makela, S. M., et al. (2010). Maternal mortality for 181 countries, 1980–2008: A systematic analysis of progress towards Millennium Development Goal 5. *Lancet*.

Holland, W. W., Florey, C., & Olsen, J. (2007). *The development of modern epidemiology: Personal stories from those who were there*. New York, NY: Oxford University Press.

Hollinghurst, S., Bevan, G., & Bowie, C. (2000). Estimating the "avoidable" burden of disease by Disability Adjusted Life Years (DALYs). *Health Care Management Science, 3*(1), 9–21.

Honeycutt, A. A., Segel, J. E., Hoerger, T. J., & Finkelstein, E. A. (2009). Comparing cost-of-illness estimates from alternative approaches: An application to diabetes. *Health Services Research*, *44*(1), 303–320.

Huang, W. Y., Winn, D. M., Brown, L. M., Gridley, G., Bravo-Otero, E., Diehl, S. R., et al. (2003). Alcohol concentration and risk of oral cancer in Puerto Rico. *American Journal of Epidemiology*, *157*, 881–887.

Jager, K. J., Zoccali, C., Kramer, R., & Dekker, F. W. (2007). Measuring disease occurrence. *Kidney International, 72*(4), 412–415.

Janssen, M. F., Birnie, E., & Bonsel, G. J. (2008). Quantification of the level descriptors for the standard EQ-5D three-level system and a five-level version according to two methods. *Qual Life Res, 17*(3), 463–473.

Javitz, H. S., Ward, M. M., Watson, J. B., & Jaana, M. (2004). Cost of illness of chronic angina. *American Journal of Managed Care, 10*(11 Suppl), S358–S369.

Johnson, F. R., Fries, E. E., & Banzhas, H. S. (1997). Valuing morbidity: An integration of the willingness-to-pay and health-status index literatures. *Journal of Health Economics, 16*(6), 641–665.

Khanna, D., Hays, R. D., Park, G. S., et al. (2007). Development of a preliminary scleroderma gastrointestinal tract 1.0 (SSC-GIT 1.0) quality of life instrument. *Arthritis Care and Research*, *57*, 1280–1286.

Kirschstein, Ruth. (2000, February). *Disease-specific estimates of direct and indirect costs of illness and NIH support*. Department of Health and Human Services, National Institutes of Health, Office of the Director. Retrieved February 12, 2010, from http://ospp.od.nih.gov/ecostudies/coireportweb.htm.

Knapp, M., & Mangalore, R. (2007). The trouble with QALYs.... *Epidemiologia e psichiatria sociale, 16*(4), 289–293.

Knies, S., Evers, S. M., Candel, M. J., Severens, J. L., & Ament, A. J. (2009). Utilities of the EQ-5D: Transferable or not? *Pharmacoeconomics, 27*(9), 767–779.

Kobelt, G., Berg, J., Atherly, D., & Hadjimichael, O. (2006). Costs and quality of life in multiple sclerosis: A cross-sectional study in the United States. *Neurology, 66*(11), 1696–1702.

Koopmanschap, M. A., & van Ineveld, B. M. (1992). Towards a new approach for estimating indirect costs of disease. *Social Science & Medicine,* 34(9), 1005–1010.

Koopmanschap, M. A., Rutten, F. F., et al. (1995). The friction cost method for measuring indirect costs of disease. *Journal of Health Economics, 14*(2), 171–189.

Krabbe, P. F., Essink-Bot, M. L., & Bonsel, G. J. (1997). The comparability and reliability of five health-state valuation methods. *Social Science & Medicine, 45*(11), 1641–1652.

Landefeld, J. S., & Seskin, E. P. (1982). The economic value of life: Linking theory to practice. *American Journal of Public Health, 72*(6), 555–566.

Law, A. V., Pathak, D. S., & McCord, M. R. (1998). Health status utility assessment by standard gamble: A comparison of the probability equivalence and the lottery equivalence approaches. *Pharamceutical Appoaches*, *15*(1), 105–109.

Lee, E. T., & Wang, J. W. (2003). *Statistical methods for survival data analysis*. 3rd ed. Hoboken, NJ: Wiley-Interscience.

Lilienfeld, A. M., & Lilienfeld, D. E. (1980). *Foundations of epidemiology*, 2nd ed. 4. New York: Oxford University Press.

Lipscomb, J., Barnett, P. G., Brown, M. L., Lawrence, W., & Yabroff, K. R. (2009). Advancing the science of health care costing. *Medical Care, 47*(7 Suppl 1), S120–S126.

Lipscomb, J., Yarbrouff, K. R., Brown, M. L., Lawrence, W. & Barnett, P. G. (2009). Health care costing: Data, methods, current applications. *Medical Care, 47*(7 Suppl 1), S1–S6.

Litwin, M., Hays, R. D., Fink, A., Ganz, P. A., Leake, B., Leach, G. E., & Brook, R. H. (1995). Quality of life outcomes in men treated for localized prostate cancer. *Journal of the American Medical Association*, *273*(2), 129–135.

Liu, J. L., Maniadakis, N., Gray, A., & Rayner, M. (2002). The economic burden of coronary heart disease in the UK. *Heart, 88*(6), 597–603.

Lopez, A. D. (2005). The evolution of the global burden of disease framework for disease, injury and risk factor quantification: Developing the evidence base for national, regional and global public health action. *Global Health, 1*(1), 5.

Lua, P. L., Salek, S., Finlay, I., & Lloyd-Richards, C. (2005). The feasibility, reliability and validity of the McGill Quality of Life Questionnaire-Cardiff Short Form (MQOL-CSF) in palliative care population. *Quality of Life Research, 14*(7), 1669–1681.

Luce, B. R., Manning, W. G., Siege, J. E., & Lipscomb, J. (1996). Estimating costs in cost-effectiveness analysis. In Gold, M. R., Siegel, J. E., Russell, L. B., & Weinstein, M. C. (Eds.), *Cost-effectiveness in health and medicine*. New York: Oxford University Press.

Luo, N., Johnson, J. A., Shaw, J. W., & Coons, S. J. (2009). Relative efficiency of the EQ-5D, HUI2, and HUI3 index scores in measuring health burden of chronic medical conditions in a population health survey in the United States. *Medical Care, 47*(1), 53–60.

Lyons, R. A., Towner, E. E., Kendrick, D. Christie, N., Brophy, S., Phillips, C. J., et al. (2007). The UK burden of injury study—A protocol. *BMC Public Health 7*, 317.

Maciosek, M. V., Coffield, A. B., Edwards, N. M., Flottemesch, T. J., Goodman, M. J., & Solberg, L. I. (2006). Priorities among effective clinical preventive services: results of a systematic review and analysis. *American Journal of Preventive Medicine, 31*(1), 52–61.

Mahadevia, P. J., Fleisher, L. A., Frick, K. D., Eng, J., Goodman, S. N., & Powe, N. R. (2003). Lung cancer screening with helical computed tomography in older adult smokers: A decision and cost-effectiveness analysis. *Journal of the American Medical Association, 289*(3), 313–322.

Mangione, C. M., Lee, P. P., Pitts, J., Gutierrez, P., Berry S., & Hays, R. D**.** (1998). Psychometric properties of the National Eye Institute Visual Function Questionnaire, the NEI-VFQ. *Archives of Ophthalmology*, *116*, 1496–1504.

Mangione, C.M., Lee, P.P., Gutierrez, P.R., et al. (2001). Development of the 25-item National Eye Institute Visual Function Questionnaire. *Archives of Ophthalmology*, *119*, 1050–1058.

Manning, W. G. (1998). The logged dependent variable, heteroscedasticity, and the retransformation problem. *Journal of Health Economics, 17*(3), 283–295.

Manning, W. G., & Mullahy, J. (2001). Estimating log models: To transform or not to transform? *Journal of Health Economics, 20*(4), 461–494.

Manuel, D. G., & Schultz, S. E. (2004). Health-related quality of life and health-adjusted life expectancy of people with diabetes in Ontario, Canada, 1996–1997. *Diabetes Care, 27*(2), 407–414.

Mathers, C. D., & Loncar, D. (2006). Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med, 3*(11), e442.

Mathers, C. D., Lopez, A. D. & Murray, C. J. (2006). The burden of disease and mortality by condition: Data, methods, and results for 2001. *Global burden of disease and risk factors* (pp. 45–93). New York: Oxford University Press.

Mausner, J. S., & Kramer, S. (1985). *Epidemiology: An introductory text*, 2nd ed. Philadelphia: W.B. Saunders & Co.

McDaid, D. (2001). Estimating the costs of informal care for people with Alzheimer's disease: Methodological and practical challenges. *International Journal of Geriatric Psychiatry, 16*, 400–405.

McDonough, C. M., & Tosteson, A. N. (2007). Measuring preferences for cost-utility analysis: How choice of method may influence decision-making. *Pharmacoeconomics, 25*(2), 93–106.

McKenna, M. T., & Zohrabian, A. (2009). U.S. burden of disease—Past, present and future. *Annals of Epidemiology, 19*(3), 212–219.

McKenzie, L., & van der Pol, M. (2008). Mapping the EORTC QLQ C-30 onto the EQ-5D Instrument: The potential to estimate QALYs without generic preference data. *Value Health*, *12*(1) 167–171.

Mokdad, A. H., Ford, E. S., Bowman, B. A., Dietz, W. H., Vinicor, F., Bales, V. S., et al. (2003). Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. *Journal of the American Medical Association, 289*(1), 76–79.

Morabia, A. (2007). Chapter 3. Epidemiological methods and concepts in the nineteenth century and their influences on the twentieth century. In Holland, W.W., Olsen, J. & Florey, C. V. (Ed.), *The development of modern epidemiology: Personal stories from those who were there*. New York, NY: Oxford University Press.

Morgenstern, H., Kleinbaum, D. G., & Kupper, L. L. (1980). Measure of disease incidence used in epidemiological research. *International Journal of Epidemiology, 9*, 97–104.

Mrozek, J. R., & Taylor, L. O. (2002). What determines the value of life? A meta-analysis. *Journal of Policy Analysis and Management, 21*(2), 253–270.

Murray, C. J. (1994). Quantifying the burden of disease: The technical basis for disability-adjusted life years. *Bulletin of the World Health Organization, 72*(3), 429–445.

Murray, C. J., & Frenk, J. (2008). Health metrics and evaluation: Strengthening the science. *Lancet, 371*(9619), 1191–1199.

Osman, L., & Silverman, M. (1996). Measuring quality of life for young children with asthma and their families. *European Respiratory Journal Supplement, 21,* 35s–41s.

Parker, S. L., Tong, T., Bolden, S., & Wingo, P. A. (1996). Cancer statistics, 1996. *CA: A Cancer Journal for Clinicians, 46*(1), 5–27.

Parkin, D. M., Pisani, P., & Ferlay, J. (1999). Global cancer statistics. *CA: A Cancer Journal for Clinicians, 49*(1), 33–64.

Pauly, M. V., Nicholson, S., Polsky, D., Berger, M. L., & Sharda, C. (2008). Valuing reductions in on-the-job illness: 'Presenteeism' from managerial and economic perspectives. *Health Economics, 17*(4), 469–485.

Paz, S. H., Liu, H., Fongwa, M. N., Morales, L. S., & Hays, R. D. (2009). Readability estimates for commonly used health-related quality of life surveys. *Quality of Life Research, 18*(7), 889–900.

Porta, M. (2008). *A dictionary of epidemiology*, 5th ed. (pp. 81–82). New York: Oxford University Press.

Portney, P. R. (1994). The contingent valuation debate: Why economists should care. *The Journal of Economic Perspectives, 8*(4), 3–17.

Raisch, D. W. (2000). Understanding quality-adjusted life years and their application to pharmacoeconomic research. *Annals of Pharmacotherapy, 34*(7-8), 906–914.

Rector, T., Cubo, S., & Cohn, J. (1987). Patient's self assessment of their congestive heart failure. Part 2: Content, reliability and validity of a new measure, the Minnesota Living with Heart Failure Questionnaire. *Heart Failure*, *3*, 192–196.

Rice, D. P., & Hodgson, T. A. (1982). The value of human life revisited. *American Journal of Public Health, 72*(6), 536–538.

Rice, D. P., & Miller, L. S. (1998). Health economics and cost implications of anxiety and other mental disorders in the United States. *British Journal of Psychiatry, 173*(suppl. 34), 4–9.

Rockhill, B. (2005). Theorizing about causes at the individual level while estimating effects at the population level: Implications for prevention. *Epidemiology, 16*(1), 124–129.

Rosen, A. B., & Cutler, D. M. (2007). Measuring medical care productivity. *Survey of Current Business*, *87*(6), 54–58.

Rosen, A. B., & Cutler, D. M. (2009). Challenges in building disease-based national health accounts. *Medical Care*, 47(7 Suppl 1), S7–S13.

Rothman, K., Greenland, S., & Lash, T. L. (2008). *Modern epidemiology*. Philadelphia, PA: Lippincott, Williams & Wilkins.

Rowen, D., Brazier, J., & Roberts, J. (2009). Mapping SF-36 onto the EQ-5D index: How reliable is the relationship? *Health and Quality of Life Outcomes, 7,* 27.

Sculpher, M. (2001). Chapter 5. The role and estimation of productivity costs in economic evaluation. In M. F. Drummond & A. McGuire (Eds.), *Economic evaluation in health care: Merging theory with practice.* New York: Oxford University Press.

Segel, J. E. (2006). *Cost-of-illness studies—A primer*. RTI International. Retrieved August 2, 2009, from http://www.rti.org/pubs/COI_Primer.pdf.

Seymour, J., McNamee, P., Scott, A., & Tinelli, M. (2009). Shedding new light onto the ceiling and floor? A quantile regression approach to compare EQ-5D and SF-6D responses. *Health Economics.* [Epub ahead of print].

Shaw, J. W., Johnson, J. A., & Coons, S. J. (2005). US Valuation of the EQ-5D health states. *Medical Care*, 43(3), 203-220.

Spasoff, R. A. (1999). *Epidemiological methods for health policy*. Oxford University Press.

Thacker, S. B., Stroup D. F, Carande-Kulis, V., Marks, J. S., Roy, K., & Gerberding, J. L. (2006). Measuring the public's health. *Public Health Reports, 121*(1), 14–22.

Thompson, M. L., Myers, J. E., & Kriebel, D. (1998). Prevalence odds ratio or prevalence ratio in the analysis of cross sectional data: What is to be done? *Occup Environ Med, 55*(4), 272–277.

Trogdon, J. G., Finkelstein, E. A., & Hoerger, T. J. (2008). Use of econometric models to estimate expenditure shares. *Health Services Research, 43*(4), 1442–1452.

Tu, K., Chen, Z., & Lipscombe, L. L. (2008). Prevalence and incidence of hypertension from 1995 to 2005: A population-based study. *Canadian Medical Association Journal*, *178*, 1429–1435.

U.S. Department of Health and Human Services, Center for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System, *National Vital Statistics Reports*, *58*(10), March 3, 2010.

U.S. Environmental Protection Agency (EPA). (2000). *Guidelines for preparing Economic Analysis Report 240-R-00-003*. Washington, DC: Office of the Administrator: U.S. Environmental Protection Agency.

United States Renal Data System (2009). *USRDS.org*. Retrieved February 1, 2010, from http://www.usrds.org/.

van Baal, P. H., Hoogenveen, R. T., de Wit, G. A., & Boshuizen, H. C. (2006). Estimating health-adjusted life expectancy conditional on risk factors: Results for smoking and obesity. *Population Health Metrics, 4*, 14.

Van Houtven, G. L., Honeycutt, A. A., Gilman, B., McCall, N., Throneburg, W. W., & Sykes, K. E. (2008). *Costs of illness among older adults: An analysis of six major health conditions with significant environmental risk factors*. Research Triangle Park, NC: RTI Press, RR-0002-0809.

Van Houtven, G., Rousu, M., Yang, J. C., Pringle, C., Wagstaff, W., & DePlatchett, J. (2003). *Valuation of morbidity losses: Meta-analysis of willingness-to-pay and health status measures*. Prepared for the Food and Drug Administration.

Varni, J. W., Sherman, S. A., Burwinkle, T. M., Dickinson, P. E., & Dixon, P. (2004). The PedsQL Family Impact Module: Preliminary reliability and validity. *Health and Quality of Life Outcomes, 2*, 55.

Viscusi, K. W. (2003). The value of a statistical life: A critical review of market estimates throughout the world. *The Journal of Risk and Uncertainty, 27*(1), 5–76.

Viscusi, W. K. Value of life. *The New Palgrave Dictionary of Economics*. 2nd ed. Retrieved February 8, 2010, from http://www.dictionaryofeconomics.com/article?id= pde2008_V000005> doi:10.1057/9780230226203.1784.

Ward, M. M., Javitz, H. S., Smith, W.M., & Bakst, A. (2000). A comparison of three approaches for attributing hospitalizations to specific diseases in cost analyses. *Int J Technol Assess Health Care, 16*(1), 125–136.

Ware, J. E. The SF Community−SF-36. *The SF Community−offering information and discussion on health outcomes*. Retrieved February 2, 2010, from http://www.sf-36.org/tools/sf36.shtml.

Whynes, D. K. (2008). Correspondence between EQ-5D health state classifications and EQ VAS scores. *Health Qual Life Outcomes, 6*, 94.

World Health Organization (2004). Measurement and Health Information. Estimated deaths per 100,000 population by cause, and Member State, 2002 and Age-standardized death rates per 100,000 by cause, and Member State, 2002. http://www.who.int/healthinfo/statistics/bodgbddeathdalyestimates.xls

World Health Organization (2006). Under-5 mortality rate in 2006 per 1000 live births. Life tables for WHO Member States. Geneva, World Health Organization, 2006. http://www.who.int/whosis/database/life_tables/life_tables.cfm, accessed 18 March 2008.

World Health Organization (2007). Maternal mortality ratio in 2005 per 100,000 live births: estimates developed by WHO, UNICEF, UNFPA and the World Bank. Geneva: World Health Organization. http://www.who.int/reproductive-health/publications/ maternal_mortality_2005/mme_2005.pdf. Accessed March 18, 2008.

World Health Organization (2008). Life Tables for WHO Member States, WHO Statistical Information System (WHOSIS), World Health Statistics, 2008. Geneva: World Health Organization. http://apps.who.int/whosis/database/life_tables/life_tables.cfm.

Yang, P., Cehran, J. R., & Vierkant, R. A. (2002). Adenocarcinoma of the lung is strongly associated with cigarette smoking: Further evidence from a prospective study of women. *American Journal of Epidemiology*, *156*, 1114–1122.

Yelin, E., Herrndorf, A., Trupin, L., & Sonneborn, D. (2001). A national study of medical care expenditures for musculoskeletal conditions. *Arthritis and Rheumatism, 44*(5), 1160–1169.

Zhang, X. H., Xie, F., Wee, H. L., Thumboo, J., & Li, S. C. (2008). Applying the expectancy-value model to understand health values. *Value Health, 11*(Suppl 1), S61–S68.

ATTACHMENT 1:
BIBLIOGRAPHY WITH ABSTRACTS

1. Epidemiology Measures: Mortality Prevalence and Methodologies

Beck, J. R., Kassirer, J. P., & Pauker, S. G. (1982). A convenient approximation of life expectancy (the "DEALE"). I. Validation of the method. *American Journal of Medicine, 73*(6), 883–888.

The physician developing a treatment plan for a particular patient often needs to know the life expectancy associated with the outcomes of therapeutic choices. Currently available methods for estimating life expectancy are cumbersome and of limited clinical use. We describe a simple approximation of life expectancy (the "DEALE") that is based on the assumption that survival follows a simple declining exponential function. In this approach, the reciprocal of the age-, sex-, and race-adjusted life expectancy is used to estimate the mortality rate of a healthy person. The life expectancy of a person who also has one or more diseases is obtained by adding disease-specific mortalities to the age-, sex-, and race-adjusted mortality rate and taking the reciprocal of that sum. In this paper we show that this approximation estimates life expectancy accurately for the great majority of clinical problems.

Behrens, T., Taeger, D., Wellmann, J., & Kiel, U. (2004). Different methods to calculate effect estimates in cross-sectional studies. A comparison between prevalence odds ratio and prevalence ratio. *Methods of Information in Medicine*, *43*(5), 505–509.

OBJECTIVES: According to results from the epidemiological literature, it can be expected that the prevalence odds ratio (POR) and the prevalence ratio (PR) differ with increasing disease prevalence. We illustrate different concepts to calculate these effect measures in cross-sectional studies and discuss their advantages and weaknesses, using actual data from the ISAAC Phase III cross-sectional survey in Munster, Germany. METHODS: We analyzed data on the association between self-reported traffic density and wheeze and asthma by means of the POR, obtained from a logistic regression, and the PR, which was estimated from a log-linear binomial model and from different variants of a Poisson regression. RESULTS: The analysis based on the less frequent disease (i.e., asthma with an overall prevalence of 7.8%) yielded similar results for all estimates. When wheezing with a prevalence of 17.5% was analyzed, the POR produced the highest estimates with the widest confidence intervals. While the point estimates were similar in the log-binomial model and Poisson regression, the latter showed wider confidence intervals. When we calculated the Poisson regression with robust variances, confidence intervals narrowed. CONCLUSIONS: Since cross-sectional studies often deal with frequent diseases, we encourage analyzing cross-sectional data based on log-linear binomial models, which is the natural method for estimating prevalence ratios. If algorithms fail to converge, a useful alternative is to define

appropriate starting values or, if models still do not converge, to calculate a Poisson regression with robust estimates to control for overestimation of errors in the binomial data.

Carson, S. S., & Bach, P. B. (2002). The epidemiology and costs of chronic critical illness. *Critical Care Clinics*, *18*(3), 461–476.

CCI patients are patients who have suffered acute illness or injury and require life support or care in an ICU setting for periods of weeks or months. These patients account for between 5% and 10% of ICU admissions, and they appear to be increasing in number. Over half of the patients are over age 65. Patients with underlying premorbid conditions who suffer complications of acute illness are at highest risk for becoming CCI. These patients have poor short-term and long-term survival, although survival may be improving for some types of CCI patients as the medical system adapts to their specific needs. Long-term survival is associated with age and premorbid condition or functional status. Survivors have significant functional limitations, but their reported quality of life is generally good. CCI patients consume a disproportionate share of ICU and hospital resources, and significant additional resources are required for continued recovery or care after discharge. Specialized units have been evolving to manage these patients at lower costs than in acute ICUs, and with similar outcomes. Further refinement of the definition of CCI is an important objective and should pave the way to better design of outcomes studies. Efforts should continue to learn how to identify patients at high risk for CCI and poor outcome so that expensive resources can be managed effectively, and patient-provider decision making can be better informed.

Cella, D., Young, S., Rothrock, N., Gershon, R., Cook, K., Reeve, B., et al. (2007). The Patient-Reported Outcomes Measurement Information System (PROMIS): Progress of an NIH Roadmap cooperative group during its first two years. *Medical Care*, *45*(5 Suppl 1): S3–S11.

BACKGROUND: The National Institutes of Health (NIH) Patient-Reported Outcomes Measurement Information System (PROMIS) Roadmap initiative (www.nihpromis.org) is a 5-year cooperative group program of research designed to develop, validate, and standardize item banks to measure patient-reported outcomes (PROs) relevant across common medical conditions. In this article, we will summarize the organization and scientific activity of the PROMIS network during its first 2 years. DESIGN: The network consists of six primary research sites (PRSs), a statistical coordinating center (SCC), and NIH research scientists. Governed by a steering committee, the network is organized into functional subcommittees and working groups. In the first year, we created an item library and activated three interacting protocols: Domain Mapping, Archival Data Analysis, and Qualitative Item Review (QIR). In the second year, we developed and initiated testing of item banks covering five broad domains of self-reported health. RESULTS: The domain mapping process is built on the World Health Organization (WHO) framework of physical, mental, and social health. From this framework, pain, fatigue, emotional distress, physical

functioning, social role participation, and global health perceptions were selected for the first wave of testing. Item response theory (IRT)-based analysis of 11 large data sets supplemented and informed item-level qualitative review of nearly 7,000 items from available PRO measures in the item library. Items were selected for rewriting or creation with further detailed review before the first round of testing in the general population and target patient populations. CONCLUSIONS: The NIH PROMIS network derived a consensus-based framework for self-reported health, systematically reviewed available instruments, and data sets that address the initial PROMIS domains. Qualitative item research led to the first wave of network testing, which began in the second year.

Dockery, D. W., Pope, C. A., Xu, X, Spengler, J. D., Ware, J. H., Fay, M. E., et al. (1993). An association between air pollution and mortality in six U.S. cities. *New England Journal of Medicine, 329*, 1753–1759.

BACKGROUND: Recent studies have reported associations between particulate air pollution and daily mortality rates. Population-based, cross-sectional studies of metropolitan areas in the United States have also found associations between particulate air pollution and annual mortality rates, but these studies have been criticized, in part because they did not directly control for cigarette smoking and other health risks. METHODS: In this prospective cohort study, we estimated the effects of air pollution on mortality, while controlling for individual risk factors. Survival analysis, including Cox proportional-hazards regression modeling, was conducted with data from a 14- to 16-year mortality follow-up of 8,111 adults in six U.S. cities. RESULTS: Mortality rates were most strongly associated with cigarette smoking. After adjusting for smoking and other risk factors, we observed statistically significant and robust associations between air pollution and mortality. The adjusted mortality-rate ratio for the most polluted of the cities as compared with the least polluted was 1.26 (95% confidence interval, 1.08 to 1.47). Air pollution was positively associated with death from lung cancer and cardiopulmonary disease but not with death from other causes considered together. Mortality was most strongly associated with air pollution with fine particulates, including sulfates. CONCLUSIONS: Although the effects of other, unmeasured risk factors cannot be excluded with certainty, these results suggest that fine-particulate air pollution, or a more complex pollution mixture associated with fine particulate matter, contributes to excess mortality in certain U.S. cities.

Dybul, M., Fauci, A. S., Bartlett, J. G., Kaplan, J. E., & Pau, A. K. (2002). Guidelines for using antiretroviral agents among HIV-infected adults and adolescents. *Annals of Internal Medicine*, *137*(5 Pt 2), 381–433.

The availability of an increasing number of antiretroviral agents and the rapid evolution of new information have introduced substantial complexity into treatment regimens for persons infected with human immunodeficiency virus (HIV). In 1996, the Department of Health and Human Services and the Henry J. Kaiser Family Foundation convened the Panel on Clinical Practices for the Treatment of HIV to develop guidelines for clinical management

of HIV-infected adults and adolescents (CDC. Report of the NIH Panel to Define Principles of Therapy of HIV Infection and Guidelines for the use of antiretroviral agents in HIV-infected adults and adolescents. MMWR. 1998;47[RR-5]:1-41). This report, which updates the 1998 guidelines, addresses (1) using testing for plasma HIV ribonucleic acid levels (i.e., viral load) and CD4+ T cell count; (2) using testing for antiretroviral drug resistance; (3) considerations for when to initiate therapy; (4) adherence to antiretroviral therapy; (5) considerations for therapy among patients with advanced disease; (6) therapy-related adverse events; (7) interruption of therapy; (8) considerations for changing therapy and available therapeutic options; (9) treatment for acute HIV infection; (10) considerations for antiretroviral therapy among adolescents; (11) considerations for antiretroviral therapy among pregnant women; and (12) concerns related to transmission of HIV to others. Antiretroviral regimens are complex, have serious side effects, pose difficulty with adherence, and carry serious potential consequences from the development of viral resistance because of nonadherence to the drug regimen or suboptimal levels of antiretroviral agents. Patient education and involvement in therapeutic decisions are critical. Treatment should usually be offered to all patients with symptoms ascribed to HIV infection. Recommendations for offering antiretroviral therapy among asymptomatic patients require analysis of real and potential risks and benefits. In general, treatment should be offered to persons who have <350 CD4+ T cells/mm3 or plasma HIV ribonucleic acid (RNA) levels of >55,000 copies/mL (by b-deoxyribonucleic acid [bDNA] or reverse transcriptase-polymerase chain reaction [RT-PCR] assays). The recommendation to treat asymptomatic patients should be based on the willingness and readiness of the person to begin therapy; the degree of existing immunodeficiency as determined by the CD4+ T cell count; the risk for disease progression as determined by the CD4+ T cell count and level of plasma HIV RNA; the potential benefits and risks of initiating therapy in an asymptomatic person; and the likelihood, after counseling and education, of adherence to the prescribed treatment regimen. Treatment goals should be maximal and durable suppression of viral load, restoration and preservation of immunologic function, improvement of quality of life, and reduction of HIV-related morbidity and mortality. Results of therapy are evaluated through plasma HIV RNA levels, which are expected to indicate a 1.0 log10 decrease at 2 to 8 weeks and no detectable virus (<50 copies/mL) at 4 to 6 months after treatment initiation. Failure of therapy at 4–6 months might be ascribed to nonadherence, inadequate potency of drugs or suboptimal levels of antiretroviral agents, viral resistance, and other factors that are poorly understood. Patients whose therapy fails in spite of a high level of adherence to the regimen should have their regimen changed; this change should be guided by a thorough drug treatment history and the results of drug-resistance testing. Because of limitations in the available alternative antiretroviral regimens that have documented efficacy, optimal changes in therapy might be difficult to achieve for patients in whom the preferred regimen has failed. *These decisions are further confounded by problems with adherence, toxicity, and resistance. For certain patients, participating in a clinical trial with or without access to*

*new drugs or using a regimen that might not achieve complete suppression of viral replication might be preferable. Because concepts regarding HIV management are evolving rapidly, readers should check regularly for additional information and updates at the HIV/AIDS Treatment Information Service website (*http://www.hivatis.org*).*

Eisenberg, M. S., Cummings, R. O., & Larsen, M. P. (1991). Numerators, denominators, and survival rates: Reporting survival from out-of-hospital cardiac arrest. *American Journal of Emergency Medicine*, *9*(6), 544–546.

This study demonstrates the effect of different denominators on the survival rate from out-of-hospital cardiac arrest. We retrospectively analyzed data from a cardiac arrest surveillance system in King County, Washington, during the years 1976 to 1988, and calculated survival rates using eight different definitions of denominators. The eight survival rates ranged from 16% to 49% discharge from hospital. The denominator for the lowest survival rate included all cases of cardiac arrest for whom emergency medical services personnel started cardiopulmonary resuscitation. The denominator for the highest survival rate included all cases of presumed cardiac etiology; first recorded rhythm was ventricular fibrillation; collapse witnessed; cardiopulmonary resuscitation started by bystanders within 4 minutes; and definitive care provided within 8 minutes. The definition of cases included in the denominator can dramatically affect the resultant survival rate. There must be national and international agreement about definitions of denominators for valid cross-community comparisons.

Gregg, E. W., Cadwell, B. L., Cheng, Y. J., Cowie, C. C., Williams, D. E., Geiss, L., et al. (2004). Trends in the prevalence and ratio of diagnosed to undiagnosed diabetes according to obesity levels in the U.S. *Diabetes Care, 27*(12), 2806–2812.

OBJECTIVE: To examine trends in the prevalence of diagnosed and undiagnosed diabetes and the proportion of total cases previously diagnosed, according to obesity status in the United States over the past 40 years. RESEARCH DESIGN AND METHODS: We assembled data from five consecutive cross-sectional national surveys: National Health Examination Survey I (1960–1962), National Health and Nutrition Examination Survey (NHANES) I (1971–1974), NHANES II (1976–1980), NHANES III (1988–1994), and NHANES 1999–2000. Diagnosed diabetes was ascertained, and height and weight were measured in adults aged 20–74 years in all surveys. In NHANES II, NHANES III, and NHANES 1999–2000, a fasting glucose level ≥126 mg/dl was used to identify cases among individuals not reporting diabetes. Design-based analyses and Bayesian models estimate the probability that prevalence of diabetes increased within four BMI groups (<25, 25–29, 30–34, and ≥35 kg/m2). RESULTS: In the U.S. population aged 20–74 years between 1976–1980 and 1999–2000, significant increases in the prevalence of diagnosed diabetes (3.3–5.8%, probability >99.9%) were accompanied by nonsignificant increases in undiagnosed diabetes (2.0–2.4%, 66.6%). This resulted in an increase in total diabetes (5.3–8.2%, >99.9%) and a modest nonsignificant increase in the proportion of cases that were diagnosed (62–70%,

62.4%). However, these trends varied considerably by BMI level. In individuals with BMI ≥35 kg/m2, diagnosed diabetes increased markedly (from 4.9% in 1960, to 8.6% during 1976–1980, to 15.1% in 1999–2000; probability >99.9%), whereas undiagnosed diabetes declined considerably (12.5% during 1976–1980 to 3.2% in 1999–2000, probability of increase 4.5%) Therefore, the proportion of total diabetes cases that were diagnosed increased from 41 to 83% (probability 99.9%) among individuals with BMI ≥35 kg/m2. By comparison, changes in prevalence within BMI strata <35 kg/m2 were modest, and there was no increase in the percentage of total cases that were diagnosed. CONCLUSIONS: National surveys over the past several decades have found large increases in diagnosed diabetes, particularly in overweight and obese individuals, but this has been accompanied by large decreases in undiagnosed diabetes only among individuals with BMI ≥35 kg/m2. This suggests that improvements in diabetes awareness and detection are most prominent among this subgroup.

Gregg, E. W., Gu, Q., Cheng, Y. J., Narayan, K. M., & Cowie C. C. (2007). Mortality trends in men and women with diabetes, 1971 to 2000. *Annals of Internal Medicine*, *147*(3), 149–155.

BACKGROUND: Whether mortality rates among diabetic adults or excess mortality associated with diabetes in the United States has declined in recent decades is not known. OBJECTIVE: To examine whether all-cause and cardiovascular disease mortality rates have declined among the U.S. population with and without self-reported diabetes. DESIGN: Comparison of three consecutive, nationally representative cohorts. SETTING: Population-based health surveys (National Health and Nutrition Examination Surveys I, II, and III) with mortality follow-up assessment. PATIENTS: Survey participants age 35 to 74 years with and without diabetes. MEASUREMENTS: Diabetes was determined by self-report for each survey (1971–1975, 1976–1980, and 1988–1994), and mortality rates were determined through 1986, 1992, and 2000 for the three surveys, respectively. RESULTS: Among diabetic men, the all-cause mortality rate decreased by 18.2 annual deaths per 1000 persons (from 42.6 to 24.4 annual deaths per 1000 persons; P = 0.03) between 1971 to 1986 and 1988 to 2000, accompanying decreases in the nondiabetic population. Trends for cardiovascular disease mortality paralleled those of all-cause mortality, with 26.4 annual deaths per 1000 persons in 1971 to 1986 and 12.8 annual deaths per 1,000 persons in 1988 to 2000 (P = 0.06). Among women with diabetes, however, neither all-cause nor cardiovascular disease mortality declined between 1971 to 1986 and 1988 to 2000, and the all-cause mortality rate difference between diabetic and nondiabetic women more than doubled (from a difference of 8.3 to 18.2 annual deaths per 1,000 persons). The difference in all-cause mortality rates by sex among people with diabetes in 1971 to 1986 were essentially eliminated in 1988 to 2000. LIMITATIONS: Diabetes was assessed by self-report, and statistical power to examine the factors explaining mortality trends was limited. CONCLUSIONS: Progress in reducing

mortality rates among persons with diabetes has been limited to men. Diabetes continues to greatly increase the risk for death, particularly among women.

Grodstein, F., Goldman, M. B., & Cramer, D. W. (1993). Relation of tubal infertility to history of sexually transmitted diseases. *American Journal of Epidemiology, 137*(5), 577–584.

We studied the history of sexually transmitted diseases in 283 nulliparous women diagnosed with infertility due to tubal adhesions or occlusion and 3,833 women admitted for delivery at seven collaborating hospitals. The adjusted risks of tubal infertility associated with the history of each sexually transmitted disease were estimated by the odds ratios obtained by multiple logistic regression. Women who reported prior infection with gonorrhea were at a significantly increased risk of tubal infertility (relative odds = 2.4, 95% confidence interval 1.3–4.4). In addition, the risk of tubal infertility was almost twice as high in women who recalled previous trichomoniasis compared with women with no such infection (relative odds = 1.9, 95% confidence interval 1.3–2.8). Furthermore, there was a trend of increasing risk with an increasing number of episodes of gonorrhea or trichomoniasis.

Hajjar, I., Kotchen, J. M., & Kotchen, T. A. (2006). Hypertension: Trends in prevalence, incidence, and control. *Annual Review of Public Health, 27*, 465–490.

