Assessment of Approaches to Evaluating Telemedicine

Final Report

Prepared for:
Office of the Assistant Secretary
for Planning and Evaluation,
Department of Health and Human Services

Contract Number: HHS-10-97-0012

December 2000

The Lewin Group, Inc.

Assessment of Approaches to Evaluating Telemedicine

Final Report

Prepared for:
Office of the Assistant Secretary
for Planning and Evaluation,
Department of Health and Human Services

Contract Number: HHS-10-97-0012

December 2000

The Lewin Group, Inc.

This report was prepared by The Lewin Group for the Office of the Assistant Secretary for Planning and Evaluation, Department of Health and Human Services, Contract Number HHS-10-97-0012. The OASPE Task Monitor was Thomas Hertz, Ph.D. Staff of The Lewin Group contributing to the report included: Clifford Goodman, Eric Gemmen, Scott Hines, Jennifer Klitenic, and Anna Wolenski. For further information, contact: Clifford Goodman, Ph.D., The Lewin Group, 3130 Fairview Park Drive, Suite 800, Falls Church, Virginia, 22042; tel. 703-269-5500, e-mail: clifford.goodman@lewin.com.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Telemedicine is part of the expanding use of communications technology in health care, or "telehealth," being used in prevention, disease management, home health care, long-term care, emergency medicine, and other applications. The diversification of such applications and continued advances in communications technologies, including the Internet, are raising expectations for telemedicine. However, the considerable attention focused on the technological aspects of telemedicine during the last decade has been accompanied by a lack of validated or well-demonstrated approaches for evaluating telemedicine. For program funding and policy making, there is increasing need to develop and adapt evaluative frameworks for telemedicine.

In the mid-1990s, the National Library of Medicine (NLM) requested that the Institute of Medicine (IOM) develop a broad framework for telemedicine evaluation. In 1996, based on the deliberations of a 15-member expert committee, the IOM released its report, *Telemedicine: A Guide to Assessing Telecommunications in Health Care*. The report presented a framework built upon five main evaluation elements: 1) quality of care and health outcomes, 2) access to care, 3) health care costs and cost-effectiveness, 4) patient perceptions, and 5) clinician perceptions (IOM 1996).

Since 1996, the field of telemedicine has continued to evolve and mature. Recently, the DHHS Office of the Assistant Secretary for Planning and Evaluation (ASPE) contracted with The Lewin Group to assess current approaches to evaluating telemedicine. In particular, ASPE requested that Lewin extend or otherwise update the 1996 IOM framework for telemedicine evaluation as it applies to teleconsultations. The purpose of this study is not to *evaluate* telemedicine, but rather to identify the different kinds of issues on which telemedicine evaluations can focus, and the kinds of information that such evaluations can yield. This report is intended to guide future evaluators and policy makers in selecting the questions that they want to answer regarding the value of telemedicine programs, and in designing evaluations that will best serve their interests and purposes.

This report confirms and provides examples of many of the points raised by the original 1996 IOM framework. In some cases, however, this report provides greater depth or complexity, identifies supplemental issues, and diminishes the importance of ones included in the IOM framework. Some of the differences between this report and the IOM's derive from an additional four years of experience with telemedicine, including practical findings about the barriers to acceptance and use of telemedicine. The following are the main findings of this report.

1. A fundamental consideration in evaluating a telemedicine application is specifying the purpose, target audience, and the scope or focus of evaluation. Although these often are not straightforward decisions, each evaluation should specify a minimum set of elements.

2. Patient satisfaction with telemedicine has consistently been demonstrated to be high. As such, resources for future evaluations may be better allocated to areas of higher priority.

3. Lack of reimbursement for telemedicine services has been a significant confounder in past evaluations of telemedicine. Future evaluation efforts (e.g., demonstration projects) should seek to establish comparable reimbursement environments for telemedicine and the usual care comparators whenever differences in reimbursement might affect study results.

4. The findings and utility of a telemedicine evaluation are likely to be influenced by the selection of economic perspective(s) of evaluation. To be of practical use, evaluations should account for one or more of multiple relevant economic perspectives, e.g., of clinicians, patients, hospitals, payers, or society-at-large.

5. Telemedicine comprises an evolving portfolio of technologies and applications. As such, any prospective evaluation must allow for and be prepared to assess the impact (on efficacy or effectiveness, cost, cost-effectiveness, etc.) of applications that may not have been foreseen during the evaluation design.

6. Plans for evaluation of telemedicine programs should make explicit their assumptions regarding the relationship between the timing of evaluation and the maturity of the telemedicine program, and the evaluations should be designed accordingly.

7. Given the need to minimize the influence of known as well as unknown sources of bias in comparative studies involving telemedicine, it is desirable to use randomized designs whenever possible. Depending upon the investigation, it may be appropriate to randomize one or more of patients, physicians, or delivery sites. However, randomization is often impractical or impossible for evaluating telemedicine applications.

8. A recurrent weakness in telemedicine evaluations has been the lack of clearly defined control groups. In general, a comparator should be the standard or level of care that would be provided in the absence of the telemedicine intervention.

9. The time horizon for a telemedicine evaluation should be sufficiently long to capture the stream of relevant health and economic effects and to detect any differences in these effects between the intervention and control groups.

10. In order to be successful, telemedicine must be integrated as smoothly as possible into existing, routine clinical and administrative functions, including facilities, scheduling and appointments, patient records, coding, and billing.

11. Independent financial viability of a telemedicine program will increase its prospects for integration into the health care mainstream and for long-term success.

INTRODUCTION

> *Telemedicine is the use of electronic communication and information technologies to provide health care when distance separates the medical professional from the patient. It also includes educational and administrative uses of these technologies in support of health care, such as distance learning and administrative videoconferencing.* — Association of Telehealth Service Providers (2000)

Telemedicine is part of the expanding use of communications technology in health care, or "telehealth," being used in prevention, disease management, home health care, long-term care, emergency medicine, and other applications. The diversification of such applications and continued advances in communications technologies, including the Internet, are raising expectations for telemedicine. However, the considerable attention focused on the technological aspects of telemedicine during the last decade has been accompanied by a lack of validated or well-demonstrated approaches for evaluating telemedicine. Indeed, although the feasibility of various telemedicine applications have been tested for more than 30 years, reliable data on costs, effectiveness, and other impacts of telemedicine remain limited (Grigsby, Kaehny, et al. 1995). For program funding and policy making, there is increasing need to develop and adapt evaluative frameworks for telemedicine.

In the mid-1990s, the National Library of Medicine (NLM) recognized the limited number and rigor of telemedicine evaluations. The NLM requested that the Institute of Medicine (IOM) develop a broad framework for telemedicine evaluation. For the purposes of its report, the IOM defined telemedicine as the use of electronic information in communications technologies to provide and support health care when distance separates the participants. In 1996, based on the deliberations of a 15-member expert committee, the IOM released its report, *Telemedicine: A Guide to Assessing Telecommunications in Health Care*. The report was intended to encourage evaluations that would guide policymakers, reassure patients and clinicians, inform health policy managers, and help those who had invested in telemedicine to identify shortcomings in, and improve upon, their programs. The report presented a framework built upon five main evaluation elements: 1) quality of care and health outcomes, 2) access to care, 3) health care costs and cost-effectiveness, 4) patient perceptions, and 5) clinician perceptions (IOM 1996).

Since 1996, the field of telemedicine has continued to evolve and mature. Recently, the DHHS Office of the Assistant Secretary for Planning and Evaluation (ASPE) contracted with The Lewin Group to assess current approaches to evaluating telemedicine. In particular, ASPE requested that Lewin extend or otherwise update the 1996 IOM framework for telemedicine evaluation as it applies to telemedical consultations. ASPE specified that the scope of this study be focused on telemedical consultations between physicians and patients. Other applications of telemedicine, such as professional and patient education, or the electronic transfer of medical information not involving consultation, are beyond the scope of this study.

The objective of this report, *Assessment of Approaches to Evaluating Telemedicine,* is to identify areas in which telemedicine evaluation is likely to be most useful in informing future policy and program decisions. Lewin's effort entailed integrating findings from a literature review, gathering information on evaluations of telemedicine activities funded by HHS, and conducting interviews with representatives of telemedicine programs and other experts in the field.

This report describes the study methods, summarizes the study findings, and addresses how future evaluations could provide the most useful information on telemedicine activities. Based on these analyses, Lewin offers a set of main findings for guiding the design of future evaluations of telemedicine programs.

BACKGROUND

A. Approach

This study had four tasks: 1) creation of an advisory committee, 2) analysis of recent relevant literature and information on HHS-supported telemedicine evaluations, 3) interviews with experts in various aspects of telemedicine, and 4) interviews and site visits with providers. A more detailed review of the project workplan is outlined below.

1. Create Advisory Committee

Drawing upon input from the ASPE Task Monitor on appropriate candidates, The Lewin Group formed a small Advisory Committee to provide guidance for the study, to suggest experts and programs to be interviewed, and to review and comment on draft deliverables and the final report. This Committee consisted of recognized national experts on telemedicine programs, including:

- Rashid Bashshur, Ph.D., University of Michigan Health System,
- Jim Grigsby, Ph.D., University of Colorado Health Sciences Center, and
- Susan Horn, Ph.D., International Severity Information Systems, Inc., and University of Utah.

2. Conduct Literature Review and Collect Available Information on HHS Evaluations

In consultation with the Task Monitor and Advisory Committee, Lewin identified the types and sources of information available on telemedicine evaluations supported by HHS. To support these efforts, Lewin conducted preliminary discussions with federal officials and experts to refine the IOM study questions and identify additional information sources and reports for review.

Using the IOM report as the initial framework for telemedicine evaluation assessment, Lewin prepared a summary and analysis of pertinent articles published since the release of the report in 1996. Consistent with the scope of the study, the search focused on clinical encounters and consultations in telemedicine, excluding such areas as teleradiology, telepathology, and reviews of specific technologies or equipment. Articles were selected based on the relevance to refining or expanding IOM's conceptual framework for evaluating telemedicine activities. Lewin summarized elements of the IOM framework and incorporated information from the literature review as appropriate.

The second part of this task involved collecting information on HHS telemedicine programs and evaluations. The goal of this effort was to identify the areas in which

current and planned evaluations are likely to be most useful in informing future policy and program decisions, and to guide future evaluation designs in directions that will be most relevant to the field of telemedicine. Agencies within HHS that support telemedicine initiatives were contacted and researched. These agencies included: 1) Agency for Healthcare Research and Quality (AHRQ, formerly the Agency for Health Care Policy and Research); 2) Health Care Financing Administration (HCFA); 3) Health Resources and Services Administration (HRSA) Office of Rural Health Policy (ORHP) and Office for the Advancement of Telehealth (OAT); 4) National Library of Medicine (NLM); and 5) Indian Health Service

3. Collect and Analyze New Information from Telemedicine Experts

Lewin conducted 15 telephone interviews with experts in various aspects of telemedicine to obtain first-hand information regarding telemedicine evaluation. Interviewees were selected based on recommendations by the Advisory Committee and the review of recent literature on telemedicine evaluation. These experts helped Lewin to identify areas in which current and future evaluations are likely to be most useful in informing future policy and program decisions. (These experts are listed in Appendix D.) The discussions were structured informally. Respondents were asked open-ended questions that addressed the areas of quality of care, health outcomes, access to care, costs, patient and clinician perceptions, and reimbursement for telemedicine services. Respondents were questioned in greater depth on issues about which they were particularly knowledgeable. Examples of the types of questions raised to the interviewees included the following.

- On what area of evaluation do you think (telemedicine) evaluators should be focusing? Quality of care and health outcomes? Access to care? Costs and cost-benefit? Patient perceptions? Clinician perceptions?
- What specific questions should evaluators of telemedicine programs be asking (with respect to the above areas of evaluation) to ensure effective evaluations?
- To what degree do you feel reimbursement drives and/or directs use of telemedicine services, and subsequent evaluations of such programs?
- What do you see as the emerging issues (in terms of policy and evaluation) within the field of telemedicine?

4. Collect and Analyze New Information from Telemedicine Providers

Lewin conducted site visits and moderated telephone discussions to assess the views of providers on the issues and areas where telemedicine evaluation findings would be most useful to them, their patients, and the health care system as a whole. Lewin staff visited telemedicine sites at Allina Health Systems (Minneapolis), University of Missouri Health Sciences Center (Columbia), and the University of Arizona (Tucson). Two 90-minute telephone "site-visits" were made to Medical College of Georgia (Augusta) and East Carolina University (Greenville, NC). These site visits were supplemented by literature

searches, Internet searches, and other information provided by the programs. The interview protocol created for these discussions is provided in Appendix C.

B. IOM Report

As noted above, limitations in the number and rigor of evaluations in the field of telemedicine prompted the NLM to request that the IOM develop a broad framework for such evaluation. In 1996, based on the deliberations of a 15-member expert committee, the IOM released a report intended to encourage evaluations that would guide policymakers, reassure patients and clinicians, inform health policy managers, and help those who had invested in telemedicine to identify shortcomings in, and improve upon, their programs.

The IOM framework identified the following four main components as being essential to the design of a telemedicine evaluation.

I. Evaluation principles

II. Steps for evaluation planning

III. Elements of an evaluation

IV. Evaluation questions (five categories)

The key points and questions of the each of the main components of the IOM framework are listed below.

1. Evaluation Principles

- Evaluation should be viewed as an integral part of program design, implementation, and redesign.
- Evaluation should be understood as a cumulative and forward-looking process for building useful knowledge and as guidance for program or policy improvement rather than as an isolated exercise in project assessment.
- The benefits and costs of specific telemedicine applications should be compared with those of current practices or reasonable alternatives.
- The potential benefits and costs of telemedicine should be broadly construed to promote the identification and measurement of unexpected and possibly unwanted effects and to encourage an assessment of overall effects on all significant strategies.
- The accent should be on identifying the least costly and most practical ways of achieving desired results rather than investigating the most exciting or advanced telemedicine options.

- By focusing on the clinical, financial, and social objectives and needs of those who may benefit or suffer from telemedicine, evaluations can avoid excessive preoccupation with the characteristics and demands of individual technologies.

2. Steps for Evaluation Planning

1. Establish evaluation objectives.

2. Set priorities for the selection of specific applications to be evaluated.

3. Assess the probable feasibility of an evaluation, including the availability of adequate funding and the likelihood of adequate cooperation from relevant parties.

4. Identify the particular intervention to be evaluated, the alternatives to which it will be compared, the outcomes of interest, and the level and timing of evaluation.

5. Specify the expected relationships between interventions and outcomes and the other factors that might affect these relationships.

6. Develop an evaluation strategy that includes a credible and feasible research design and analysis plan.

3. Elements of an Evaluation

- *Project description and research question(s).* The description identifies the application being evaluated and the alternative to which it is being compared. Research questions are to serve as the link between the program intervention and desired outcomes.
- *Strategic objectives.* State the intended effects of the project on the organization's or sponsor's goals and how the evaluation strategy relates to these goals.
- *Clinical objectives.* State the intended effects of the project on the individual or population health by changing the quality, accessibility, or cost of care.
- *Project management plan or business plan.* A management plan functions to outline project's leadership and management structures, its workplan and schedule, and its budget; while a business plan is ideally more extensive and incorporates a detailed financial analysis and appraisal of the program's fit with the organization's strategic plan.
- *Level and perspective of evaluation.* Perspectives may be clinical, institutional, or system/societal.
- *Research design and analysis plan.* These evaluation elements must take into consideration the following: (1) characteristics of experimental and comparison groups; (2) technical, clinical, and administrative processes; (3) measurable outcomes; and (4) sensitivity analysis.