Hypertension is the leading cause of cardiovascular disease worldwide. Prior to 1990, population data suggest that hypertension prevalence was decreasing; however, recent data suggest that it is again on the rise. In 1999–2002, 28.6% of the U.S. population had hypertension. Hypertension prevalence has also been increasing in other countries, and an estimated 972 million people in the world are suffering from this problem. Incidence rates of hypertension range between 3% and 18%, depending on the age, gender, ethnicity, and body size of the population studied. Despite advances in hypertension treatment, control rates continue to be suboptimal. Only about one-third of all hypertensives are controlled in the United States. Programs that improve hypertension control rates and prevent hypertension are urgently needed.

Heikkinen, T., & Jarvinen, A. (2003). The common cold. *Lancet*, *361*(9351), 51–59.

Despite great advances in medicine, the common cold continues to be a great burden on society in terms of human suffering and economic losses. Of the several viruses that cause the disease, the role of rhinoviruses is most prominent. About a quarter of all colds are still without proven cause, and the recent discovery of human metapneumovirus suggests that other viruses could remain undiscovered. Research into the inflammatory mechanisms of the common cold has elucidated the complexity of the virus-host relation. Increasing evidence is also available for the central role of viruses in predisposing to complications. New antivirals for the treatment of colds are being developed, but optimum use of these agents would require rapid detection of the specific virus causing the infection. Although

vaccines against many respiratory viruses could also become available, the ultimate prevention of the common cold seems to remain a distant aim.

Hogan, M. C., Foreman, K. J., Naghavi, M., Ahn, S. Y., Wang, M., Makela, S. M., et al. (2010). Maternal mortality for 181 countries, 1980–2008: A systematic analysis of progress towards Millennium Development Goal 5. *Lancet*.

BACKGROUND: Maternal mortality remains a major challenge to health systems worldwide. Reliable information about the rates and trends in maternal mortality is essential for resource mobilization, and for planning and assessment of progress towards Millennium Development Goal 5 (MDG 5), the target for which is a 75% reduction in the maternal mortality ratio (MMR) from 1990 to 2015. We assessed levels and trends in maternal mortality for 181 countries. METHODS: We constructed a database of 2651 observations of maternal mortality for 181 countries for 1980–2008, from vital registration data, censuses, surveys, and verbal autopsy studies. We used robust analytical methods to generate estimates of maternal deaths and the MMR for each year between 1980 and 2008. We explored the sensitivity of our data to model specification and show the out-of-sample predictive validity of our methods. FINDINGS: We estimated that there were 342,900 (uncertainty interval 302,100–394,300) maternal deaths worldwide in 2008, down from 526 300 (446,400–629,600) in 1980. The global MMR decreased from 422 (358–505) in 1980 to 320 (272–388) in 1990, and was 251 (221–289) per 100,000 live births in 2008. The yearly rate of decline of the global MMR since 1990 was 1.3% (1.0–1.5). During 1990–2008, rates of yearly decline in the MMR varied between countries, from 8.8% (8.7–14.1) in the Maldives to an increase of 5.5% (5.2–5.6) in Zimbabwe. More than 50% of all maternal deaths were in only six countries in 2008 (India, Nigeria, Pakistan, Afghanistan, Ethiopia, and the Democratic Republic of the Congo). In the absence of HIV, there would have been 281 500 (243,900–327,900) maternal deaths worldwide in 2008. INTERPRETATION: Substantial, albeit varied, progress has been made toward MDG 5. Although only 23 countries are on track to achieve a 75% decrease in MMR by 2015, countries such as Egypt, China, Ecuador, and Bolivia have been achieving accelerated progress. FUNDING: Bill & Melinda Gates Foundation.

Huang, W. Y., Winn, D. M., Brown, L. M., Gridley, G., Bravo-Otero, E., Diehl, S. R., et al. (2003). Alcohol concentration and risk of oral cancer in Puerto Rico. *American Journal of Epidemiology, 157*, 881–887.

Alcohol consumption is a major risk factor for cancers of the mouth and pharynx (oral cancer), but the differential risks by beverage type are unclear. In this 1992–1995 study, the authors examined oral cancer risk in Puerto Rico, comparing alcohol intake among 286 male cases aged 21 to 79 years and 417 population-based male controls, frequency matched by age. Heavy consumers of liquor (≥43 drinks per week) had strongly increased risks of oral cancer (odds ratio = 6.4, 95% confidence interval: 2.4, 16.8); beer/wine showed only modest effects. Among liquor drinkers, risks were consistently greater for

those who drank straight (undiluted) liquor than for those who usually drank mixed (diluted) liquor (odds ratio = 4.0, 95% confidence interval: 2.4, 6.7). Risks associated with combined exposure to tobacco were also more pronounced when subjects drank liquor straight. The elevated risks associated with drinking homemade rum were similar to those for other types of liquor. These results suggest that alcohol concentration is a risk factor for oral cancer independent of the total quantity of alcohol consumed.

Jager, K. J., Zoccali, C., Kramer, R., & Dekker, F. W. (2007). Measuring disease occurrence. *Kidney International, 72*(4), 412–415.

Different measures may be used to describe how often disease (or another health event) occurs in a population. Incidence expresses the development of new cases and is mostly used against the background of prevention, to assess disease etiology or to determine the risk factors of disease. Depending on the specific study question, incidence may be reported as risk or as incidence rate. This paper discusses that it is preferable to use incidence rate in case of a dynamic population or in cases where the observation period is sufficiently long for competing risks or loss to follow-up to play a significant role. Prevalence is the number of existing cases, which is affected by both the number of incident cases and the length of disease time. It reflects the burden of disease on a population that may, among others, be measured in terms of costs or morbidity. Knowledge about this burden can be used for the planning of health-care facilities. This paper discusses the different measures of disease occurrence using a number of examples taken from the nephrology literature.

Manuel, D. G., & Schultz, S. E. (2004). Health-related quality of life and health-adjusted life expectancy of people with diabetes in Ontario, Canada, 1996–1997. *Diabetes Care*, *27*(2), 407–414.

OBJECTIVE: To estimate the burden of illness from diabetes using a population health survey linked to a population-based diabetes registry. RESEARCH DESIGN AND METHODS: Measures of health-related quality of life (HRQoL) from the 1996/1997 Ontario Health Survey (n = 35,517) were combined with diabetes prevalence and mortality data from the Ontario Diabetes Database (n = 487,576) to estimate the impact of diabetes on life expectancy, health-adjusted life expectancy (HALE), and HRQoL. RESULTS: Life expectancy of people with diabetes was 64.7 and 70.7 years for men and women, respectively—12.8 and 12.2 years less than that for men and women without diabetes. Diabetes had a large impact on instrumental and basic activities of daily living, more so than on functional health. HALE was 58.3 and 62.7 years, respectively, for men and women—11.9 and 10.7 years less than that of men and women without diabetes. Eliminating diabetes would increase Ontario life expectancy by 2.8 years for men and 2.6 years for women; HALE would increase by 2.7 and 3.2 years for men and women, respectively. CONCLUSIONS: The burden of illness from diabetes in Ontario is considerable. Efforts to reduce diabetes would likely result in a "compression of morbidity." An approach of estimating diabetes burden using linked data sources provides a robust approach for the surveillance of diabetes.

Mathers, C. D., Sadana, R., Salomon, J. A., Murray, C. J., & Lopez, A. D. (2001). Healthy life expectancy in 191 countries, 1999. *Lancet*, *357*(9269), 1685–1691.

We describe here the methods used to produce the first estimates of healthy life expectancy (DALE) for 191 countries in 1999. These were based on estimates of the incidence, prevalence, and disability distributions for 109 disease and injury causes by age group, sex, and region of the world, and an analysis of 60 representative health surveys across the world. We used Sullivan's method to compute healthy life expectancy for men and women in each WHO member country. Japan had the highest average healthy life expectancy of 74.5 years at birth in 1999. The bottom 10 countries are all in sub-Saharan Africa, where the HIV-AIDS epidemic is most prevalent, resulting in DALE at birth of less than 35 years. Years of healthy life lost due to disability represent 18% of total life expectancy in the bottom countries and decreases to around 8% in the countries with the highest healthy life expectancies. Globally, the male-female gap is lower for DALE than for total life expectancy. Healthy life expectancy increases across countries at a faster rate than total life expectancy, suggesting that reductions in mortality are accompanied by reductions in disability. Although women live longer, they spend a greater amount of time with disability. As average levels of health expenditure per capita increase, healthy life expectancy increases at a greater rate than total life expectancy.

McKenna, M. T., & Zohrabian, A. (2009). U.S. burden of disease—Past, present and future. *Annals of Epidemiology, 19*(3), 212–219.

PURPOSE: To review the history and challenges of "burden of disease" studies, how these are dependent on robust epidemiologic data as well as complex conceptual constructions, and to identify the public health policy issues these studies can most usefully inform. METHODS: The emergence of the concept of the "burden of disease" in the public health literature is reviewed, with a focus on the results of an analysis of data from the United States that used the methodology presented in the Global Burden of Disease Study. RESULTS: The systematic analysis of public health mortality data to identify major health problems was conducted by Graunt in 16th-century London. He found that many of the predominant sources of mortality were not the focus of public attention. Today, despite refinements in epidemiologic measurement methods designed to capture the impact of non-fatal health conditions, there are similar incongruities between the major public health problems and expenditures on prevention interventions. CONCLUSIONS: Controversies surrounding the interpretation of "burden of disease" studies are not new. Particularly in developed countries, these studies appear more useful for setting research priorities than for allocating resources to support prevention efforts. Such investigations are not possible without ongoing support for systematic collection and analysis of descriptive epidemiologic data.

Miller, B. G., & Hurley, J. F. (2003). Life table methods for quantitative impact assessments in chronic mortality. *Journal of Epidemiology and Community Health*, *57*(3), 200–206.

Quantitative health impact assessments of chronic mortality, where the impacts are expected to be observed over a number of years, are complicated by the link between death rates and surviving populations. A general calculation framework for quantitative impact assessment is presented, based on standard life table calculation methods, which permits consistent future projections of impacts on mortality from changes in death rates. Implemented as a series of linked spreadsheets, the framework offers complete flexibility in the sex-specific, age-specific, and year-specific patterns of baseline mortality death rates; in the predicted impacts upon these; in the weights or values placed on gains in life; and in the summary measures of impact. Impacts can be differential by cause of death. Some examples are given of predictions of the impacts of reductions in chronic mortality in the populations of England and Wales and of Scotland.

Mokdad, A. H., Ford, E. S., Bowman, B. A., Dietz, W. H., Vinicor, F., Bales, V. S., et al. (2003). Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. *Journal of the American Medical Association, 289*(1), 76–79.

CONTEXT: Obesity and diabetes are increasing in the United States. OBJECTIVE: To estimate the prevalence of obesity and diabetes among U.S. adults in 2001. DESIGN, SETTING, AND PARTICIPANTS: Random-digit telephone survey of 195,005 adults aged 18 years or older residing in all states participating in the Behavioral Risk Factor Surveillance System in 2001. MAIN OUTCOME MEASURES: Body mass index (BMI), based on self-reported weight and height and self-reported diabetes. RESULTS: In 2001, the prevalence of obesity (BMI ≥30) was 20.9% vs. 19.8% in 2000, an increase of 5.6%. The prevalence of diabetes increased to 7.9% vs. 7.3% in 2000, an increase of 8.2%. The prevalence of BMI of 40 or higher in 2001 was 2.3%. Overweight and obesity were significantly associated with diabetes, high blood pressure, high cholesterol, asthma, arthritis, and poor health status. Compared with adults with normal weight, adults with a BMI of 40 or higher had an odds ratio (OR) of 7.37 (95% confidence interval [CI], 6.39–8.50) for diagnosed diabetes, 6.38 (95% CI, 5.67–7.17) for high blood pressure, 1.88 (95% CI,1.67–2.13) for high cholesterol levels, 2.72 (95% CI, 2.38–3.12) for asthma, 4.41 (95% CI, 3.91–4.97) for arthritis, and 4.19 (95% CI, 3.68–4.76) for fair or poor health. CONCLUSIONS: Increases in obesity and diabetes among U.S. adults continue in both sexes, all ages, all races, all educational levels, and all smoking levels. Obesity is strongly associated with several major health risk factors.

Morgenstern, H., Kleinbaum, D. G., & Kupper, L. L. (1980). Measure of disease incidence used in epidemiological research. *International Journal of Epidemiology, 9*, 97–104.

This paper distinguishes between two concepts for measuring the incidence of disease: risk and rate. Alternative procedures for estimating these measures from epidemiologic data are

reviewed and illustrated. An attempt is made to integrate statistical principles with epidemiologic methods while minimizing the use of higher mathematics. Several theoretical and practical criteria are discussed for choosing the appropriate incidence measure in the planning of a study and for selecting the best method of estimation in the analysis.

Murray, C. J., & Lopez, A. D. (1999). On the comparable quantification of health risks: lessons from the Global Burden of Disease Study. *Epidemiology*, *10*(5), 594–605.

Extensive discussion and comments on the Global Burden of Disease Study findings have suggested the need to examine more carefully the basis for comparing the magnitude of different health risks. Attributable burden can be defined as the difference between burden currently observed and burden that would have been observed under an alternative population distribution of exposure. Population distributions of exposure may be defined over many different levels and intensities of exposure (such as systolic or diastolic blood pressure on a continuous scale), and the comparison distribution of exposure need not be zero. Avoidable burden is defined as the reduction in the future burden of disease if the current levels of exposure to a risk factor were reduced to those specified by the counterfactual distribution of exposure. Choosing the alternative population distribution for a variable, the counterfactual distribution of exposure, is the critical step in developing a more general and standardized concept of comparable, attributable, or avoidable burden. We have identified four types of distributions of exposure that could be used as the counterfactual distributions: theoretical minimum risk, plausible minimum risk, feasible minimum risk, and cost-effective minimum risk. Using tobacco and alcohol as examples, we explore the implications of using these different types of counterfactual distributions to define attributable and avoidable burden. The 10 risk factor assessments included in the Global Burden of Disease Study reflect a range of methods and counterfactual distributions. We recommend that future assessments should focus on avoidable and attributable burden based on the plausible minimum risk counterfactual distribution of exposure.

Murray, C. J., Salomon, J. A., & Mathers, C. (2000). A critical examination of summary measures of population health. *Bulletin of the World Health Organization*, *78*(8), 981–994.

In the past decade, interest has been rising in the development, calculation, and use of summary measures of population health, which combine information on mortality and non-fatal health outcomes. This paper reviews the issues and challenges in the design and application of summary measures and presents a framework for evaluating different alternatives. Summary measures have a variety of uses, including comparisons of health in different populations and assessments of the relative contributions of different diseases, injuries, and risk factors to the total disease burden in a population. Summary measures may be divided into two broad families: health expectancies and health gaps. Within each family, there are many different possible measures, but they share a number of inputs, including information on mortality, non-fatal health outcomes, and health state valuations.

Other critical points include calculation methods and a range of conceptual and methodological issues regarding the definition, measurement, and valuation of health states. This paper considers a set of basic criteria and desirable properties that may lead to rejection of certain summary measures and the development of new ones. Despite the extensive developmental agenda that remains, applications of summary measures cannot await the final resolution of all methodological issues, so they should focus on those measures that satisfy as many basic criteria and desirable properties as possible.

Parker, S. L., Tong, T., Bolden, S., & Wingo, P. A. (1996). Cancer statistics, 1996. *Canadian Cancer Journal for Clinicians, 46*(1), 5–27.

The American Cancer Society's Department of Epidemiology and Statistics reports its 30th annual compilation of cancer incidence, survival, and mortality data for the United States and around the world.

Parkin, D. M., Pisani, P., & Ferlay, J. (1999). Global cancer statistics. *Canadian Cancer Journal for Clinicians, 49*(1), 33–64.

Statistics are given for global patterns of cancer incidence and mortality for males and females in 23 regions of the world.

Pinto-Prades, J. L. & Abellan-Perpinan, J. M. (2005). Measuring the health of populations: The veil of ignorance approach. *Health Economics*, *14*(1), 69–82.

We report the results from two surveys designed to explore whether an application of Harsanyi's principle of choice form behind a veil of ignorance (VEI) can be used to measure the health of populations. This approach was tentatively recommended by Murray et al. (Bull World Health Organ, 2000; 78:981–994; Summary Measures of population health: Concepts, Ethics, Measurement and Applications, WHO, 2002) as an appropriate way of constructing summary measures of population health (SMPH) for comparative purposes. The operationalization of the VEI approach used in this paper was suggested by Nord (Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications, WHO, 2002). We test if VEI and person trade-off (PTO) methods generate similar quality-of-life weights. In addition, we compare VEI and PTO weights with individual utilities estimated by means of the conventional standard gamble (SG) and a variation of it we call double gamble. Finally, psychometric properties like feasibility, reliability, and consistency are examined. Our main findings are next: (1) VEI and PTO approaches generate very different weights; (2) it seems that differences between PTO and VEI are not due to the rule of rescue; (3) the VEI resembled more a DG than a classical SG; (4) PTO, VEI, and DG exhibited good feasibility, reliability, and consistency.

Robine, J. M., & Ritchie, K. (1991). Healthy life expectancy: evaluation of global indicator of change in population health. *British Medical Journal*, *302*(6774), 457–460.

OBJECTIVE: To review and evaluate the usefulness of healthy life expectancy as a global indicator of changes in a population's health. DESIGN: Review of all known studies to date from the United States, mainland Europe, Canada, and the United Kingdom that have used Sullivan's method of calculating disability free life expectancy. MAIN OUTCOME MEASURES: Life expectancy and disability free life expectancy. RESULTS: Over the past decade, the average healthy life expectancy was 60 years for men and 64 for women, with the proportion of years of disability ranging from 11% to 21% in men and from 14% to 24% in women. At the age of 65, men could expect 8 years of disability free life and women 10, with the life expectancy being 14 and 19 years, respectively. The difference between the wealthiest and poorest income quintiles was 6.3 years in life expectancy and 14.3 in disability free life expectancy for men and 2.8 and 7.6 respectively for women. These results suggest that disparities in health are greater between social groups than between the sexes. Diseases affect mortality and morbidity differently. The order of importance for affecting life expectancy was circulatory disease, cancer, and accidents and for disability free life expectancy, circulatory disease, locomotor disorders, and respiratory disorders. CONCLUSIONS: Healthy life expectancy is a valuable index for the appreciation of changes in both the physical and the mental health states of the general population, for allocating resources, and for measuring the success of political programmes. Future calculations should also take into account the probability of recovery and thus extend the applicability of the indicator to populations in poor health rather than focusing on the well population.

Rockhill, B. (2005). Theorizing about causes at the individual level while estimating effects at the population level: Implications for prevention. *Epidemiology, 16*(1), 124–129.

The dominant philosophy of modern epidemiology is individualism, despite the limitations of epidemiologic tools and methods when considering the individual level. We pursue information on increasingly reductionist causes in our search for knowledge of causes of specific cases. Philosophical reasoning and empiric evidence suggest that this search may not be as fruitful as proponents claim. I argue that using individualism to guide our search for causes of disease hinders our effectiveness in contributing to disease prevention, because the positive predictive values of most established genetic and environmental risk factors for noninfectious diseases are too low to be quantitatively convincing to an individual.

Thompson, M. L., Myers, J. E., & Kriebel, D. (1998). Prevalence odds ratio or prevalence ratio in the analysis of cross sectional data: What is to be done? *Occupational and Environmental Medicine, 55*(4), 272–277.

OBJECTIVES: To review the appropriateness of the prevalence odds ratio (POR) and the prevalence ratio (PR) as effect measures in the analysis of cross-sectional data and to

evaluate different models for the multivariate estimation of the PR. METHODS: A system of linear differential equations corresponding to a dynamic model of a cohort with a chronic disease was developed. At any point in time, a cross-sectional analysis of the people then in the cohort provided a prevalence-based measure of the effect of exposure on disease. This formed the basis for exploring the relations between the POR, the PR, and the incidence rate ratio (IRR). Examples illustrate relations for various IRRs, prevalences, and differential exodus rates. Multivariate point and interval estimation of the PR by logistic regression is illustrated and compared with the results from proportional hazards regression (PH) and generalised linear modelling (GLM). RESULTS: The POR is difficult to interpret without making restrictive assumptions and the POR and PR may lead to different conclusions with regard to confounding and effect modification. The PR is always conservative relative to the IRR and, if PR > 1, the POR is always > PR. In a fixed cohort and with an adverse exposure, the POR is always ≥ IRR, but in a dynamic cohort with sufficient underlying follow up the POR may overestimate or underestimate the IRR, depending on the duration of follow up. Logistic regression models provide point and interval estimates of the PR (and POR) but may be intractable in the presence of many covariates. Proportional hazards and generalised linear models provide statistical methods directed specifically at the PR, but the interval estimation in the case of PH is conservative and the GLM procedure may require constrained estimation. CONCLUSIONS: The PR is conservative, consistent, and interpretable relative to the IRR and should be used in preference to the POR. Multivariate estimation of the PR should be executed by means of generalised linear models or, conservatively, by proportional hazards regression.

Tu, K., Chen, Z., & Lipscombe, L. L. (2008). Prevalence and incidence of hypertension from 1995 to 2005: A population-based study. *Canadian Medical Association Journal, 178*, 1429–1435.

BACKGROUND: We have reported that the prevalence of diagnosed hypertension increased by 60% from 1995 to 2005 in Ontario. In the present study, we asked whether this increase is explained by a decrease in the mortality rate. METHODS: We performed a population-based cohort study using linked administrative data for Ontario, a Canadian province with over 12 million residents. We identified prevalent cases of hypertension using a validated case-definition algorithm for hypertension, and we examined trends in mortality from 1995 to 2005 among adults aged 20 or older with hypertension. RESULTS: The age- and sex-adjusted mortality among patients with hypertension decreased from 11.3 per 1,000 people in 1995 to 9.6 per 1,000 in 2005 ($p < 0.001$), which is a relative reduction of 15.5%. We found that the relative decrease in age-adjusted mortality was higher among men than among women (−22.2% vs. −7.3%, $p < 0.001$). INTERPRETATION: Mortality rates among patients with hypertension have decreased. Along with an increasing incidence, decreased mortality rates may contribute to the increased prevalence of diagnosed hypertension. Sex-related discrepancies in the reduction of mortality warrant further investigation.

Tugwell, P., de Savigny, D., Hawker, G., & Robinson, V. (2006). Applying clinical epidemiological methods to health equity: The equity effectiveness loop. *British Medical Journal*, *332*(7537), 358–361.

Focusing on the average effects of interventions on health may miss important differences within populations. Examining these effects across gradients in wealth allows the identification of the interventions most likely to reduce health inequalities.

Yang, P., Cehran, J. R., & Vierkant, R. A. (2002). Adenocarcinoma of the lung is strongly associated with cigarette smoking: Further evidence from a prospective study of women. *American Journal of Epidemiology, 156*, 1114–1122.

In a prospective cohort of 41,836 Iowa women aged 55 to 69 years with 13 years of follow-up from 1986 through 1998, the authors examined the association between cigarette smoking history and three common histologic subtypes of lung cancer (123 small cell, 115 squamous cell, and 234 adenocarcinoma). Using Cox proportional hazards and additive Poisson regression analysis, they estimated four epidemiologic measures of effect: age-adjusted incidence rate, relative risk, excess risk (or risk difference), and population attributable risk. Of the three major lung cancer subtypes, the excess risk for heavy smokers compared with never smokers was higher for adenocarcinoma (excess risk = 206) than for squamous cell (excess risk = 122) and small cell (excess risk = 104) carcinomas. Adenocarcinoma of the lung is more strongly associated with tobacco smoke exposure than previously recognized.

2. Economic Measures

Abelson, P. (2003). The value of life and health for public policy. *The Economic Record*, *79*, S2–S13.

Expenditure on health and safety is a substantial part of gross domestic product (GDP), but public agencies in many countries, including Australia, have only qualitative views about the value of life and health. Also, despite considerable work by economists on the value of life and health in recent years, some important issues, such as the value of a healthy life-year, remain unresolved. This paper presents a framework for valuing life and health. It then draws on international and Australian research to estimate possible values for life, healthy life-years, and various chronic and acute health states for public policy purposes in Australia.

Akobundu, E., Ju, J., Blatt, L., & Mullins, C. D. (2006). Cost-of-illness studies: A review of current methods. *Pharmacoeconomics, 24*(9), 869–890.

The number of cost-of-illness (COI) studies has expanded considerably over time. One outcome of this growth is that the reported COI estimates are inconsistent across studies, thereby raising concerns over the validity of the estimates and methods. Several factors have been identified in the literature as reasons for the observed variation in COI estimates.

To date, the variation in the methods used to calculate costs has not been examined in great detail even though the variations in methods are a major driver of variation in COI estimates. The objective of this review was to document the variation in the methodologies employed in COI studies and to highlight the benefits and limitations of these methods. The review of COI studies was implemented following a four-step procedure: (i) a structured literature search of MEDLINE, JSTOR, and EconLit; (ii) a review of abstracts using predefined inclusion and exclusion criteria; (iii) a full-text review using predefined inclusion and exclusion criteria; and (iv) classification of articles according to the methods used to calculate costs. This review identified four COI estimation methods (Sum_All Medical, Sum_Diagnosis Specific, Matched Control and Regression) that were used in categorizing articles. Also, six components of direct medical costs and five components of indirect/non-medical costs were identified and used in categorizing articles; 365 full-length articles were reflected in the current review following the structured literature search. The top five cost components were emergency room/inpatient hospital costs, outpatient physician costs, drug costs, productivity losses, and laboratory costs. The dominant method, Sum_Diagnosis Specific, was a total costing approach that restricted the summation of medical expenditures to those related to a diagnosis of the disease of interest. There was considerable variation in the methods used within disease subcategories. In several disease subcategories (e.g., asthma, dementia, diabetes mellitus), all four estimation methods were represented, and in other cases (e.g., HIV/AIDS, obesity, stroke, urinary incontinence, schizophrenia), three of the four estimation methods were represented. There was also evidence to suggest that the strengths and weaknesses of each method were considered when applying a method to a specific illness. Comparisons and assessments of COI estimates should consider the method used to estimate costs both as an important source of variation in the reported COI estimates and as a marker of the reliability of the COI estimate.

Barnett, P. G. (2009). An improved set of standards for finding costs for cost effectiveness analysis. *Medical Care*, *47*(Suppl 7), S76–S81.

BACKGROUND: Guidelines have helped standardize methods of cost-effectiveness analysis, allowing different interventions to be compared and enhancing the generalizability of study findings. There is agreement that all relevant services be valued from the societal perspective using a long-term time horizon and that more exact methods be used to cost services most affected by the study intervention. Guidelines are not specific enough with respect to costing methods, however. METHOD: The literature was reviewed to identify the problems associated with the four principal methods of cost determination. FINDINGS: Microcosting requires direct measurement and is ordinarily reserved to cost novel interventions. Analysts should include nonwage labor cost; person-level and institutional overhead; and the cost of development, set-up activities, supplies, space, and screening. Activity-based cost systems have promise of finding accurate costs of all services provided, but are not widely adopted. Quality must be evaluated, and the generalizability of cost

estimates to other settings must be considered. Administrative cost estimates, chiefly cost-adjusted charges, are widely used, but the analyst must consider items excluded from the available system. Gross costing methods determine quantity of services used and employ a unit cost. If the intervention will affect the characteristics of a service, the method should not assume that the service is homogeneous. CONCLUSIONS: Questions are posed for future reviews of the quality of costing methods. The analyst must avoid inappropriate assumptions, especially those that bias the analysis by exclusion of costs that are affected by the intervention under study.

Basu, A., Arondekar, B. V., & Rathouz, P. J. (2006). Scale of interest versus scale of estimation: Comparing alternative estimators for the incremental costs of a comorbidity. *Health Economics*, *15*(10), 1091–1107.

We investigate how the scale of estimation in risk-adjustment models for health-care costs affects the covariate effect, where the scale of interest for the covariate effect may be different from the scale of estimation. As an illustrative example, we use claims data to estimate the incremental costs associated with heart failure within 1 year subsequent to myocardial infarction. Here, the scale of interest for the effect of heart failure on costs is additive. However, traditional methods for modeling costs use predetermined scale of estimation—for example, ordinary least squares (OLS) regression assumes an additive scale, while log-transformed OLS and generalized linear models with log-link assume a multiplicative scale of estimation. We compare these models with a new flexible model that lets the data determine the appropriate scale of estimation. We use a variety of goodness-of-fit measures along with a modified Copas test to assess robustness, lack of fit, and over-fitting properties of the alternative estimators. Biases up to 19% in the scale of interest are observed due to the misrepresentation of the scale of estimation. The new flexible model is found to appropriately represent the scale of estimation and less susceptible to over-fitting despite estimating additional parameters in the link and the variance functions.

Basu, A., & Manning, D. G. (2009). Issues for the next generation of health care cost analyses. *Medical Care, 47*(7 Suppl 1), S109–S114.

BACKGROUND: Given the characteristics of health care expenditure/cost data, a mass of observations at zero, and skewed positive expenditures, various alternative estimators have been developed that can address the analytical issues these characteristics raise. The field continues to develop new approaches and to evaluate the performance of the existing ones. OBJECTIVES: We discuss the strengths and limitations in existing methods for estimation and for model specification and checking. We suggest some areas that need fuller development or a better understanding of how the estimation approach performs when the outcome exhibits the skewness and heavy right tails that are typical of health care data. We also address various other aspects of cost analysis that include dealing with induced censoring, estimating casual effects, and generating reliable predictions that may apply to many studies. RESULTS: No current method is optimal or dominant for all cost applications.

Many of the diagnostics used in choosing among alternatives have limitations that need more careful study. Several avenues in modeling cost data remain unexplored. CONCLUSIONS: Taken together, we hope that this essay would serve as a guide to the choice among methods and to the next generation of methodological research in this field.

Bloom, B. S., Bruno, D. J., Maman, D. Y, & Jayadevappa, R. (2001). Usefulness of U.S. cost-of-illness studies in healthcare decision making. *Pharmacoeconomics, 19*(2), 207–213.

OBJECTIVE: Cost-of-illness studies have been completed on scores of diseases over the past 30 years. The goal of this study was to review published cost-of-illness studies on U.S. populations in order to evaluate the potential usefulness of the results in decision making. METHODS: Medline and related databases were searched using diagnosis and economic terms. The bibliographies of the articles found were reviewed visually to identify further studies. Inclusion criteria required a specified diagnosis, the study to be published between 1 January 1985 and 30 April 1999 in an English-language peer-reviewed journal, a clearly defined U.S. sample or national population, available and recent epidemiological data on prevalence and incidence of diagnosis, and money estimates of direct and/or indirect costs. Three readers reviewed each study. The senior reviewer settled all differences. RESULTS: Searches found 1725 published studies; only 110 (6.4%) met all inclusion criteria. Main reasons for rejection were insufficient cost data (80%), insufficient information on data sources and aggregation or estimation methods (56%), inadequate sector data (e.g., hospitalizations or work loss) (48%), study of value, not cost, of illness (44%), not a U.S. population (30%) and insufficient population detail (19%). There were 80 diagnosis categories, 28 of which had more than one study. Only 5 diagnoses had ≥ 5 studies: Alzheimer's dementia, depression, diabetes mellitus, mental illness and stroke. Multifold cost variations were found among studies within diagnosis categories, even with the same method and data sources. The more narrowly defined diagnoses, depression and stroke, had the smallest cost variation at 41.7 and 17.2%, respectively. A generalized linear regression model found that a significant portion of total and direct cost variance could be explained only for Alzheimer's dementia. CONCLUSIONS: The wide variation of cost estimates for the same diagnosis raises serious questions of comparability, accuracy, validity, and usefulness of all studies. Implementing guidelines to standardize methods and study design for cost-of-illness studies would be a worthwhile first step. The advantages and disadvantages of using money or another metric such as disability-adjusted life-years as the prime outcome measure should also be publicly discussed.

Buntin, M. B., & Zaslavsky, A. M. (2004). Too much ado about two-part models and transformation? Comparing methods of modeling Medicare expenditures. *Journal of Health Economics, 23*(3), 525–542.

Many methods for modeling skewed health care cost and use data have been suggested in the literature. This paper compares the performance of eight alternative estimators,

including OLS and GLM estimators and one- and two-part models, in predicting Medicare costs. It finds that four of the alternatives produce very similar results in practice. It then suggests an efficient method for researchers to use when selecting estimators of health care costs.

Diehr, Yanez, P., Ash, A. Hornbrook, M., & Lin, D. Y. (1999). Methods for analyzing health care utilization and costs. *Annual Review of Public Health, 20*, 125–144.

Important questions about health care are often addressed by studying health care utilization. Utilization data have several characteristics that make them a challenge to analyze. In this paper, we discuss sources of information, the statistical properties of utilization data, common analytic methods including the two-part model, and some newly available statistical methods including the generalized linear model. We also address issues of study design and new methods for dealing with censored data. Examples are presented.

Gilleskie, D. B., & Mroz, T. A. (2004). A flexible approach for estimating the effects of covariates on health expenditures. *Journal of Health Economics, 23*(2), 391–418.

Our estimation strategy uses sequences of conditional probability functions, similar to those used in discrete time hazard rate analyses, to construct a discrete approximation to the density function of an outcome of interest conditional on exogenous explanatory variables. Once the conditional density function has been constructed, we can examine expectations of arbitrary functions of the outcome of interest and evaluate how these expectations vary with observed exogenous covariates. We demonstrate the features and precision of the conditional density estimation method (and compare it to other commonly used methods) through Monte Carlo experiments and an application to health expenditures using the RAND Health Insurance Experiment data. Overall, we find that the approximate conditional density estimator provides accurate and precise estimates of derivatives of expected outcomes for a wide range of types of explanatory variables.

Goeree, R., O'Brien, B. J., Blackhouse, G., Agro, K., & Goering, P. (1999). The valuation of productivity costs due to premature mortality: A comparison of the human-capital and friction-cost methods for schizophrenia. *Canadian Journal of Psychiatry, 44*(5), 455–463.

OBJECTIVE: To compare productivity-cost estimates for schizophrenia-related premature mortality in Canada in 1996 using the human-capital (HC) approach and friction-cost (FC) method. METHODS: The number of deaths directly attributed to schizophrenia was combined with the estimated number of deaths attributable to schizophrenia from suicide. These premature deaths were valued using two methods: (1) the traditional HC approach, based on "potential" lost output to normal age of retirement; and (2) the FC method, based on finding a replacement worker. RESULTS: In 1996, there were 342 male and female preretirement deaths attributed to schizophrenia, directly or indirectly by suicide, in Canada. Most deaths were in males (78%) and by suicide (97%). The productivity cost of these

deaths was estimated to be $105 million using the HC approach but only $1.53 million using the FC method. CONCLUSIONS: Productivity-cost estimates from the HC approach are substantially higher than those obtained from the FC method (69 times higher). In circumstances of unemployment, the HC approach is an overestimate of future productivity losses for premature mortality.

Grosse, S. D., Krueger, K. V., & Mvundura, M. (2009). Economic productivity by age and sex: 2007 estimates for the United States. *Medical Care, 47*(7 Suppl 1), S94–S103.

BACKGROUND: Human capital estimates of labor productivity are often used to estimate the economic impact of diseases and injuries that cause incapacitation or death. OBJECTIVES: Estimates of average hourly, annual, and lifetime economic productivity, both market and household, were calculated in 2007 U.S. dollars for 5-year age groups for men, women, and both sexes in the United States. RESEARCH DESIGN: Data from the American Time Use Survey were used to estimate hours of paid work and household services and hourly and annual earnings and household productivity. Present values of discounted lifetime earnings were calculated for each age group using the 2004 U.S. life tables and a discount rate of 3% per year and assuming future productivity growth of 1% per year. SUBJECTS: The estimates of hours and productivity were calculated using the time diaries of 72,922 persons included in the American Time Use Survey for the years 2003 to 2007. RESULTS: The present value of lifetime productivity is approximately $1.2 million in 2007 dollars for children under 5 years of age. For adults in their 20s and 30s, it is approximately $1.6 million and then it declines with increasing age. Productivity estimates are higher for males than for females, more for market productivity than for total productivity. CONCLUSIONS: Changes in hours of paid employment and household services can affect economic productivity by age and sex. This is the first publication to include estimates of household services based on contemporary time use data for the U.S. population.

Heffler, S., Nuccio, O., & Freeland, M. (2009). An overview of the NHEA with implications for cost analysis researchers. *Medical Care, 47*(7 Suppl 1), S37–S43.

BACKGROUND/OBJECTIVE: The National Health Expenditure Accounts (NHEA) are the official government estimates of aggregate U.S. health care spending. We summarize the data sources, methods, strengths, limitations, and applications of the NHEA. METHODS: To compile this article, we provide background on the NHEA; a description of the data sources and methods used to produce them; some recent findings that the NHEA produced; and a discussion of their strengths, limitations, and applications drawn from several different sources, both internal and external to the Centers for Medicare & Medicaid Services. RESULTS: The NHEA have a multitude of applications, including comparison with other economic data such as the Gross Domestic Product, reconciliation with other health spending data sources, and use in predictive and analytic models. The NHEA adhere to national income accounting standards and are comprehensive, mutually exclusive,

multidimensional, and consistent over time. The NHEA do not contain micro-level detailed data and are subject to both sampling and nonsampling errors during the interim census years, although this is the case for all available data sources. CONCLUSIONS: Determining the correct method for measuring health care costs depends on one's purpose, and analysis of health care cost data that requires aggregate-level statistics should consider use of the NHEA.