- Documentation of methods and results.

4. Evaluation Questions

- Evaluating Quality of Care and Health Outcomes
  - What were the effects of the telemedicine application on the clinical process of care compared to the alternative(s)?
  - What were the effects of the telemedicine application on immediate, intermediate, or long-term health outcomes compared to the alternative(s)?
- Evaluating Access to Care
  - Did telemedicine affect the use of services or the level or appropriateness of care compared to the alternative(s)?
  - Did the application affect the timeliness of care or the burden of obtaining care compared to the alternative(s)?
- Evaluating Health Care Costs and Cost-Effectiveness
  - What were the costs of the telemedicine application for participating health care providers or health plans compared to the alternative(s)?
  - What were the costs of the telemedicine application for patients and families compared to the alternative(s)?
  - What were the costs for society overall compared to the alternative(s)?
  - How did the cost of the application relate to the benefits of the telemedicine application compared to the alternative(s)?
- Evaluating Patient Perceptions
  - Were patients satisfied with the telemedicine service compared to the alternative(s)?
- Evaluating Clinician Perceptions
  - Were attending and/or consulting clinicians satisfied with the telemedicine application compared to the alternative(s)?

Our report uses the IOM framework as a base and attempts to build on its points. Of the four main IOM components shown above, our report devotes greatest attention to the fourth one, as requested by the Task Monitor.

RESULTS

This section summarizes the major approaches, issues, and questions that can be addressed in evaluations of telemedicine. The results of this study reflect the literature since the 1996 IOM study, and the results of interviews with the 15 experts and the five programs. Most of the interviewees and other experts with whom we consulted were familiar with the IOM framework, although only a few referred to it in detail. The results are organized into two main dimensions of analysis: (1) evaluation properties and impacts and (2) evaluation methodology issues. The first section is a consideration of evaluation properties and impacts as they relate to telemedicine. The second section focuses on methodological issues related specifically to evaluation of telemedicine. Evaluation methodology was not addressed in-depth in the IOM Framework.

A. Evaluation Properties and Impacts

This section addresses the following properties and impacts of the evaluation of health care technologies, interventions, and systems:

- access;
- technical properties;
- safety;
- efficacy and effectiveness;
- cost and other economic impacts;
- appropriateness of the technology;
- clinician acceptance;
- patient satisfaction; and
- integration into the mainstream of care.

Many respondents noted "quality" an important evaluation attribute of telemedicine systems or programs. However, when describing quality, such respondents usually described it in terms of one or more of technical properties, efficacy, effectiveness, or appropriateness of telemedicine. Respondents emphasized tradeoffs and other interdependence among multiple evaluation attributes, for example, among cost (or other economic impacts) and technical properties, access, or effectiveness. As a result, many of the points mentioned below appear in more than one area of evaluation. In general, respondents considered that access and efficacy should be accorded the highest priority in telemedicine evaluation.

1. Access

Access refers to the ability of a patient to avail himself or herself of appropriate health care in a timely manner. As suggested by the IOM (1996), access can be enhanced by increased availability of health information, allowing patients or other consumers to learn more about health problems, care options, and prevention strategies. This study focused on the former definition of access to care. Among our respondents, access to health care was generally regarded as the greatest advantage that telemedicine affords, and it was given highest priority (together with quality of care and health outcomes) with respect to areas of telemedicine evaluation on which to focus.

Among the main purposes of implementing a telemedicine program is to improve access to care by lowering geographical and temporal barriers. A telemedicine application may provide care that would otherwise not have been provided, i.e., yielding a net increase in care.

Telemedicine remains underutilized in the views of many observers. According to the 1999 Report on U.S. Telemedicine Activity of the Association of Telehealth Service Providers (ATSP), there were an estimated 41,740 telemedicine consults in 1997 and 52,223 in 1998, with 75,000 projected for 1999 (based on first quarter data). (These estimates did not include radiology and home health consults.) The most active specialties in telemedicine are mental health, dermatology, cardiology, orthopedics, and radiology. According to ATSP, given underreporting of visits, the actual numbers of telemedicine consultations may have been 40-100% higher. In comparison, there were some 750 million in-person patient-provider visits in 1998 (ATSP 2000).

The types of factors that may affect access to health care services include:

- geographic proximity of a service provider;
- financial status and insurance coverage;
- motivation and care-seeking behavior;
- convenience (timing, availability of transportation, etc.); and
- socioeconomic status.

Most of our respondents cited geographic barriers as a primary factor that limits access to care in rural settings and that may be overcome with telemedicine. Systemic (i.e., related to health care delivery or organization) barriers such as lack of inner-city health care services or inadequate health care services were more often cited as limiting access in urban settings. It was also pointed out that urban programs may be focused more on efficiency (i.e., removing internal systemic or bureaucratic barriers to treating existing patients) rather than overcoming geographic barriers to access.

In evaluation, utilization is often used as a proxy for access to care. For example, in one network's telepsychiatry program, 46% of those patients taking part in the program were seeing a psychiatrist for the first time, suggesting that psychiatric assistance was not

available to these individuals before it was offered through telemedicine. It is important to note, however, that an initial surge in telemedicine utilization may reflect pent-up demand and may subside once this consultation backlog is handled. That is, an evaluation of access may reveal a spike in patient volume at the onset of a telemedicine program as patients who have yet to seek care may have their initial appointment via telemedicine. Following these initial visits, the immediate needs of the population have been met and thus the number of visits may drop until a steady, maintainable level is reached. Further, any estimate of the rate of patients seeing a provider for the first time in a telemedicine program should be compared to the rate for patients in conventional settings.

In evaluating telemedicine, it is not sufficient to compare its effectiveness against conventional care. It also is important to identify ways in which telemedicine provides care that would not be available through conventional means. For example, telemedicine may improve access by coordinating care in a way that would otherwise not have occurred, as in an instance recounted by one of our study sites. A boy was involved in a motor vehicle accident early in 1999, during which he sustained a traumatic brain injury. He returned to school in late summer after sufficient recovery. However, the boy's injuries had behavioral effects, and it became apparent that both he and his teachers would need assistance in dealing with sometimes very disruptive behavioral problems. Through the use of the telemedicine facility, a real-time conference with two of the boy's clinical providers, his mother, his classroom teacher, the school counselor, and the school speech therapist was held. The teachers and counselors were able to express their observations and concerns, and the clinical staff was able to explain the changes in the boy's behavior and provide the educational team with some guidance on what future behaviors to expect. During the course of a two-hour teleconference, they drafted a plan for how to proceed and best allow the boy to continue to function in a regular classroom. Real-time conferences of this sort rarely occur at a single location given the difficulty of having a team of local providers (e.g., teachers, parents, and therapists) travel to a larger health care center, or having specialists from the health care center travel to a remote location.

2. Technical Properties

Evaluation of telemedicine systems can focus on a variety of technical properties, including data transmission speed or bandwidth, data quality (e.g., resolution), system functions and features, ease of use, reliability, and service or maintenance requirements. These properties are at the core of one of the most challenging aspects of telemedicine evaluation, i.e., the "moving target problem," where many technologies used in telemedicine are undergoing continual change. Technical properties such as bandwidth and resolution are steadily improving, while the costs to achieve given levels of technical performance are decreasing. In many cases, the technology is improving on a yearly basis, improving the ability of health care providers to make accurate diagnoses via telemedicine (for example in conducting dermatology consultations), and making the application more user friendly (for example with use in home health care).

When the lifecycle of certain key component technologies of telemedicine is shorter than the evaluation cycle, the findings of such an evaluation can be outdated and misleading.

The evolution of technology involves diversification as well as substitution of new for old. According to a recent buyer's guide, categories of videoconferencing technology alone include: interactive video room systems and rollabouts, specialized telemedicine rollabout units, computer-based desktop videoconferencing units, videophones, laptops, set-tops, hand-held mobile, and wireless communications and data systems (Telemedicine 2000 Buyer's Guide & Directory, 2000). This "moving target" problem is not unique to telemedicine. Telemedicine evaluations should be designed to account for these moving targets.

Among our interviewees, the two most widely cited of the technical issues were greater bandwidth and the impact of the Internet on telehealth (including the accompanying security and confidentiality issues). Bandwidth refers to the amount of data that can be transmitted in a fixed amount of time. Thus, greater bandwidth allows for more data to be transmitted more quickly. As demand and use of bandwidth increase in all areas of telecommunication, associated costs of each individual area of use will decrease. As other applications use bandwidth, the cost burden on any particular application, including telemedicine, will be reduced. Greater bandwidth enables greater resolution, use of real-time vs. store-forward images, full-motion imaging, and other properties that will expand the technical capacity of telemedicine.

The Internet has considerable potential as a medium for teleconsultations, monitoring patient condition, and other unforeseen applications in telemedicine. Use of the Internet for teleconsultations and other telemedicine applications will move these applications into the mainstream of other communications used by physicians and other health care providers, decreasing the need for separate facilities (equipment, space, etc.), procedures, and telecommunications standards for telemedicine. As many of our interviewees emphasized, any developments that reduce the "separateness" of telemedicine from other parts of the health care system will improve its acceptance and efficiency.

As noted by the Association of Telehealth Services Providers, the potential impacts of the Internet and greater bandwidth in advancing the technical properties of telemedicine are linked:

> *The Internet has become the common standard for transmission of nearly all types of data, including web-based data transfer, audio, and video. The reason that we don't use the Internet more for all of these things is that the bandwidth and switching capacity is not there. These will clearly grow in time, however, making the Internet Protocol the lingua franca of data transmission of all types. In the next ten years, virtually all telehealth transmissions will happen using Internet Protocol, whether or not the transmissions happen over the Internet. As Internet capacity grows, we expect that nearly all telehealth transactions will be done via the Internet.* — Association of Telehealth Service Providers (2000)

3. Safety

Safety is a judgment of the acceptability of the health risk (e.g., due to complications or adverse effects) associated with using a technology. Experience to date indicates that telemedicine is associated with little or no adverse health risk. Respondents made very little mention of safety. As Bashshur argues, "[t]he issue of safety should be put to rest. It should be clear that information technology, as now incorporated into clinical practice, does not present significant health hazards or risks to the patient or the provider any more than conventional patient assessment and treatment techniques" (Bashshur 1998). When addressed, safety was defined more as a function of clinician judgment (in deciding whether to use the telemedicine technology for a particular case) than with the technology itself.

4. Efficacy and Effectiveness

The distinction between efficacy and effectiveness poses a challenge to telemedicine evaluation. Efficacy refers to the benefit of using a technology for a particular health problem in ideal conditions of use, for example, in a strict protocol of a randomized controlled trial conducted at a "center of excellence." Effectiveness is the benefit of using a technology for a particular health problem in general or routine conditions of use, for example, in a community setting. In most health care applications, efficacy and effectiveness comparisons present tradeoffs between internal and external validity.

The carefully controlled, ideal circumstances of an efficacy trial tends to provide findings with stronger internal validity concerning the causal relationship between a health care intervention and outcomes of interest. However, the findings of an efficacy trial may have only limited external validity, or generalizability, to other settings. On the other hand, the less controlled, routine circumstances of an effectiveness trial may provide more generalizable findings, but may have been less able to account for factors that may have confounded the causal relationship between an intervention and outcomes of interest. For many types of technologies, efficacy trials are conducted initially. If the technology is shown to be efficacious, it is then tried in other circumstances (different settings, patient group, different providers) to determine if it is effective more broadly.

Reports of the findings of telemedicine demonstrations or other studies are often made by "champions" or "early adopters" who tend to be advocates using the telemedicine applications in carefully chosen settings. As such, it may be difficult to generalize findings of individual telemedicine studies or demonstrations to general or routine circumstances. That is, while efficacy may be established in these studies, it may be more difficult or impractical to demonstrate effectiveness.

Getting a "fix" on effectiveness is complicated by the "moving target" nature of the field. Even as this initial experience is gained with a telemedicine application, its component technologies, their configurations, or other aspects of the application are evolving. As such the findings of a study may be outdated by the time a report appears in the literature.

In order to evaluate the effectiveness of telemedicine, the application of the technology needs to be considered. By focusing on a specific application of the technology (and/or on a specific setting and condition treated), an efficacy evaluation may achieve greater internal validity. Within the field of teleconsultation, applications include a number of activities (Grigsby et al. 1994):

- Supervision and consultation for primary care encounters in sites where a physician is not available.
- Routine diagnostic evaluations based on history, physical exam findings, and available test data.
- Extended diagnostic work-ups or short-term management of self-limited conditions.
- Medical and surgical follow-up and medication checks.
- Management of chronic diseases and conditions requiring a specialist not available locally.
- Initial urgent evaluation of patients, triage decisions, and pretransfer arrangements.

An all-encompassing evaluation of "telemedicine," per se, is not necessary to demonstrate the application's effectiveness. If an application is effective consistently across a representative set of indicators/applications, it is not necessary to evaluate for all indications. An illustration of this point is provided by considering the case of antibiotics. It is commonly understood that antibiotics are effective – as a treatment class, they do not need to be evaluated every time they are used. It remains, however, necessary to demonstrate that a particular antibiotic is effective at destroying a particular infection. Similarly, Grigsby et al. (1994) suggests narrowing the scope of evaluation, by selecting certain conditions to serve as indicators of the effectiveness of telemedicine. The accuracy of the diagnosis (specificity and sensitivity) for these conditions would demonstrate the effectiveness of this mode of health care delivery. The degree of accuracy required for a given condition depends not only on the seriousness of the condition, but on the nature of its progression as well. Grigsby illustrates his point by comparing the diagnostic process for Chronic Obstructive Pulmonary Disease (COPD) to that of hantavirus pulmonary syndrome. A missed diagnosis in the early stages of a progressive, chronic disease like COPD may not result in adverse health outcomes in the patient. On the other hand, for a condition such as hantavirus that becomes life-threatening very quickly, accuracy in initial diagnosis is critical.

A specific example of measuring effectiveness is provided by the clinical evaluation of Parkinsonian tremor via a teleconsultation. If the patient were to be evaluated over a telemedicine connection that allowed for too few screens per second, the tremor could not be appropriately evaluated by the clinician, and this technology would be ineffective for this particular indication.

Currently, health outcomes data for telemedicine applications are limited, as small sample sizes limit the ability to derive meaningful results from an evaluation. Our

interviewees emphasize that the dearth of outcomes data stems, in part, from the limited funding for effectiveness studies of telemedicine.

5. Cost and Other Economic Impacts

Health care technologies can have a wide range of microeconomic and macroeconomic effects or impacts. At the microeconomic level, the cost of a technology may be determined by formal cost accounting, or by such proxies as prices, charges, and payment levels. Other microeconomic impacts are measured in terms of comparisons of resource requirements and the outcomes or benefits of a technology for particular applications through such analyses as cost-minimization analysis, cost-effectiveness analysis, cost-utility analysis, or cost-benefit analyses. Macroeconomic impacts include the impact of technology on national health care costs and the effect of technology on resource allocation among different health programs or among health care and other sectors of the economy.