Hodgson, T. A. (1994). Costs of illness in cost-effectiveness analysis. A review of the methodology. *Pharmacoeconomics*, *6*(6), 536–552.

Costs of illness are an important input in cost-effectiveness analysis (CEA). Reviews of the literature have found that many CEAs are of low technical quality and fail to take account of costs of illness appropriately. The costs of illness and disease averted by an intervention, indirect costs, and medical care costs in added years of life are topics that present methodological issues and are not handled consistently in CEAs. Costs of illness and disease averted may be estimated by prevalence- or incidence-based methods; the correct conceptual paradigm depends on the nature of the disease. Incidence costs may be estimated by modeling the disease process, or directly from prevalence costs, the choice being determined by the extent and quality of data available. Regardless of the method, in forward-looking CEAs potential technological change must be taken into account so that incidence-based lifetime costs estimated from current treatment practices will not be biased. Whether to include indirect costs is an important issue, because indirect costs may be large and have a significant impact on the cost-effectiveness ratio. In the pure CEA model, indirect costs are excluded on ethical grounds and to prevent incursion of elements of cost-benefit analysis into CEA. The modified CEA model accepts enhanced productivity as an economic benefit made possible by, but distinct from, the health effect of an intervention. Indirect costs are included when appropriate, depending on the perspective of the analysis, the measure of effectiveness, and who bears the costs. When medical care extends life, expenditures will be incurred in the added years for illness and disease unrelated to the intervention. As with indirect costs, the pure CEA considers unrelated downstream costs an indirect consequence of the health benefit of the intervention and excludes them from CEAs with the societal perspective. The modified CEA treats unrelated downstream costs as an economic effect of the change in health due to the intervention and includes them in order to have a more complete accounting of the cost of the intervention.

Hodgson, T. A., & Meiners, M. R. (1982). Cost-of-illness methodology: A guide to current practices and procedures. *Milbank Memorial Fund Quarterly Health and Society, 60*(3), 429–462.

Estimating the cost burdens of illness in society is essential to governmental and private health policy decisions; cost-benefit and cost-effectiveness analyses, technology assessment, and agency budgets rely on such estimates. Direct cost, loss in output cost,

and psychosocial cost have no standard measurements or methods of estimation. Alternative approaches are reviewed and improvements suggested.

Hodgson, T. A., & Cohen, A. J. (1999). Medical care expenditures for diabetes, its chronic complications, and its comorbidities. *Preventive Medicine, 29*, 173–186.

BACKGROUND: Medical expenditures for diabetes are estimated, including expenditures for chronic complications of diabetes, unrelated conditions for which diabetics are at higher risk, and various comorbidities that raise the cost of medical care. METHODS: A variety of national data sources are used to disaggregate the Health Care Financing Administration's national health expenditures in 1995 by sex, age, and diagnosis. Expenditures for chronic complications and other unrelated conditions for which diabetics have higher rates of utilization are determined by analysis of attributable risks. Additional expenditures generated by extra hospital inpatient days and higher charges for nursing home and home health care for comorbidities are estimated by regression analyses. Sensitivity analysis is used to calculate a range of estimated expenditures. RESULTS: Total expenditures attributed to diabetes are $47.9 billion in 1995, including $18.8 billion for first listed diabetes, $18.7 billion for chronic complications, $8.5 billion for unrelated conditions, and $1.9 billion for comorbidities. The range of total expenditures is $34.3 to $63.7 billion. CONCLUSIONS: Comprehensive accounting of expenditures more accurately assesses the economic burden of diabetes and potential savings from prevention, especially of chronic complications. This analysis is illustrative for other chronic illnesses.

Honeycutt, A. A., Segel, J. E., Hoerger, T. J., & Finkelstein, E. A. (2009). Comparing cost-of-illness estimates from alternative approaches: An application to diabetes. *Health Services Research, 44*(1), 303–320.

OBJECTIVE: To compare disease cost estimates from two commonly used approaches. DATA SOURCE: Pooled Medical Expenditure Panel Survey (MEPS) data for 1998–2003. STUDY DESIGN: We compared regression-based (RB) and attributable fraction (AF) approaches for estimating disease-attributable costs with an application to diabetes. The RB approach used results from econometric models of disease costs, while the AF approach used epidemiologic formulas for diabetes-attributable fractions combined with the total costs for seven conditions that result from diabetes. DATA EXTRACTION: We used SAS version 9.1 to create a data set that combined data from six consecutive years of MEPS. PRINCIPAL FINDINGS: The RB approach produced higher estimates of diabetes-attributable medical spending ($52.9 billion in 2004 dollars) than the AF approach ($37.1 billion in 2004 dollars). RB model estimates may in part be higher because of the challenges of implementing the two approaches in a similar manner, but may also be higher because they capture the costs of increased treatment intensity for those with the disease. CONCLUSIONS: We recommend using the RB approach for estimating disease costs whenever individual-level data on health care spending are available and when the presence of the disease affects treatment costs for other conditions, as in the case of diabetes.

Javitz, H. S., Ward, M. M., Watson, J. B., Jaana, M. (2004). Cost of illness of chronic angina. *American Journal of Managed Care, 10*(11 Suppl), S358–S369.

BACKGROUND: Angina pectoris is one of the principal manifestations of coronary artery disease (CAD). Chronic angina is a debilitating condition that affects millions of people in the United States. OBJECTIVE: The objective of the study is to estimate, from a societal perspective, the direct costs of chronic angina in the year 2000. METHODS: Data on medical utilization related to chronic angina were extracted from National Center for Health Statistics public-use databases and from IMS databases on medications (nitrates, beta-blockers, and calcium channel blockers). National average Medicare reimbursement rates were used to estimate costs. We identified medical utilization related to chronic angina based on International Classification of Diseases, Ninth Revision (ICD-9) codes. When ICD-9 codes that do not explicitly identify angina are used in medical databases, people with chronic angina may be coded as having CAD only. To address this, we developed upper- and lower-boundary estimates of the costs of chronic angina. The lower-boundary estimate is based on diagnoses that narrowly define the presence of chronic angina, and is termed "narrowly defined chronic angina." The upper-boundary estimate is based on diagnoses of CAD. RESULTS: The lower boundary on the cost of chronic angina is the estimated direct medical cost of narrowly defined chronic angina ($1.9 billion when it is the first-listed diagnosis and $8.9 billion when it is listed in any position). The upper boundary on the cost of chronic angina is the estimated total direct medical cost of CAD, which is $33 billion when it is the first-listed diagnosis and $75 billion when it is listed in any position. CONCLUSION: These analyses capture the range of direct costs that might be attributed to the care of chronic angina in the United States for the year 2000. Some components of care were not available, and estimated costs will be significantly higher if private payer reimbursement rates are used.

Johnson, F. R., Fries, E. E., & Banzhas, H. S. (1997). Valuing morbidity: An integration of the willingness-to-pay and health-status index literatures. *Journal of Health Economics, 16*(6), 641–665.

Placing dollar values on human health has long been a controversial aspect of policy analysis and remains difficult given the relatively small number of morbidity-valuation studies available. By combining both the economic and health literature, this paper offers an alternative approach to morbidity valuation and provides estimates for a wide range of short-term health conditions.

Kobelt, G., Berg, J., Atherly, D., & Hadjimichael, O. (2006). Costs and quality of life in multiple sclerosis: A cross-sectional study in the United States. *Neurology, 66*(11), 1696–1702.

OBJECTIVE: To estimate current costs and quality of life (utility) of patients treated with disease modifying drugs (DMDs) in the United States, and to investigate the effect of disease severity on costs and utility. METHODS: Data on demographics, disease

information, resource utilization (including work capacity), and utility were collected directly from patients in a cross-sectional postal survey. Patients were selected randomly among participants in the North American Committee on Multiple Sclerosis Patient Registry (NARCOMS) receiving DMDs. Mean annual costs per patient (2004 USD) and mean utilities for the sample and for different levels of functional disability are estimated from the societal perspective. RESULTS: The target answer rate of 50% was reached and 1,909 patients were included in the analysis. The mean age was 49 years and time since diagnosis was 13 years. A total of 10.5% of patients had primary progressive, 47.6% relapsing-remitting, and 33.3% secondary progressive disease. A total of 28.8% of patients indicated to have experienced a relapse during the past 3 months. Total average costs are estimated at 47,215 dollars per patient and year. Of these, 53% were for direct medical and nonmedical costs, 37% for production losses, and 10% for informal care. Mean utility in the sample was 0.698. CONCLUSIONS: Consistent with other studies, costs and utilities were significantly correlated with functional capacity. The proportion of costs represented by disease modifying drugs is considerably lower than estimated in other studies, as all costs related to the disease are included.

Koopmanschap, M. A., & van Ineveld, B. M. (1992). Towards a new approach for estimating indirect costs of disease. *Social Science & Medicine, 34*(9), 1005–1010.

Many researchers in the field of evaluation of health care doubt the usefulness of estimates of indirect costs of disease in setting priorities in health care. This paper attempts to meet part of the criticism on the concept of indirect costs, which are defined as the value of production lost to society due to disease. Thus far in cost of illness studies and cost-effectiveness analyses, the potential indirect costs of disease were calculated. In the following, a first step will be taken toward a new method for estimating indirect costs which are expected to be effectuated in reality: the friction cost method. This method explicitly takes into account short- and long-run processes in the economy that reduce the production losses substantially as compared with the potential losses. According to this method, production losses will be confined to the period needed to replace a sick worker: the so-called friction period. The length of this period and the resulting indirect costs depend on the situation on the labor market. Some preliminary results are presented for the indirect costs of the incidence of cardiovascular disease in the Netherlands for 1988, both for the friction costs and the potential costs. The proposed methodology for estimating indirect costs is promising, but needs further development. The consequences of illness in people without a paid job need to be incorporated in the analysis. Also the relation between internal labor reserves and costs of disease should be further investigated. Next to this, more refined labor market assumptions, allowing for diverging situations on different segments of the labor market are necessary.

Koopmanschap, M. A., Rutten, F. F., van Ineveld, B. M., & van Roijen, L. (1995). The friction cost method for measuring indirect costs of disease. *Journal of Health Economics*, *14*(2), 171–189.

A new approach for estimating the indirect costs of disease, which explicitly considers economic circumstances that limit production losses due to disease, is presented (the friction cost method). For the Netherlands, the short-term friction costs in 1990 amount to 1.5–2.5% of net national income (NNI), depending on the extent to which short-term absence from work induces production loss and costs. The medium-term macro-economic consequences of absence from work and disability reduce NNI by an additional 0.8%. These estimates are considerably lower than estimates based on the traditional human capital approach, but they better reflect the economic impact of illness.

Landefeld, J. S., & Seskin, E. P. (1982). The economic value of life: Linking theory to practice. *American Journal of Public Health, 72*(6), 555–566.

Human capital estimates of the economic value of life have been routinely used in the past to perform cost-benefit analyses of health programs. Recently, however, serious questions have been raised concerning the conceptual basis for valuing human life by applying these estimates. Most economists writing on these issues tend to agree that a more conceptually correct method to value risks to human life in cost-benefit analyses would be based on individuals' "willingness to pay" for small changes in their probability of survival. Attempts to implement the willingness-to-pay approach using survey responses or revealed-preference estimates have produced a confusing array of values fraught with statistical problems and measurement difficulties. As a result, economists have searched for a link between willingness to pay and standard human capital estimates and have found that for most individuals a lower bound for valuing risks to life can be based on their willingness to pay to avoid the expected economic losses associated with death. However, while these studies provide support for using individual's private valuation of forgone income in valuing risks to life, it is also clear that standard human capital estimates cannot be used for this purpose without reformulation. After reviewing the major approaches to valuing risks to life, this paper concludes that estimates based on the human capital approach—reformulated using a willingness-to-pay criterion—produce the only clear, consistent, and objective values for use in cost-benefit analyses of policies affecting risks to life. The paper presents the first empirical estimates of such adjusted willingness-to-pay/human capital values.

Lazaro, A. (2002). Theoretical arguments for the discounting of health consequences: Where do we go from here? *Pharmacoeconomics*, *20*(14), 943–961.

Despite the theoretical arguments presented in the literature regarding discounting over the last 25 years, no satisfactory reply has yet been offered to the question of whether health consequences have to be discounted at the same rate as monetary consequences in the economic evaluation of health programs or interventions designed to improve health.

Against this background, the main objective of this paper was to review and systemize these theoretical arguments, with the aim of determining whether any of the positions identified can be accepted without reservation. Having determined that this is not possible, we investigated the rationality of discounting in the literature and, on this basis, propose a potential way to resolve the problem. Thus, we argue that the relationship between the discount of monetary and health consequences has to be determined in an indirect manner, by reference to the relationship maintained by the individual time preference rates for health and money in the context of private and social choice. Although this proposal moves the debate into the empirical field, its advantages must be weighed against the difficulties associated with the estimation of the time preferences.

Lipscomb, J., Barnett, P. G., Brown, M. L., Lawrence, W., & Yabroff, K. R. (2009). Advancing the science of health care costing. *Medical Care, 47*(7 Suppl 1), S120–S126.

The preceding articles in this volume amply illustrate and critically discuss the major issues in health care costing. This concluding article has two purposes. First, we synthesize and evaluate the main findings. Second, we identify the elements of a research agenda for improving the scientific soundness and relevance of health cost analyses for decision making.

Lipscomb, J., Yabroff, K. R., Brown, M. L., Lawrence, W. & Barnett, P. G. (2009). Health care costing: Data, methods, current applications. *Medical Care, 47*(7 Suppl 1), S1–S6.

Health care costs continue to grow rapidly, straining budgets and raising questions about whether consumers are getting good value for the money spent. There has never been a more pressing need for conceptually sound and empirically accurate estimates of health care costs, for a variety of applications. For example, cost estimates are pivotal in the setting of public and private health care budgets at all levels and establishing reimbursement rates; in cost-effectiveness analyses and other economic evaluations; and in assessing the impact of investments in research to prevent, detect, and treat disease.

Manning, W. G., & Mullahy, J. (2001). Estimating log models: To transform or not to transform? *Journal of Health Economics, 20*(4), 461–494.

Health economists often use log models to deal with skewed outcomes, such as health utilization or health expenditures. The literature provides a number of alternative estimation approaches for log models, including ordinary least squares on ln(y) and generalized linear models. This study examines how well the alternative estimators behave econometrically in terms of bias and precision when the data are skewed or have other common data problems (e.g., heteroscedasticity, heavy tails). No single alternative is best under all conditions examined. The paper provides a straightforward algorithm for choosing among the alternative estimators. Even if the estimators considered are consistent, there can be major losses in precision from selecting a less appropriate estimator.

Mullahy, J. (2009). Econometric modeling of health care costs and expenditures: A survey of analytical issues and related policy considerations. *Medical Care*, *47*(7 Suppl 1), S104–108.

BACKGROUND: Econometric modeling of health care costs and expenditures has become an important component of decision making across a wide array of real-world settings. OBJECTIVES: The objective of this article is to provide a brief summary of important conceptual and analytical issues involved in econometric health care cost modeling. To this end, the article explores outcome measures typically analyzed in such work; the decision maker's perspective in econometric cost modeling exercises; specific analytical issues in econometric model specification; statistical goodness-of-fit testing; empirical implications of "upper tail" (or "high cost") phenomena; and issues relating to the reporting of findings. DATA: Some of the concepts explored here are illustrated in light of samples drawn from the 2005 Medical Expenditure Panel Survey and the 2005 Nationwide Inpatient Sample. RESULTS AND CONCLUSIONS: Analysts of health care cost data have at their disposal an increasingly sophisticated tool kit for analyzing such data that can in principle and in fact yield increasingly interesting insights into data structures. Yet for such analyses to usefully inform policy decisions, the manner in which such studies are designed, undertaken, and reported must accommodate considerations relevant to the decision-making community. The article concludes with some preliminary thoughts on how such bridges might be constructed.

Mrozek, J. R., & Taylor, L. O. (2002). What determines the value of life? A meta-analysis. *Journal of Policy Analysis and Management, 21*(2), 253–270.

A large literature has developed in which labor market contracts are used to estimate the value of a statistical life (VSL). Reported estimates of the VSL vary substantially, from less than $100,000 to more than $25 million. This research uses meta-analysis to quantitatively assess the VSL literature. Results from existing studies are pooled to identify the systematic relationships between VSL estimates and each study's particular features, such as the sample composition and research methods. This meta-analysis suggests that a VSL range of approximately $1.5 million to $2.5 million (in 1998 dollars) is what can be reasonably inferred from past labor-market studies when "best practice" assumptions are invoked. This range is considerably below many previous qualitative reviews of this literature.

O'Brien, B. & Viramontes, J. L. (1994). Willingness to pay: A valid and reliable measure of health state preference? *Medical Decision Making*, *14*(3), 289–297.

The development of methods to measure willingness to pay (WTP) has renewed interest in cost-benefit analysis (CBA) for the economic evaluation of health care programs. The authors studied the construct validity and test-retest reliability of WTP as a measure of health state preferences in a survey of 102 persons (mean age 62 years; 54% male) who had chronic lung disease (forced expiratory volume < 70%). Interview measurements

included self-reported symptoms, the oxygen-cost diagram for dyspnea, Short-Form 36 for general health status, rating scale and standard gamble for value and utility of current health state relative to death and healthy lung functioning, and WTP for a hypothetical intervention offering a 99% chance of healthy lung functioning and a 1% chance of death. WTP was elicited by a simple bidding game. To test for starting-point bias, the respondents were randomly assigned to one of five starting bids. All health status and preference measurements except WTP (controlling for income) showed significant ($p < 0.05$) difference between disease-severity groups (mild/moderate/severe). WTP was significantly ($p = 0.01$) associated with household income, but other health status and preference measure were not. The measure most highly correlated with WTP was standard gamble ($r = -0.46$). There was no association between starting bid and mean WTP adjusted for income and health status. The test-retest reliability of WTP was acceptable ($r = 0.66$) but lower than that for the standard gamble ($r = 0.82$).

Ofman, J. J., Sullivan, S. D., Neumann, P. J., Chiou, C. F., Henning, J. M., Wade, S. W., & et al. (2003). Examining the value and quality of health economic analyses: implications of utilizing the QHES. *Journal of Managed Care Pharmacy*, *9*(1), 53–61.

OBJECTIVE: To examine the increasing use of health economic studies and practical implications of evaluating their quality utilizing the Quality of Health Economic Studies (QHES) instrument. METHODS: We first reviewed secondary references to examine ways in which health economic analyses are used in different health care settings, the manner in which these data are appraised and evaluated, and their relevance and value in decision making. The QHES, a new instrument designed to support fast, accurate initial assessments of study quality, was then introduced and validated. A case study was performed using the QHES to score the quality of 30 cost-effectiveness studies in gastroesophageal reflux disease (GERD) published since 1985. Areas where additional research could guide efforts to identify and enhance the use of higher-quality cost-effectiveness studies were suggested. RESULTS: Results from the published validation study of the QHES demonstrated the validity of this new instrument. The resulting QHES scores in the case study of GERD papers ranged from 43 to 91 with a mean of 63.6 (SD=14.7). Approximately 27% of the studies rated had scores less than 50, and 27% had scores above or equal to 75. All 30 studies made conclusions and recommendations and justified them based on their study results. Most studies used appropriate cost and health outcome measures. Very few studies stated the perspective of their analysis and reasons for its selection. The majority of the studies did not perform incremental analysis. CONCLUSION: An examination of the QHES validation study and the case study in GERD suggests that there is a rationale and potential utility to use a quality scoring system for cost-effectiveness studies. The QHES may play an important role in discriminating higher-quality cost-effectiveness information to enhance decision making. The QHES can also serve as a guideline for conducting and reporting future cost-effectiveness studies, as an aid in the editorial process, and for stratification in

systematic reviews. Complex decisions regarding resource allocation rarely rely solely on economic considerations but do increasingly use health economic analyses. To the extent that such analyses are used, the QHES may help ensure that higher-quality analyses receive more analytic attention and greater weight in the decision-making process.

Pauly, M. V., Nicholson, S., Polsky, D., Berger, M. L., & Sharda, C. (2008). Valuing reductions in on-the-job illness: 'Presenteeism' from managerial and economic perspectives. *Health Economics, 17*(4), 469–485.

This paper reports on a study of manager perceptions of the cost to employers of on-the-job employee illness, sometimes termed "presenteeism," for various types of jobs. Using methods developed previously, the authors analyzed data from a survey of more than 800 U.S. managers to determine the characteristics of various jobs and the relationship of those characteristics to the manager's view of the cost to the firm of absenteeism and presenteeism. Jobs with characteristics that suggest unusually high cost (relative to wages) were similar in terms of their absenteeism multipliers and their presenteeism multipliers. Jobs with high values of team production, high requirements for timely output, and high difficulties of substitution for absent or impaired workers had significantly higher indicators of cost for both absenteeism and presenteeism, although substitution was somewhat less important for presenteeism.

Portney, P. R. (1994). The contingent valuation debate: Why economists should care. *The Journal of Economic Perspectives, 8*(4), 3–17.

The contingent valuation method involves the use of sample surveys (questionnaires) to elicit the willingness of respondents to pay for (generally) hypothetical projects or programs. The name of the method refers to the fact that the values revealed by respondents are contingent upon the constructed or simulated market presented in the survey. A spirited (and occasionally mean-spirited) battle over such methods is currently being waged, involving competing factions within the federal government, economists and lawyers representing business and environmental groups, and interested academics as well. At issue is a seemingly quite specific question: should environmental regulations currently under development at both the Department of the Interior and the Department of Commerce sanction the use of the contingent valuation method in estimating the damage done by spills of oil, chemicals, or other substances covered by federal law? More generally, the debate raises broad questions about what economists have to say about the values that individuals place on public or private goods.

Powers, C. A., Meyer, C. M., Roebuck, M. C., & Vaziri, B. (2005). Predictive modeling of total healthcare costs using pharmacy claims data: A comparison of alternative econometric cost modeling techniques. *Medical Care*, *43*(11), 1065–1072.

OBJECTIVE: We sought to evaluate several statistical modeling approaches in predicting prospective total annual health costs (medical plus pharmacy) of health plan participants

using Pharmacy Health Dimensions (PHD), a pharmacy claims-based risk index. METHODS: We undertook a 2-year (baseline year/follow-up year) longitudinal analysis of integrated medical and pharmacy claims. Included were plan participants younger than 65 years of age with continuous medical and pharmacy coverage (n = 344,832). PHD drug categories, age, gender, and pharmacy costs were derived across the baseline year. Annual total health costs were calculated for each plan participant in follow-up year. Models examined included ordinary least squares (OLS) regression, log-transformed OLS regression with smearing estimator, and 3 two-part models using OLS regression, log-OLS regression with smearing estimator, and generalized linear modeling (GLM), respectively. A 10% random sample was withheld for model validation, which was assessed via adjusted r, mean absolute prediction error, specificity, and positive predictive value. RESULTS: Most PHD drug categories were significant independent predictors of total costs. Among models tested, the OLS model had the lowest mean absolute prediction error and highest adjusted r. The log-OLS and two-part log-OLS models did not predict costs accurately as the result of issues of log-scale heteroscedasticity. The two-part model using GLM had lower adjusted r but similar performance in other assessment measures compared with the OLS or 2-part OLS models. CONCLUSION: The PHD system derived solely from pharmacy claims data can be used to predict future total health costs. Using PHD with a simple OLS model may provide similar predictive accuracy in comparison to more advanced econometric models.

Prosser, L. A., Hammitt, J. K., & Keren, R. (2007). Measuring health preferences for use in cost-utility and cost-benefit analyses of interventions in children: Theoretical and methodological considerations. *Pharmacoeconomics*, *25*(9), 713–726.

Valuing the health of children for cost-utility or cost-benefit analysis poses a number of additional challenges when compared with valuing adult health. Some of these challenges relate to the inability of young children to value changes in health directly and the potential biases associated with using proxy respondents. Other challenges arise from children not being able to perform as independent economic actors, but dependent on others for care and decision making. In addition, illness in children may affect parent/caregiver quality of life, further complicating the measurement of value associated with a change in a child's health status. We review the most common approaches (QALYs and willingness-to-pay values) for valuing health in economic evaluations and consider the methodological and practical issues associated with measuring child health using each framework. Recommendations for advancing the field of valuing child health for economic evaluations will vary by age; a one size fits all approach does not readily fit. Although limitations exist for all of the methods considered for valuing child health, the currently recommended approach for infants and preschoolers is direct valuation by a proxy respondent. For school-age children and adolescents, existing multi-attribute instruments can be applied in some situations but direct valuation may be required for others. Future research should focus on minimizing bias from proxy respondents, consideration of a family- or household-based

approach to valuing health effects, and development of generic instruments with domains that are appropriate to children and that vary with age.

Rice, D. P., & Hodgson, T. A. (1982). The value of human life revisited. *American Journal of Public Health, 72*(6), 536–538.

It has been almost 70 years since the Journal published its first article on "The Value of Human Life" by Dr. Charles V. Chapin. In 1913, Dr. Chapin stated forcefully that it was unwise to emphasize the financial or monetary side of public health by placing a money value on life. Considerable progress has been made since that time. We no longer argue about whether we should attach a value to human life for cost benefit analyses of health programs. The debate now is on the method to be used as illustrated in the article in this issue of the Journal by Landefeld and Seskin, in which the various valuation methods are reviewed and a new measure is proposed. The current debate centers around two methodologies for calculating benefits of reductions in loss of life: the human capital (HC) and willingness-to-pay (WTP) approaches. Having been involved for some years in refining the HC method, we enter this discussion to clarify our views that HC and WTP are conceptually different, serve different purposes, and measure different aspects of threats to health.

Rice, D. P., & Miller, L. S. (1998). Health economics and cost implications of anxiety and other mental disorders in the United States. *British Journal of Psychiatry, 173*(Suppl. 34), 4–9.

BACKGROUND: Mental disorders impose a multibillion dollar burden on the economy each year; translating the burden into economic terms is important to facilitate formulating policies about the use of resources. METHODS: For direct costs, data were obtained from national household interview and provider surveys; for morbidity costs, a timing model was used that measures the lifetime effect on current income of individuals with mental disorders, taking into account the timing of onset and the duration of these disorders, based on regression analysis of Epidemiologic Catchment Area study data. RESULTS: The total economic costs of mental disorders amounted to US$147.8 billion in 1990. Anxiety disorders are the most costly, amounting to $46.6 billion, or 31.5% of the total; schizophrenic disorders accounted for $32.5 billion, affective disorders for $30.4 billion, and other mental disorders for $38.4 billion. CONCLUSIONS: Mental illnesses, especially anxiety disorders, are costly to society. Although anxiety disorders have a higher prevalence than affective disorders and schizophrenia, use of medical care services is lowest for anxiety disorders. Anxiety disorders appear to be under-recognized and untreated even though treatment interventions have been shown to be effective and can be delivered in a cost-efficient manner.

Rosen, A. B., & Cutler, D. M. (2009). Challenges in building disease-based national health accounts. *Medical Care, 47*(7 Suppl 1), S7–S13.

BACKGROUND: Measuring spending on diseases is critical to assessing the value of medical care. OBJECTIVE: To review the current state of cost of illness estimation methods, identifying their strengths, limitations, and uses. We briefly describe the current National Health Expenditure Accounts and then discuss the addition of cost of illness estimation to the National Health Expenditure Accounts. CONCLUSION: Recommendations are made for future research aimed at identifying the best methods for developing and using disease-based national health accounts to optimize the information available to policy makers as they struggle with difficult resource allocation decisions.

Russell, L. B. (1999). Modelling for cost-effectiveness analysis. *Statistics in Medicine*, *18*(23), 3235–3244.

A model creates the framework for a cost-effectiveness analysis, allowing decision makers to explore the implications of using an intervention in different ways and under different conditions. To serve its purpose, a model must produce accurate predictions and allow for substantial variation in the factors that influence costs and effects. This paper considers three aspects of modeling: validating effectiveness estimates, modeling costs, and the implications of common statistical forms. Validation procedures similar to those for effectiveness estimates are proposed for costs. Modelers need to pay more attention to ensuring that the pathway of events described by a model represents costs as well as it does effects. Modelers can also help improve the epidemiological and clinical research on which cost-effectiveness analyses depend by showing the implications for resource allocation of the statistical forms conventionally used in these fields.

Smith, R. D. (2000). The discrete-choice willingness-to-pay question format in health economics: Should we adopt environmental guidelines? *Medical Decision Making*, *20*(2), 194–206.

The use of willingness to pay (WTP) in valuing the benefits of health care programs is increasing. Although such values have been derived using open-ended, bidding, or payment-card techniques, recently discrete-choice questionnaires have been advocated, particularly following the report of the National Oceanographic and Atmospheric Administration concerning the validity of using WTP to estimate environmental benefits. It is argued that discrete-choice questions offer a more realistic market, and will therefore lead to more valid responses and yield higher response rates through reduced mental demands. The author reviews these issues in a critical assessment of discrete-choice questions.

Trogdon, J. G., Finkelstein, E. A., & Hoerger, T. J. (2008). Use of econometric models to estimate expenditure shares. *Health Services Research, 43*(4), 1442–1452.

OBJECTIVE: To investigate the use of regression models to calculate disease-specific shares of medical expenditures. DATA SOURCES/STUDY SETTING: Medical Expenditure Panel

Survey (MEPS), 2000–2003. STUDY DESIGN: Theoretical investigation and secondary data analysis. DATA COLLECTION/EXTRACTION METHODS: Condition files used to define the presence of 10 medical conditions. PRINCIPAL FINDINGS: Incremental effects of conditions on expenditures, expressed as a fraction of total expenditures, cannot generally be interpreted as shares. When the presence of one condition increases treatment costs for another condition, summing condition-specific shares leads to double-counting of expenditures. CONCLUSIONS: Condition-specific shares generated from multiplicative models should not be summed. We provide an algorithm that allows estimates based on these models to be interpreted as shares and summed across conditions.

Viscusi, K. W. (2003). The value of a statistical life: A critical review of market estimates throughout the world. *The Journal of Risk and Uncertainty, 27*(1), 5–76.

A substantial literature over the past 30 years has evaluated tradeoffs between money and fatality risks. These values in turn serve as estimates of the value of a statistical life. This article reviews more than 60 studies of mortality risk premiums from 10 countries and approximately 40 studies that present estimates of injury risk premiums. This critical review examines a variety of econometric issues, the role of unionization in risk premiums, and the effects of age on the value of a statistical life. Our meta-analysis indicates an income elasticity of the value of a statistical life from about 0.5 to 0.6. The paper also presents a detailed discussion of policy applications of these value of a statistical life estimates and related issues, including risk-risk analysis.

Ward, M. M., H. S. Javitz, Smith, W. M., & Bakst, A. (2000). A comparison of three approaches for attributing hospitalizations to specific diseases in cost analyses. *International Journal of Technology Assessment in Health Care*, *16*(1), 125–136.

OBJECTIVES: Calculations of health care costs rarely disclose the specific approach used to allocate the cost of hospitalizations by diagnosis. However, the type of approach used can have a major impact on the findings in the case of significant comorbidities. The present analyses compared three approaches for attributing Medicare DRG reimbursements (which were used as surrogates for average costs) for hospitalization by diagnosis. METHODS: Medical resource utilization data from the National Hospital Discharge Survey were analyzed using each of three allocation approaches: (a) attributing 100% of the cost of hospitalization to the disease when it was the first-listed diagnosis; (b) attributing a portion of the cost of hospitalization to the disease, depending on its position in the list of diagnoses and the relevance of any comorbidities; and (c) an incremental analysis of cost based upon the hospitalization experiences of an age and gender matched cohort. These three approaches were applied to the cost of hospitalization for chronic obstructive pulmonary disease (COPD). RESULTS: The first approach projected 206,098 hospitalizations at $3,449 per hospitalization for a projected U.S. annual total of $711 million. The second approach projected 681,547 hospitalizations at $3,205 per hospitalization for a projected U.S. annual total of $2.2 billion. The third approach also projected 681,547 hospitalizations, but at

$2,361 per hospitalization, for a projected U.S. annual total of $1.6 billion. CONCLUSIONS: Expanding from the example on COPD, the limitations of each approach are described and their applications to other conditions are presented.

Weil, D. (2001). Valuing the economic consequences of work injury and illness: A comparison of methods and findings. *American Journal of Industrial Medicine*, *40*(4), 418–437.

BACKGROUND: Workplace injuries and fatalities in the United States create significant economic costs to society. Although economic costs should measure the opportunity cost to society arising from injuries and fatalities, estimating them often proves difficult as a practical matter. This leads to a range of estimates for valuing these costs. METHODS: This paper compares methods of economic valuation, focusing in particular on how different methods diverge to varying degrees from measuring the "true" economic costs of injuries and illnesses. In so doing, it surveys the literature that has arisen in the past 25 years to measure different aspects of economic consequences. RESULTS: Estimates of the costs of injuries and fatalities tend to understate the true economic costs from a social welfare perspective, particularly in how they account for occupational fatalities and losses arising from work disabilities. CONCLUSION: Although data availability often makes estimation of social welfare costs difficult, researchers should attempt to more fully integrate such approaches into estimation procedures and interpretation of their results.

A.3 Health Status and Quality of Life (QoL) Measures

Albrecht G. L. & Devlieger, P. J. (1999). The disability paradox: High quality of life against all odds. *Social Science and Medicine*, *48*, 977–988.

This paper builds on the work of Sol Levine to examine a disability paradox: Why do many people with serious and persistent disabilities report that they experience a good or excellent quality of life when to most external observers these individuals seem to live an undesirable daily existence? The paper uses a qualitative approach to develop an explanation of this paradox using semistructured interviews with 153 persons with disabilities. 54.3% of the respondents with moderate to serious disabilities reported having an excellent or good quality of life confirming the existence of the disability paradox. Analysis of the interviews reveals that for both those who report that they have a good and those who report that they have a poor quality of life, quality of life is dependent upon finding a balance between body, mind, and spirit in the self and on establishing and maintaining an harmonious set of relationships within the person's social context and external environment. A theoretical framework is developed to express these relationships. The findings are discussed for those with and without disabilities and directions are given for future research.

Bayliss, E. A., Ellis, J. L., & Steiner, J. F. (2005). Subjective assessments of comorbidity correlate with quality of life health outcomes: Initial validation of a comorbidity assessment instrument. *Health and Quality of Life Outcomes*, *3*, 51.

BACKGROUND: Interventions to improve care for persons with chronic medical conditions often use quality of life (QOL) outcomes. These outcomes may be affected by coexisting (comorbid) chronic conditions as well as the index condition of interest. A subjective measure of comorbidity that incorporates an assessment of disease severity may be particularly useful for assessing comorbidity for these investigations. METHODS: A survey including a list of 25 common chronic conditions was administered to a population of HMO members age 65 or older. Disease burden (comorbidity) was defined as the number of self-identified comorbid conditions weighted by the degree (from 1 to 5) to which each interfered with their daily activities. We calculated sensitivities and specificities relative to chart review for each condition. We correlated self-reported disease burden, relative to two other well-known comorbidity measures (the Charlson Comorbidity Index and the RxRisk score) and chart review, with our primary and secondary QOL outcomes of interest: general health status, physical functioning, depression screen and self-efficacy. RESULTS: 156 respondents reported an average of 5.9 chronic conditions. Median sensitivity and specificity relative to chart review were 75% and 92%, respectively. QOL outcomes correlated most strongly with disease burden, followed by number of conditions by chart review, the Charlson Comorbidity Index and the RxRisk score. CONCLUSION: Self-report appears to provide a reasonable estimate of comorbidity. For certain QOL assessments, self-reported disease burden may provide a more accurate estimate of comorbidity than existing measures that use different methodologies, and that were originally validated against other outcomes. Investigators adjusting for comorbidity in studies using QOL outcomes may wish to consider using subjective comorbidity measures that incorporate disease severity.

Barnett, D. B. (1991). Assessment of quality of life. *American Journal of Cardiology*, *67*(12), 41C–44C.

Assessment of quality of life has emerged in recent years as an important part of the overall evaluation of drug therapy and health care in general. Measurement techniques for this difficult assessment range from simple unqualified questions on patient well-being to complex statistical analyses of a wide range of lifestyle and activity variables. The factors that influence quality of life during chronic drug therapy differ in the treatment of symptomatic (e.g., heart failure) vs. asymptomatic (e.g., hypertension) disease, and include drug side effects, relief of symptoms, improved prognosis, return to work, physical activity, and the need for further hospital treatments. The manifestation of quality of life varies for different people leading to lack of agreement on the precise definition. The absence of standardization of methods of measurement also contributes to this and leads to lack of comparability of studies and unreasonable claims by some drug manufacturers. Further complicating issues in multicenter trials across countries include language problems

and interethnic differences in "sickness" behavior. The recently introduced quality-adjusted life year (QALY) index, designed to take account of both the quality and duration of life in assessing the outcome of treatments, may avoid some of these problems. By classifying illness states (the Rosser index) on the basis of disability and distress, and comparing outcomes in terms of improved prognosis, QALYs have already been used for cost/benefit analyses of a number of new and expensive therapies. Like other methods, QALYs have problems related to variability in individual appreciation of life values. To date, a perfect method of quality of life assessment remains elusive.

Bleichrodt, H., & Johanneson, M. (1997). Standard gamble, time trade-off and rating scale: Experimental results on the ranking properties of QALYs. *Journal of Health Economics, 16*, 155–175.

This paper compares the relative performance of quality-adjusted life years (QALYs) based on quality weights elicited by rating scale (RS), time trade-off (TTO) and standard gamble (SG). The standard against which relative performance is assessed is individual preference elicited by direct ranking. The correlation between predicted and direct ranking is significantly higher for TTO-QALYs than for RS-QALYs and SG-QALYs. This holds both based on mean Spearman rank correlation coefficients calculated per individual and based on two social choice rules: the method of majority voting and the Borda rule. Undiscounted TTO-QALYs are more consistent with direct ranking than discounted TTO-QALYs.

Brazier, J., Deverill, M., & Green, C. (1999). A review of the use of health status measures in economic evaluation. *Journal of Health Services Research & Policy, 4*(3), 174–184.

OBJECTIVE: To review the use of measures of health status in the assessment of benefits in economic evaluation, whether or not the measures were designed for this purpose. METHODS: The review was based on a systematic search of the literature. It provides a comprehensive assessment of the evidence where it exists and a balanced overview of opinion otherwise. RESULTS: Over 3,000 papers were identified, of which 632 were found to be relevant. The review provides a set of recommendations, including a checklist of questions for selecting a measure for use in economic evaluation, a list of circumstances in which nonpreference-based measures can be used, and recommendations surrounding the use of health state valuation techniques and multi-attribute utility scales. CONCLUSION: These recommendations should help to identify poor economic evaluations and hence guard against inefficient conclusions being drawn regarding the provision of health services.

Cella, D., & Webster, K. (1997). Linking outcomes management to quality-of-life measurement. *Oncology*, 232–235.

Health-care workers now accept quality of life (QOL) as an important outcome to evaluate in clinical research and as a useful measure of quality care. Indeed, current demand for QOL assessment in clinical practice has outpaced the availability of valid, streamlined, cost-effective methods for carrying out such assessment, although new tools are in the offing.

This paper will highlight some of the major challenges facing outcomes management and outcomes research, with particular focus on the development of a QOL instrument to evaluate and manage anemia and fatigue in cancer patients—the Functional Assessment of Cancer Therapy-Anemia (FACT-An). The newest version of broader QOL assessment system, the Functional Assessment of Chronic Illness Therapy (FACIT), will also be described.

Charlson, M., Szatrowski, T. P., Peterson, J., & Gold, J. (1994). Validation of a combined comorbidity index. *Journal of Clinical Epidemiology, 47*(11), 1245–1251.

The basic objective of this paper is to evaluate an age-comorbidity index in a cohort of patients who were originally enrolled in a prospective study to identify risk factors for peri-operative complications. Two-hundred and twenty-six patients were enrolled in the study. The participants were patients with hypertension or diabetes who underwent elective surgery between 1982 and 1985 and who survived to discharge. Two-hundred and eighteen patients survived until discharge. These patients were followed for at least 5 years post-operatively. The estimated relative risk of death for each comorbidity rank was 1.4 and for each decade of age was 1.4. When age and comorbidity were modeled as a combined age-comorbidity score, the estimated relative risk for each combined age-comorbidity unit was 1.45. Thus, the estimated relative risk of death from an increase of one in the comorbidity score proved approximately equal to that from an additional decade of age. The combined age-comorbidity score may be useful in some longitudinal studies to estimate relative risk of death from prognostic clinical covariates.

Cheak-Zamora, N. C., Wyrwich, K. W., & McBride, T. D. (2009). Reliability and validity of the SF-12v2 in the medical expenditure panel survey. *Quality of Life Research, 18*(6), 727–735.

OBJECTIVE: Evaluate the reliability and validity of the Medical Outcomes Study Short-Form version 2 (SF-12v2) in the 2003–2004 Medical Expenditure Panel Survey (MEPS). RESEARCH DESIGN: Data were collected in the self-administered mail-out questionnaire and face-to-face interviews of the MEPS (n = 20,661). Internal consistency and test-retest reliability and construct, discriminate, predictive and concurrent validity were tested. The EQ-5D, perceived health and mental health questions were used to test construct and discriminate validity. Self-reported work, physical and cognitive limits tested predictive validity and number of chronic conditions assessed concurrent validity. RESULTS: Both Mental Component Summary Scores (MCS) and Physical Component Summary Scores (PCS) were shown to have high internal consistency reliability (alpha > .80). PCS showed high test-retest reliability (ICC = .78) while MCS demonstrated moderate reliability (ICC = .60). PCS had high convergent validity for EQ-5D items (except self-care) and physical health status (r > .56). MCS demonstrated moderate convergent validity on EQ-5D and mental health items (r > .38). PCS distinguish between groups with different physical and work limitations. Similarly, MCS distinguished between groups with and without cognitive limitations. The MCS and PCS showed perfect dose response when variations in scores were

examined by participant's chronic condition status. CONCLUSIONS: Both component scores showed adequate reliability and validity with the 2003–2004 MEPS and should be suitable for use in a variety of proposes within this database.

Chuang, L. H., & Kind, P. (2009). Converting the SF-12 into the EQ-5D: An empirical comparison of methodologies. *Pharmacoeconomics, 27*(6), 491–505.

BACKGROUND: For cost-utility analysis, analysts need a measure that summarizes health-status utilities in a single index of health-related quality of life (HR-QOL). It is common to find in clinical studies that only an HR-QOL profile measure such as the SF-36 is included, but not the summary HR-QOL index. Therefore, the economist's usual practice is to reprocess the profile data into a single index format. Several 'after-market' tools are available to convert the SF-36 or SF-12 into a single form with or without utility-weighting metric property. However, there has been no consensus with regard to a regression method that should be recommended for such a mapping task. OBJECTIVE: To report on the performance of different regression methods that have previously been applied to the conversion of SF-12 data in the analysis of a single common dataset. The mapping between the SF-12 and EQ-5D is the focus. METHODS: The data were adopted from the Medical Expenditure Panel Survey 2003 where 19,678 adults completed both EQ-5D and SF-12 questionnaires. Four econometric techniques, namely ordinary least squares (OLS), censored least absolute deviation, multinomial logit model and two-part model regressions were investigated together with two main types of model specifications: item-based and summary score-based. The performance of each examined model was judged by various criteria, including its estimated mean, the size of mean absolute error and the number of errors. RESULTS: Among four compared econometric techniques, OLS regression was the most accurate model in estimating the group mean. Models with item-based model specification performed better than those with summary score-based regardless of the chosen econometric technique. Nevertheless, the accuracy of OLS deteriorates in older and less healthy subgroups. The results also suggested that the two-part model, which addresses the heterogeneity issue, performs better in these vulnerable subgroups. CONCLUSIONS: None of the mapping methods included in the current study are suitable for estimating at the individual level. The methodology exemplified here has wider applicability and might just as readily be applied to other members of the SF family or indeed to other profile measures of HR-QOL. However, it is recommended that a preference-based, single index measure of HR-QOL should be included in the clinical studies for the purpose of economic evaluation.

Cleeland, C. S. (2007). Symptom burden: Multiple symptoms and their impact as patient-reported outcomes. *Journal of the National Cancer Institute Monographs*, *37*, 16–21.

Cancer and its treatment produce multiple symptoms that significantly distress patients and impair function. Symptoms caused by treatment may delay treatment or lead to premature

treatment termination, and residual treatment-related symptoms often complicate posttreatment rehabilitation. When treatment is no longer possible, symptom control becomes the focus of cancer care. Patient ratings of symptom severity and impact are important patient-reported outcomes (PROs) in cancer clinical trials and comprise a subset of a larger domain of PROs generally referred to as health-related quality of life (HRQoL). Symptoms rarely occur in isolation; rather, there is now ample evidence that symptoms frequently occur in clusters. The impact of these multiple symptoms upon the patient can be described as "symptom burden," a concept that encompasses both the severity of the symptoms and the patient's perception of the impact of the symptoms. The distress caused by symptoms is a subject of much investigation, and several validated measures of the severity and impact of multiple symptoms are now available. Symptom measures are generally brief, thereby reducing respondent burden, and can be administered repeatedly during a trial to give a relatively fine-grained picture of the patient's status across time. In many instances, information on trial-related changes in symptom burden, or comparison of symptom burden between arms in a clinical trial, may provide sufficient self-report data for clinical trial consumers (patients, clinicians, and regulators) to make treatment choices or to evaluate new therapies, without measuring other HRQoL domains.

Cohen, S. R., Mount, B. M., Strobel, M. G., & Bui, F. (1995). The McGill Quality of Life Questionnaire: A measure of quality of life appropriate for people with advanced disease. A preliminary study of validity and acceptability. *Palliative Medicine*, *9*(3), 207–219.

This is the first report on the McGill Quality of Life Questionnaire (MQOL), a questionnaire relevant to all phases of the disease trajectory for people with a life-threatening illness. This questionnaire differs from most others in three ways: the existential domain is measured, the physical domain is important but not predominant, and positive contributions to quality of life are measured. This study was conducted in a palliative care setting. Principal components analysis suggests four subscales: physical symptoms, psychological symptoms, outlook on life, and meaningful existence. Construct validity of the subscales is demonstrated through the pattern of correlations with the items from the Spitzer Quality of Life Index. The importance of measuring the existential domain is highlighted by the finding that, of all the MQOL subscales and Spitzer items, only the meaningful existence subscale correlated significantly with a single item scale rating overall quality of life.

Cook, K. F., Ashton, C. M., Byrne, M. M., Brody, B., Geracj, J., Giesler, R. B, et al. (2001). A psychometric analysis of the measurement level of the rating scale, time trade-off, and standard gamble. *Social Science & Medicine*, *53*(10), 1275–1285.

A fundamental assumption of utility-based analyses is that patient utilities for health states can be measured on an equal-interval scale. This assumption, however, has not been widely examined. The objective of this study was to assess whether the rating scale (RS), standard gamble (SG), and time trade-off (TTO) utility elicitation methods function as equal-interval

level scales. We wrote descriptions of eight prostate-cancer-related health states. In interviews with patients who had newly diagnosed, advanced prostate cancer, utilities for the health states were elicited using the RS, SG, and TTO methods. At the time of the study, 77 initial and 73 follow-up interviews had been conducted with a consecutive sample of 77 participants. Using a Rasch model, the boundaries (Thurstone Thresholds) between four equal score sub-ranges of the raw utilities were mapped onto an equal-interval logit scale. The distance between adjacent thresholds in logit units was calculated to determine whether the raw utilities were equal-interval. None of the utility scales functioned as interval-level scales in our sample. Therefore, since interval-level estimates are assumed in utility-based analyses, doubt is raised regarding the validity of findings from previous analyses based on these scales. Our findings need to be replicated in other contexts, and the practical impact of non-interval measurement on utility-based analyses should be explored. If cost-effectiveness analyses are not found to be robust to violations of the assumption that utilities are interval, serious doubt will be cast upon findings from utility-based analyses and upon the wisdom of expending millions in research dollars on utility-based studies.

Craig, B. M., Busschbach, J. J., & Salomon, J. A. (2009). Modeling ranking, time trade-off, and visual analog scale values for EQ-5D health states: A review and comparison of methods. *Medical Care, 47*(6), 634–641.

BACKGROUND: There is rising interest in eliciting health state valuations using rankings. Due to their relative simplicity, ordinal measurement methods may offer an attractive practical alternative to cardinal methods, such as time trade-off (TTO) and visual analog scale (VAS). In this article, we explore alternative models for estimating cardinal health state values from rank responses in a unique multicountry database. We highlight an estimation challenge pertaining to health states just below perfect health (the "nonoptimal gap") and propose an analytic solution to ameliorate this problem. METHODS: Using a standardized protocol developed by the EuroQol Group, rank, VAS, and TTO responses were collected for 43 health states in 8 countries: Slovenia, Argentina, Denmark, Japan, Netherlands, Spain, United Kingdom, and United States, yielding a sample of 179,431 state responses from 11,483 subjects. States were described using the EQ-5D system, which allows for three different possible levels on five different dimensions of health. We estimated conditional logit and probit regression models for rank responses. The regressions included 17 health state attribute variables reflecting specific levels on each dimension and counts of different levels across dimensions. This flexible specification accommodates previously published valuation models, such as models applied in the United Kingdom and United States. In addition to fitting standard conditional logit and probit models, which assume equal variance across health states (homoscedasticity), we examined a heteroscedastic probit model that assumes no variance for the two points anchoring the scale ("optimal health" and "dead") and relaxes the equal-variance assumption for all other states. Rank-

based predictions for the 243 unique states defined by the EQ-5D system were compared with predictions from conventional linear models fitted to TTO and VAS responses. RESULTS: By construction, the TTO and VAS models assume no variance around the anchoring states of optimal health and dead. Mimicking this assumption in the probit rank models helps dissolve the nonoptimal gap. For all other states, variances in TTO and VAS were negatively associated with mean values, which contradict the assumption of homoscedasticity. Estimated health state values from the heteroscedastic probit model for the ranking data were highly correlated with predictions from both TTO and VAS models for the 243 EQ-5D states. Between VAS and rank-based estimates, Lin's rho, a measure of agreement, was over 0.98 with a mean absolute difference of 0.028. Corresponding measures of agreement between rank and TTO estimates were 0.96 and 0.12, which is similar to the agreement between VAS and TTO. CONCLUSIONS: Rank-based valuation techniques, which offer advantages of flexibility, generalizability, and ease of administration, may be attractive substitutes for TTO and VAS in the measurement of societal values for health outcomes.

Diener, A., O'Brien, B., & Gafni, A. (1998). Health care contingent valuation studies: A review and classification of the literature. *Health Economics*, *7*(4), 313–326.

PURPOSE: The contingent valuation method (CVM) is a survey-based approach for eliciting consumer's monetary valuations for program benefits for use in cost-benefit analysis (CBA). We used the conceptual framework of O'Brien and Gafni (1996) to classify and critically appraise health care CVM studies. METHODS: Search of computerized health care and economic citation databases (e.g. MEDLINE, ECONLIT) and manual search for papers published between 1984 1996 reporting primary data valuing health program benefits in monetary units by CVM using willingness-to-pay (WTP) or accept (WTA). We classified studies using both empirical (i.e., who was surveyed and how) and conceptual criteria (i.e., which measure of consumer utility was measured and why). RESULTS: 48 CVM studies were retrieved; the majority (42) undertook money valuation in the context of CBA, with the remainder being pricing/demand studies. Among the 42 CBA studies, the consumer utility being measured (i.e., compensating [CV] vs. equivalent variation [EV]) was explicitly stated in only three (7%) studies. WTP was measured in 95% of studies and WTA in 5%. By cross-tabulation, 42 (91%) studies were designed as WTP/CV, two (4%) were WTP/EV, two (4%) were WTA/CV, and no studies used WTA/EV. Most studies were administered by mail (52%) with 38% being in-person interviews. Value elicitation techniques included open-ended questions (38%), payment cards (19%), discrete choice questions (26%), or bidding games (29%). Some form of construct validation tests, particularly associations between WTP and income, were done in 21 studies (50%). CONCLUSIONS: (i) The number of health care CVM studies is growing rapidly and the majority are done in the context of CBA; (ii) there is wide variation among health care CVM studies in terms of the types of questions being posed and the elicitation formats being used; (iii) classification and appraisal of the literature is difficult

because reporting of methods and their relationship with the conceptual framework of CBA is poor; and (iv) the applicability to health care of the CVM guidelines issued by the National Oceanic and Atmospheric Administration (NOAA) panel for environmental economics is unclear.

Dolan, P., Gudex, C., Kind, P., & Williams, A. Valuing health states: A comparison of methods. *Journal of Health Economics*, *15*(2), 209–231.

In eliciting health state valuations, two widely used methods are the standard gamble (SG) and the time trade off (TTO). Both methods make assumptions about individual preferences that are too restrictive to allow them to act as perfect proxies for utility. Therefore, a choice between them might instead be made on empirical grounds. This paper reports on a study that compared a "props" (using specifically-designed boards) and a "no props" (using self-completion booklets) variant of each method. The results suggested that both no props variants might be susceptible to framing effects and that TTP props outperformed SG props.

Ferreira, P. L., Ferreira, L. N., & Pereira, L. N. (2008). How consistent are health utility values? *Quality of Life Research, 17*(7), 1031–1042.

The use of preference-based generic instruments to measure the health-related quality of life of a general population or of individuals suffering from a specific disease has been increasing. However, there are several discrepancies between instruments in terms of utility results. This study compares SF-6D and EQ-5D when administered to patients with cataracts and aims at explaining the differences. Agreement between EQ-5D and SF-6D health state classifications was assessed by correlation coefficients. Simple correspondence analysis was used to assess the agreement among the instrument's descriptive systems and to investigate similarities between dimensions' levels. Cluster analysis was used to classify SF-6D and EQ-5D levels into homogeneous groups. There was evidence of floor effects in SF-6D and ceiling effects in EQ-5D. Comparisons of means showed that SF-6D values exceeded EQ-5D values. Agreement between both instruments was high, especially between similar dimensions. However, different valuation methods and scoring algorithms contributed to the main differences found. We suggest that one or both instruments should be revised, in terms of their descriptive systems or their scoring algorithms, in order to overcome the weakness found.

Fryback, D. G., Palta, M., Cherepanov, D., Bolt, D. & Kim, J. S. (2010). Comparison of 5 health-related quality-of-life indexes using item response theory analysis. *Medical Decision Making, 30*, 5–15.

BACKGROUND: Five health-related quality-of-life (HRQoL) indexes—EQ-5D, HUI2, HUI3, QWB-SA, and SF-6D—are each used to assign community-based utility scores to health states, although these scores differ. OBJECTIVE: The authors transform these indexes to a common scale to understand their interrelationships. METHODS: Data were from the National Health Measurement Study, a telephone survey of 3,844 U.S. adults. The five

indexes were analyzed using item response theory analysis to estimate scores on an underlying construct of summary health. Unidimensionality was evaluated using nonlinear principal components analysis. Index scores were plotted against the estimated scores on the common underlying construct. In addition, scores on the Health and Activities Limitation Index (HALex), the Centers for Disease Control and Prevention Healthy Days questions, and self-rated health on a 5-category scale ranging from excellent to poor were plotted. RESULTS: SF-6D and QWB-SA are nearly linear across the range but with a shallow slope; EQ-5D, HUI2, and HUI3 are linear with a steep slope from low (poor health) into midrange, then approximately linear with a less steep slope for higher (health just below to well above average), although the inflection points differ by index. CONCLUSION: Simple linear functions may serve as crosswalks among these indexes only for lower health states, albeit with low precision. Ceiling effects make crosswalks among most of the indexes ill specified above a certain level of health. Although each index measures generic health on a utility scale, these indexes are not identical but are relatively simply, if imprecisely, related.

Gardner, J. W., & Sanborn, J. S. (1990). Years of potential life lost (YPLL)—what does it measure? *Epidemiology*, *1*(4), 322–329.

The concept of years of potential life lost (YPLL) involves estimating the average time a person would have lived had he or she not died prematurely. This measure is used to help quantify social and economic loss owing to premature death, and it has been promoted to emphasize specific causes of death affecting younger age groups. YPLL inherently incorporates age at death, and its calculation mathematically weights the total deaths by applying values to death at each age. The method of calculating YPLL varies from author to author, each producing different rankings of leading causes of premature death. One can choose between heart disease, cancer, or accidents as the leading cause of premature death, depending on which method is used. Confusion in the use of this measure stems from a misunderstanding of the value system inherent in the calculation, as well as from differing views as to values that should be applied to each age at death.

Garratt, A. M., Hutchinson, A., Russell, L. (2000). Patient-assessed measures of health outcome in asthma: A comparison of four approaches. *Respiratory Medicine*, *94*(6), 597–606.

The study compares the psychometric properties of four different approaches to patient-assessed health outcomes in asthma: the Asthma Quality of Life Questionnaire (AQLQ), Newcastle Asthma Symptoms Questionnaire (NASQ), SF-12, and EuroQol. The instruments were administered by means of a self-completed postal questionnaire to 394 patients recruited from general practices in the North East of England. Patients completed a follow-up questionnaire at 6 months. The levels of missing data were assessed and instrument scores compared using correlational analysis. Scores were related to self-reports of smoking behaviour, socioeconomic status, and health transition. Responsiveness was assessed using standardized response means. Two hundred and thirty-five patients took part in the study

giving a response rate of 59.6%. There was a relatively large amount of missing data for the individualized section of the AQLQ. Correlational analysis provided evidence of convergent validity between the specific instruments; the largest correlation was found between NASQ scores and the asthma symptoms scale of the AQLQ (r = 0.84). The NASQ was found to be the most powerful at discriminating between smokers and non-smokers. All four instruments were linearly related to self-reported asthma transition (P<0.05); the specific instruments having the strongest association. The specific instruments showed good levels of responsiveness with the NASQ producing a large SRM of 0.82. SRMs for the AQLQ were of a moderate to large size (0.32–0.77) and the SRMs for the SF-12 and EuroQol were of a small size. The two specific instruments are capable of greater levels of discrimination between groups of patients and are more responsive to changes in health than the generic SF-12 and EuroQol. The greater responsiveness of the NASQ is probably due to its focus being restricted to symptoms of asthma compared to the broader focus of the AQLQ domains. The NASQ has a strong relationship with the AQLQ and is a more practical instrument that is more acceptable to patients. However, the AQLQ does measure broader patient concerns. The SF-12 and EuroQol have greater potential to capture side-effects and have wider scope for application in economic evaluation.

Garster, N. C., Palta, M. Sweitzer, N. K., Kaplan, R. M., & Fryback, D. G. (2009). Measuring health-related quality of life in population-based studies of coronary heart disease: Comparing six generic indexes and a disease-specific proxy score. *Quality of Life Research*, *18*(9), 1239–1247.

PURPOSE: To compare HRQoL differences with CHD in generic indexes and a proxy CVD-specific score in a nationally representative sample of U.S. adults. METHODS: The National Health Measurement Study, a cross-sectional random-digit-dialed telephone survey of adults aged 35 to 89, administered the EQ-5D, QWB-SA, HUI2, HUI3, SF-36v2 (yielding PCS, MCS, and SF-6D), and HALex. Analyses compared 3,350 without CHD (group 1), 265 with CHD not taking chest pain medication (group 2), and 218 with CHD currently taking chest pain medication (group 3), with and without adjustment for demographic variables and comorbidities. Data on 154 patients from heart failure clinics were used to construct a proxy score utilizing generic items probing CVD symptoms. RESULTS: Mean scores differed between CHD groups for all indexes with and without adjustment (P < 0.0001 for all except MCS P = 0.018). Unadjusted group 3 versus 1 differences were about three times larger than for group 2 versus 1. Standardized differences for the proxy score were similar to those for generic indexes, and were about 1.0 for all except MCS for group 3 versus 1. CONCLUSIONS: Generic indexes capture differences in HRQoL in population-based studies of CHD similarly to a score constructed from questions probing CVD-specific symptoms.

Glasziou, P., Alexander, J., Beller, E., & Clarke, P. (2007). Which health-related quality of life score? A comparison of alternative utility measures in patients with type 2 diabetes in the ADVANCE trial. *Health and Quality of Life Outcomes, 5*, 21.

BACKGROUND: Diabetes has a high burden of illness both in life years lost and in disability through related co-morbidities. Accurate assessment of the non-mortality burden requires appropriate health-related quality of life and summary utility measures of which there are several contenders. The study aimed to measure the impact of diabetes on various health-related quality of life domains and compare several summary utility measures. METHODS: In the ADVANCE (Action in Diabetes and Vascular Disease: Preterax and Diamicron MR Controlled Evaluation) study, 978 Australian patients with type 2 diabetes completed two health-related quality of life questionnaires at baseline: the EQ-5D and the SF-36v2, from which nine summary utility measures were calculated, and compared. The algorithms were grouped into four classes: (i) based on the EQ-5D, (ii) using fewer items than those in the SF-12, (iii) using the items in the SF-12, and (iv) using all items of the SF-36. RESULTS: Overall health-related quality of life of the subjects was good (mean utility ranged from 0.68 (±0.08) to 0.85(±0.14) over the nine utility measures) and comparable to patients without diabetes. Summary indices were well correlated with each other (r = 0.76 to 0.99), and showed lower health-related quality of life in patients with major diabetes-related events such as stroke or myocardial infarction. Despite the smaller number of items used in the scoring of the EQ-5D, it generally performed at least as well as SF-36 based methods. However, all utility measures had some limitation such as limited range or ceiling effects. CONCLUSION: The summary utility measures showed good agreement, and showed good discrimination between major and minor health state changes. However, EQ-5D based measures performed as well and are generally simpler to use.

Gold, M. R., Stevenson, D. & Fryback, D. G. (2002). HALYS and QALYS and DALYS, Oh my: Similarities and differences in summary measures of population health. *Annual Review of Public Health, 23*, 115–134.

Health-adjusted life years (HALYs) are population health measures permitting morbidity and mortality to be simultaneously described within a single number. They are useful for overall estimates of burden of disease, comparisons of the relative impact of specific illnesses and conditions on communities, and in economic analyses. Quality-adjusted life years (QALYs) and disability-adjusted life years (DALYs) are types of HALYs whose original purposes were at variance. Their growing importance and the varied uptake of the methodology by different U.S. and international entities makes it useful to understand their differences as well as their similarities. A brief history of both measures is presented and methods for calculating them are reviewed. Methodological and ethical issues that have been raised in association with HALYs more generally are presented. Finally, we raise concerns about the practice of using different types of HALYs within different decision-making contexts and urge action that builds and clarifies this useful measurement field.

Greenley, J. R., Greenberg, J. S., & Brown, R. (1997). Measuring quality of life: A new and practical survey instrument. *Journal of Social Work*, *42*(3), 244–254.

Despite increasing recognition of the importance of measuring the quality of life (QOL) of people with severe mental illness, such assessments are seldom carried out because of the lack of an efficient, easy-to-use, and valid measurement instrument. To facilitate the gathering of QOL information from clients in evaluation, program improvement, or other efforts, the authors present a new short (24-item) self-administered questionnaire called the Quality of Life Questionnaire (QLQ). The questionnaire assesses subjective QOL in seven areas. Evidence for the reliability and validity of the QLQ is based on data gathered from 971 clients with serious mental illness who were receiving publicly funded mental health services at the time of the study. The results of a confirmatory factor analysis using a random split-half procedure indicated that a seven-factor solution fit the data well. Scores on the QLQ also correlated significantly with the client's functioning and satisfaction with services, providing support for the validity of the QLQ. The advantages of the QLQ over existing measures include low-cost administration and some superior psychometric properties.

Gyrd-Hansen, D. (2005). Willingness to pay for a QALY: Theoretical and methodological issues. *Pharmacoeconomics*, *23*(5), 423–432.

What is a QALY worth in monetary units? This paper presents the main arguments in the literature regarding the obstacles involved in establishing one unique willingness to pay (WTP) estimate for the value of a QALY. To directly translate QALYs into monetary units, and in this manner translate existing and forthcoming cost-effectiveness analyses (CEA) to cost-benefit analyses (CBA), it is necessary that one unique WTP per QALY can be established irrespective of context-specific characteristics such as severity of illness, magnitude of health gain, patient characteristics, etc. Because CEA and CBA are two methods of economic evaluation that are based on two very different normative perceptions of the role of health versus other goods in society, the task of performing a linear translation from QALYs to WTP is theoretically unattainable. CBA is based on the welfarist perception that the welfare associated with health is measured by way of individual preferences for health outcomes relative to other goods in society. In contrast, CEA is based on the extra-welfarist notion, which focuses on maximizing health and not welfare, and suppresses any variation across income/social groups in utility derived from improvements in health. Another obstacle to one unique WTP per QALY value is that marginal utility of income is non-constant, and a function of income level and possibly health status. When marginal utility of income varies across individuals as well as contexts, measuring the value of health in monetary units may result in valuations of health increments that are very different from valuations retrieved had another unit of measure been applied. In conclusion, from a theoretical point of view, establishing one unique WTP cannot be attained. Applying one sole WTP per QALY value will entail overriding individual preferences such as

diminishing marginal utility of health and potential differences in the value of incremental health across population groups. However, one problem that can, and should, be overcome when seeking to establish a monetary value for a QALY is the problem of variance in the marginal utility of income. The importance of applying the appropriate perspective when formulating WTP questions to ensure that the marginal utility of income of the respondents equals that of the financiers of the costs invested to produce the health gains should not be overlooked.

Harrison, M. J., Boonen, A., Tugwell, P., & Symmons, D. P. (2009). Same question, different answers: A comparison of global health assessments using visual analogue scales. *Quality of Life Research*, *18*(10), 1285–1292.

PURPOSE: To compare responses to two global health VAS of patients with rheumatoid arthritis at the same assessment within the same questionnaire. METHODS: Secondary analysis of randomized controlled trial data. Patients completed the patient global assessment VAS (PG-VAS) (horizontal 10 cm scale, left (0), right (100), no incremental markers) and EQ-5D-VAS (EQ-VAS) (vertical 20 cm scale, 100 at the top, markers at each increment of 10). Both asked "how good or bad is your health today, in your opinion, from 100 'Best imaginable health state' to 0 'Worst imaginable health state'." Agreement was assessed using intra-class correlation coefficients (ICC) and Bland-Altman plots. RESULTS: Four hundred and forty-eight patients reported median PG-VAS 66 (IQR 51, 77) and EQ-VAS 65 (IQR 50, 80) scores. Correlation of the VAS scales was moderate at baseline (ICC 0.564) and longitudinally (ICC 0.503). Bland-Altman plots suggested poor concordance of the PG-VAS and EQ-VAS; the limits of agreement were ±32.3 on a 0–100 scale. PG-VAS scores were evenly distributed; EQ-VAS scores clustered at increments of 10; rounding did not improve agreement. CONCLUSIONS: The EQ-VAS and PG-VAS scores are not interchangeable at the individual level. The EQ-VAS correlated more strongly with disease-specific and health-related quality of life measures, therefore, appears preferable.

Heyworth, I. T., Hazell, M. L., Linehan, M. F., & Frank, T. L. (2009). How do common chronic conditions affect health-related quality of life? *British Journal of General Practice*, *59*(568), e353–358.

BACKGROUND: Comorbidity is common. National Institute for Health and Clinical Excellence (NICE) guidelines have been created to make best use of resources to improve patients' quality of life but do not currently take account of comorbidity. The effect of multiple chronic conditions with regard to health-related quality of life (HRQoL) is poorly researched. Criticisms of previous research have been due to patient-defined chronic conditions, lack of quantification of the effects of confounding factors, selection of affected patients only, small sample sizes, and upper age limits. AIM: This study aims to address these issues, looking into the impact of combinations of chronic conditions on HRQoL. DESIGN OF THE STUDY: Participants filled in a questionnaire containing general health information, specific respiratory questions, and the EQ-5D measure of HRQoL. The questionnaires were then

matched up to their GP records to obtain their disease status for six common chronic diseases (asthma, chronic obstructive pulmonary disease, ischaemic heart disease, hypertension, diabetes, and cerebrovascular disease). METHOD: Data from a mailed questionnaire were analyzed from 5169 patients aged >16 years from two general practices in Wythenshawe, Manchester, in 2004. Completion of the questionnaire was taken to indicate consent to participate. RESULTS: Significant correlations were found between a lower HRQoL and increasing numbers of chronic conditions (P<0.001), increasing age, possible obstructive airway disease, lack of higher education, smoking, and female sex. These all remained significant following regression, except for sex, with number of chronic conditions being a strong predictor of the weighted health state index score, EQ-5D(index) (coefficient = −0.079, P<0.001). CONCLUSION: Increasing numbers of chronic conditions have a strong negative effect on HRQoL.

Hilditch, J. R., Lewis, J., Peter, A., van Maris, B., Ross, A., Franssen, E., Guyatt, G. H., et al. (2008). A menopause-specific quality of life questionnaire: Development and psychometric properties. *Maturitas, 61*(1-2), 107–121.

OBJECTIVE: To develop a condition-specific quality of life questionnaire for the menopause with documented psychometric properties, based on women's experience. METHODS: *Subjects*: Women 2–7 years post-menopause with a uterus and not currently on hormone replacement therapy. *Questionnaire development*: A list of 106 menopause symptoms was reduced using the importance score method. Replies to the item-reduction questionnaire from 88 women resulted in a 30-item questionnaire with four domains (vasomotor, physical, psychosocial, and sexual) and a global quality of life question. *Psychometric properties*: A separate sample of 20 women was used to determine face validity, and a panel of experts was used to confirm content validity. Reliability, responsiveness and construct validity were determined within the context of a randomized controlled trial. Construct validation involved comparison with the Neugarten and Kraines' Somatic, Psychosomatic and Psychologic subscales, the reported intensity of hot flushes, the General Well-Being Schedule, Channon and Ballinger's Vaginal Symptoms Score and Libido Index, and the Life Satisfaction Index. RESULTS: The face validity score was 4.7 out of a possible 5. Content validity was confirmed. Test-retest reliability measures, using intraclass correlation coefficients were 0.81, 0.79, 0.70 and 0.55 for the physical, psychosocial, sexual domains and the quality of life question. The intraclass correlation coefficient for the vasomotor domain was 0.37 but there is evidence of systematic change. Discriminative construct validity showed correlation coefficients of 0.69 for the physical domain, 0.66 and 0.40 for the vasomotor domain, 0.65 and −0.71 for the psychosocial domain, 0.48 and 0.38 for the sexual domain, and 0.57 for the quality of life question. Evaluative construct validity showed correlation coefficients of 0.60 for the physical domain, 0.28 for the vasomotor domain, 0.55 and –0.54 for the psychosocial domain, 0.54 and 0.32 for the sexual domain, and 0.12 for the quality of life question. Responsiveness scores ranged from 0.78 to 1.34. CONCLUSIONS: The MENQOL

(Menopause-Specific Quality of Life) questionnaire is a self-administered instrument which functions well in differentiating between women according to their quality of life and in measuring changes in their quality of life.

Hirth, R. A., Chernew, M. E., Miller, E. & Fendrick, A. M. (2000). Willingness to pay for a quality-adjusted life year: In search of a standard. *Medical Decision Making, 20*(3), 332–342.

Cost-benefit analysis (CBA) provides a clear decision rule: undertake an intervention if the monetary value of its benefits exceed its costs. However, due to a reluctance to characterize health benefits in monetary terms, users of cost-utility and cost-effectiveness analyses must rely on arbitrary standards (e.g., < $50,000 per QALY) to deem a program "cost-effective." Moreover, there is no consensus regarding the appropriate dollar value per QALY gained upon which to base resource allocation decisions. To address this, the authors determined the value of a QALY as implied by the value-of-life literature and compared this value with arbitrary thresholds for cost-effectiveness that have come into common use. A literature search identified 42 estimates of the value of life that were appropriate for inclusion. These estimates were classified by method: human capital (HK), contingent valuation (CV), revealed preference/job risk (RP-JR), and revealed preference/non-occupational safety (RP-S), and by U.S. or non-U.S. origin. After converting these value-of-life estimates to 1997 U.S. dollars, the life expectancy of the study population, age-specific QALY weights, and a 3% real discount rate were used to calculate the implied value of a QALY. An ordinary least-squares regression of the value of a QALY on study type and national origin explained 28.4% of the variance across studies. Most of the explained variance was attributable to study type; national origin did not significantly affect the values. Median values by study type were $24,777 (HK estimates), $93,402 (RP-S estimates), $161,305 (CV estimates), and $428,286 (RP-JR estimates). With the exception of HK, these far exceed the "rules of thumb" that are frequently used to determine whether an intervention produces an acceptable increase in health benefits in exchange for incremental expenditures.

Janssen, M. F., Birnie, E., & Bonsel, G. J. (2008). Quantification of the level descriptors for the standard EQ-5D three-level system and a five-level version according to two methods. *Qual Life Res, 17*(3), 463–473.