Some of the commonly recognized types of economic impact of telemedicine applications are costs associated with: patient time and productivity; transportation; capital (equipment, space, etc.), maintenance, and communications; utilization of health care services; and staffing levels and productivity of health professionals. As is the case for other types of technology, introduction of telemedicine can prompt various cost tradeoffs. For example, changes in utilization of health care services may appear in different forms. By lowering barriers to access, telemedicine may increase near-term utilization of services and related health care costs. However, costs of earlier care for patients who otherwise may have delayed care in the absence of telemedicine may be offset by savings from reducing or obviating the need for downstream medical costs for treating what would have been progressively worse conditions. More well-designed longer-term studies of these cost tradeoffs are needed to demonstrate the health and economic value of telemedicine. Even so, as described below, the shorter-term costs may be overestimated because of the start-up costs associated with establishing a telemedicine program, particularly if these are determined based on per-patient costs where patient utilization is low for start-up programs.

Assessing the economic attributes of teleconsultations and other telemedicine applications raises some special challenges. Among these are the following.

- Low utilization in start-up or pilot telemedicine programs yields high levels of cost-per-patient or cost-per-consultation that may be misleading compared to steady-state utilization levels.
- Cost structures change with rapidly evolving technologies.
- Cost accounting may be complex for a telemedicine system that is shared by different services, departments, or institutions.
- The establishment of a telemedicine program may lead to expanded or unanticipated application (Ohinmaa et al. 1999; Sisk and Sanders 1998).

Changing cost structures can be dramatic, with the potential to alter evaluation results. Staff of one of our site visit programs informed us that the price of an interactive video system was $45,000 in 1998; the price of this same system was approximately $17,000 just one year later. The changing cost structure of telemedicine and low utilization in start-up programs can complicate analyses of cost tradeoffs presented by telemedicine.

The main types of cost analysis used in technology assessment include the following.

- Cost of Illness Analysis: economic impact of illness/condition, including treatment costs.
- Cost Minimization Analysis: least costly among alternatives that produce equivalent outcomes.
- Cost Effectiveness Analysis (CEA): costs in monetary units, outcomes in quantitative non-monetary units, e.g., reduced mortality, morbidity; life-years saved; ratio is calculated.
- Cost Consequence Analysis: form of CEA, but without aggregating or weighting across costs or outcomes; ratio is not calculated.
- Cost Utility Analysis: form of CEA, with outcomes in terms of utility or quality of life, e.g., quality-adjusted life-years (QALYs); ratio is calculated.
- Cost Benefit Analysis: costs and outcomes in monetary units, both of which are quantified in common monetary units; ratio or difference is calculated.

Cost-of-illness studies are used to quantify the magnitude of a health care problem, providing some context for the importance or potential of a new technology to have a meaningful effect on this problem. Cost-effectiveness analyses in telemedicine are still scarce (Ohinmaa et al. 1999). Although the term "cost-effectiveness" is used frequently in the literature, very few studies collect data on both costs and effectiveness. Instead, many studies assume that a telemedicine program and usual care are equally effective, and simply determine which alternative is less costly, i.e., a cost-minimization analysis. Cost-consequence analyses are increasingly used in other forms of technology evaluation when there are multiple relevant economic perspectives for a technological intervention, so that presenting the array of costs and outcomes in a disaggregated format allows particular stakeholders to use those that accrue to them. By accounting for patient utilities or preferences rather than more specific natural health care units as in cost-effectiveness analysis, cost-utility analysis enables comparisons across different types of health problems. There are few if any cost-utility analyses or cost-benefit analyses reported in the literature (Ohinmaa et al. 1999).

The approaches to accounting for costs and outcomes or benefits in cost analyses of telemedicine applications can vary in a number of important respects, including the following.

- perspective (e.g., society, payer, provider, patient);

- direct costs (medical and non-medical);
- indirect costs (e.g., loss of productivity);
- actual costs vs. charges or prices;
- time horizon (short-term or long-term);
- marginal costs vs. average costs;
- discounting;
- correction for inflation; and
- sensitivity analysis.

The *perspective* of a cost analysis refers to the standpoint from which costs and benefits (or other outcomes or impacts) are realized, e.g., clinician, patient, hospital, payer, or society at large. Our respondents identified economic perspective as a critical evaluation issue. Many respondents stressed the importance of the impact of telemedicine to costs experienced by patients, although these costs often are not adequately evaluated compared to costs to institutions or payers. Economic perspective must be made explicit in teleconsultation evaluations. (Perspective is discussed in greater detail below.)

Evaluations should identify *direct costs* and *indirect costs* of telemedicine applications. Direct costs including direct medical care costs for clinicians and other staff, capital equipment, facilities costs, communications, maintenance, etc. Direct non-medical costs include care provided by family members and transportation to and from the site of care. Indirect costs usually include the cost of time lost from work and decreased productivity for patients.

Instead of accounting for *actual costs* (of physician services and other health care services), many analyses use readily available health care charges or payments. However, charges (as well as actual payments) tend to reflect provider cost shifting and other factors that decrease their validity for representing the true costs of providing care. In the telemedicine literature, the types of costs analyzed and the methods for accounting for these vary widely, making study-to-study comparisons of costs or cost-effectiveness impractical (Ohinmaa et al. 1999).

Interpretation of cost analyses must consider that the *time horizon* of a study is likely to affect the findings regarding the relative magnitude of the costs and outcomes of a health care intervention. Costs and outcomes usually do not accrue in steady streams over time. Comparisons of costs and outcomes after one year may yield much different findings than comparisons made after 5, 10, or 25 years. Of course, studies with longer time horizons typically require more data collection and may be more costly. (Time horizon is discussed in greater detail below.)

Evaluations should make clear whether *average costs* or *marginal costs* are being used in the analysis. Whereas average cost analysis considers the total costs and outcomes of a telemedicine program (e.g., total program costs per patient consultation or per diagnosis),

marginal cost analysis considers the additional costs and outcomes for the next service (e.g., costs per additional consultation or per next diagnosis), which may provide more information about how to use resources efficiently. For example, marginal cost analysis may reveal how per-consultation costs change with increased utilization.

Cost analyses should account for the effect of the passage of time on the value of costs and outcomes. Costs and outcomes that occur in the future usually have less present value than costs and outcomes realized today. Thus, costs and outcomes should be *discounted* relative to their present value (e.g., at a rate of five percent per year). Analysis should also correct for the effects of *inflation* (which is different from discounting), such as when costs or cost-effectiveness for one year are compared to another year.

In any evaluation, there is some uncertainty associated with the estimates of certain costs, outcomes, and other variables used. Therefore, *sensitivity analysis* should be performed to determine if plausible variations in the estimates of these variables affect the results of the analysis. For example, for teleconsultations, sensitivity analysis can be used to determine how anticipated improvements in technical specifications of video conferencing systems might improve physician acceptance, what level of utilization would be required to meet certain levels of cost-effectiveness (i.e., a "break-even" analysis), or how feasible decreases in communications technology costs would affect marginal cost of consultations.

One form of cost-effectiveness analysis that can be performed across a given time horizon, whether prospectively or retrospectively, is a net present value (NPV) analysis. This type of analysis calculates the long-term return of an investment in a program by subtracting the total costs from the total returns of the investment. If the NPV is positive, future cash flows will exceed current investment and therefore the investment should be made; if NPV is negative, the investment should not be made. NPV analysis accounts for time horizon, the cost of capital (or discount rate), and economic perspective, since the costs and returns of a program accrue differently to different stakeholders, as described above (Rendina 2000).

Given the different ways in which costs and outcomes may be determined, all studies should make clear their methodology with respect to economic perspective, accounting for direct and indirect costs, and the other aspects noted above.

At issue in cost evaluation for telemedicine is determining which of the various types of cost analysis are most appropriate for the telemedicine program or application being evaluated. Few of our study respondents had specific preferences among types of cost analysis; however, their comments as a whole offer some insight into approaches to cost evaluation of telemedicine.

Of all the areas of telemedicine evaluation, respondents provided the most commentary on cost evaluation. Even so, in part because of the difficulty of carrying out cost evaluations in telemedicine, the respondents generally considered cost and related

economic attributes to be lesser priorities in conducting evaluations of telemedicine.[1] Consistent with the literature, respondents generally observed that cost evaluations have been inadequate. Several respondents did acknowledge the particular importance for telemedicine evaluation of perspective of analysis (especially accounting for patient perspectives) and specification of appropriate time horizons, e.g., capturing utilization and cost data at more mature stages rather than just during project start-up, and lasting long enough to capture downstream effects of early interventions.

6. Appropriateness

The appropriateness of a technology refers to a judgment about whether the technology should be used in particular circumstances. Appropriateness is a function of other evaluation attributes such as access, safety, effectiveness, and cost in a particular situation. For example, the appropriateness of a teleconsultation system may depend on its accessibility and effectiveness compared to alternative available interventions for a particular patient indication, geographic setting, and prevailing resource constraints.

Uncertainty about the appropriateness pertains to many new technologies, where multiple indications are feasible and clinicians are learning about the advantages and disadvantages of technologies when applied under various circumstances. Medical lasers have been applied for a variety of indications, including ones that persist (e.g., laser keratotomy), ones that are used rarely (e.g., laser coronary revascularization), and others under investigation (e.g., lasers used in intervertebral disc surgery). While minimally invasive surgery has flourished for some indications (e.g., laparoscopic cholecystectomy) it has been tried but largely abandoned for others (e.g., laparoscopic appendectomy).

As with many new technologies that have the potential for multiple applications, health care providers and administrators have been trying to determine the most appropriate applications of telemedicine. This is of particular importance for teleconsultations given the costs associated with implementing new telemedicine systems, the potential for teleconsultations to replace face-to-face consultations in instances where doing so might compromise the quality of care, and concerns among health care payers that providing coverage to telemedicine services could lead to costly, unnecessary use of these services.

Providers and administrators need guidance regarding the circumstances in which the use of telemedicine is appropriate. Various forms of "triage" may be necessary to minimize the potential for inappropriate use of telemedicine technology (Bashshur 1998). This should include establishment of specific telemedicine-related protocols to reduce arbitrary or unnecessary use of the technology. In an evolving field such as telemedicine, an important role of evaluation is to determine when new applications yield clinically significant gains in accessibility and effectiveness for a given indication, and are cost-

[1] Our first question asked respondents to informally "rank" areas that they felt would benefit most from further evaluation. Given our small sample size (N=15) and the expressed sentiment that these areas are interwoven (and therefore difficult to rank), we are able only to make qualitative statements about respondents' responses.

effective relative to alternative interventions. For example, full-motion video, as opposed to still images, may be appropriate for some types of teleconsultation, but superfluous for others. While a powerful televideo system may provide clinically important information for certain indications, a simple telephone conversation may suffice for others. Findings of these evaluations can be incorporated into the guidance or protocols for appropriate use of telemedicine, as noted above. By the same token, a determination that a technology is inappropriate (e.g., is not safe, effective, or cost-effective) for a given indication or circumstance is not necessarily generalizable to other indications or circumstances.

As is the case for other attributes of telemedicine, evaluation of appropriateness must account for the changing nature of the technology and costs. Greater bandwidth and lower costs can convert what was an inappropriate application to an appropriate one. As clinicians and administrators gain familiarity with telemedicine, they will continue to experiment with it and otherwise push the envelope of applications. As such, determinations of appropriateness is a continual process.

7. Clinician Acceptance

Acceptance of telemedicine by physicians, nurses, and other health personnel was cited by the respondents as being of moderate to high importance in telemedicine evaluation. If clinicians are not comfortable with the technology, or judge that the technology decreases their control over patient care, they may avoid using it, thereby precluding other benefits of telemedicine. Clinical acceptance of a telemedicine application may depend on the degree of confidence the clinician has in his or her clinical findings (e.g., diagnosis) from using the application as well as the clinician's satisfaction with the encounter in the absence of proximate, tactile interaction with the patient.

Evaluation instruments used to measure physician satisfaction with telemedicine have asked questions such as the following:[2]

- How would this situation have been handled without telemedicine?
- How was the patient's care affected by this encounter?
- What is the next step for the patient in terms of future care for this problem (e.g., continue with current care, referral, admission)?
- Did current experience make it more or less likely that you would use telemedicine in the future?

Five-point Likert scales may be used for the following questions:

- Overall, how satisfied were you with this telemedicine session?
- How essential was visual contact with the other site?

[2] Missouri Telehealth Network (http://telehealth.muhealth.org/eval)

- How essential was it to have full-motion video (as opposed to still images) in this encounter?
- How well did the telemedicine equipment work?

Attempts to gauge clinician satisfaction can be confounded by selection bias. Clinicians who are asked about their satisfaction with a telemedicine application are most likely to be those who are currently using it, including those who may have volunteered to participate in a demonstration project. This excludes those clinicians who may have used the application but are no longer doing so, as well as those who did not choose to participate at all. Furthermore, even among clinicians who are current users, those who choose to respond to inquiries about satisfaction may have different perceptions from those that chose not to respond. Evaluations that do not account for selection bias can provide misleading findings. By not tapping the perceptions of clinicians who no longer use the technology or who have decided not to use it at all, evaluators miss out on learning what aspects of acceptance affect the diffusion of the technology into broader, mainstream practice.

Our interviewees stressed that clinician acceptance may depend on factors that extend beyond the clinical aspects of individual patient interactions, to practice patterns and broader delivery and financing issues. For example, the acceptance of telemedicine may depend upon the patient load and capacity of a clinician, and whether the clinician is a generalist or a specialist. For an overextended local GP, it may remain preferable simply to refer a patient to a specialist rather than to take up appointment slots with telemedical consultations with the specialist. Further, the local GP provider may feel less confident performing procedures onsite or otherwise managing a patient when these functions might be better performed by an offsite specialist. On the other hand, a specialist who requires a large population base to stay viable (e.g., a hand surgeon) may welcome the opportunity to expand access to a larger population pool. Other types of users whose acceptance may affect the success of a telemedicine program are administrative and business staff, instructors, and students.

Of course, clinician or institutional acceptance of telemedicine also may be tied to reimbursement status, as well as other financial factors. Hospitals in rural or otherwise isolated areas can be at risk given declining populations or lower occupancy rates. Failure of a hospital can affect the viability of other businesses in a community. With the ability to offer teleconsultations with off-site specialists and other interactions with off-site health care resources, a hospital can increase its attractiveness and utility in the local community. As such, telemedicine can be quite acceptable to clinicians and institutions.

However, as revealed in one of our site visits to a rural telemedicine program, establishing a telemedicine program can have undesirable consequences among clinicians. In this instance, clinicians were being paid a substantial retainer (approximately $100,000 per year) by a hospital in a rural community to be on call during off hours. When a teleconsultation system was implemented successfully, the requirement for local on-call physician coverage decreased, and the hospital lowered the retainer by approximately two-thirds. This was unacceptable to the local physicians, who

countered that they would no longer be on-call. Consequently, the telemedicine program was forced to discontinue its use of teleconsultations during off hours.

Clearly, evaluation of clinician satisfaction with telemedicine must account for selection bias and must consider the broader professional, delivery, and financial context of health care. Evaluations may take measures to avoid selection bias in a manner analogous to "intention to treat" analysis used in clinical trials; that is, satisfaction data can be collected from all clinicians who were offered, or who initiated but did not necessarily sustain involvement in a telemedicine program. Evaluation should not be limited to satisfaction derived from individual patient interactions, but should account for factors such as the impact of telemedicine on patient load, adequacy of reimbursement for telemedicine-based services, and the viability of professional practice and institutional status.