OBJECTIVES: Our aim was to compare the quantitative position of the level descriptors of the standard EQ-5D three-level system (3L) and a newly developed, experimental five-level version (5L) using a direct and a vignette-based indirect method. METHODS: Eighty-two respondents took part in the study. The direct method represented a visual analog scale (VAS) rating of the nonextreme level descriptors for each dimension and each instrument separately. The indirect method required respondents to score 15 health scenarios with 3L, 5L, and a VAS scale. Investigated were (1) equidistance (Are 3L and 5L level descriptors distributed evenly over the VAS continuum?); (2) isoformity (Do the identical level descriptors on 3L and 5L yield similar results?); and (3) consistency between dimensions

(Do the positions of similar level descriptors differ across dimensions within instruments?). RESULTS: Equidistance without transformation was rejected for all dimensions for both 3L and 5L but satisfied for 5L after transformation. Isoformity gave mixed results. Consistency between dimensions was satisfied for both instruments and both methods. DISCUSSION: The level descriptors have similar distributions across comparable dimensions within each system, but the pattern differs between 3L and 5L. This methodological study provides evidence of increased descriptive power and a broadened measurement continuum that encourages the further development of an official five-level EQ-5D.

Klose, T. (2003). A utility-theoretic model for QALYs and willingness to pay. *Health Economics*, *12*(1), 17–31.

Despite the widespread use of quality-adjusted life years (QALY) in economic evaluation studies, their utility-theoretic foundation remains unclear. A model for preferences over health, money, and time is presented in this paper. Under the usual assumptions of the original QALY-model, an additive separable presentation of the utilities in different periods exists. In contrast to the usual assumption that QALY-weights do solely depend on aspects of health-related quality of life, wealth-standardized QALY-weights might vary with the wealth level in the presented extension of the original QALY-model resulting in an inconsistent measurement of QALYs. Further assumptions are presented to make the measurement of QALYs consistent with lifetime preferences over health and money. Even under these strict assumptions, QALYs and WTP (which also can be defined in this utility-theoretic model) are not equivalent preference-based measures of the effects of health technologies on an individual level. The results suggest that the individual WTP per QALY can depend on the magnitude of the QALY-gain as well as on the disease burden, when health influences the marginal utility of wealth. Further research seems to be indicated on this structural aspect of preferences over health and wealth and to quantify its impact.

Knapp, M., & Mangalore, R. (2007). The trouble with QALYs.... *Epidemiologia e Psichiatria Sociale, 16*(4), 289–293.

This paper summarizes the use of QALYs in evaluating changes in mental health states, highlighting the benefits and challenges of their use in this field. The general principles underlying the QALY measure and the most common methods of measuring QALYs are discussed briefly. Evidence of the usefulness and problems of using this generic measure of health-related quality of life are provided from a sample of recent studies relating to depression, schizophrenia, attention deficit hyperactivity disorder, and dementia. In each case, attempts were made to use QALYs to measure changes in health states. While in principle, the QALY is enormously attractive, its suitability for measuring changes in many mental health conditions remains open to doubt as existing tools for generating QALY scores such as the EQ-5D have tended not to perform sufficiently well in reflecting changes in many mental health states. New developmental work is needed to construct better QALY-

measuring tools for use in the mental health field. Both the conceptualization and measurement of QALYs need to be built on a valid, comprehensive model of quality of life specific to a mental health disorder, to ensure that the resultant tool is sensitive enough to pick up changes that would be expected and seen as relevant in the course of the illness.

Knies, S., Evers, S. M., Candel, M. J., Severens, J. L., & Ament, A. J. (2009). Utilities of the EQ-5D: Transferable or Not? *Pharmacoeconomics*, *27*(9), 767–779.

Within the framework of economic evaluations, the transferability of utility scores between jurisdictions remains unclear. The EQ-5D is a generic instrument for measuring health-related quality of life in economic evaluations, which can be used for comparing utility scores across countries. At present, the EQ-5D has several national value sets or tariffs. Nevertheless, utility estimates from foreign studies are often used directly for cost-effectiveness estimates, without adapting by applying the appropriate national value set. It is unclear if this practice is advisable, due to dissimilarities between the national value sets. To examine the effects of differences in national EQ-5D value sets on absolute and marginal utilities of health states, and determine to what degree these differences can be explained by methodological factors. First, the relative importance of the EQ-5D domains for the utility estimates was compared across the 15 value sets. Second, two hypothetical health states for a depressed patient and a pain patient (21232 and 33321) were selected for additional analysis, by comparing the utilities as scored by the value sets. The marginal influence of a one-level deterioration in a domain of these health states on the utility estimate was then determined. Third, the differences between the value sets were examined in more detail by using multilevel analysis to examine the role of methodological differences in the valuation studies. Differences can be perceived between the national value sets of the EQ-5D in the preferences for the domains. The utilities of the two hypothetical health states show that the value sets differ substantially. Furthermore, the differences between the marginal values of the deteriorations are large, which can be explained partly by the type of valuation method. Other methodological differences also influence the value sets. All results indicate that the differences between the EQ-5D value sets are considerable and should not be ignored. The differences can largely be explained by methodological differences in the valuation studies. The remaining differences may reflect cultural dissimilarities between countries. Therefore, further research should focus on investigating the transferability of utilities across countries or agreeing on a standard to perform valuation studies. For the time being, transferring utilities from one country to another without any adjustment is not advisable.

Krabbe, P. F., Essink-Bot, M. L., & Bonsel, G. J. (1997). The comparability and reliability of five health-state valuation methods. *Social Science & Medicine, 45*(11), 1641–1652.

The objective of the study was to consider five methods for valuing health states with respect to their comparability (convergent validity, value functions) and reliability. Valuation

tasks were performed by 104 student volunteers using five frequently used valuation methods: standard gamble (SG), time trade-off (TTO), rating scale (RS), willingness-to-pay (WTP), and the paired comparisons method (PC). Throughout the study, the EuroQol classification system was used to construct 13 health-state descriptions. Validity was investigated using the multitrait-multimethod (MTMM) methodology. The extent to which results of one method could be predicted by another was examined by transformations. Reliability of the methods was studied parametrically with Generalizability Theory (an ANOVA extension), as well as non-parametrically. Mean values for SG were slightly higher than TTO values. The RS could be distinguished from the other methods. After a simple power transformation, the RS values were found to be close to SG and TTO. Mean values of WTP were linearly related to SG and TTO, except at the extremes of the scale. However, the reliability of WTP was low and the number of inconsistencies substantial. Valuations made by the RS proved to be the most reliable. Paired comparisons did not provide stable results. In conclusion, the results of the parametric transformation function between RS and SG/TTO provide evidence to justify the current use of RS (with transformations) not only for reasons of feasibility and reliability but also for reasons of comparability. A definite judgment on PC requires data of a complete design. Due to the specific structure of the correlation matrix which is inherent in valuing health states, we believe that full MTMM is not applicable for the standard analysis of health-state valuations.

Law, A. V., Pathak, D. S., McCord, M. R. (1998). Health status utility assessment by standard gamble: A comparison of the probability equivalence and the lottery equivalence approaches. *Pharamceutical Appoaches, 15*(1), 105–109.

PURPOSE: Utility values obtained with the standard gamble (SG) method using the probability equivalence approach (PE) have a reported bias due to the "certainty effect." This effect causes individuals to overvalue a positive outcome when it occurs under certainty. Researchers in the decision sciences have proposed an alternative, "lottery equivalence" (LE) approach, using paired gambles, to eliminate this bias. The major objective of the current study was to investigate the certainty effect in health status utility measures and to test our hypothesis that the certainty effect would act in a reverse direction for negatively valued outcomes. METHODS: Fifty-four subjects completed the study by assessing preferences for three health states by rating scale and then by SG using PE as well as LE approaches with assessment lotteries of 0.5 and 0.75. RESULTS: The results from 41 useable responses point towards possible existence of the certainty effect in health in the hypothesized direction: utility values obtained with the PE were significantly lower than with the LEs. There was no significant difference between the LE values indicating elimination of the bias. CONCLUSIONS: The results have important implications since the SG using PE is thought be the "gold standard" in health status utility measurements.

Lua, P. L., Salek, S., Finlay, I., & Lloyd-Richards, C. (2005). The feasibility, reliability and validity of the McGill Quality of Life Questionnaire-Cardiff Short Form (MQOL-CSF) in palliative care population. *Quality of Life Research, 14*(7), 1669–1681.

In terminally-ill patients, effective measurement of health-related quality of life (HRQoL) needs to be done while imposing minimal burden. In an attempt to ensure that routine HRQoL assessment is simple but capable of eliciting adequate information, the McGill Quality of Life Questionnaire-Cardiff Short Form (MQOL-CSF: 8 items) was developed from its original version, the McGill Quality of Life Questionnaire (MQOL: 17 items). Psychometric properties of the MQOL-CSF were then tested in palliative care patients consisting of 55 out-patients, 48 hospice patients and 86 in-patients: The MQOL-CSF had little respondent burden (mean completion time = 3.3 min) and was evaluated as "very clear" or "clear" (98.2%), comprehensive (74.5%), and acceptable (96.4%). The internal consistency reliability was moderate to high (Cronbach's alpha = 0.462–0.858), and test-retest reliability (Spearman's r[s]) ranged from 0.512–0.861. Correlation was moderate to strong (0.478–0.725) between items in the short form and their analogous domains in the MQOL. Most MQOL-CSF items showed strong associations with their own domain (r(s) > or = 0.40). Scores from MQOL-CSF significantly differentiated between patients with differing haemoglobin levels ($p < 0.05$). Construct validity was overall supported by principal component analysis. It is concluded that the MQOL-CSF is a feasible tool with favorable psychometric properties for routine HRQoL assessment in the palliative care population.

Luo, N., Johnson, J. A., Shaw, J. W., & Coons, S. J. (2009). Relative efficiency of the EQ-5D, HUI2, and HUI3 index scores in measuring health burden of chronic medical conditions in a population health survey in the United States. *Medical Care, 47*(1), 53–60.

OBJECTIVE: We sought to compare the ability of the EQ-5D, Health Utilities Index Mark 2 (HUI2), and HUI Mark 3 (HUI3) index scores to discriminate between respondents based on the presence or absence of chronic medical conditions in a population health survey. METHODS: Secondary analyses were conducted with data from a probability sample (n = 3480, mean age: 42.5 years, male: 42.4%, Hispanic: 28.6%) of the 2001 noninstitutionalized U.S. general adult population. F-statistic ratios were used to evaluate the relative efficiency of the EQ-5D, HUI2, and HUI3 in differentiating respondents with or without each of 18 chronic medical conditions, and differentiating respondents with low- or high-burden conditions. RESULTS: In comparing respondents with and without chronic medical conditions, the F-statistic values of these three indices were not significantly different, except for EQ-5D versus HUI2 (mean F-statistic ratio: 0.79, 95% confidence interval [CI]: 0.59–0.98). In comparing respondents with a low-burden condition with those with a high-burden condition, the F-statistic values of EQ-5D and HUI2 index scores were similar, while those for EQ-5D versus HUI3 (mean: 0.79; 95% CI: 0.66–0.92) and for HUI2 versus HUI3 (mean: 0.83; 95% CI: 0.71–0.95) were significantly less than 1.0. The overall ceiling effects of the EQ-5D, HUI2, and HUI3 index scores were 48.9%, 15.4%, and 15.3%,

respectively. CONCLUSIONS: Although the EQ-5D seems to be marginally less informative, the EQ-5D, HUI2, and HUI3 index scores were generally comparable in determining health burden of chronic medical conditions in this population health survey data.

Maciosek, M. V., Coffield, A. B., Edwards, N. M., Flottemesch, T. J., Goodman, M. J., & Solberg, L. I. (2006). Priorities among effective clinical preventive services: Results of a systematic review and analysis. *American Journal of Preventive Medicine, 31*(1), 52–61.

BACKGROUND: Decision makers at multiple levels need information about which clinical preventive services matter the most so that they can prioritize their actions. This study was designed to produce comparable estimates of relative health impact and cost effectiveness for services considered effective by the U.S. Preventive Services Task Force and Advisory Committee on Immunization Practices. METHODS: The National Commission on Prevention Priorities (NCPP) guided this update to a 2001 ranking of clinical preventive services. The NCPP used new preventive service recommendations up to December 2004, improved methods, and more complete and recent data and evidence. Each service received 1 to 5 points on each of two measures—clinically preventable burden and cost effectiveness—for a total score ranging from 2 to 10. Priorities for improving delivery rates were established by comparing the ranking with what is known of current delivery rates nationally. RESULTS: The three highest-ranking services each with a total score of 10 are discussing aspirin use with high-risk adults, immunizing children, and tobacco-use screening and brief intervention. High-ranking services (scores of 6 and above) with data indicating low current utilization rates (around 50% or lower) include: tobacco-use screening and brief intervention, screening adults aged 50 and older for colorectal cancer, immunizing adults aged 65 and older against pneumococcal disease, and screening young women for Chlamydia. CONCLUSION: This study identifies the most valuable clinical preventive services that can be offered in medical practice and should help decision-makers select which services to emphasize.

Mahadevia, P. J., Fleisher, L. A., Frick, K. D., Eng, J., Goodman, S. N., & Powe, N. R. (2003). Lung cancer screening with helical computed tomography in older adult smokers: A decision and cost-effectiveness analysis. *Journal of the American Medical Association, 289*(3), 313–322.

CONTEXT: Encouraged by direct-to-consumer marketing, smokers and their physicians are contemplating lung cancer screening with a promising but unproven imaging procedure, helical computed tomography (CT). OBJECTIVE: To estimate the potential benefits, harms, and cost-effectiveness of lung cancer screening with helical CT in various efficacy scenarios. DESIGN, SETTING, AND POPULATION: Using a computer-simulated model, we compared annual helical CT screening to no screening for hypothetical cohorts of 100 000 current, quitting, and former heavy smokers, aged 60 years, of whom 55% were men. We simulated efficacy by changing the clinical stage distribution of lung cancers so that the screened

group would have fewer advanced-stage cancers and more localized-stage cancers than the nonscreened group (ie, a stage shift). Our model incorporated known biases in screening programs such as lead time, length, and overdiagnosis bias. MAIN OUTCOME MEASURES: We measured the benefits of screening by comparing the absolute and relative difference in lung cancer-specific deaths. We measured harms by the number of false-positive invasive tests or surgeries per 100 000 and incremental cost-effectiveness in US dollars per quality-adjusted life-year (QALY) gained. RESULTS: Over a 20-year period, assuming a 50% stage shift, the current heavy smoker cohort had 553 fewer lung cancer deaths (13% lung cancer-specific mortality reduction) and 1186 false-positive invasive procedures per 100 000 persons. The incremental cost-effectiveness for current smokers was $116 300 per QALY gained. For quitting and former smokers, the incremental cost-effectiveness was $558 600 and $2 322 700 per QALY gained, respectively. Other than the degree of stage shift, the most influential parameters were adherence to screening, degree of length or overdiagnosis bias in the first year of screening, quality of life of persons with screen-detected localized lung cancers, cost of helical CT, and anxiety about indeterminate nodule diagnoses. In 1-way sensitivity analyses, none of these parameters was sufficient to make screening highly cost-effective for any of the cohorts. In multiway sensitivity analyses, a program screening current smokers was $42 500 per QALY gained if extremely favorable estimates were used for all of the influential parameters simultaneously. CONCLUSION: Even if efficacy is eventually proven, screening must overcome multiple additional barriers to be highly cost-effective. Given the current uncertainty of benefits, the harms from invasive testing, and the high costs associated with screening, direct-to-consumer marketing of helical CT is not advisable.

Manuel, D. G., & Schultz, S. E. (2004). Health-related quality of life and health-adjusted life expectancy of people with diabetes in Ontario, Canada, 1996–1997. *Diabetes Care, 27*(2), 407–414.

OBJECTIVE: To estimate the burden of illness from diabetes using a population health survey linked to a population-based diabetes registry. RESEARCH DESIGN AND METHODS: Measures of health-related quality of life (HRQoL) from the 1996/97 Ontario Health Survey (n = 35,517) were combined with diabetes prevalence and mortality data from the Ontario Diabetes Database (n = 487,576) to estimate the impact of diabetes on life expectancy, health-adjusted life expectancy (HALE), and HRQoL. RESULTS: Life expectancy of people with diabetes was 64.7 and 70.7 years for men and women, respectively—12.8 and 12.2 years less than that for men and women without diabetes. Diabetes had a large impact on instrumental and basic activities of daily living, more so than on functional health. HALE was 58.3 and 62.7 years, respectively, for men and women—11.9 and 10.7 years less than that of men and women without diabetes. Eliminating diabetes would increase Ontario life expectancy by 2.8 years for men and 2.6 years for women; HALE would increase by 2.7 and 3.2 years for men and women, respectively. CONCLUSIONS: The burden of illness from

diabetes in Ontario is considerable. Efforts to reduce diabetes would likely result in a "compression of morbidity." An approach of estimating diabetes burden using linked data sources provides a robust approach for the surveillance of diabetes.

McDonough, C. M., & Tosteson, A. N. (2007). Measuring preferences for cost-utility analysis: How choice of method may influence decision-making. *Pharmacoeconomics, 25*(2), 93–106.

Preferences for health are required when the economic value of health-care interventions are assessed within the framework of cost-utility analysis. The objective of this paper was to review alternative methods for preference measurement and to evaluate the extent to which the method may affect health care decision-making. Two broad approaches to preference measurement that provide societal health state values were considered: (i) direct measurement, and (ii) preference-based health state classification systems. Among studies that compared alternative preference-based systems, the EQ-5D tended to provide larger change scores and more favorable cost-effectiveness ratios than the Health Utilities Index (HUI) −2 and −3, while the SF-6D provided smaller change scores and less favorable ratios than the other systems. However, these patterns may not hold for all applications. Empirical evidence comparing systems and decision-making impact suggests that preferences will have the greatest impact on economic analyses when chronic conditions or long-term sequelae are involved. At present, there is no clearly superior method, and further study of cost-effectiveness ratios from alternative systems is needed to evaluate system performance. Although there is some evidence that incremental cost-effectiveness ratio (ICER) thresholds (e.g. $US50,000 per QALY gained) are used in decision making, they are not strictly applied. Nonetheless, as ICERs rise, the probability of acceptance of a new therapy is likely to decrease, making the differences in QALYs obtained using alternative methods potentially meaningful. It is imperative that those conducting cost-utility analyses characterize the impact that uncertainty in health state values has on the economic value of the interventions studied. Consistent reporting of such analyses would provide further insight into the policy implications of preference measurement.

McKenzie, L., & van der Pol, M. (2008). Mapping the EORTC QLQ C-30 onto the EQ-5D Instrument: The potential to estimate QALYs without generic preference data. *Value Health, 12*(1), 167–171.

OBJECTIVES: The aim of this article is to map the European Organization for Research and Treatment of Cancer (EORTC) QLQ C-30 onto the EQ-5D measure to enable the estimation of health state values based on the EORTC QLQ C-30 data. The EORTC QLQ C-30 is of interest because it is the most commonly used instrument to measure the quality of life of cancer patients. METHODS: Regression analysis is used to establish the relationship between the two instruments. The performance of the model is assessed in terms of how well the responses to the EORTC QLQ C-30 predict the EQ-5D responses for a separate data set. RESULTS: The results showed that the model explaining EQ-5D values predicted well.

All of the actual values were within the 95% confidence intervals of the predicted values. More importantly, predicted difference in quality-adjusted life-years (QALYs) between the arms of the trial was almost identical to the actual difference. CONCLUSION: There is potential to estimate EQ-5D values using responses to the disease-specific EORTC QLQ C-30 measure of quality of life. Such potential implies that in studies that do not include disease-specific measures, it might still be possible to estimate QALYs.

Osman, L., & Silverman, M. (1996). Measuring quality of life for young children with asthma and their families. *European Respiratory Journal Supplement, 21*, 35s–41s.

In assessing therapeutic interventions in asthma we have become increasingly aware of the importance of measuring outcomes which relate to patient experience of illness and its impact on their lives. These patient oriented assessments are usually defined as "quality of life" measures. There are good reasons for wanting to measure the impact of disease on people with asthma. Quality of life is likely to be related to health behavior, such as adherence to therapy and use of health resources. Quality of life may be a stronger predictor of these behaviors than objective symptoms. Thus, in evaluating health-related interventions, quality of life is an important dimension of outcome measurement. Adult measures are now frequently used but there has been less development of measures suitable for children. Important issues in measuring quality of life for children include the development of age-appropriate scales, measurement of the impact of illness on the whole family, and the relationship between child's report and proxy report. Recently developed asthma specific scales for children include the Child Asthma Questionnaire (CAQ), and the Paediatric Asthma Quality of Life Questionnaire (PAQLQ). The PAQLQ comprises a form directly assessing child quality of life and a form assessing pediatric caregivers (usually a parent). The impact of child asthma on the family as a whole may be particularly important not only for comparing benefits of different interventions, but also for predicting outcomes, such as medical help-seeking.

Paz, S. H., Liu, H., Fongwa, M. N., Morales, L. S., & Hays, R. D. (2009). Readability estimates for commonly used health-related quality of life surveys. *Quality of Life Research, 18*(7), 889–900.

PURPOSE: To estimate readability of seven commonly used health-related quality of life instruments: SF-36, HUI, EQ-5D, QWB-SA, HALex, Minnesota Living with Heart Failure Questionnaire (MLHFQ), and the NEI-VFQ-25. METHODS: The Flesch-Kincaid (F-K) and Flesch Reading Ease (FRE) formulae were used to estimate readability for every item in each measure. RESULTS: The percentage of items that require more than 5 years of formal schooling according to F-K was 50 for the EQ-5D, 53 for the SF-36, 80 for the VFQ-25, 85 for the QWB-SA, 100 for the HUI, HALex, and the MLHFQ. The percentage of items deemed harder than "easy" according to FRE was 50 for the SF-36, 67 for the EQ-5D, 79 for the QWB-SA, 80 for the VFQ-25, 100 for the HUI, HALex, and the MLHFQ. CONCLUSIONS: All

seven surveys have a substantial number of items with high readability levels that may not be appropriate for the general population.

Raisch, D. W. (2000). Understanding quality-adjusted life years and their application to pharmacoeconomic research. *Annals of Pharmacotherapy, 34*(7-8), 906–914.

OBJECTIVE: To provide a basic overview of quality-adjusted life years (QALYs) and their application in cost-effectiveness analysis (CEA), compare and contrast QALYs with other health-related quality-of-life (HRQoL) assessments, describe current controversies regarding QALYs, and provide comparisons between QALY instruments. METHODS: The literature regarding HRQoL and QALYs was reviewed and key issues are summarized. RESULTS: QALYs provide relative preferences of patients for different health states. They range from 0, representing death, to 1.0, representing optimal health. QALYs are distinguished from other HRQoL assessments in that they provide a summary measurement that incorporates quantity of life in addition to HRQoL. When QALYs are used as the outcome measure in CEA, a cost per QALY is calculated. The Panel on Cost Effectiveness in Health and Medicine has recommended that QALYs be used as outcome measures in CEA; when QALYs are used in CEA, comparisons between treatments for different illnesses as well as within an illness are possible. The three most commonly used preference measurement techniques in determining QALYs are visual analog scales, time trade-off, and standard gamble. Controversies regarding QALYs include which preference measurement technique is most appropriate, whether QALY assessments should be obtained from patients or the community, and how to address states, such as coma, that individuals sometimes assess as worse than death. QALY instruments can be compared regarding preference measurement technique, HRQoL domains assessed, ease of administration, validity, reliability, and sensitivity. CONCLUSIONS: When used appropriately, QALYs provide valuable outcome measures for pharmacoeconomic research.

Rector, T., Cubo, S. & Cohn, J. (1987) Patient's self assessment of their congestive heart failure. Part 2: Content, reliability and validity of a new measure, the Minnesota Living with Heart Failure Questionnaire. *Heart Failure*, *3*, 192–196.

To determine the reliability and validity of a patient outcome questionnaire for chronic heart failure, a randomized, double-blind, placebo-controlled, 3-month trial of pimobendan, an investigational medication with inotropic and vasodilator activities, was performed. Evaluated were 198 ambulatory patients with primarily New York Heart Association (NYHA) class III heart failure from 20 referral centers. Baseline therapy included digoxin, diuretics and, in 80%, a converting enzyme inhibitor. Oral pimobendan at 2.5 (n = 49), 5.0 (n = 51), or 10 (n = 49) mg daily or matching placebo (n = 49) was administered. The Minnesota Living with Heart Failure (LIhFE) questionnaire was a primary outcome measure, along with an exercise test. Inter-item correlations identified subgroups of questions representing physical and emotional dimensions. Repeated baseline scores were highly correlated (r = 0.93), as were the physical (r = 0.89) and emotional (r = 0.88) dimension scores. Placebo

did not have a significant effect with median (25th, 75th percentile) changes from baseline scores of 1 (−3, 5), 1 (−2, 3), and 0 (−1, 2), respectively (all *p* values > 0.10). The 5 mg dose significantly improved the total score, 7.5 (0, 18; $p = 0.01$) and the physical dimension, 4 (0, 8; $p = 0.01$), compared with placebo. Changes in the total ($r = 0.33$; $p < 0.01$) and physical ($r = 0.35$; $p < 0.01$) scores were weakly related to changes in exercise times, but corresponded well with changes in patients' ratings of dyspnea and fatigue. These data suggest that the LIhFE questionnaire was a reliable and valid patient self-assessment of the therapeutic benefit from pimobendan.

Robertson, C., Langston, A. L., Stapley, S., McColl, E., Campbell, M. K., Maclennan, G., et al. (2009). Meaning behind measurement: Self-comparisons affect responses to health-related quality of life questionnaires. *Quality of Life Research*, *18*(2), 221–230.

PURPOSE: The subjective nature of quality of life is particularly pertinent to the domain of health-related quality of life (HRQoL) research. The extent to which participants' responses are affected by subjective information and personal reference frames is unknown. This study investigated how an elderly population living with a chronic metabolic bone disorder evaluated self-reported quality of life. METHODS: Participants (n = 1,331) in a multi-centre randomized controlled trial for the treatment of Paget's disease completed annual HRQoL questionnaires, including the SF-36, EQ-5D and HAQ. Supplementary questions were added to reveal implicit reference frames used when making HRQoL evaluations. Twenty-one participants (11 male, 10 female, aged 59–91 years) were interviewed retrospectively about their responses to the supplementary questions, using cognitive interviewing techniques and semistructured topic guides. RESULTS: The interviews revealed that participants used complex and interconnected reference frames to promote response shift when making quality of life evaluations. The choice of reference frame often reflected external factors unrelated to individual health. Many participants also stated that they were unclear whether to report general or disease-related HRQoL. CONCLUSIONS: It is important, especially in clinical trials, to provide instructions clarifying whether quality of life refers to disease-related HRQoL. Information on self-comparison reference frames is necessary for the interpretation of responses to questions about HRQoL.

Rowen, D., Brazier, J., & Roberts, J. (2009). Mapping SF-36 onto the EQ-5D index: How reliable is the relationship? *Health and Quality of Life Outcomes, 7*, 27.

BACKGROUND: Mapping from health status measures onto generic preference-based measures is becoming a common solution when health state utility values are not directly available for economic evaluation. However, the accuracy and reliability of the models employed is largely untested, and there is little evidence of their suitability in patient datasets. This paper examines whether mapping approaches are reliable and accurate in terms of their predictions for a large and varied UK patient dataset. METHODS: SF-36 dimension scores are mapped onto the EQ-5D index using a number of different model

specifications. The predicted EQ-5D scores for subsets of the sample are compared across inpatient and outpatient settings and medical conditions. This paper compares the results to those obtained from existing mapping functions. RESULTS: The model including SF-36 dimensions, squared and interaction terms estimated using random effects GLS has the most accurate predictions of all models estimated here and existing mapping functions as indicated by MAE (0.127) and MSE (0.030). Mean absolute error in predictions by EQ-5D utility range increases with severity for our models (0.085 to 0.34) and for existing mapping functions (0.123 to 0.272). CONCLUSION: Our results suggest that models mapping the SF-36 onto the EQ-5D have similar predictions across inpatient and outpatient setting and medical conditions. However, the models overpredict for more severe EQ-5D states; this problem is also present in the existing mapping functions.

Russell, L. B., Gold, M. R., Siegel, J. E., Daniels, N., & Weinstein, M. C. (1996). The role of cost-effectiveness analysis in health and medicine. Panel on Cost-Effectiveness in Health and Medicine. *Journal of American Medical Association*, *276*(14), 1172–1177.

OBJECTIVE: To develop consensus-based recommendations guiding the conduct of cost-effectiveness analysis (CEA) to improve the comparability and quality of studies. The recommendations apply to analyses intended to inform the allocation of health care resources across a broad range of conditions and interventions. This article, first in a three-part series, discusses how this goal affects the conduct and use of analyses. The remaining articles will outline methodological and reporting recommendations, respectively. PARTICIPANTS: The Panel on Cost-Effectiveness in Health and Medicine, a nonfederal panel with expertise in CEA, clinical medicine, ethics, and health outcomes measurement, was convened by the U.S. Public Health Service (PHS). EVIDENCE: The panel reviewed the theoretical foundations of CEA, current practices, and alternative procedures for measuring and assigning values to resource use and health outcomes. CONSENSUS PROCESS: The panel met 11 times during 2 1/2 years with PHS staff and methodologists from federal agencies. Working groups brought issues and preliminary recommendations to the full panel for discussion. Draft recommendations were circulated to outside experts and the federal agencies prior to finalization. CONCLUSIONS: The panel's recommendations define a "reference case" cost-effectiveness analysis, a standard set of methods to serve as a point of comparison across studies. The reference case analysis is conducted from the societal perspective and accounts for benefits, harms, and costs to all parties. Although CEA does not reflect every element of importance in health care decisions, the information it provides is critical to informing decisions about the allocation of health care resources.

Salomon, J. A., & Murray, C. J. (2004). A multi-method approach to measuring health-state valuations. *Health Economics*, *13*(3), 281–290.

Existing techniques for eliciting health-state valuations incorporate both strength of preferences for health states and other values such as risk aversion or time preference. This paper presents a new methodological approach that allows estimation of a set of core

underlying health-state values based on responses elicited through multiple measurement techniques. A study was undertaken in which respondents completed the visual analogue (VAS) scale, time trade-off (TTO), standard gamble (SG) and person trade-off (PTO) for a range of states. By specifying flexible parametric functions to explain responses on each measurement technique, we estimated both the underlying strength of preference values for the health states in the study and the values for a set of auxiliary parameters characterizing risk attitudes, discount rates, distributional concerns and scale distortion effects in the group of respondents. This study demonstrates that it is possible to understand responses on these four different measurement techniques based on a consistent set of core values. The approach presented here can provide insights into different sources of observed variation in VAS, TTO, SG and PTO responses and facilitate appropriate adjustment of valuations elicited through different methods for use in summary health measures and economic analyses.

Seymour, J., McNamee, P., Scott, A., & Tinelli, M. (2009). Shedding new light onto the ceiling and floor? A quantile regression approach to compare EQ-5D and SF-6D responses. *Health Economics*.

An important issue in the measurement of health status concerns the extent to which an instrument displays lack of sensitivity to changes in health status at the extremes of the distribution, known as floor and ceiling effects. Previous studies use relatively simple methods that focus on the mean of the distribution to examine these effects. The aim of this paper is to determine whether quantile regression using longitudinal data improves our understanding of the relationship between quality of life instruments. The study uses EQ-5D and SF-36 (converted to SF-6D values) instruments with both baseline and follow-up data. Relative to ordinary least least-squares (OLS), a first difference model shows much lower association between the measures, suggesting that OLS methods may lead to biased estimates of the association, due to unobservable patient characteristics. The novel finding, revealed by quantile regression, is that the strength of association between the instruments is different across different parts of the health distribution, and is dependent on whether health improves or deteriorates. The results suggest that choosing one instrument at the expense of another is difficult without good prior information surrounding the expected magnitude and direction of health improvement related to a health-care intervention.

van Baal, P. H., Hoogenveen, R. T., de Wit, G. A., & Boshuizen, H. C. (2006). Estimating health-adjusted life expectancy conditional on risk factors: Results for smoking and obesity. *Population Health Metrics, 4*, 14.

BACKGROUND: Smoking and obesity are risk factors causing a large burden of disease. To help formulate and prioritize among smoking and obesity prevention activities, estimations of health-adjusted life expectancy (HALE) for cohorts that differ solely in their lifestyle (e.g. smoking vs. non smoking) can provide valuable information. Furthermore, in combination with estimates of life expectancy (LE), it can be tested whether prevention of obesity and

smoking results in compression of morbidity. METHODS: Using a dynamic population model that calculates the incidence of chronic disease conditional on epidemiological risk factors, we estimated LE and HALE at age 20 for a cohort of smokers with a normal weight (BMI < 25), a cohort of nonsmoking obese people (BMI>30) and a cohort of "healthy living" people (i.e., nonsmoking with a BMI < 25). Health state valuations for the different cohorts were calculated using the estimated disease prevalence rates in combination with data from the Dutch Burden of Disease study. Health state valuations are multiplied with life years to estimate HALE. Absolute compression of morbidity is defined as a reduction in unhealthy life expectancy (LE-HALE) and relative compression as a reduction in the proportion of life lived in good health (LE-HALE)/LE. RESULTS: Estimates of HALE are highest for a healthy living cohort (54.8 years for men and 55.4 years for women at age 20). Differences in HALE compared to healthy living men at age 20 are 7.8 and 4.6, respectively, for smoking and obese men. Differences in HALE compared to healthy living women at age 20 are 6.0 and 4.5, respectively, for smoking and obese women. Unhealthy life expectancy is about equal for all cohorts, meaning that successful prevention would not result in absolute compression of morbidity. Sensitivity analyses demonstrate that although estimates of LE and HALE are sensitive to changes in disease epidemiology, differences in LE and HALE between the different cohorts are fairly robust. In most cases, elimination of smoking or obesity does not result in absolute compression of morbidity but slightly increases the part of life lived in good health. CONCLUSION: Differences in HALE between smoking, obese and healthy living cohorts are substantial and similar to differences in LE. However, our results do not indicate that substantial compression of morbidity is to be expected as a result of successful smoking or obesity prevention.

Varni, J. W., Sherman, S. A., Burwinkle, T. M., Dickinson, P. E., & Dixon, P. (2004). The PedsQL Family Impact Module: Preliminary reliability and validity. *Health and Quality of Life Outcomes, 2*, 55.

BACKGROUND: The PedsQL Measurement Model was designed to measure health-related quality of life (HRQoL) in children and adolescents. The PedsQL 4.0 Generic Core Scales were developed to be integrated with the PedsQL Disease-Specific Modules. The newly developed PedsQL Family Impact Module was designed to measure the impact of pediatric chronic health conditions on parents and the family. The PedsQL Family Impact Module measures parent self-reported physical, emotional, social, and cognitive functioning, communication, and worry. The Module also measures parent-reported family daily activities and family relationships. METHODS: The 36-item PedsQL Family Impact Module was administered to 23 families of medically fragile children with complex chronic health conditions who either resided in a long-term care convalescent hospital or resided at home with their families. RESULTS: Internal consistency reliability was demonstrated for the PedsQL Family Impact Module Total Scale Score (alpha = 0.97), Parent HRQoL Summary Score (alpha = 0.96), Family Functioning Summary Score (alpha = 0.90), and Module

Scales (average alpha = 0.90, range = 0.82 – 0.97). The PedsQL Family Impact Module distinguished between families with children in a long-term care facility and families whose children resided at home. CONCLUSIONS: The results demonstrate the preliminary reliability and validity of the PedsQL Family Impact Module in families with children with complex chronic health conditions. The PedsQL Family Impact Module will be further field tested to determine the measurement properties of this new instrument with other pediatric chronic health conditions.

Whynes, D. K. (2008). Correspondence between EQ-5D health state classifications and EQ VAS scores. *Health and Quality of Life Outcomes, 6*, 94.

The EQ-5D health-related quality of life instrument comprises a health state classification followed by a health evaluation using a visual analogue scale (VAS). The EQ-5D has been employed frequently in economic evaluations, yet the relationship between the two parts of the instrument remains ill-understood. In this paper, we examine the correspondence between VAS scores and health state classifications for a large sample, and identify variables which contribute to determining the VAS scores independently of the health states as classified. METHODS: A UK trial of management of low-grade abnormalities detected on screening for cervical pre-cancer (TOMBOLA) provided EQ-5D data for over 3,000 women. Information on distress and multi-dimensional health locus of control had been collected using other instruments. A linear regression model was fitted, with VAS score as the dependent variable. Independent variables comprised EQ-5D health state classifications, distress, locus of control, and sociodemographic characteristics. Equivalent EQ-5D and distress data, collected at 12 months, were available for over 2,000 of the women, enabling us to predict changes in VAS score over time from changes in EQ-5D classification and distress. RESULTS: In addition to EQ-5D health state classification, VAS score was influenced by the subject's perceived locus of control, and by her age, educational attainment, ethnic origin, and smoking behavior. Although the EQ-5D classification includes a distress dimension, the independent measure of distress was an additional determinant of VAS score. Changes in VAS score over time were explained by changes in both EQ-5D severities and distress. Women allocated to the experimental management arm of the trial reported an increase in VAS score, independently of any changes in health state and distress. CONCLUSION: In this sample, EQ VAS scores were predictable from the EQ-5D health state classification, although there also existed other group variables that contributed systematically and independently toward determining such scores. These variables comprised psychological disposition, sociodemographic factors such as age and education, clinically important distress, and the clinical intervention itself.