8. Patient Satisfaction

Among the various attributes of impacts of telemedicine, satisfaction of patients (and sometimes patients' relatives) has been the one evaluated most often. Aspects of patient satisfaction that typically are evaluated are: convenience, comfort during a consult, comparison to in-person consultation, privacy concerns, and willingness to use telemedicine in the future. Past patient satisfaction instruments have rated patient responses to such questions such as:[3]

- Overall, how satisfied are you with today's telemedicine session?
- How easy was it to talk with the provider on the other end of the telemedicine connection?
- Are you comfortable that the provider was able to understand what your health problem was?
- How much did the telemedicine provider seem to care about you as a person?
- Did you feel relaxed or tense during the telemedicine session?
- Did the telemedicine make it easier for you to get care today?
- Do you think telemedicine improves your medical care?
- Do you think your telemedicine session was as good as a regular in-person visit?
- How well did the telemedicine equipment work today?
- Would you use telemedicine again?

Our respondents generally indicated that the area least in need of further evaluation was patient satisfaction. The results from past evaluations demonstrate that patient satisfaction has been nearly universally high, to the extent that the lack of variation in satisfaction limits an evaluator's ability to discern the sensitivity of satisfaction to other

[3] Missouri Telehealth Network (http://telehealth.muhealth.org/eval)

factors (Gustke et al. 2000).[4] Telemedicine project staff and other experts concur that consistently high levels of patient satisfaction have demonstrated that further patient satisfaction evaluation is not a priority and that evaluation resources may be spent more wisely on other areas. Indeed, telemedicine staff indicate that the effort to complete patient satisfaction forms, and the length of the forms themselves, may be regarded as a nuisance to patients as well as clinical staff.

9. Integration into the Mainstream of Care

Unlike most new technologies that diffuse smoothly into health care delivery, implementing telemedicine systems, and teleconsultations in particular, often presents departures from standard means of health care delivery, administration, and financing. Most new medications, medical devices, and medical procedures are delivered within already existing systems. Some technologies have necessitated special arrangements, such as magnetic resonance imaging (MRI) and positron emission tomography (PET) scanners that require special rooms and related facilities, but they are otherwise similarly captured in medical records systems, billing systems, and other standard processes.

Although teleconsultations can improve health care access, efficiency, and other attributes, in certain important respects it has been necessary to conduct teleconsultations outside of the health care mainstream. Given the need to accommodate the necessary video and related telecommunications technology, special rooms have been designated for teleconsultations that may be apart from the regular clinic traffic flow. Processes for making teleconsultation appointments have been separate from routine appointment systems. Because teleconsultations have not been reimbursed in the manner as other physician visits, billing and related coding and payment of teleconsultations has been conducted in separate, parallel systems. Furthermore, they may hinder data collection and evaluation. As one respondent stated, "In the eyes of the reimbursement bureaucracy, no services are being provided if no CPT code is assigned to those services." These differences represent departures from the health care mainstream that clinicians, patients, and other participants are seeking to diminish.

In order to be successful, telemedicine must be integrated as smoothly as possible into existing, routine clinical and administrative functions. This does not preclude some adaptation on the part of these existing functions if the net result is more efficient health care delivery overall. Therefore, it is critical for telemedicine evaluations to distinguish between any inefficiencies or lack of acceptance that are inherent in telemedicine applications themselves, as opposed to those that derive from an awkward fit between the telemedicine application and the clinical mainstream.

Integration of telemedicine into the clinical and administrative mainstream was well demonstrated during our site visits. At to the University of Missouri, dermatology

[4] One respondent pointed out that this may not be the case in a more urban or suburban setting, if a telemedical consult represented one more hurdle for the patient to overcome in obtaining treatment if in-person clinical care is readily available.

consultations are held in one of the standard examination rooms, which is equipped with the necessary telecommunications equipment that links the dermatologist in the exam room to a patient and accompanying nurse in an exam room at a remote site. The visits are conducted as routine visits of approximately 15 minutes. Scheduling and appointments, patient records, coding, and billing for teleconsultations are handled within the same systems as other types of encounters.

Aside from physical proximity, evaluation of telemedical consultations should inquire how or to what extent these consultations cause clinicians and health care management to depart from routine care. The further this departure, experts argue, the less likely it is that the program will succeed. As described above regarding clinician acceptance, implementing a telemedicine program may disrupt physician practice patterns and income streams. Interviewees at our site visits indicated that they have tried to ensure that referring physicians remain integral to teleconsultations and that patients (and income) are not taken away from them. The implication for evaluation is that gauging the integration of teleconsultations into the health care mainstream should consider their impact on practice patterns, patient flow, and income streams, along with any resulting physician or institutional resistance to the program. These findings can contribute to modifications toward achieving better integration of programs into delivery environments.

The views of our expert interviewees and demonstrations in our site visits indicate that ongoing integration of telemedicine into the health care mainstream may be a defining criterion for success. As stated by Grigsby, "Success [of a telemedicine program should be] measured by the extent to which it is no longer a stand-alone application" (Grigsby, Schlenker et al. 1995). Concurring is Michael Ackerman, of the Lister Hill National Center for Biomedical Communications of the NLM (Interview, May 3, 2000), who asserts that the most effective programs are those that are most seamlessly integrated into current clinical and business practice and that can operate on their own in the absence of outside funding.

Independent financial viability of a telemedicine program will increase its prospects for integration into the health care mainstream and long-term success. Indeed, the single most important evaluation criterion for any telemedicine program may be its ability to achieve independent financial viability. The ability of most programs to achieve financial viability will depend on their ability to secure appropriate coverage and adequate third-party payment for their services. Proposed telemedicine programs should include multi-year business plans that describe how the program will progress toward financial viability as outside funding from grants or other temporary sources diminishes. Such plans should be considered in any grants review process, and financial status with respect to these plans should be regularly monitored.

B. Evaluation Methodology Issues

The literature on telemedicine evaluation expresses concern about the rigor and consistency of methods used in the field. In a field where large, prospective randomized clinical trials (RCTs) are the methodological gold standard for evaluating the safety and

efficacy of pharmaceuticals and other medical interventions, teleconsultations and other telemedicine applications present numerous evaluative challenges.

Among the shortcomings cited in the literature of telemedicine evaluations are small sample sizes, flawed and poorly implemented study designs, and inaccurate and imprecise measurement (Bashshur 1998). Specific recommendations for improving the methodology include pooling of data across programs, using RCTs, and using case control studies with relevant meta-analyses (Yellowlees 1998).

A recent effort to conduct a meta-analysis of the costs associated with telemedicine is instructive regarding the methodological strength of the available body of telemedicine. Drawing from a comprehensive literature search, the investigators identified 551 non-duplicative, English language articles reporting the findings of studies of the costs of telemedicine. Of these only 38 articles had usable quantitative cost data. Among these, so many were inadequately designed or conducted that it was not possible to perform a traditional meta-analysis. A large proportion of the studies had such severe methodological flaws as omission of the number of consultations or patients, minimal longitudinal data, and lack of uniformity in cost analysis. As a result, the investigators concluded that "it is premature for any statements to be made, either positive or negative, regarding the cost-effectiveness of telemedicine in general" (Whitten et al. 2000).

Nitzen et al. (1997) attempted to ensure methodological rigor by establishing a gold standard, requiring that each patient be examined by multiple physicians, conducting the in-person and teleconsultations within a very short time span, conducting matched-pair analyses on all study data, and by calculating kappa coefficients, both for comparison of their findings with other studies and as a check on their success in reducing bias in the study design.

In response to inquiries about the need to improve the rigor of telemedicine evaluations, several of our expert interviewees acknowledged shortcomings but also noted that many technologies in widespread clinical use have not been subjected to high standards of evidence.

Based on our review of the literature, expert interviews, and site visits, we have organized prevailing evaluation methodology issues into the following categories:

- technological maturity;
- focus of evaluation;
- perspective of evaluation;
- comparator (control group/intervention);
- randomization; and
- time horizon (i.e., study duration or follow-up).