Zhang, X. H., Xie, F., Wee, H. L., Thumboo, J., & Li, S. C. (2008). Applying the expectancy-value model to understand health values. *Value Health, 11*(Suppl 1), S61–S68.

OBJECTIVES: Expectancy-Value Model (EVM) is the most structured model in psychology to predict attitudes by measuring attitudinal attributes (AAs) and relevant external variables. Because health value could be categorized as attitude, we aimed to apply EVM to explore its usefulness in explaining variances in health values and investigate underlying factors. METHODS: Focus group discussion was carried out to identify the most common and significant AAs toward 5 different health states (coded as 11111, 11121, 21221, 32323, and 33333 in EuroQol Five-Dimension (EQ-5D) descriptive system). AAs were measured in a sum of multiplications of subjective probability (expectancy) and perceived value of attributes with 7-point Likert scales. Health values were measured using visual analog scales (VAS, range 0–1). External variables (age, sex, ethnicity, education, housing, marital status, and concurrent chronic diseases) were also incorporated into survey questionnaire distributed by convenience sampling among eligible respondents. Univariate analyses were used to identify external variables causing significant differences in VAS. Multiple linear regression model (MLR) and hierarchical regression model were used to investigate the explanatory power of AAs and possible significant external variable(s) separately or in combination, for each individual health state and a mixed scenario of five states, respectively. RESULTS: Four AAs were identified, namely, "worsening your quality of life in terms of health" (WQoL), "adding a burden to your family" (BTF), "making you less independent" (MLI) and "unable to work or study" (UWS). Data were analyzed based on 232 respondents (mean [SD] age: 27.7 [15.07] years, 49.1% female). Health values varied significantly across five health states, ranging from 0.12 (33333) to 0.97 (11111). With no significant external variables identified, EVM explained up to 62% of the variances in health values across five health states. The explanatory power of four AAs were found to be between 13% and 28% in separate MLR models ($P < 0.05$). When data were analyzed for each health state, variances in health values became small and explanatory power of EVM was reduced to a range between 8% and 23%. CONCLUSION: EVM was useful in explaining variances of health values and predicting important factors. Its power to explain small variances might be restricted due to limitations of 7-point Likert scale to measure AAs accurately. With further improvement and validation of a compatible continuous scale for more accurate measurement, EVM is expected to explain health values to a larger extent.

A.4 Burden of Illness/General

Barlow, W. E. (2009). Overview of methods to estimate the medical costs of cancer. *Medical Care*, *47*(7 Suppl 1), S33–36.

BACKGROUND: Methods to estimate the direct medical costs of cancer care have evolved into several commonly used methods. OBJECTIVES: We describe the different estimation techniques briefly to contrast these approaches and provide a framework for other articles

in this monograph. MEASURES AND RESULTS: One can estimate costs for all individuals with a specific cancer in a fixed calendar period (prevalent costs) or describe costs starting at the point of diagnosis and estimate immediate and long-term costs (incident costs). A variant of the incidence approach is to divide cancer care into initial, continuing, and terminal care phases and apply these phase-specific cost estimates to survival probabilities. The additional burden because of the cancer may be computed using cancer services (attributable costs) or by subtracting costs of healthy matched individuals (net costs). CONCLUSIONS: The strengths and weaknesses of these approaches are illustrated to show that the most appropriate choice will depend on whether the goal is to plan for health care costs, set public policy, or assess impact of potential interventions.

Brown, M. L., Lipscomb, J., & Snyder, C. (2001). The burden of illness of cancer: Economic cost and quality of life. *Annual Review of Public Health, 22*, 91–113.

Cancer is a major public health issue and represents a significant burden of disease. In this chapter, we analyze the main measures of burden of disease as relate to cancer. Specifically, we review incidence and mortality, years of life lost from cancer, and cancer prevalence. We also discuss the economic burden of cancer, including cost of illness, phase-specific and long-term costs, and indirect costs. We then examine the impact of cancer on health-related quality of life as measured in global terms (disability-adjusted life years and quality-adjusted life years) and using evaluation-oriented applications of health-related quality of life scales. Throughout, we note the relative strengths and weaknesses of the various approaches to measuring the burden of cancer as well as the methodologic challenges that persist in burden-of-illness research. We conclude with a discussion of the research agenda to improve our understanding of the burden of cancer and of illness more generally.

Cella, D., Yount, S., Rothrock, N., Gershon, R., Cook, K., Reeve, B., et al. (2007). The patient-reported outcomes measurement information system (PROMIS): Progress of an NIH roadmap cooperative group during its first two years. *Medical Care, 45*(5 Suppl 1), S3–S11.

BACKGROUND: The National Institutes of Health (NIH) Patient-Reported Outcomes Measurement Information System (PROMIS) Roadmap initiative (www.nihpromis.org) is a 5-year cooperative group program of research designed to develop, validate, and standardize item banks to measure patient-reported outcomes (PROs) relevant across common medical conditions. In this article, we will summarize the organization and scientific activity of the PROMIS network during its first 2 years. DESIGN: The network consists of six primary research sites (PRSs), a statistical coordinating center (SCC), and NIH research scientists. Governed by a steering committee, the network is organized into functional subcommittees and working groups. In the first year, we created an item library and activated three interacting protocols: Domain Mapping, Archival Data Analysis, and Qualitative Item Review (QIR). In the second year, we developed and initiated testing of

item banks covering five broad domains of self-reported health. RESULTS: The domain mapping process is built on the World Health Organization (WHO) framework of physical, mental, and social health. From this framework, pain, fatigue, emotional distress, physical functioning, social role participation, and global health perceptions were selected for the first wave of testing. Item response theory (IRT)-based analysis of 11 large data sets supplemented and informed item-level qualitative review of nearly 7,000 items from available PRO measures in the item library. Items were selected for rewriting or creation with further detailed review before the first round of testing in the general population and target patient populations. CONCLUSIONS: The NIH PROMIS network derived a consensus-based framework for self-reported health, systematically reviewed available instruments and datasets that address the initial PROMIS domains. Qualitative item research led to the first wave of network testing which began in the second year.

Chang, M. N., Guess, H., Heyse, J. F. (1994). Reduction in burden of illness: A new efficacy measure for prevention trials. *Statistics in Medicine*, *13*(18), 1807–1814.

A new efficacy measure is developed for use in prevention trials of interventions which may affect both disease incidence and disease severity. We assign a severity score to each incident case and sum severity scores over all incident cases within each treatment group to create a burden-of-illness score for each treatment group. Efficacy is evaluated by the difference between the burden-of-illness per randomized subject in the two randomized treatment groups. Since the numbers of summands in each burden-of-illness score is a random variable, standard methods of analysis are not directly applicable. The asymptotic distribution and sampling properties of the net reduction in the burden-of-illness score are derived for trials designed to stop either after a fixed length of follow-up or after the occurrence of a fixed number of cases. We illustrate the method with data from a clinical trial of a human rotavirus vaccine.

Cleeland, C. S. (2007). Symptom burden: Multiple symptoms and their impact as patient-reported outcomes. *Journal of the National Cancer Institute Monographs*, *37*, 16–21.

Cancer and its treatment produce multiple symptoms that significantly distress patients and impair function. Symptoms caused by treatment may delay treatment or lead to premature treatment termination, and residual treatment-related symptoms often complicate posttreatment rehabilitation. When treatment is no longer possible, symptom control becomes the focus of cancer care. Patient ratings of symptom severity and impact are important patient-reported outcomes (PROs) in cancer clinical trials and comprise a subset of a larger domain of PROs generally referred to as health-related quality of life (HRQoL). Symptoms rarely occur in isolation; rather, there is now ample evidence that symptoms frequently occur in clusters. The impact of these multiple symptoms upon the patient can be described as "symptom burden," a concept that encompasses both the severity of the symptoms and the patient's perception of the impact of the symptoms. The distress caused

by symptoms is a subject of much investigation, and several validated measures of the severity and impact of multiple symptoms are now available. Symptom measures are generally brief, thereby reducing respondent burden, and can be administered repeatedly during a trial to give a relatively fine-grained picture of the patient's status across time. In many instances, information on trial-related changes in symptom burden, or comparison of symptom burden between arms in a clinical trial, may provide sufficient self-report data for clinical trial consumers (patients, clinicians, and regulators) to make treatment choices or to evaluate new therapies, without measuring other HRQoL domains.

Corso, P., Finkelstein, E., Miller, T., Fiebelkorn, I., & Zaloshnja, E. (2006). Incidence and lifetime costs of injuries in the United States. *Injury Prevention, 12*(4), 212–218.

BACKGROUND: Standardized methodologies for assessing economic burden of injury at the national or international level do not exist. OBJECTIVE: To measure national incidence, medical costs, and productivity losses of medically treated injuries using the most recent data available in the United States, as a case study for similarly developed countries undertaking economic burden analyses. METHOD: The authors combined several data sets to estimate the incidence of fatal and non-fatal injuries in 2000. They computed unit medical and productivity costs and multiplied these costs by corresponding incidence estimates to yield total lifetime costs of injuries occurring in 2000. MAIN OUTCOME MEASURES: Incidence, medical costs, productivity losses, and total costs for injuries stratified by age group, sex, and mechanism. RESULTS: More than 50 million Americans experienced a medically treated injury in 2000, resulting in lifetime costs of 406 billion dollars; 80 billion dollars for medical treatment and 326 billion dollars for lost productivity. Males had a 20% higher rate of injury than females. Injuries resulting from falls or being struck by/against an object accounted for more than 44% of injuries. The rate of medically treated injuries declined by 15% from 1985 to 2000 in the United States. For those aged 0–44, the incidence rate of injuries declined by more than 20%; while persons aged 75 and older experienced a 20% increase. CONCLUSIONS: These national burden estimates provide unequivocal evidence of the large health and financial burden of injuries. This study can serve as a template for other countries or be used in intercountry comparisons.

Engelgau, M. M., Geiss, L. S., Saaddine, J. B., Boyle, J. P., Benjamin, S. M., Gregg, E. W., et al. (2004). The evolving diabetes burden in the United States. *Annals of Internal Medicine*, *140*(11), 945–950.

A diabetes epidemic emerged during the 20th century and continues unchecked into the 21st century. It has already taken an extraordinary toll on the U.S. population through its acute and chronic complications, disability, and premature death. Trend data suggest that the burden will continue to increase. Efforts to prevent or delay the complications of diabetes or, better yet, to prevent or delay the development of diabetes itself are urgently needed.

England, R. (2007). Are we spending too much on HIV? *British Medical Journal*, *334*(7589), 344.

Billions of pounds are being spent on the fight against AIDS in developing countries. Roger England believes that much of the money could be better used elsewhere, whereas Paul de Lay and colleagues argue that current spending is not enough.

Flegal, K. M., Graubard, B. I., Williamson, D. F., & Gail, M. H. (2005). Excess deaths associated with underweight, overweight, and obesity. *Journal of the American Medical Association, 293*(15), 1861–1867.

CONTEXT: As the prevalence of obesity increases in the United States, concern over the association of body weight with excess mortality has also increased. OBJECTIVE: To estimate deaths associated with underweight (body mass index [BMI] <18.5), overweight (BMI 25 to <30), and obesity (BMI ≥30) in the United States in 2000. DESIGN, SETTING, AND PARTICIPANTS: We estimated relative risks of mortality associated with different levels of BMI (calculated as weight in kilograms divided by the square of height in meters) from the nationally representative National Health and Nutrition Examination Survey (NHANES) I (1971–1975) and NHANES II (1976–1980), with follow-up through 1992, and from NHANES III (1988–1994), with follow-up through 2000. These relative risks were applied to the distribution of BMI and other covariates from NHANES 1999–2002 to estimate attributable fractions and number of excess deaths, adjusted for confounding factors and for effect modification by age. MAIN OUTCOME MEASURES: Number of excess deaths in 2000 associated with given BMI levels. RESULTS: Relative to the normal weight category (BMI 18.5 to <25), obesity (BMI ≥30) was associated with 111,909 excess deaths (95% confidence interval [CI], 53,754–170,064) and underweight with 33,746 excess deaths (95% CI, 15,726–51,766). Overweight was not associated with excess mortality (−86,094 deaths; 95% CI, −161,223 to −10,966). The relative risks of mortality associated with obesity were lower in NHANES II and NHANES III than in NHANES I. CONCLUSIONS: Underweight and obesity, particularly higher levels of obesity, were associated with increased mortality relative to the normal weight category. The impact of obesity on mortality may have decreased over time, perhaps because of improvements in public health and medical care. These findings are consistent with the increases in life expectancy in the United States and the declining mortality rates from ischemic heart disease.

Franks, P., Hanmer, J., & Fryback, D. G. (2006). Relative disutilities of 47 risk factors and conditions assessed with seven preference-based health status measures in a national U.S. sample: Toward consistency in cost-effectiveness analyses. *Medical Care, 44*(5), 478–485.

BACKGROUND: Preference-based health measures yield summary scores that are compatible with cost-effectiveness analyses. There is limited comparative information, however, about how different measures weight health conditions in the U.S. population. METHODS: We examined data from 11,421 adults in the 2000 Medical Expenditure Panel

Survey, a nationally representative sample of the U.S. general population, using information on sociodemographics (age, gender, race/ethnicity, income, and education), health status (EQ-5D, EQ-VAS, and SF-12), 4 risk factors (smoking, overweight, obesity, and lacking health insurance), and 43 conditions. From the EQ-5D, we derived summary scores using U.K. [EQ(UK)] and U.S. weights. From the SF-12 we derived SF-6D, and regression-predicted EQ-5D (U.S. and U.K. weights) and Health Utility Index scores. Each of the 7 preference measures was regressed on each of the 47 problems (risk factors and conditions) to determine the disutility associated with the problem, adjusting for sociodemographics. RESULTS: The adjusted disutilities averaged across the 47 problems for the 7 preference measures ranged from 0.059 for the SF-6D to 0.104 for the EQ(UK). Correlations between each of the measures of the adjusted disutilities ranged from 0.85–1.0. Standardization, using linear regression, attenuated between measure differences in disutilities. CONCLUSIONS: Absolute incremental cost-effectiveness analyses of a given problem would likely vary depending on the measure used, whereas the relative ordering of incremental cost-effectiveness analyses of a series of problems would likely be similar regardless of the measure chosen, as long as the same measure is used in each series of analyses. Absolute consistency across measures may be enhanced by standardization.

Hofstetter, P., & Hammitt, J. K. (2002). Selecting human health metrics for environmental decision-support tools. *Risk Analysis, 22*(5), 965–983.

Environmental decision-support tools often predict a multitude of different human health effects due to environmental stressors. The accounting and aggregating of these morbidity and mortality outcomes is critical to support decision making and can be accomplished by different methods that we call human health metrics. This article attempts to answer two questions: (1) Does it matter which metric is chosen?; and (2) What are the relevant characteristics of these metrics in environmental applications? Three metrics (quality adjusted life years [QALYs], disability adjusted life years [DALYs], and willingness to pay [WTP]) have been applied to the same diverse set of health effects due to environmental impacts. In this example, the choice of metric mattered for the ranking of these environmental impacts and it was found for this example that WTP was dominated by mortality outcomes. Further, QALYs and DALYs are sensitive to mild illnesses that affect large numbers of people and the severity of these mild illnesses are difficult to assess. Eight guiding questions are provided in order to help select human health metrics for environmental decision-support tools. Since health metrics tend to follow the paradigm of utility maximization, these metrics may be supplemented with a semi-quantitative discussion of distributional and ethical aspects. Finally, the magnitude of age-dependent disutility due to mortality for both monetary and nonmonetary metrics may bear the largest practical relevance for future research.

Hollinghurst, S., Bevan, G., & Bowie, C. (2000). Estimating the "avoidable" burden of disease by Disability Adjusted Life Years (DALYs). *Health Care Management Science, 3*(1), 9–21.

The World Bank's Global Burden of Disease Study pioneered the use of Disability Adjusted Life Years (DALYs). In this paper, we distinguish between the total and the "avoidable" burden of disease. We identify different ways of measuring DALYs: incidence-based DALYs are appropriate where the means of reducing the burden of disease is by prevention; prevalence-based DALYs are appropriate when a disease cannot be prevented but effective treatment is available. The methods of estimating each are explained and we describe how we have applied these methods to seven causes of death and disability in the South and West Region. We discuss the relevance of this work for monitoring the health of populations and deciding how best to use scarce resources to improve health.

Hyder, A. A., Rotllant, G., Morrow, R. H. (1998). Measuring the burden of disease: Healthy life-years. *American Journal of Public Health*, *88*(2), 196−202.

OBJECTIVES: This paper presents the background and rationale for a composite indicator, healthy life-year (HeaLY), that incorporates mortality and morbidity into a single number. HeaLY is compared with the disability-adjusted life year (DALY) indicator to demonstrate the relative simplicity and ease of use of the former. METHODS: Data collected by the Ghana Health Assessment team from census records, death certificates, medical records, and special studies were used to create a spreadsheet. HeaLYs lost as a result of premature mortality and disability from 56 conditions were estimated. RESULTS: Two-thirds of HeaLYs lost in Ghana were from maternal and communicable diseases and were largely preventable. The age weighting in DALYs leads to a higher value placed on deaths at younger ages than in HeaLYs. This spreadsheet can be used as a template for assessing changes in health status attributable to interventions. CONCLUSIONS: HeaLY can aid in setting health priorities and identifying disadvantaged groups. The disaggregated approach of the HeaLY spreadsheet tool is simpler for decision makers and useful for country application.

Jemal, A., Ward, E., Hao, Y., & Thun, M. (2005). Trends in the leading causes of death in the United States, 1970–2002. *Journal of American Medical Association*, *294*(10), 1255–1259.

CONTEXT: The decrease in overall death rates in the United States may mask changes in death rates from specific conditions. OBJECTIVE: To examine temporal trends in the age-standardized death rates and in the number of deaths from the six leading causes of death in the United States. DESIGN AND SETTING: Analyses of vital statistics data on mortality in the United States from 1970 to 2002. MAIN OUTCOME MEASURE: The age-standardized death rate and number of deaths (coded as underlying cause) from each of the six leading causes of death: heart disease, stroke, cancer, chronic obstructive pulmonary disease, accidents (i.e., related to transportation [motor vehicle, other land vehicles, and water, air, and space] and not related to transportation [falls, fire, and accidental poisoning]), and

diabetes mellitus. RESULTS: The age-standardized death rate (per 100,000 per year) from all causes combined decreased from 1242 in 1970 to 845 in 2002. The largest percentage decreases were in death rates from stroke (63%), heart disease (52%), and accidents (41%). The largest absolute decreases in death rates were from heart disease (262 deaths per 100,000), stroke (96 deaths per 100,000), and accidents (26 deaths per 100,000).The death rate from all types of cancer combined increased between 1970 and 1990 and then decreased through 2002, yielding a net decline of 2.7%. In contrast, death rates doubled from chronic obstructive pulmonary disease over the entire time interval and increased by 45% for diabetes since 1987. Despite decreases in age-standardized death rates from four of the six leading causes of death, the absolute number of deaths from these conditions continues to increase, although these deaths occur at older ages. CONCLUSIONS: The absolute number of deaths and age at death continue to increase in the United States. These temporal trends have major implications for health care and health care costs in an aging population.

Klis, S., Vingerhoets, A. J., de Wit, M. Zandebelt, N., & Snoek F. J. (2008). Pictorial Representation of Illness and Self Measure Revised II (PRISM-RII): A novel method to assess perceived burden of illness in diabetes patients. *Health and Quality of Life Outcomes*, *6*, 104.

BACKGROUND: The Pictorial Representation of Illness and Self Measure (PRISM) has been introduced as a visual measure of suffering. We explored the validity of a revised version, the PRISM-RII, in diabetes patients as part of the annual review. METHODS: Participants were 308 adult outpatients with either type 1 or type 2 diabetes. Measures: (1) the PRISM-RII, yielding Self-Illness Separation (SIS) and Illness Perception Measure (IPM); (2) the Problem Areas in Diabetes (PAID) scale, a measure of diabetes-related distress; (3) the WHO-5 Well-Being Index; (4) and a validation question on suffering (SQ). In addition, patients' complication status, comorbidity, and glycemic control values (HbA1c) were recorded. RESULTS: Patients with complications had marginally significant higher scores on IPM, compared to patients without complications. Type 2 patients had higher IPM scores than Type 1 patients. SIS and IPM showed low intercorrelation ($r = -.25$; $p < .01$). Convergent validity of PRISM-RII was demonstrated by significant correlations between IPM and PAID ($r = 0.50$; $p < 0.01$), WHO-5 ($r = -.26$; $p < 0.01$), and SQ ($r = 0.36$; $p < 0.01$). SIS showed only significant correlations with PAID ($r = -0.28$; $p < 0.01$) and SQ ($r = -0.22$; $p < 0.01$). Neither IPM nor SIS was significantly associated with HbA1c. The PRISM-RII appeared easy to use and facilitated discussion with care providers on coping with the burden of diabetes. CONCLUSION: PRISM-RII appears a promising additional tool to assess the psychological burden of diabetes.

Liu, J. L., Maniadakis, N., Gray, A., & Rayner, M. (2002). The economic burden of coronary heart disease in the UK. *Heart, 88*(6), 597–603.

OBJECTIVE: To estimate the economic burden of coronary heart disease in the UK using both direct and indirect costs. DESIGN AND SETTING: A prevalence based approach was used to assess coronary heart disease related costs from the societal perspective. PATIENTS: All UK residents in 1999 with coronary heart disease (ICD 9 codes 410–414 and ICD10 codes I20–I25). MAIN OUTCOME MEASURES: Direct health care costs were estimated from spending on prevention, accident and emergency care, hospital care, rehabilitation, and drug treatment. Direct non-health service costs were estimated from data on informal care. "Friction period" adjusted productivity costs were estimated using the human capital approach from lost earnings attributable to coronary heart disease related mortality and morbidity. The friction period is the period of employees' absence from work before the employer replaces them with other workers. Failure to adjust for this factor would overstate production loss. RESULTS: Coronary heart disease cost pound 1.73 billion to the UK health care system in 1999: pound 2.42 billion in informal care and pound 2.91 billion in friction period adjusted productivity loss; 24.1% of production losses were attributable to mortality and 75.9% to morbidity. The total annual cost of all coronary heart disease related burdens was pound 7.06 billion, the highest of all diseases in the UK for which comparable analyses have been done. CONCLUSIONS: Coronary heart disease is a leading public health problem in the UK in terms of the economic burden from disease. Cost estimates would be substantially understated if informal care/productivity costs were excluded.

Lopez, A. D. (2005). The evolution of the global burden of disease framework for disease, injury and risk factor quantification: Developing the evidence base for national, regional and global public health action. *Global Health, 1*(1), 5.

Reliable, comparable information about the main causes of disease and injury in populations, and how these are changing, is a critical input for debates about priorities in the health sector. Traditional sources of information about the descriptive epidemiology of diseases, injuries and risk factors are generally incomplete, fragmented and of uncertain reliability and comparability. Lack of a standardized measurement framework to permit comparisons across diseases and injuries, as well as risk factors, and failure to systematically evaluate data quality have impeded comparative analyses of the true public health importance of various conditions and risk factors. As a consequence the impact of major conditions and hazards on population health has been poorly appreciated, often leading to a lack of public health investment. Global disease and risk factor quantification improved dramatically in the early 1990s with the completion of the first Global Burden of Disease Study. For the first time, the comparative importance of over 100 diseases and injuries, and ten major risk factors, for global and regional health status could be assessed using a common metric (disability-adjusted life years) which simultaneously accounted for both premature mortality and the prevalence, duration and severity of the non-fatal

consequences of disease and injury. As a consequence, mental health conditions and injuries, for which non-fatal outcomes are of particular significance, were identified as being among the leading causes of disease/injury burden worldwide, with clear implications for policy, particularly prevention. A major achievement of the study was the complete global descriptive epidemiology, including incidence, prevalence and mortality, by age, sex and Region, of over 100 diseases and injuries. National applications, further methodological research and an increase in data availability have led to improved national, regional and global estimates for 2000, but substantial uncertainty around the disease burden caused by major conditions, including HIV, remains. The rapid implementation of cost-effective data collection systems in developing countries is a key priority if global public policy to promote health is to be more effectively informed.

Loza, E., Abasolo, L, Jover, J. A., & Carmona L. (2008). Burden of disease across chronic diseases: A health survey that measured prevalence, function, and quality of life. *Journal of Rheumatology*, *35*(1), 159–165.

OBJECTIVE: To assess health related quality of life (HRQoL) and functional ability across groups of chronic diseases in Spain. METHODS: A national health survey was conducted during 1999–2000. Participants were randomly selected from city censuses among persons aged over 20 years. All 2192 participants (response rate 73%) completed generic instruments measuring functional ability in activities of daily living [Health Assessment Questionnaire (HAQ)] and HRQoL [Short-Form 12 (SF-12)]. Chronic diseases were defined by self-report and elicited from two specific questions: "Have you ever been told you have a chronic disease by a physician?" and "Are you taking any chronic medication?" Only diagnoses present for ≥ 3 months were included as chronic. We estimated mean HAQ and SF-12 scores for the different groups of chronic diseases. We then adjusted the scores for covariates and compared them between diseases by multiple linear regressions. RESULTS: Over half the population had at least one chronic disease [n = 1276 (58.2%)], and 22.6% had any rheumatic disease. Rheumatic diseases have an adverse effect on daily functioning [HAQ beta-coefficient 0.11 (95% CI 0.06–0.15)] and HRQoL [SF-12 physical beta-coefficient −5.78 (95% CI −6.27 to −4.28); SF-12 mental beta-coefficient −2.61 (95% CI −3.79 to −1.41)]. Thus, the influence of the rheumatic diseases is greater when their prevalence is taken into account. CONCLUSION: When the definition of burden of disease includes a measure of function and HRQoL that is weighted by disease prevalence, rheumatic diseases as a group can be ranked alongside neurological, cardiac, or pulmonary conditions as a major disease.

Lyons, R. A., Towner, E. E., Kendrick, D. Christie, N., Brophy, S., Phillips, C. J., et al. (2007). The UK burden of injury study—A protocol. *BMC Public Health, 7*, 317.

BACKGROUND: Globally and nationally large numbers of people are injured each year, yet there is little information on the impact of these injuries on people's lives, on society and on health and social care services. Measurement of the burden of injuries is needed at a global,

national and regional level to be able to inform injured people of the likely duration of impairment; to guide policy makers in investing in preventative measures; to facilitate the evaluation and cost effectiveness of interventions and to contribute to international efforts to more accurately assess the global burden of injuries. METHODS/DESIGN: A prospective, longitudinal multi-centre study of 1,333 injured individuals, attending. Emergency Departments or admitted to hospital in four UK areas: Swansea, Surrey, Bristol and Nottingham. Specified quotas of patients with defined injuries covering the whole spectrum will be recruited. Participants (or a proxy) will complete a baseline questionnaire regarding their injury and pre-injury quality of life. Follow up occurs at 1, 4, and 12 months post injury or until return to normal function within 12 months, with measures of health service utilization, impairment, disability, and health related quality of life. National estimates of the burden of injuries will be calculated by extrapolation from the sample population to national and regional computerized hospital in-patient, emergency department and mortality data. DISCUSSION: This study will provide more detailed data on the national burden of injuries than has previously been available in any country and will contribute to international collaborative efforts to more accurately assess the global burden of injuries. The results will be used to advise policy makers on prioritization of preventive measures, support the evaluation of interventions, and provide guidance on the likely impact and degree of impairment and disability following specific injuries.

Manuel, D. G., S. E. Schultz, et al. (2002). Measuring the health burden of chronic disease and injury using health adjusted life expectancy and the Health Utilities Index. *Journal of Epidemiology and Community Health, 56*(11), 843–850.

OBJECTIVES: To estimate the burden of illness from chronic disease and injury using a population based health survey, which contains both measures of chronic disease and a utility based health related quality of life (HRQoL) measure. DESIGN: An adapted Sullivan method was used to calculate cause deleted health adjusted life expectancies for chronic conditions. SETTING: Ontario, Canada, 1996/97. SUBJECTS: The 1996/97 Ontario Health Survey (n=35 527) was used to estimate the prevalence of chronic conditions. A cause deleted approach was used to estimate the impact of these conditions on the Health Utilities Index (HUI). Cause deleted probabilities of dying were derived with the cause eliminated life table technique and death data from vital statistics for Ontario 1996/97 (n=156 610). RESULTS: Eliminating cardiovascular disease and cancer will cause an "expansion of morbidity," while eliminating mental conditions and musculosketal disorders will result in a "contraction of morbidity." The HUI score varies depending on chronic condition, age, and sex-most of which were assumed not to vary in previous summary measures of population health. CONCLUSIONS: Health adjusted life expectancy estimated for chronic conditions using a utility based measure of health related quality of life from population health surveys addresses several limitations of previous studies that estimate the burden of disease using

either a categorical measure of disability or expert opinion and related epidemiological evidence.

Mathers, C. D., & Loncar, D. (2006). Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med, 3*(11), e442.

BACKGROUND: Global and regional projections of mortality and burden of disease by cause for the years 2000, 2010, and 2030 were published by Murray and Lopez in 1996 as part of the Global Burden of Disease project. These projections, which are based on 1990 data, continue to be widely quoted, although they are substantially outdated; in particular, they substantially underestimated the spread of HIV/AIDS. To address the widespread demand for information on likely future trends in global health, and thereby to support international health policy and priority setting, we have prepared new projections of mortality and burden of disease to 2030 starting from World Health Organization estimates of mortality and burden of disease for 2002. This paper describes the methods, assumptions, input data, and results. METHODS AND FINDINGS: Relatively simple models were used to project future health trends under three scenarios—baseline, optimistic, and pessimistic—based largely on projections of economic and social development, and using the historically observed relationships of these with cause-specific mortality rates. Data inputs have been updated to take account of the greater availability of death registration data and the latest available projections for HIV/AIDS, income, human capital, tobacco smoking, body mass index, and other inputs. In all three scenarios there is a dramatic shift in the distribution of deaths from younger to older ages and from communicable, maternal, perinatal, and nutritional causes to noncommunicable disease causes. The risk of death for children younger than 5 y is projected to fall by nearly 50% in the baseline scenario between 2002 and 2030. The proportion of deaths due to noncommunicable disease is projected to rise from 59% in 2002 to 69% in 2030. Global HIV/AIDS deaths are projected to rise from 2.8 million in 2002 to 6.5 million in 2030 under the baseline scenario, which assumes coverage with antiretroviral drugs reaches 80% by 2012. Under the optimistic scenario, which also assumes increased prevention activity, HIV/AIDS deaths are projected to drop to 3.7 million in 2030. Total tobacco-attributable deaths are projected to rise from 5.4 million in 2005 to 6.4 million in 2015 and 8.3 million in 2030 under our baseline scenario. Tobacco is projected to kill 50% more people in 2015 than HIV/AIDS, and to be responsible for 10% of all deaths globally. The three leading causes of burden of disease in 2030 are projected to include HIV/AIDS, unipolar depressive disorders, and ischaemic heart disease in the baseline and pessimistic scenarios. Road traffic accidents are the fourth leading cause in the baseline scenario, and the third leading cause ahead of ischaemic heart disease in the optimistic scenario. Under the baseline scenario, HIV/AIDS becomes the leading cause of burden of disease in middle- and low-income countries by 2015. CONCLUSIONS: These projections represent a set of three visions of the future for population health, based on certain explicit assumptions. Despite the wide uncertainty ranges around future projections, they enable us to appreciate

better the implications for health and health policy of currently observed trends, and the likely impact of fairly certain future trends, such as the ageing of the population, the continued spread of HIV/AIDS in many regions, and the continuation of the epidemiological transition in developing countries. The results depend strongly on the assumption that future mortality trends in poor countries will have a relationship to economic and social development similar to those that have occurred in the higher-income countries.

Mathers, C. D., Vos, E. T., Stevenson, C. E., & Begg, S. J. (2001). The burden of disease and injury in Australia. *Bulletin of the World Health Organization*, *79*(11), 1076–1084.

An overview of the results of the Australian Burden of Disease (ABD) study is presented. The ABD study was the first to use methodology developed for the Global Burden of Disease study to measure the burden of disease and injury in a developed country. In 1996, mental disorders were the main causes of disability burden, responsible for nearly 30% of total years of life lost to disability (YLD), with depression accounting for 8% of the total YLD. Ischaemic heart disease and stroke were the main contributors to the disease burden disability-adjusted life years (DALYs), together causing nearly 18% of the total disease burden. Risk factors such as smoking, alcohol consumption, physical inactivity, hypertension, high blood cholesterol, obesity and inadequate fruit and vegetable consumption were responsible for much of the overall disease burden in Australia. The lessons learnt from the ABD study are discussed, together with methodological issues that require further attention.

McDaid, D. (2001). Estimating the costs of informal care for people with Alzheimer's disease: Methodological and practical challenges. *International Journal of Geriatric Psychiatry, 16*, 400–405.

Although Alzheimer's disease and related disorders may have a heavy impact on informal caregivers, estimates of informal care costs have been neglected and when included in cost of illness studies, valuations have been highly variable. Although these variations are in part due to differences in samples and the difficulty in measuring caregiving time, this illustrates the need to standardize the methodology not only for valuing formal, but also informal care costs. Methods used for valuing informal care are identified, together with theoretical and practical challenges in measurement. In particular the measurement of time and its associated satisfaction or utility is complex and valuations of time need to consider aspects of the caregiving experience which influence the marginal valuation of the time spent caring. More empirical work is required to elicit information on both the positive and negative satisfaction associated with caregiving and to incorporate this into valuations of the costs related to informal care.

McKenna, M., Michaud, C. M., Murray, C. J. L, & Marks, J. S. (2005). Assessing the burden of disease in the US using DALYs. *American Journal of Preventative Medicine*, *28*, 415–423.

PURPOSE: To review the history and challenges of "burden of disease" studies, how these are dependent on robust epidemiologic data as well as complex conceptual constructions, and to identify the public health policy issues these studies can most usefully inform. METHODS: The emergence of the concept of the "burden of disease" in the public health literature is reviewed, with a focus on the results of an analysis of data from the United States that used the methodology presented in the Global Burden of Disease Study. RESULTS: The systematic analysis of public health mortality data to identify major health problems was conducted by Graunt in 16th-century London. He found that many of the predominant sources of mortality were not the focus of public attention. Today, despite refinements in epidemiologic measurement methods designed to capture the impact of non-fatal health conditions, there are similar incongruities between the major public health problems and expenditures on prevention interventions. CONCLUSIONS: Controversies surrounding the interpretation of "burden of disease" studies are not new. Particularly in developed countries, these studies appear more useful for setting research priorities rather than allocating resources to support prevention efforts. Such investigations are not possible without ongoing support for systematic collection and analysis of descriptive epidemiologic data.

Michaud, C. M., Murray, C. J., & Bloom, B. R. (2001). Burden of disease—Implications for future research. *Journal of American Medical Association*, *285*(5), 535–539.

One overall challenge for public health and medicine in the future is to allocate available resources effectively to reduce major causes of disease burden globally and to decrease health disparities between poor and affluent populations. The major risk factors for death and disability worldwide are malnutrition; poor water supply, sanitation, and personal and domestic hygiene; unsafe sexual behavior; tobacco use; alcohol use; occupational hazards; hypertension; physical inactivity; illicit drugs; and air pollution. The challenge for research in the 21st century is to maintain and improve life expectancy and the quality of life that was achieved for most of the world's population during the 20th century.

Mokdad, A. H., Marks, J. S., Stroup, D. F., & Gerberding, J. L. (2004). Actual causes of death in the United States, 2000. *Journal of American Medical Association*, *291*(10), 1238–1245.

CONTEXT: Modifiable behavioral risk factors are leading causes of mortality in the United States. Quantifying these will provide insight into the effects of recent trends and the implications of missed prevention opportunities. OBJECTIVES: To identify and quantify the leading causes of mortality in the United States. DESIGN: Comprehensive MEDLINE search of English-language articles that identified epidemiological, clinical, and laboratory studies linking risk behaviors and mortality. We used 2000 mortality data reported to the Centers

for Disease Control and Prevention to identify the causes and number of deaths. The estimates of cause of death were computed by multiplying estimates of the cause-attributable fraction of preventable deaths with the total mortality data. MAIN OUTCOME MEASURES: Actual causes of death. RESULTS: The leading causes of death in 2000 were tobacco (435,000 deaths; 18.1% of total U.S. deaths), poor diet and physical inactivity (365,000 deaths; 15.2%) [corrected], and alcohol consumption (85,000 deaths; 3.5%). Other actual causes of death were microbial agents (75,000), toxic agents (55 000), motor vehicle crashes (43,000), incidents involving firearms (29,000), sexual behaviors (20,000), and illicit use of drugs (17,000). CONCLUSIONS: These analyses show that smoking remains the leading cause of mortality. However, poor diet and physical inactivity may soon overtake tobacco as the leading cause of death. These findings, along with escalating health care costs and aging population, argue persuasively that the need to establish a more preventive orientation in the U.S. health care and public health systems has become more urgent.

Murray, C. J., & Frenk, J. (2008). Health metrics and evaluation: Strengthening the science. *Lancet, 371*(9619), 1191–1199.

With the growing importance of health in the global agenda comes the responsibility to develop a scientific foundation of metrics and evaluation. The scope of this emerging field can be viewed in terms of key topics, including health outcomes, other social outcomes related to health systems, health services, resource inputs, evaluations of programs and systems, and analyses to support policy choice. It can also be defined in terms of key activities that are needed to strengthen the scientific basis of the field: development of new methods, instruments, software, and hardware; setting global norms and standards for data collection; increasing the availability of high-quality primary data; systematic analysis and synthesis of existing datasets; strengthening national capacity to obtain, analyze, and use data; and reporting and disseminating results. We explore in depth topics with major scientific challenges and institutional and cultural barriers that are slowing the development of the field. Cutting across the various topical areas and disciplinary approaches to these problems are some common scientific issues, including limited comparability of measurement, uncorrected known biases in data, no standard approach to missing data, unrealistic uncertainty estimates, and the use of disease models that have not been properly validated. Only through concerted action will it be possible to assure the production, reproduction, and use of knowledge that is crucial to the advancement of global health.

Murray, C. J., & Lopez, A. D. (1997). Alternative projections of mortality and disability by cause 1990–2020: Global Burden of Disease Study. *Lancet*, *349*(9064), 1498–1504.

BACKGROUND: Plausible projections of future mortality and disability are a useful aid in decisions on priorities for health research, capital investment, and training. Rates and

patterns of ill health are determined by factors such as socioeconomic development, educational attainment, technological developments, and their dispersion among populations, as well as exposure to hazards such as tobacco. As part of the Global Burden of Disease Study (GBD), we developed three scenarios of future mortality and disability for different age-sex groups, causes, and regions. METHODS: We used the most important disease and injury trends since 1950 in nine cause-of-death clusters. Regression equations for mortality rates for each cluster by region were developed from gross domestic product per person (in international dollars), average number of years of education, time (in years, as a surrogate for technological change), and smoking intensity, which shows the cumulative effects based on data for 47 countries in 1950–90. Optimistic, pessimistic, and baseline projections of the independent variables were made. We related mortality from detailed causes to mortality from a cause cluster to project more detailed causes. Based on projected numbers of deaths by cause, years of life lived with disability (YLDs) were projected from different relation models of YLDs to years of life lost (YLLs). Population projections were prepared from World Bank projections of fertility and the projected mortality rates. FINDINGS: Life expectancy at birth for women was projected to increase in all three scenarios; in established market economies to about 90 years by 2020. Far smaller gains in male life expectancy were projected than in females; in formerly socialist economies of Europe, male life expectancy may not increase at all. Worldwide mortality from communicable maternal, perinatal, and nutritional disorders was expected to decline in the baseline scenario from 17.2 million deaths in 1990 to 10.3 million in 2020. We projected that non-communicable disease mortality will increase from 28.1 million deaths in 1990 to 49.7 million in 2020. Deaths from injury may increase from 5.1 million to 8.4 million. Leading causes of disability-adjusted life years (DALYs) predicted by the baseline model were (in descending order): ischaemic heart disease, unipolar major depression, road-traffic accidents, cerebrovascular disease, chronic obstructive pulmonary disease, lower respiratory infections, tuberculosis, war injuries, diarrhoeal diseases, and HIV. Tobacco-attributable mortality is projected to increase from 3.0 million deaths in 1990 to 8.4 million deaths in 2020. INTERPRETATION: Health trends in the next 25 years will be determined mainly by the ageing of the world's population, the decline in age-specific mortality rates from communicable, maternal, perinatal, and nutritional disorders, the spread of HIV, and the increase in tobacco-related mortality and disability. Projections, by their nature, are highly uncertain, but we found some robust results with implications for health policy.

Murray, C. J. (1994). Quantifying the burden of disease: The technical basis for disability-adjusted life years. *Bulletin of the World Health Organization, 72*(3), 429–445.

Detailed assumptions used in constructing a new indicator of the burden of disease, the disability-adjusted life year (DALY), are presented. Four key social choices in any indicator of the burden of disease are carefully reviewed. First, the advantages and disadvantages of various methods of calculating the duration of life lost due to a death at each age are

discussed. DALYs use a standard expected-life lost based on model life-table West Level 26. Second, the value of time lived at different ages is captured in DALYs using an exponential function which reflects the dependence of the young and the elderly on adults. Third, the time lived with a disability is made comparable with the time lost due to premature mortality by defining six classes of disability severity. Assigned to each class is a severity weight between 0 and 1. Finally, a 3% discount rate is used in the calculation of DALYs. The formula for calculating DALYs based on these assumptions is provided.

Rodgers, A., Ezzati, M. Vander Hoorn, S., Lopez, A. D., Lin, R. B., & Murray, C. J. (2004). Distribution of major health risks: Findings from the Global Burden of Disease study. *Public Library of Science-Medicine, 1*(1), e27.

BACKGROUND: Most analyses of risks to health focus on the total burden of their aggregate effects. The distribution of risk-factor-attributable disease burden, for example by age or exposure level, can inform the selection and targeting of specific interventions and programs, and increase cost-effectiveness. METHODS AND FINDINGS: For 26 selected risk factors, expert working groups conducted comprehensive reviews of data on risk-factor exposure and hazard for 14 epidemiological subregions of the world, by age and sex. Age-sex-subregion-population attributable fractions were estimated and applied to the mortality and burden of disease estimates from the World Health Organization Global Burden of Disease database. Where possible, exposure levels were assessed as continuous measures, or as multiple categories. The proportion of risk-factor-attributable burden in different population subgroups, defined by age, sex, and exposure level, was estimated. For major cardiovascular risk factors (blood pressure, cholesterol, tobacco use, fruit and vegetable intake, body mass index, and physical inactivity) 43%–61% of attributable disease burden occurred between the ages of 15 and 59 y, and 87% of alcohol-attributable burden occurred in this age group. Most of the disease burden for continuous risks occurred in those with only moderately raised levels, not among those with levels above commonly used cut-points, such as those with hypertension or obesity. Of all disease burden attributable to being underweight during childhood, 55% occurred among children 1 to 3 standard deviations below the reference population median, and the remainder occurred among severely malnourished children, who were three or more standard deviations below median. CONCLUSIONS: Many major global risks are widely spread in a population, rather than restricted to a minority. Population-based strategies that seek to shift the whole distribution of risk factors often have the potential to produce substantial reductions in disease burden.

Stein, C., T. Kuchenmuller, Hendrickx, S. Pruss-Ustun, A., Wolfson, L., Engels, D., & Schlundt, J. (2007). The Global Burden of Disease Assessments—WHO Is Responsible? *Public Library of Science-Neglected Tropical Diseases, 1*(3), e161.

The Global Burden of Disease (GBD) concept has been used by the World Health Organization (WHO) for its reporting on health information for nearly 10 years. The GBD approach results in a single summary measure of morbidity, disability, and mortality, the

so-called disability-adjusted life year (DALY). To ensure transparency and objectivity in the derivation of health information, WHO has been urged to use reference groups of external experts to estimate burden of disease. Under the leadership and coordination of WHO, expert groups have been appraising and abstracting burden of disease information. Examples include the Child Health Epidemiology Reference Group (CHERG), the Malaria Monitoring and Evaluation Reference Group (MERG), and the recently established Foodborne Disease Burden Epidemiology Reference Group (FERG). The structure and functioning of and lessons learnt by these groups are described in this paper. External WHO expert groups have provided independent scientific health information while operating under considerable differences in structure and functioning. Although it is not appropriate to devise a single "best practice" model, the common thread described by all groups is the necessity of WHO's leadership and coordination to ensure the provision and dissemination of health information that is to be globally accepted and valued.

Szucs, T. D. (1999). Influenza. The role of burden-of-illness research. *Pharmacoeconomics*, *16*(Suppl 1), 27–32.

Influenza is associated with a significant economic burden on both society and the individual, resulting in considerable healthcare costs and loss of productivity, as well as intangible costs such as suffering, grief and social disruption. The incidence and severity of influenza infection depend, at least in part, on the age and health status of the population. For example, the incidence of influenza is relatively high among children and young adults, but serious complications are much more likely to occur among the very young (< 1 year of age) and the elderly (> 65 years of age). School absenteeism tends to peak in the first half of a typical 6- to 8-week influenza epidemic, followed by workplace absenteeism in the latter half as school-aged children pass the infection to family members. Cost-of-illness studies are used by policy makers to justify budgets and set priorities for prevention programs, research and other expenditures. On the basis of German Sickness Fund data, recent estimates indicate that the cost of an influenza epidemic in that country is approximately 2 billion Deutschmarks (approximately $US1.4 billion). The bulk of these costs reflects indirect costs associated with lost productivity, a finding also noted in an earlier French cost-of-illness study of influenza. Thus, the main economic burden of influenza falls on infected individuals, their employers and their relatives. Methodology used in cost-of-illness studies can be quite variable. For example, two main approaches are used in measuring indirect costs (human capital and willingness to pay), although there is controversy as to which is the preferred method. Thus, investigators involved in cost-of-illness studies must be explicit regarding study methodology in order to allow for appropriate interpretation of study results by interested parties.

Thacker, S. B., Stroup D. F, Carande-Kulis, V., Marks, J. S., Roy, K., & Gerberding, J. L. (2006). Measuring the public's health. *Public Health Reports, 121*(1), 14–22.

Allocation of public health resources should be based, where feasible, on objective assessments of health status, burden of disease, injury, and disability, their preventability, and related costs. In this article, we first analyze traditional measures of the public's health that address the burden of disease and disability and associated costs. Second, we discuss activities that are essential to protecting the public's health but whose impact is difficult to measure. Third, we propose general characteristics of useful measures of the public's health. We contend that expanding the repertoire of measures of the public's health is a critical step in targeting attention and resources to improve health, stemming mounting health care costs, and slowing declining quality of life that threatens the nation's future.

Thun, M., Jemal, A., Descantis, C., Blackard, D., & Ward, E. (2009). An overview of the cancer burden for primary care physicians. *Primary Care*, *36*(3), 439–454.

Primary care physicians and other caregivers are uniquely positioned to communicate with patients about their real risks of developing or dying from cancer and actions that can reduce these risks. This article discusses the statistics used to measure the cancer burden in a manner intended to help primary caregivers communicate more effectively with patients about cancer. The basic terms used to measure incidence, mortality, and relative survival, and considerations that influence the interpretation of cancer trends are described; opportunities to accelerate progress in reducing cancer incidence and death rates are identified. Although integrating effective prevention measures into standard clinical care will require changes in health care policy and in clinical practice, the combination of these approaches is essential to prevent the massive anticipated increase in the number of cancer cases, due to growth and aging of the population.

Ubel, P. A., Nord, E., Gold, M., Menzel, P., Prades, J. L., & Richardson, J. (2000). Improving value measurement in cost-effectiveness analysis. *Medical Care*, *38*(9), 892–901.

OBJECTIVE: Before cost-effectiveness analysis (CEA) can fulfill its promise as a tool to guide health care allocation decisions, the method of incorporating societal values into CEA may need to be improved. DESIGN: The study design was a declarative exposition of potential fallacies in the theoretical underpinnings of CEA. Two values held by many people-preferences for giving priority to severely ill patients and preferences to avoid discrimination against people who have limited treatment potential because of disability or chronic illness-that are not currently incorporated into CEA are discussed. CONCLUSIONS: Traditional CEA, through the measurement of quality-adjusted life years (QALYs), is constrained because of a "QALY trap." If, for example, saving the life of a person with paraplegia is equally valuable as saving the life of a person without paraplegia, then current QALY methods force us to conclude that curing paraplegia brings no benefit. Basing cost-effectiveness measurement on societal values rather than QALYs may allow us to better capture public rationing

preferences, thereby escaping the QALY trap. CEA can accommodate a wider range of such societal values about fairness in its measurements by amending its methodology.

Valderas, J. M., & Alonso (2008). Patient reported outcome measures: A model-based classification system for research and clinical practice. *Quality of Life Research*, *17*(9), 1125–1135.

PURPOSE: The umbrella term Patient Reported Outcomes (PRO) has been successfully proposed for instruments measuring perceived health outcomes, but its relationship to current conceptual models remains to be established. Our aim was to develop a classification system for PRO measures based on a valid conceptual model. METHODS: We reviewed models and classification schemes of health outcomes and integrated them in a common conceptual framework, based on the models by Wilson and Cleary and the International Classification of Functioning (ICF). We developed a cross-classification system based on the minimum common set of consistent concepts identified in previous classifications, and specified categories based on the WHO International Classifications (ICD-10, and ICF). We exemplified the use of the classification system with selected PRO instruments. RESULTS: We identified three guiding concepts: (1) construct (the measurement object), (2) population (based on age, gender, condition, and culture), and (3) measurement model (dimensionality, metric, and adaptability). The application of the system to selected PRO measures demonstrated the feasibility of its use and showed that most of them actually assess more than one construct. CONCLUSION: This classification system of PRO measures, based on a valid integrated conceptual model, should allow the classification of most currently used instruments and may facilitate a more adequate selection and application of these instruments.

Williams, A. (1999). Calculating the global burden of disease: Time for a strategic reappraisal? *Health Economics*, *8*(1), 1–8.

Large organizations responsible for health care face formidable problems in gathering and deploying data relevant to their principal tasks, which are to monitor the health of the communities they serve, and to decide where resources could most effectively be used to improve things. In the short run, health improvements come by choosing and supporting the right health care activities. In the long run, they also come by choosing and supporting the right research activities. The ambitions of the Global Burden of Disease Study are to contribute to all three of these important objectives. Broadly speaking, the argument is that if we knew the impact of each disease or injury upon people's life expectancy and upon the (health-related) quality of their lives, and if we knew the incidence and prevalence of each disease or injury, we could use this information to monitor population health, to establish priorities between interventions, and to guide research priorities. In what follows, these general claims are considered from three different viewpoints. In the first phase, I will abstract from practical difficulties and ask whether, in principle, calculating the global burden of disease is the best way to approach each of these problems. I conclude that it is

not. This will lead into the second stage, in which I ask whether there is anything useful that could be extracted from the global burden of disease calculations *as they are actually performed*, that would help with the three major problems that have to solved. I conclude that there is a little, but that it is not worth the cost, and that resources could, and should, be better targeted on discovering those things that we really need to know. This leads into my third and final section, where I aim to be rather more constructive, by offering an alternative strategy that would achieve the above objectives more straightforwardly.

Yelin, E., Herrndorf, A., Trupin, L., & Sonneborn, D. (2001). A national study of medical care expenditures for musculoskeletal conditions. *Arthritis and Rheumatism, 44*(5), 1160–1169.

OBJECTIVE: To provide estimates of all medical care expenditures on behalf of persons with musculoskeletal conditions in the United States in 1996, to estimate the increment in expenditures attributable to the musculoskeletal conditions among such persons, and to ascertain the impact of the presence or absence of health insurance and/or managed care on such expenditures. METHODS: The estimates were derived from the Medical Expenditure Panel Survey (MEPS), a national probability sample of 9,488 households, which includes responses from 21,571 persons. In the MEPS, respondents are surveyed every 6 months to report on medical care utilization and health care expenditures. Of the 21,571 persons surveyed, 4,161 reported having 1 or more musculoskeletal conditions. After weighting the data, these 4,161 individuals were inferred to represent 53.935 million persons in the nation as a whole. We tabulated all medical care expenditures of these individuals, stratified by comorbidity status, and then compared their expenditures with those among persons with chronic conditions other than musculoskeletal disease or with no chronic conditions. We then used regression techniques to estimate the increment of health care expenditures attributable to the musculoskeletal conditions. Finally, we used regression to estimate the impact of health insurance status and managed care status on the health care expenditures of the persons with musculoskeletal conditions. RESULTS: Per capita medical care expenditures in 1996 averaged $3,578 among persons with musculoskeletal conditions, for a national total of $193 billion, the equivalent of 2.5% of the Gross Domestic Product in that year. The largest components were hospital admissions (37%), physician visits (23%), and prescriptions (16%). Estimates of the per capita increment in total medical care expenditures attributable to musculoskeletal conditions ranged from a high of $723 when controlling for the other medical conditions present, to $364 when controlling for these variables and demographics. Persons with musculoskeletal conditions ages 16–64 who lacked health insurance reported total expenditures of $793, versus $3,249 among those with insurance ($P < 0.0001$). Among such persons with insurance, expenditures did not differ significantly between those in fee-for-service plans and those in managed care health plans. CONCLUSION: Persons with musculoskeletal conditions and health insurance experienced high total expenditures for medical care and high expenditures attributable to

the musculoskeletal conditions. Insurance coverage under a managed care plan had no effect on the magnitude of these total expenditures, but lack of insurance coverage did have a significant effect among persons with musculoskeletal conditions.

ATTACHMENT 2: NATIONAL HEALTH EXPENDITURES AGGREGATE, PER CAPITA AMOUNTS, PERCENT DISTRIBUTION, AND AVERAGE ANNUAL PERCENT GROWTH, BY SOURCE OF FUNDS: SELECTED CALENDAR YEARS 1960–2008

Item	1960	1970	1980	1990	1993	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
								Billions of Dollars									
National Health Expenditures	$27.5	$74.9	$253.4	$714.2	$912.5	$1,125.1	$1,190.0	$1,265.2	$1,352.9	$1,469.2	$1,602.4	$1,735.2	$1,855.4	$1,982.5	$2,112.5	$2,239.7	$2,338.7
Private	20.7	46.8	147.0	427.4	512.3	613.5	662.1	709.2	756.5	807.0	880.6	956.6	1015.5	1082.8	1136.8	1201.0	1232.0
Public	6.7	28.1	106.4	286.8	400.2	511.6	527.9	556.0	596.4	662.2	721.8	778.6	839.9	899.8	975.7	1038.7	1106.7
Federal	2.9	17.7	71.6	193.9	279.4	365.3	372.3	389.9	417.6	464.9	509.1	552.0	599.8	641.4	709.6	755.3	816.9
State and Local	3.9	10.4	34.8	92.9	120.8	146.3	155.6	166.1	178.8	197.3	212.7	226.5	240.2	258.4	266.1	283.4	289.8
Millions																	
U.S. Population[1]	186	210	230	254	263	274	277	280	283	285	288	291	293	296	299	302	304
								Billions of Dollars									
Gross Domestic Product[2]	$526	$1,038	$2,788	$5,801	$6,667	$8,332	$8,794	$9,354	$9,952	$10,286	$10,642	$11,142	$11,868	$12,638	$13,399	$14,078	$14,441
Per Capita Amount in Dollars																	
National Health Expenditures	$148	$356	$1,100	$2,814	$3,468	$4,103	$4,295	$4,522	$4,789	$5,150	$5,564	$5,973	$6,327	$6,701	$7,071	$7,423	$7,681
Private	111	222	638	1,684	1,947	2,237	2,390	2,535	2,678	2,829	3,058	3,293	3,463	3,660	3,805	3,980	4,046
Public	36	134	462	1,130	1,521	1,866	1,905	1,987	2,111	2,321	2,506	2,680	2,864	3,041	3,266	3,443	3,635
Federal	15	84	311	764	1,062	1,332	1,344	1,393	1,478	1,629	1,768	1,900	2,045	2,168	2,375	2,503	2,683
State and Local	21	49	151	366	459	533	562	594	633	692	738	780	819	873	891	939	952
								Percent Distribution									
National Health Expenditures	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Private	75.5	62.5	58.0	59.8	56.1	54.5	55.6	56.1	55.9	54.9	55.0	55.1	54.7	54.6	53.8	53.6	52.7
Public	24.5	37.5	42.0	40.2	43.9	45.5	44.4	43.9	44.1	45.1	45.0	44.9	45.3	45.4	46.2	46.4	47.3
Federal	10.4	23.7	28.2	27.2	30.6	32.5	31.3	30.8	30.9	31.6	31.8	31.8	32.3	32.4	33.6	33.7	34.9
State and Local	14.1	13.8	13.7	13.0	13.2	13.0	13.1	13.1	13.2	13.4	13.3	13.1	12.9	13.0	12.6	12.7	12.4
								Percent of Gross Domestic Product									
National Health Expenditures	5.2	7.2	9.1	12.3	13.7	13.5	13.5	13.5	13.6	14.3	15.1	15.6	15.6	15.7	15.8	15.9	16.2
								Average Annual Percent Growth from Previous Year Shown									
National Health Expenditures		10.5	13.0	10.9	8.5	5.4	5.8	6.3	6.9	8.6	9.1	8.3	6.9	6.9	6.6	6.0	4.4
Private		8.5	12.1	11.3	6.2	4.6	7.9	7.1	6.7	6.7	9.1	8.6	6.1	6.6	5.0	5.6	2.6
Public		15.3	14.2	10.4	11.7	6.3	3.2	5.3	7.3	11.0	9.0	7.9	7.9	7.1	8.4	6.5	6.5
Federal		20.0	15.0	10.5	12.9	6.9	1.9	4.7	7.1	11.3	9.5	8.4	8.6	6.9	10.6	6.4	8.2
State and Local		10.3	12.9	10.3	9.1	4.9	6.4	6.7	7.7	10.3	7.8	6.5	6.0	7.6	3.0	6.5	2.2
U.S. Population[1]		1.2	0.9	1.0	1.2	1.0	1.0	1.0	1.0	1.0	0.9	0.9	0.9	0.9	1.0	1.0	0.9
Gross Domestic Product[2]		7.0	10.4	7.6	4.8	5.7	5.5	6.4	6.4	3.4	3.5	4.7	6.5	6.5	6.0	5.1	2.6

[1] Census resident-based population less armed forces overseas and population of outlying areas. Source: U.S. Bureau of the Census.

[2] U.S. Department of Commerce, Bureau of Economic Analysis.

Note: Numbers and percentages may not add to totals because of rounding. Dollar amounts shown are in current dollars.

Source: Centers for Medicare & Medicaid Services, Office of the Actuary, National Health Statistics Group; U.S. Department of Commerce, Bureau of Economic Analysis; and U.S. Bureau of the Census.

ATTACHMENT 3: NATIONAL HEALTH EXPENDITURES, BY SOURCE OF FUNDS AND TYPE OF EXPENDITURE: CALENDAR YEARS 2003–2008

Year and Type of Expenditure	Total	Private					Public		
		All Private Funds	Consumer			Other	Total	Federal	State and Local
			Total	Out-of-Pocket Payments	Private Health Insurance				
2003					Billions of Dollars				
National Health Expenditures	$1,735.2	$956.6	$829.3	$224.7	$604.6	$127.3	$778.6	$552.0	$226.5
Health Services and Supplies	1,623.5	894.0	829.3	224.7	604.6	64.7	729.5	517.3	212.2
Personal Health Care	1,447.5	810.2	746.7	224.7	522.0	63.5	637.3	486.2	151.1
Hospital Care	527.4	229.7	205.0	17.2	187.7	24.7	297.7	241.0	56.7
Professional Services	543.0	352.2	318.6	84.7	233.9	33.6	190.8	143.4	47.4
Physician and Clinical Services	366.7	241.7	215.7	37.6	178.1	26.0	125.1	102.6	22.5
Other Professional Services	49.0	33.4	30.8	12.8	18.0	2.6	15.6	10.9	4.7
Dental Services	76.9	72.2	72.1	34.2	37.9	0.0	4.7	2.8	1.9
Other Personal Health Care	50.4	4.9	—	—	—	4.9	45.4	27.1	18.4
Nursing Home and Home Health	148.5	54.7	49.6	35.5	14.1	5.1	93.8	66.0	27.8
Home Health Care	38.0	11.3	10.4	5.0	5.4	0.9	26.7	20.0	6.7
Nursing Home Care	110.5	43.4	39.2	30.5	8.7	4.2	67.1	46.0	21.1
Retail Outlet Sales of Medical Products	228.7	173.5	173.5	87.3	86.2	0.0	55.1	35.8	19.3
Prescription Drugs	174.2	127.5	127.5	44.1	83.4	0.0	46.6	27.8	18.9
Other Medical Products	54.5	46.0	46.0	43.2	2.8	0.0	8.5	8.1	0.4
Durable Medical Equipment	22.4	15.8	15.8	13.0	2.8	0.0	6.6	6.2	0.4
Other Non-Durable Medical Products	32.1	30.2	30.2	30.2	0.0	0.0	1.9	1.9	0.0
Government Administration and Net Cost of Private Health Insurance	122.3	83.8	82.6	—	82.6	1.2	38.5	22.2	16.3
Government Public Health Activities	53.6	—	—	—	—	—	53.6	8.9	44.7
Investment	111.7	62.6	—	—	—	62.6	49.1	34.8	14.3
Research	35.5	3.3	—	—	—	3.3	32.1	27.9	4.2
Structures and Equipment	76.3	59.3	—	—	—	59.3	17.0	6.9	10.1
2004									
National Health Expenditures	1,855.4	1,015.5	881.0	234.8	646.1	134.5	839.9	599.8	240.2
Health Services and Supplies	1,733.6	947.6	881.0	234.8	646.1	66.6	786.0	561.9	224.1
Personal Health Care	1,549.9	860.5	795.1	234.8	560.3	65.3	689.4	527.3	162.1
Hospital Care	566.5	246.0	220.7	18.6	202.1	25.3	320.5	259.7	60.7
Professional Services	581.2	375.9	341.2	89.9	251.3	34.8	205.3	155.3	50.0
Physician and Clinical Services	393.6	258.0	231.4	40.1	191.4	26.5	135.7	112.0	23.7
Other Professional Services	52.9	36.2	33.2	13.8	19.4	2.9	16.7	12.1	4.6
Dental Services	81.5	76.6	76.5	36.0	40.5	0.0	4.9	3.0	2.0
Other Personal Health Care	53.3	5.3	—	—	—	5.3	48.0	28.3	19.7
Nursing Home and Home Health	157.9	55.6	50.3	36.2	14.1	5.3	102.3	72.4	29.9
Home Health Care	42.7	11.7	10.7	5.3	5.5	1.0	31.0	23.3	7.7
Nursing Home Care	115.2	43.9	39.6	30.9	8.7	4.3	71.3	49.1	22.2

(continued)

ATTACHMENT 3: NATIONAL HEALTH EXPENDITURES, BY SOURCE OF FUNDS AND TYPE OF EXPENDITURE: CALENDAR YEARS 2003–2008 (CONTINUED)

Year and Type of Expenditure	Total	Private					Public		
		All Private Funds	Consumer			Other	Total	Federal	State and Local
			Total	Out-of-Pocket Payments	Private Health Insurance				
Retail Outlet Sales of Medical Products	244.3	182.9	182.9	90.1	92.9	0.0	61.4	39.8	21.5
Prescription Drugs	188.8	136.2	136.2	46.2	90.0	0.0	52.5	31.4	21.1
Other Medical Products	55.5	46.7	46.7	43.9	2.8	0.0	8.8	8.4	0.4
Durable Medical Equipment	22.8	16.0	16.0	13.2	2.8	0.0	6.8	6.4	0.4
Other Non-Durable Medical Products	32.7	30.7	30.7	30.7	0.0	0.0	2.0	2.0	0.0
Government Administration and Net Cost of Private Health Insurance	129.8	87.1	85.8	—	85.8	1.3	42.7	25.6	17.1
Government Public Health Activities	53.8	—	—	—	—	—	53.8	9.0	44.9
Investment	121.8	67.9	—	—	—	67.9	53.9	37.8	16.1
Research	38.9	3.4	—	—	—	3.4	35.5	31.0	4.5
Structures and Equipment	83.0	64.5	—	—	—	64.5	18.4	6.8	11.6
2005					Billions of Dollars				
National Health Expenditures	$1,982.5	$1,082.8	$938.5	$247.5	$691.0	$144.3	$899.8	$641.4	$258.4
Health Services and Supplies	1,851.9	1,008.5	938.5	247.5	691.0	70.0	843.4	601.4	242.0
Personal Health Care	1,655.2	916.0	847.3	247.5	599.8	68.7	739.3	562.3	176.9
Hospital Care	607.5	262.4	235.5	19.8	215.7	26.9	345.1	277.5	67.6
Professional Services	621.5	403.3	367.1	96.3	270.8	36.2	218.2	164.7	53.5
Physician and Clinical Services	422.4	279.0	251.6	43.8	207.8	27.4	143.3	118.3	25.1
Other Professional Services	55.9	37.6	34.4	14.3	20.1	3.1	18.4	13.7	4.6
Dental Services	86.3	81.1	81.1	38.2	42.8	0.1	5.2	3.1	2.1
Other Personal Health Care	56.9	5.6	—	—	—	5.6	51.3	29.6	21.7
Nursing Home and Home Health	168.8	57.1	51.6	37.1	14.5	5.5	111.7	78.8	32.9
Home Health Care	48.1	12.3	11.1	5.5	5.7	1.1	35.8	27.0	8.9
Nursing Home Care	120.7	44.8	40.4	31.6	8.9	4.4	75.8	51.9	24.0
Retail Outlet Sales of Medical Products	257.4	193.1	193.1	94.4	98.8	0.0	64.3	41.4	22.9
Prescription Drugs	199.7	144.6	144.6	48.8	95.8	0.0	55.1	32.6	22.5
Other Medical Products	57.7	48.6	48.6	45.6	3.0	0.0	9.2	8.8	0.4
Durable Medical Equipment	23.8	16.7	16.7	13.8	3.0	0.0	7.0	6.6	0.4
Other Non-Durable Medical Products	34.0	31.8	31.8	31.8	0.0	0.0	2.2	2.2	0.0
Government Administration and Net Cost of Private Health Insurance	140.3	92.6	91.2	—	91.2	1.4	47.7	29.9	17.8
Government Public Health Activities	56.4	—	—	—	—	—	56.4	9.2	47.2
Investment	130.6	74.2	—	—	—	74.2	56.4	40.0	16.4
Research	40.7	3.7	—	—	—	3.7	37.0	32.4	4.6
Structures and Equipment	90.0	70.5	—	—	—	70.5	19.4	7.6	11.8

(continued)

ATTACHMENT 3: NATIONAL HEALTH EXPENDITURES, BY SOURCE OF FUNDS AND TYPE OF EXPENDITURE: CALENDAR YEARS 2003–2008 (CONTINUED)

Year and Type of Expenditure	Total	Private					Public		
		All Private Funds	Consumer			Other	Total	Federal	State and Local
			Total	Out-of-Pocket Payments	Private Health Insurance				
2006					Billions of Dollars				
National Health Expenditures	2,112.5	1,136.8	982.5	254.9	727.6	154.3	975.7	709.6	266.1
Health Services and Supplies	1,975.4	1,058.7	982.5	254.9	727.6	76.1	916.8	668.3	248.5
Personal Health Care	1,762.9	964.1	889.5	254.9	634.6	74.6	798.8	620.1	178.7
Hospital Care	649.4	286.3	255.7	21.1	234.6	30.6	363.1	291.1	72.0
Professional Services	658.4	425.6	387.2	101.3	285.9	38.3	232.8	175.8	57.0
Physician and Clinical Services	446.5	295.3	266.3	46.3	220.0	29.0	151.2	125.5	25.7
Other Professional Services	58.4	39.1	35.8	14.8	20.9	3.3	19.3	14.8	4.6
Dental Services	90.7	85.3	85.2	40.2	45.0	0.1	5.5	3.2	2.2
Other Personal Health Care	62.7	5.9	—	—	—	5.9	56.8	32.3	24.5
Nursing Home and Home Health	178.1	58.8	53.2	38.3	14.8	5.7	119.3	85.1	34.1
Home Health Care	53.0	12.5	11.4	5.8	5.6	1.1	40.5	30.9	9.6
Nursing Home Care	125.1	46.3	41.8	32.6	9.2	4.5	78.8	54.3	24.5
Retail Outlet Sales of Medical Products	277.0	193.4	193.4	94.2	99.2	0.0	83.6	68.0	15.6
Prescription Drugs	217.0	143.1	143.1	46.9	96.2	0.0	73.9	58.7	15.2
Other Medical Products	60.0	50.3	50.3	47.3	3.0	0.0	9.7	9.3	0.4
Durable Medical Equipment	24.7	17.3	17.3	14.2	3.0	0.0	7.4	7.0	0.4
Other Non-Durable Medical Products	35.3	33.0	33.0	33.0	0.0	0.0	2.3	2.3	0.0
Government Administration and Net Cost of Private Health Insurance	152.0	94.5	93.0	—	93.0	1.5	57.4	38.7	18.7
Government Public Health Activities	60.6	—	—	—	—	—	60.6	9.5	51.1
Investment	137.1	78.2	—	—	—	78.2	58.9	41.3	17.7
Research	41.8	4.0	—	—	—	4.0	37.8	33.0	4.8
Structures and Equipment	95.3	74.2	—	—	—	74.2	21.1	8.2	12.9
2007					Billions of Dollars				
National Health Expenditures	$2,239.7	$1,201.0	$1,030.0	$270.3	$759.7	$171.0	$1,038.7	$755.3	$283.4
Health Services and Supplies	2,089.7	1,112.6	1,030.0	270.3	759.7	82.6	977.1	713.8	263.3
Personal Health Care	1,866.4	1,016.3	935.4	270.3	665.1	81.0	850.1	661.3	188.7
Hospital Care	687.6	302.4	268.4	22.3	246.1	34.0	385.1	307.6	77.5
Professional Services	697.5	450.7	410.1	107.7	302.4	40.6	246.7	186.1	60.6
Physician and Clinical Services	472.6	312.6	282.0	49.2	232.7	30.7	160.0	132.7	27.2
Other Professional Services	62.2	41.5	38.0	15.8	22.2	3.5	20.7	16.1	4.6
Dental Services	96.4	90.3	90.2	42.7	47.5	0.1	6.1	3.6	2.5
Other Personal Health Care	66.3	6.3	—	—	—	6.3	60.0	33.7	26.3
Nursing Home and Home Health	191.7	63.4	57.0	41.5	15.5	6.4	128.3	92.8	35.6
Home Health Care	59.3	12.9	11.7	6.1	5.7	1.2	46.4	35.6	10.9
Nursing Home Care	132.4	50.5	45.3	35.4	9.8	5.2	81.9	57.2	24.7

(continued)

ATTACHMENT 3: NATIONAL HEALTH EXPENDITURES, BY SOURCE OF FUNDS AND TYPE OF EXPENDITURE: CALENDAR YEARS 2003–2008 (CONTINUED)

Year and Type of Expenditure	Total	Private					Public		
		All Private Funds	Consumer			Other			
			Total	Out-of-Pocket Payments	Private Health Insurance		Total	Federal	State and Local
Retail Outlet Sales of Medical Products	289.7	199.8	199.8	98.9	100.9	0.0	89.9	74.8	15.1
Prescription Drugs	226.8	146.7	146.7	48.9	97.8	0.0	80.0	65.4	14.7
Other Medical Products	62.9	53.1	53.1	49.9	3.2	0.0	9.8	9.4	0.4
Durable Medical Equipment	25.5	18.0	18.0	14.9	3.2	0.0	7.5	7.1	0.4
Other Non-Durable Medical Products	37.4	35.1	35.1	35.1	0.0	0.0	2.4	2.4	0.0
Government Administration and Net Cost of Private Health Insurance	158.4	96.2	94.6	—	94.6	1.6	62.2	42.7	19.5
Government Public Health Activities	64.8	—	—	—	—	—	64.8	9.7	55.1
Investment	150.0	88.4	—	—	—	88.4	61.6	41.5	20.1
Research	42.5	4.3	—	—	—	4.3	38.2	33.0	5.1
Structures and Equipment	107.5	84.1	—	—	—	84.1	23.5	8.5	15.0
2008									
National Health Expenditures	2,338.7	1,232.0	1,060.9	277.8	783.2	171.1	1,106.7	816.9	289.8
Health Services and Supplies	2,181.3	1,138.1	1,060.9	277.8	783.2	77.2	1,043.1	774.0	269.1
Personal Health Care	1,952.3	1,044.5	969.0	277.8	691.2	75.5	907.8	718.0	189.8
Hospital Care	718.4	309.3	282.2	23.2	259.0	27.1	409.0	330.7	78.3
Professional Services	731.2	467.7	425.5	111.1	314.3	42.2	263.5	202.3	61.1
Physician and Clinical Services	496.2	323.9	291.9	50.1	241.8	31.9	172.3	144.6	27.7
Other Professional Services	65.7	43.3	39.8	16.4	23.4	3.5	22.4	17.9	4.5
Dental Services	101.2	93.9	93.8	44.6	49.2	0.1	7.3	4.5	2.8
Other Personal Health Care	68.1	6.6	—	—	—	6.6	61.5	35.3	26.2
Nursing Home and Home Health	203.1	65.8	59.6	43.5	16.1	6.2	137.3	101.9	35.4
Home Health Care	64.7	13.5	12.4	6.6	5.8	1.2	51.1	39.8	11.3
Nursing Home Care	138.4	52.3	47.2	36.9	10.3	5.1	86.2	62.1	24.1
Retail Outlet Sales of Medical Products	299.6	201.7	201.7	100.0	101.7	0.0	97.9	83.0	14.9
Prescription Drugs	234.1	147.0	147.0	48.5	98.5	0.0	87.0	72.5	14.5
Other Medical Products	65.5	54.6	54.6	51.4	3.2	0.0	10.9	10.5	0.4
Durable Medical Equipment	26.6	18.1	18.1	14.9	3.2	0.0	8.4	8.0	0.4
Other Non-Durable Medical Products	39.0	36.5	36.5	36.5	0.0	0.0	2.5	2.5	0.0
Government Administration and Net Cost of Private Health Insurance	159.6	93.6	92.0	—	92.0	1.7	65.9	45.5	20.4
Government Public Health Activities	69.4	—	—	—	—	—	69.4	10.5	59.0
Investment	157.5	93.9	—	—	—	93.9	63.6	43.0	20.6
Research	43.6	4.7	—	—	—	4.7	38.9	33.5	5.4
Structures and Equipment	113.9	89.2	—	—	—	89.2	24.7	9.5	15.2

Note: Research and development expenditures of drug companies and other manufacturers and providers of medical equipment and supplies are excluded from research expenditures. These research expenditures are implicitly included in the expenditure class in which the product falls, in that they are covered by the payment received for that product. Numbers may not add to totals because of rounding. The figure 0.0 denotes amounts less than $50 million. Dashes (—) indicate "not applicable."

Dollar amounts shown are in current dollars.

Source: Centers for Medicare & Medicaid Services, Office of the Actuary, National Health Statistics Group.

ATTACHMENT 4: COST-OF-ILLNESS SUMMARIES FOR SELECTED CONDITIONS, JANUARY 2006

Cost-of-Illness													
	Costs (in billions of $)				Costs Adjusted to 2004 (in billions of $)							Indirect Costs Include:	
Disease or Risk Factor	Total	Direct	Indirect	Intangible	Total	Direct	Indirect	Intangible	Reference Year of Estimate	Annual/ Lifetime	Direct Costs Include: Non-health	Mortality	Lost Work Days
Allergic rhinitis and conjunctivitis	—	5.9	—	—	—	8	—	—	1996	Annual			
Allergic rhinitis	—	—	2.4–4.6	—	—	—	3.0–5.8	—	1998	Annual			✓
Alzheimer's disease	51.4	40.4	11	—	69.6	54.9	14.7	—	1996	Annual			
Blood disease	11.5	8	3.5	—	11.5	8	3.5	—	2004	Annual	✓	✓	✓
Cancer	189.8	69.4	120.4	—	184.5	67	117.5	—	2005	Annual		✓	✓
Cancer	—	45.5	—	—	—	60.1	—	—	1997	Annual			
Bone neoplasms	9.1	1.6	7.5	—	12.6	2.3	10.3	—	1995	Annual		✓	✓
Breast cancer	—	5.1	—	—	—	5.8	—	—	2001	Annual	✓		
Colorectal cancer	—	2.6	—	—	—	3	—	—	2001	Annual			
Lung cancer	—	2.7	—	—	—	3.1	—	—	2001	Annual			
Prostate cancer	—	3	—	—	—	3.4	—	—	2001	Annual			
Respiratory malignancies	—	5	—	—	—	6.8	—	—	1996	Annual			
Cardiovascular disease	393.5	241.9	151.6	—	381.3	233.4	147.9	—	2005	Annual		✓	✓
Cardiovascular disease	368.4	226.7	141.7	—	368.4	226.7	141.7	—	2004	Annual		✓	✓
Angina	—	2.05	—	—	—	2.3	—	—	2001	Annual			
Angina	—	1.9–8.9 33–74.8	—	—	—	2.3–10.5 39.2–88.9	—	—	2000	Annual			
Cardiac dysrhythmias	—	7.2	—	—	—	9.8	—	—	1996	Annual			
Cerebrovascular disease (stroke)	—	16.3	—	—	—	21.5	—	—	1997	Annual			
Cerebrovascular disease	—	8.3	—	—	—	11.3	—	—	1996	Annual			
Cerebrovascular disease	—	20.9	—	—	—	29.4	—	—	1995	Annual			
Cerebrovascular disease	—	11.6	—	—	—	13.2	—	—	2001	Annual			
Circulatory disease	—	127.8	—	—	—	179.7	—	—	1995	Annual			
Congestive heart failure	—	5.2	—	—	—	7.1	—	—	1996	Annual			
Congestive heart failure	—	6.6	—	—	—	7.5	—	—	2001	Annual			
Congestive heart failure	—	14.2	—	—	—	20.0	—	—	1995	Annual			
Coronary heart disease	—	38.7	—	—	—	54.4	—	—	1995	Annual			
Dyslipidemia	—	9.2	—	—	—	10.5	—	—	2001	Annual			
Heart disease	—	57.5	—	—	—	76.0	—	—	1997	Annual			
Heart disease	—	75.9	—	—	—	106.7	—	—	1995	Annual			
Hypertension	—	108.8	—	—	—	139.4	—	—	1998	Annual			
Hypertension	—	18.2	—	—	—	24.1	—	—	1997	Annual			
Hypertension	—	14.8	—	—	—	20.1	—	—	1996	Annual			
Hypertension	—	24.9	—	—	—	28.3	—	—	2001	Annual			
Hypertensive disease	—	18.4	—	—	—	25.9	—	—	1995	Annual			
Ischemic heart disease	—	21.5	—	—	—	29.2	—	—	1996	Annual			
Myocardial infarction	—	10.7	—	—	—	12.2	—	—	2001	Annual			
Peripheral vascular disease	—	6.8	—	—	—	9.2	—	—	1996	Annual			
Chronic fatigue syndrome	—	—	9.1	—	—	—	9.8	—	2002	Annual			✓

(continued)

ATTACHMENT 4: COST-OF-ILLNESS SUMMARIES FOR SELECTED CONDITIONS, JANUARY 2006 (CONTINUED)

	Cost-of-Illness									
	Indirect Costs Include:			Costs Include:						
Disease or Risk Factor	Reduced Productivity of Patient	Unpaid Caregivers	Non-health	Related Conditions Beyond ICD-9	Disease as Secondary Diagnosis	Disease as Underlying Cause	Discount Rate	Part of Larger Study	Number of Deaths (2002)	Authors
Allergic rhinitis and conjunctivitis					✓				—	Ray et al.
Allergic rhinitis	✓								—	Cyrstal-Peters et al.
Alzheimer's Disease		✓			✓	✓			58,866	Leon et al.
Blood disease								✓	9,876	NHLBI
Cancer	✓								557,271	American Cancer Society
Cancer					✓	✓		✓	557,271	Cohen and Krauss
Bone neoplasms					✓	✓	4%	✓	1,644	Rice
Breast cancer					✓	✓		✓	41,883	Pfizer
Colorectal cancer					✓	✓		✓	56,741	Pfizer
Lung cancer					✓	✓		✓	157,713	Pfizer
Prostate cancer					✓	✓		✓	30,446	Pfizer
Respiratory malignancies					✓	✓		✓	157,713	Druss et al.
Cardiovascular disease	✓						3%		918,628	American Heart Association
Cardiovascular disease								✓	918,628	NHLBI
Angina					✓	✓		✓	373	Pfizer
Angina					✓	✓			373	Javitz et al.
Cardiac dysrhythmias					✓	✓	✓	✓	20,769	Druss et al.
Cerebrovascular disease (stroke)					✓	✓		✓	162,672	Cohen and Krauss
Cerebrovascular disease				✓	✓	✓		✓	162,672	Druss et al.
Cerebrovascular disease								✓	162,672	Hodgson and Cohen
Cerebrovascular disease					✓	✓		✓	162,672	Pfizer
Circulatory disease								✓	923,339	Hodgson and Cohen
Congestive heart failure					✓	✓		✓	56,494	Druss et al.
Congestive heart failure					✓	✓		✓	56,494	Pfizer
Congestive heart failure								✓	56,494	Hodgson and Cohen
Coronary heart disease								✓	494,382	Hodgson and Cohen
Dyslipidemia					✓	✓		✓	5,415	Pfizer
Heart disease					✓	✓		✓	696,947	Cohen and Krauss
Heart disease								✓	696,947	Hodgson and Cohen
Hypertension				✓	✓	✓			26,551	Hodgson and Cai
Hypertension					✓	✓		✓	26,551	Cohen and Krauss
Hypertension					✓	✓		✓	26,551	Druss et al.
Hypertension					✓	✓		✓	26,551	Pfizer
Hypertensive disease								✓	26,551	Hodgson and Cohen
Ischemic heart disease					✓	✓		✓	494,382	Druss et al.
Myocardial infarction					✓	✓		✓	179,514	Pfizer
Peripheral vascular disease					✓	✓	✓	✓	✓	Druss et al.
Chronic fatigue syndrome		✓							—	Reynolds et al.

(continued)

ATTACHMENT 4: COST-OF-ILLNESS SUMMARIES FOR SELECTED CONDITIONS, JANUARY 2006 (CONTINUED)

Disease or Risk Factor	Cost-of-Illness: Costs (in billions of $) Total	Direct	Indirect	Intangible	Costs Adjusted to 2004 (in billions of $) Total	Direct	Indirect	Intangible	Reference Year of Estimate	Annual/ Lifetime	Direct Costs Include: Non-health	Indirect Costs Include: Mortality	Lost Work Days
Dental services	—	74.3	—	—	—	77.6	—	—	2003	Annual			
Developmental								—					
Cerebral palsy	11.5	2.3	9.2	—	11.9	2.4	9.5	—	2003	Lifetime	✓	✓	✓
Hearing loss	2.1	0.77	1.3	—	2.2	0.8	1.3	—	2003	Lifetime	✓	✓	✓
Mental retardation	51.2	12.3	38.9	—	53.2	12.8	40.4	—	2003	Lifetime	✓	✓	✓
Vision impairment	2.5	0.57	1.9	—	2.6	0.6	2	—	2003	Lifetime	✓	✓	✓
Diabetes	132	91.8	39.8	—	142.6	99.7	42.9	—	2002	Annual		✓	✓
Diabetes	—	19.7	—	—	—	26.0	—	—	1997	Annual			
Diabetes	—	10.1	—	—	—	13.7	—	—	1996	Annual			
Diabetes	—	20.6	—	—	—	23.4	—	—	2001	Annual			
Diabetic peripheral neuropathy	—	10.9	—	—	—	12.4	—	—	2001	Annual			
Digestive disease	39.2	36.3	2.8	—	50	46.5	3.5	—	1998	Annual			✓
Drug-related illness	—	177.4	—	—	—	210.9	—	—	2000	Annual			
Endocrine disorders	—	9.7	—	—	—	12.8	—	—	1997	Annual			
Epilepsy	12.5	1.7	10.8	—	17.2	2.4	14.8	—	1995	Annual		✓	✓
Epilepsy	11.1	1.8	9.3	—	15.3	2.5	12.8	—	1995	Lifetime		✓	✓
Epilepsy	—	2.2–2.4	—	—	—	3.1–3.4	—	—	1995	Annual			
Hepatitis A	0.49	0.13	0.36	—	0.64	0.2	0.47	—	1997	Annual		✓	✓
Hepatitis C	5.46	1.8	3.66	—	7.1	2.4	4.7	—	1997	Annual		✓	✓
Infectious disease	—	5.8	—	—	—	7.7	—	—	1997	Annual			
Injury													
Childhood injuries	89	17	72	257	126.5	25.0	101.5	327.6	1994	Annual	✓	✓	✓
Childhood injuries	81.4	13.8	67.6	—	108.9	18.8	90.1	—	1996	Lifetime		✓	✓
Unintentional home injuries	55	22	33	—	69.5	28.2	41.3	—	1998	Lifetime	✓	✓	✓
Kidney disease	—	9.7	—	—	—	12.8	—	—	1997	Annual			
Mental disorder	—	29.7	—	—	—	39.3	—	—	1997	Annual			
Mental disorders	147.8	67	80.8	—	259.3	127.6	131.7	—	1990	Annual	✓	✓	✓
ADHD	5.3	1.6	3.7	—	6.2	1.9	4.3	—	2000	Annual			✓
Affective disorders	—	16	—	—	—	18.2	—	—	2001	Annual			
Affective disorders	30.4	19.2	11.2	—	54.9	36.6	18.3	—	1990	Annual	✓	✓	✓
Anxiety disorders	46.6	10.7	35.9	—	78.9	20.4	58.5	—	1990	Annual	✓	✓	✓
Bipolar disorder	24	13.3	10.7		30.4	17.0	13.4	—	1998	Lifetime		✓	✓
Depression	83.1	26.1	56.9	—	97.2	31.0	66.2	—	2000	Annual		✓	✓
Depression	—	—	31	—	—	—	33.4	—	2002	Annual			✓
Mood disorders	—	10.2	—	—	—	13.9	—	—	1996	Annual			
Schizophrenia	32.5	17.3	15.2	—	57.8	33.0	24.8	—	1990	Annual	✓	✓	✓
Schizophrenia + other psychoses	—	2.75	—	—	—	3.1	—	—	2001	Annual			
Motor vehicle accidents	230.6	34.1	196.5	—	269.1	40.5	228.5	—	2000	Annual	✓	✓	✓
Motor vehicle accidents	—	21.2	—	—	—	28.8	—	—	1996	Annual			

(continued)

ATTACHMENT 4: COST-OF-ILLNESS SUMMARIES FOR SELECTED CONDITIONS, JANUARY 2006 (CONTINUED)

	Cost-of-Illness									
	Indirect Costs Include:			Costs Include:						
Disease or Risk Factor	Reduced Productivity of Patient	Unpaid Caregivers	Non-health	Related Conditions Beyond ICD-9	Disease as Secondary Diagnosis	Disease as Underlying Cause	Discount Rate	Part of Larger Study	Number of Deaths (2002)	Authors
Dental services									—	National Health Statistics Group
Developmental										
Cerebral palsy					✓	✓	3%	✓	1,200	Honeycutt et al.
Hearing loss					✓	✓	3%	✓	—	Honeycutt et al.
Mental retardation					✓	✓	3%	✓	568	Honeycutt et al.
Vision impairment					✓	✓	3%	✓	—	Honeycutt et al.
Diabetes	✓		✓	✓	✓	✓			73,249	American Diabetes Association
Diabetes					✓	✓		✓	73,249	Cohen and Krauss
Diabetes					✓	✓		✓	73,249	Druss et al.
Diabetes					✓	✓		✓	73,249	Pfizer
Diabetic peripheral neuropathy				✓					785	Gordois et al.
Digestive disease	✓		✓	✓	✓	✓			149,597	Sandler et al.
Drug-related illness					✓	✓			2,843	Ernst and Grizzle
Endocrine disorders					✓	✓		✓	24,500	Cohen and Krauss
Epilepsy	✓				✓	✓	3.6%		887	Begley et al.
Epilepsy	✓				✓	✓	3.6%		887	Begley et al.
Epilepsy					✓	✓			887	Halpern et al.
Hepatitis A						✓	3%		76	Berge et al.
Hepatitis C						✓			8,000	Leigh et al.
Infectious disease					✓	✓		✓	—	Cohen and Krauss
Injury										
Childhood injuries	✓		✓		✓		2.5%		—	Danseco et al.
Childhood injuries	✓	✓	✓		✓		3%		—	Miller et al.
Unintentional home injuries		✓					3%		58,376	Zaloshnja et al.
Kidney disease					✓	✓		✓	32,038	Cohen and Krauss
Mental disorder					✓	✓		✓	—	Cohen and Krauss
Mental disorders	✓	✓	✓		✓	✓	6%	✓	—	Rice and Miller
ADHD		✓		✓	✓	✓			—	Birnbaum et al.
Affective disorders					✓	✓		✓	—	Pfizer
Affective disorders	✓	✓	✓		✓	✓	6%	✓	—	Rice and Miller
Anxiety disorders	✓	✓	✓		✓	✓	6%	✓	—	Rice and Miller
Bipolar disorder	✓				✓	✓			—	Begley et al.
Depression	✓								31,655	Greenberg et al.
Depression	✓					✓			31,655	Stewart et al.
Mood disorders					✓	✓		✓	—	Druss et al.
Schizophrenia	✓	✓	✓		✓	✓	6%	✓	419	Rice and Miller
Schizophrenia + other psychoses					✓	✓		✓	✓	Pfizer
Motor vehicle accidents	✓	✓	✓		✓	✓	4%		45,380	Blincoe et al.
Motor vehicle accidents					✓	✓		✓	45,380	Druss et al.

(continued)

ATTACHMENT 4: COST-OF-ILLNESS SUMMARIES FOR SELECTED CONDITIONS, JANUARY 2006 (CONTINUED)

Cost-of-Illness													
		Costs (in billions of $)			Costs Adjusted to 2004 (in billions of $)							Indirect Costs Include:	
Disease or Risk Factor	Total	Direct	Indirect	Intangible	Total	Direct	Indirect	Intangible	Reference Year of Estimate	Annual / Lifetime	Direct Costs Include: Non-health	Mortality	Lost Work Days
Multiple sclerosis	6.8–11.9	3.2–5.7	3.6–6.2	—	9.8–17.1	4.7–8.4	5.1–8.7	—	1994	Annual	✓	unclear	✓
Musculoskeletal conditions	—	39	—	—	—	53.0	—	—	1996	Annual			
Musculoskeletal conditions	214.9	88.7	126.2	—	297.8	124.7	173.1	—	1995	Annual	✓	✓	✓
Arthritis	—	16.3	—	—	—	21.5	—	—	1997	Annual			
Arthritis	—	21.5	—	—	—	24.4	—	—	2001	Annual			
Arthritis	82.4	21.7	60.8	—	113.9	30.5	83.4	—	1995	Annual	✓	✓	✓
Arthritis + other rheumatic cond.	86.2	51.1	35.1	—	112.9	67.5	45.4	—	1997	Annual			✓
Arthropathies	—	15.9	—	—	—	21.6	—	—	1996	Annual			
Congenital deformities	0.86	0.48	0.38	—	1.2	0.68	0.52	—	1995	Annual	✓	✓	✓
Fractures	21.3	15.1	6.2	—	29.7	21.2	8.5	—	1995	Annual	✓	✓	✓
Hip fractures	7.7	6.7	0.95	—	10.7	9.4	1.3	—	1995	Annual	✓	✓	✓
Osteoporosis	34	32	2	—	34	32.0	2	—	2004	Annual			
Obesity	—	51.5–78.5	—	—	—	66.0–100.5	—	—	1998	Annual			
Obesity	99.2	51.6	47.6	—	137.9	72.6	65.3	—	1995	Annual		✓	✓
Obesity and physical inactivity	—	70	—	—	—	98.4	—	—	1995	Annual			
Occupational injury and illness	171	65	106	—	265.6	106.0	159.6	—	1992	Annual		✓	✓
Occupational injury and illness	77.6	—	-	—	116.2	—	—	—	1993	Lifetime		✓	✓
Pain conditions	—	—	61.2	—	—	—	65.9	—	2002	Annual			✓
Back problems	—	13	—	—	—	17.2	—	—	1997	Annual			
Back problems	—	12.2	—	—	—	16.6	—	—	1996	Annual			
Back problems	—	17.7	—	—	—	20.1	—	—	2001	Annual			
Back problems	—	—	28	—	—	—	37.3	—	1996	Annual			✓
Back problems	—	26.3	—	—	—	33.7	—	—	1998	Annual			
Parkinson's disease	14.3	3.1	11.2	—	18	4.0	14	—	1998	Annual	✓		✓
Respiratory conditions	—	45.3	—	—	—	61.6	—	—	1996	Annual			
Acute respiratory infection	—	17.9	—	—	—	24.3	—	—	1996	Annual			
Asthma	12.7	7.4	5.3	—	16.1	9.5	6.6	—	1998	Annual		✓	✓
Asthma	—	5.7	—	—	—	7.7	—	—	1996	Annual			
Asthma	—	7.1	—	—	—	8.1	—	—	2001	Annual			
COPD	—	14.5	—	—	—	19.7	—	—	1996	Annual			
COPD	—	6.4	—	—	—	8.7	—	—	1996	Annual			
COPD	—	7.2	—	—	—	8.2	—	—	2001	Annual			
COPD	—	—	9.9	—	—	—	14	—	1994	Annual			✓
Lung disease	131.9	75.9	56	—	131.9	75.9	56	—	2004	Annual		✓	✓
Pneumonia	—	8.4	—	—	—	11.8	—	—	1995	Annual			
Pneumonia	—	16.3	—	—	—	21.5	—	—	1997	Annual			
Pneumonia	—	8.4	—	—	—	9.5	—	—	2001	Annual			
Pulmonary conditions	—	29	—	—	—	38.3	—	—	1997	Annual			
Respiratory infection (cold)	40	17	22	—	43.9	19.3	24.6	—	2001	Lifetime			✓

(continued)

ATTACHMENT 4: COST-OF-ILLNESS SUMMARIES FOR SELECTED CONDITIONS, JANUARY 2006 (CONTINUED)

	Cost-of-Illness									
	Indirect Costs Include:			Costs Include:						
Disease or Risk Factor	Reduced Productivity of Patient	Unpaid Caregivers	Non-health	Related Conditions Beyond ICD-9	Disease as Secondary Diagnosis	Disease as Underlying Cause	Discount Rate	Part of Larger Study	Number of Deaths (2002)	Authors
Multiple sclerosis						✓	3%		3,124	Whetten-Goldstein et al.
Musculoskeletal conditions				✓	✓	✓			—	Yelin et al.
Musculoskeletal conditions					✓	✓	4%	✓	—	Rice
Arthritis					✓	✓		✓	5,674	Cohen and Krauss
Arthritis					✓	✓		✓	5,674	Pfizer
Arthritis					✓	✓	4%	✓	5,674	Rice
Arthritis + other rheumatic cond.	✓				✓				—	CDC
Arthropathies					✓	✓		✓	—	Druss et al.
Congenital deformities					✓	✓	4%	✓	—	Rice
Fractures					✓	✓	4%	✓	—	Rice
Hip fractures					✓	✓	4%	✓	—	Rice
Osteoporosis		✓	✓			✓			1,862	Vanness and Tosteson
Obesity					✓	✓			—	Finkelstein et al.
Obesity	✓				✓	✓			—	Wolf and Colditz
Obesity and physical inactivity					✓	✓	✓		—	Colditz
Occupational injury and illness	✓			✓	✓	✓	4%		—	Leigh et al.
Occupational injury and illness	✓		✓			✓	3%		—	Waehrer et al.
Pain conditions	✓								—	Stewart et al.
Back problems					✓	✓		✓	540	Cohen and Krauss
Back problems					✓	✓		✓	540	Druss et al.
Back problems					✓	✓		✓	540	Pfizer
Back problems									540	Rizzo et al.
Back problems					✓				540	Luo et al.
Parkinson's disease		✓		✓		✓			16,959	Siderowf et al.
Respiratory conditions				✓	✓	✓			125,196	Yelin et al.
Acute respiratory infection					✓	✓		✓	—	Druss et al.
Asthma						✓			4,261	Weiss and Sullivan
Asthma					✓	✓		✓	4,261	Druss et al.
Asthma					✓	✓		✓	4,261	Pfizer
COPD					✓				16,444	Wilson et al.
COPD					✓	✓		✓	16,444	Druss et al.
COPD					✓	✓		✓	16,444	Pfizer
COPD									16,444	Sin et al.
Lung disease								✓	124,816	NHLBI
Pneumonia					✓				64,954	Niederman et al.
Pneumonia					✓	✓		✓	64,954	Cohen and Krauss
Pneumonia					✓	✓		✓	64,954	Pfizer
Pulmonary conditions					✓	✓		✓	165,159	Cohen and Krauss
Respiratory infection (cold)		✓		✓					—	Fendrick et al.

(continued)

ATTACHMENT 4: COST-OF-ILLNESS SUMMARIES FOR SELECTED CONDITIONS, JANUARY 2006 (CONTINUED)

	Cost-of-Illness												
		Costs (in billions of $)			Costs Adjusted to 2004 (in billions of $)							Indirect Costs Include:	
Disease or Risk Factor	Total	Direct	Indirect	Intangible	Total	Direct	Indirect	Intangible	Reference Year of Estimate	Annual/ Lifetime	Direct Costs Include: Non-health	Mortality	Lost Work Days
Sexually transmitted diseases	—	6.5	—	—	—	7.7	—	—	2000	Lifetime			
HIV/AIDS	—	6.7–7.8	—	—	—	9.1–10.6	—	—	1996	Annual			
Pelvic inflammatory disease	—	1.88	—	—	—	2.4	—	—	1998	Lifetime			
Sinusitis	—	5.8	—	—	—	7.9	—	—	1996	Annual			
Skin disease	35.9	34.3	1.6	—	47.4	45.3	2.1	51.9	1997	Annual			✓
Skin disease	37.2	27	10.2	51.9	37.2	27.0	10.2	—	2004	Annual		✓	✓
Skin disorders	—	8.8	—	—	—	11.6	—	—	1997	Annual			
Skin disorders	—	11.2	—	—	—	12.7	—	—	2001	Annual			
Atopic dermatitis and eczema	—	0.9–3.8	—	—	—	1.2–5.0	—	—	1997	Annual			
Smoking	157.7	75.8	81.9	—	206	100.1	105.9	—	1995–1999	Annual		✓	
Smoking	—	—	92.5	—	—	-	111.7	—	1997–2001	Annual		✓	
Smoking	—	72.7	—	—	—	111.9	—	—	1993	Annual			
Substance abuse	180.8	15.8	165	—	195	17.2	177.8	—	2002	Annual		✓	✓
Alcohol abuse	184.6	26.3	158.3	—	231.8	33.7	198.1	—	1998	Annual		✓	✓
Thyroid disorders	—	3.7	—	—	—	4.2	—	—	2001	Annual			
Trauma	—	44.2	—	—	—	58.4	—	—	1997	Annual			
Traumatic brain injury	—	5.4	—	—	—	7.3	—		1996	Annual			
Urinary								—					
Urinary incontinence	19.5	19	0.55	—	23.2	22.6	0.6	—	2000	Annual			✓
Urinary incontinence	—	16.3	—	—	—	22.9	—	—	1995	Annual			
Overactive bladder	12.6	11.7	0.827	—	14.9	13.9	1	—	2000	Annual			✓

(continued)

ATTACHMENT 4: COST-OF-ILLNESS SUMMARIES FOR SELECTED CONDITIONS, JANUARY 2006 (CONTINUED)

	Cost-of-Illness										
	Indirect Costs Include:			Costs Include:							
Disease or Risk Factor	Reduced Productivity of Patient	Unpaid Caregivers	Non-health	Related Conditions Beyond ICD-9	Disease as Secondary Diagnosis	Disease as Underlying Cause	Discount Rate	Part of Larger Study	Number of Deaths (2002)	Authors
Sexually transmitted diseases						✓	3%		—	Chesson et al.
HIV/AIDS					✓	✓			14,095	Hellinger and Fleishman
Pelvic inflammatory disease							5%		790	Rein et al.
Sinusitis				✓	✓	✓			—	Ray et al.
Skin disease	✓	✓		✓	✓				3,600	Dehkharghani et al.
Skin disease		✓			✓			✓	3,600	Lewin Group
Skin disorders					✓	✓		✓	3,600	Cohen and Krauss
Skin disorders					✓	✓		✓	3,600	Pfizer
Atopic dermatitis and eczema				✓	✓	✓			—	Ellis et al.
Smoking					✓	✓			438,000	CDC
Smoking					✓	✓			438,000	CDC
Smoking					✓	✓			438,000	Miller LS et al.
Substance abuse	✓		✓		✓	✓			—	ONDCP
Alcohol abuse	✓	✓	✓		✓	✓			—	Harwood et al.
Thyroid disorders					✓	✓		✓	1,940	Pfizer
Trauma					✓	✓		✓		Cohen and Krauss
Traumatic brain injury					✓				—	Schootman et al.
Urinary										
Urinary incontinence	✓	✓		✓	✓	✓			—	Hu et al.
Urinary incontinence				✓	✓	✓			—	Wilson et al.
Overactive bladder	✓	✓		✓	✓	✓			—	Hu et al.

Note: The number of deaths is based on data from: CDC/ NCHS, Worktable I: Deaths from Each Cause, by 5-year Age Groups, Race, Sex: United States, 2002. National Vital Statistics System, Mortality, http://www.cdc.gov/nchs/data/dvs/mortfinal2002_workipt1.pdf Deaths are categorized by ICD-10 code and unless otherwise noted are based on classifications of causes in Table 10 of National Vital Statistics Report 2004, 53(5).

The following are notes about how deaths attributable to certain illnesses were recorded: Blood disease (ICD-10 codes: D50–89, E83.1); Bone neoplasms (C40, 41,79.5, D16); Cardiac dysrhythmias (I47.0, 48, 49); Circulatory disease(I00–78, 80–99); Coronary heart disease (I20–25); Dyslipidemia (E75, 78); Myocardial infarction(I21–22); Peripheral vascular disease (I70–78); Cerebral palsy (G80); Mental retardation (E45, F70–73, 79); Diabetic peripheral neuropathy (E10.5); Drug-related illness (Y40–84, 88); Hepatitis A (B15); Home injuries (W00–X59, Y86); Digestive disease (A01–A09; B18.2; C18,20,22,25; K21–22,25–28,50–51,57–58,70,73–74,82,85–86,92.2; P78.8); Childhood injuries (V01–X59, Y85–86, for under age 15); Endocrine disorders (D80–83; E04–07, 20–29, 31, 32, 34, 43–46, 51, 53–55, 63.8, 63.9, 66.8, 66.9, 67, 70–72, 74, 78, 83–85; M10–11); Depression (suicide: U03, X60–84, Y87.0); Kidney disease (N00,01,03–05,11.1,13.3,13.4,17,18.9,19–20,25–27,28.8); Arthritis (M00,05,06,08,13,15–19,24,25,45,86); Back problems (M45–48, 50–51, 53, 54.2); Respiratory conditions (J41–47, 60–61); Skin disease/disorders (L01.0,02,03,05.0,08,10.9,13,28.0,30.3,40.1,51.9,53.8,60,70.8,72.1,81,89,92.8, 93.0,97,98.0,98.8,98.9; R61.9); Pulmonary conditions (J30,33,37–39,40–43,44.1,8,9,45–47,60–61,62.8,64,66.8,68,69,70.0,1,8, 9,81–82,84.0,1,85.3,86,93.01,1,96.9 98; R04.0,06,09.1,2,8); PID (N70–77); Thyroid disorders (E04–07); Hepatitis C (from Leigh et al. study); Smoking (from CDC study).

ATTACHMENT 5:
EXAMPLES OF HEALTH-RELATED QUALITY OF LIFE QUESTIONAIRES

Questionnaires for SF-36, EQ-5D, HUI 2&3, and QWB Index:

http://www.healthmeasurement.org/Measures.html

Questionnaire for SF-6D:

http://www.associationfornetworkcare.com/pdf/newlongitudinal.pdf

CDC Healthy Days Measures:

- Would you say that in general your health is

  a. Excellent

  b. Very good

  c. Good

  d. Fair or

  e. Poor

- Now thinking about your physical health, which includes physical illness and injury, for how many days during the past 30 days was your physical health not good?
- Now thinking about your mental health, which includes stress, depression, and problems with emotions, for how many days during the past 30 days was your mental health not good?
- During the past 30 days, for about how many days did poor physical or mental health keep you from doing your usual activities, such as self-care, work, or recreation?

ATTACHMENT 6:
RANKING OF THE CLINICAL PREVENTABLE BURDEN (CPB) AND COST-EFFECTIVENESS (CE) FOR VARIOUS SERVICES[a]

Services	Description	CPB	CE	Total
Aspirin chemoprophylaxis	Discuss the benefits/harms of daily aspirin use for the prevention of cardiovascular events with men ≥40, women ≥50, and others at increased risk.	5	5	10
Childhood immunization	Immunize children: diphtheria, tetanus, pertussis, measles, mumps, rubella, inactivated polio virus, Haemophilus influenzae type b, hepatitis B, varicella, pneumococcal conjugate, influenza.	5	5	10
Tobacco-use screening and brief intervention	Screen adults for tobacco use, provide brief counseling, and offer pharmacotherapy.	5	5	10
Colorectal cancer screening	Screen adults aged ≥50 years routinely with FOBT, sigmoidoscopy, or colonoscopy.	4	4	8
Hypertension screening	Measure blood pressure routinely in all adults and treat with antihypertensive medication to prevent incidence of cardiovascular disease.	5	3	8
Influenza immunization	Immunize adults aged ≥50 against influenza annually.	4	4	8
Pneumococcal immunization	Immunize adults aged ≥65 against pneumococcal disease with one dose for most in this population.	3	5	8
Problem drinking screening and brief counseling	Screen adults routinely to identify those whose alcohol use places them at increased risk and provide brief counseling with follow-up.	4	4	8
Vision screening—adults	Screen adults aged ≥65 routinely for diminished visual acuity with Snellen visual acuity chart.	3	5	8
Cervical cancer screening	Screen women who have been sexually active and have a cervix within 3 years of onset of sexual activity or age 21 routinely with cervical cytology (Pap smears).	4	3	7
Cholesterol screening	Screen routinely for lipid disorders among men aged ≥35 and women aged ≥45 and treat with lipid-lowering drugs to prevent the incidence of cardiovascular disease	5	2	7
Breast cancer screening	Screen women aged ≥50 routinely with mammography alone or with clinical breast examination, and discuss screening with women aged 40 to 49 to choose an age to initiate screening.	4	2	6
Chlamydia screening	Screen sexually active women aged ≤25 routinely.	2	4	6
Calcium chemoprophylaxis	Counsel adolescent and adult women to use calcium supplements to prevent fractures.	3	3	6

(continued)

ATTACHMENT 6: RANKING OF THE CLINICAL PREVENTABLE BURDEN (CPB) AND COST-EFFECTIVENESS (CE) FOR VARIOUS SERVICES[a] (CONTINUED)

Services	Description	CPB	CE	Total
Vision screening—children	Screen children aged ≤5 years routinely to detect amblyopia, strabismus, and defects in visual acuity.	2	4	6
Folic acid chemoprophylaxis	Counsel women of childbearing age routinely on use of folic acid supplements to prevent birth defects.	2	3	5
Obesity screening	Screen all adult patients routinely for obesity and offer obese patients high intensity counseling about diet, exercise, or both together with behavioral interventions for at least 1 year.	3	2	5
Depression screening	Screen adults for depression in clinical practices that have systems in place to assure accurate diagnosis, treatment, and follow-up.	3	1	4
Hearing screening	Screen for hearing impairment in adults aged ≥65 and make referrals to specialists.	2	2	4
Injury prevention counseling	Assess safety practices of parents of children aged ≤5 years and provide counseling on child safety seats, window/stair guards, pool fence, poison control, hot water temperature, and bicycle helmets.	1	3	4
Osteoporosis screening	Screen women aged ≥65 and women aged ≥60 at increased risk routinely for osteoporosis and discuss benefits and harms of treatment options.	2	2	4
Cholesterol screening—high risk	Screen men aged 20 to 35 and women aged 20 to 45 routinely for lipid disorders if they have other risk factors for coronary heart disease, and treat with lipid-lowering drugs to prevent incidence of cardiovascular disease.	1	1	2
Diabetes screening	Screen for diabetes in adults with high cholesterol or hypertension, and treat with a goal of lowering levels below conventional target values.	1	1	2
Diet counseling	Offer intensive behavioral dietary counseling to adult patients with hyperlipidemia and other known risk factors for cardiovascular and diet related chronic disease.	1	1	2
Tetanus-diphtheria booster	Immunize adults every 10 years.	1	1	2

[a]Rankings are out of 5 with 5 being the highest.

Source: Maciosek et al. (2006).