1. Technological Maturity

The evaluation of technology is usually influenced, and sometimes triggered, by its progress through its lifecycle. One set of terms that is used to describe stages of technological maturity includes "future," "experimental," "investigational," "established," and "obsolete" (The Lewin Group 2000). Future technologies are in a conceptual stage, anticipated, or in the earliest stages of development. Experimental technologies are undergoing bench or laboratory testing. Investigational technologies are undergoing initial clinical evaluation with patients. Established technologies are considered to be standard or mainstream approaches to management of a particular indication or set of clinical circumstances. Obsolete, outmoded, or abandoned technologies have been superseded by other technologies or found to be ineffective or harmful.

Technology companies, state and federal regulatory agencies, payers, clinicians, and others tend to make decisions about technologies' particular junctions in their lifecycles. Indeed, the determination of a technology's stage of development may be the primary purpose of an assessment. For payers, technologies deemed experimental or investigational are usually excluded from coverage, but those that are established and fall within the set of covered benefits are typically eligible for coverage. Other legal and regulatory requirements may affect these considerations for telemedicine as well as other services, e.g., definitions of what constitutes a physician service and licensing as it pertains to out-of-state services.

There are tradeoffs inherent in the timing of evaluation. At an early stage, evaluation may curtail diffusion of a telemedicine application that is ineffective. However, as noted elsewhere in this report, the findings of an early evaluation may not be definitive or may be misleading. An investigational telemedicine application may not yet be perfected, clinicians may not have honed their skills with the technology, and its costs may not stabilized. One telemedicine program reported to us that it took approximately 1,000 cases of telemedical pathology consults for a physician to become proficient at using the technology. The application may not have been used in enough clinical or geographical circumstances to recognize its potential benefits, and its long-term impact on health outcomes and costs may not be known.

Premature evaluation of telemedicine technology may miss evolving and/or unanticipated applications of the technology. For example, we learned from our site visit to Allina Health Systems that, although the intended primary application of its telemedicine system was to be teleconsultations, the use of the system for continuing education for clinicians and other providers and managers has emerged as a major application. As noted by various of our interviewees, ongoing program evaluation can be integrated into planning and adjusting of telemedicine programs over time.

Some observers consider that telemedicine in the form of video-based teleconsultations facilities will be eclipsed by other applications. In its 2010 forecast, the Institute for the Future anticipates that:

> *(Telemedicine) over dedicated videoconferencing facilities .... will stay a fringe activity, but some telemedicine will transmute into the use of groupware (computer communications) to share information between care teams and the use of online environments for collaboration between clinicians. In another example, the production of digital signals from imaging equipment will dramatically increase radiologists' ability to use computers to analyze and abstract information from X-rays, MR images, and other imaging devices.* — Institute for the Future (2000)

Decision makers should recognize that studies of telemedicine applications that are prototypes or are not integrated into the health care mainstream provide only interim findings about the feasibility of an application but not how well it operates as a mature application (Ohinmaa et al. 1999). Further, as noted by DeChant et al. (1996), "Methods appropriate for mature technologies may not be suitable for emerging ones, and indeed, may risk stifling their development with premature negative conclusions."

The lack of technological or programmatic maturity has profound impacts on cost evaluations in particular. Because telemedicine often is underutilized early in deployment relative to its subsequent steady-state use, the average cost (e.g., per patient or per teleconsultation) at this early stage may appear to be unacceptably high if it is taken to represent costs of the application at a more mature stage. As noted below with respect to the time horizon of analysis, this may be compounded by methods of cost accounting for the capital equipment, facilities, and staff required for the local and remote sites involved in teleconsultations. Our respondents emphasized the value of data collection beginning at the inception of a telemedicine program, which allows tracking of effectiveness, costs, satisfaction, and other parameters over time and for setting realistic expectations for other new initiatives. However, they also stressed that critical evaluations of the success of a telemedicine application should be based to the extent possible on performance at steady-state levels. This view is consistent with that of others in the literature who address the value of distinguishing between pilot and steady-state evaluations of telemedicine (Crowe 1998; Mintzer et al. 1997).

One approach to evaluating telemedicine programs that accounts for technological maturity is taking a staged approach modeled after the paradigm used for pharmaceuticals, i.e., preclinical testing followed by evaluation at phases I, II, III, and IV. DeChant et al. (1996) propose an analogous set of stages for telemedicine where, "in each stage of the analysis, the evaluation is tailored to the technology's state of development." These results would then be used to improve the technology before it is more widely diffused. The method addresses to varying degrees the three primary elements outlined in the IOM framework of quality, access, and cost. DeChant et al. argue that not all three of these components would play a role in each stage of the process, but should be considered only as appropriate. The intent is not only to adapt the evaluation to the respective stages of maturity of the technology, but to better "capture telemedicine's potential to produce system-wide change." Further work is required to develop or adapt evaluation designs that take staged approaches commensurate with technological maturity. As suggested above, this may be analogous to stages or phases of evaluation used for other types of health care technology. However, given important differences in

the nature of the technologies and their respective regulatory requirements, the particular evaluation models used for pharmaceuticals and medical devices themselves are largely inappropriate for telemedicine.

As noted elsewhere in this report, the moving target problem of technology is subject to the timing of evaluation. By the time an evaluation of a telemedicine application is conducted, reviewed, published (or otherwise disseminated), and incorporated into a clinical protocol or payment policy, its findings may be outdated by further data collection or changes in the component technologies, how the application is used, or competing technologies (Goodman 1996).

2. Focus of Evaluation

A fundamental consideration in evaluating a telemedicine application is specifying the scope or focus of evaluation. Doing so may not be as straightforward as in the case of a new drug or new medical device. In a narrow sense, an evaluation may focus on a particular store-and-forward technology or a two-way interactive television system. In a broader sense, an evaluation may encompass a full teleconsultation network, including the component technologies as well as the related facilities, protocols, staffing, reimbursement, etc. In the latter instance, telemedicine is a broadly encompassing technology, as suggested by Ohinmaa et al.:

> *The scope of telemedicine as a technology is considerably wider than the telecommunications equipment and systems that enable exchange of information at a distance. Telemedicine should be regarded in terms of the interaction of the equipment and the information transmitted with the activities of the health care professionals who use them, and the consequences for patients and others who are their clients (Ohinmaa et al. 1999).*

Telemedicine often is a means to facilitate or transmit care, or is used in conjunction with other technologies, and thus can be difficult to assess as an independent intervention. Grigsby, Schlenker, et al. (1995) developed two conceptual models that incorporate this distinction to facilitate research on telemedicine. The first is a more narrowly focused framework for studying the efficacy of telemedicine as a diagnostic medium, based on the analysis of sensitivity and specificity to establish the accuracy of telemedicine in relation to usual care. The second model is a scheme for classifying broader telemedicine applications based on processes of care rather than on specialties or disorders; this is intended to account for such variables as cost, access, acceptability, and effects on practice patterns.

Farand (1997) conducted a study designed to examine, in part, the "clinical problem-solving processes in the context of a telemedical consultation, in order to verify to what extent the technological environment preserves the characteristics of medical reasoning that are known to occur in more traditional clinical settings." This represents a more unusual approach by focusing on the clinical interaction and its consequences, rather than on the effectiveness or utility of one or more of its component technologies. Among

Farand's conclusions was that an evaluation should account for the interacting problem-solving modalities that may be encountered in the context of telemedicine consultation – that is, the reasoning that a health professional may use to make a clinical determination – and the changes, if any, in the interaction between the physician and the patient.

Another dimension of evaluation focus has to do with the extent to which the findings are intended to pertain to a particular setting or telemedicine programs in general. That is, an evaluation may be designed more to yield findings with external validity or findings with internal validity. For example, a multi-center trial may be designed to demonstrate that hub-and-spoke telepsychiatry programs can increase accessibility and reduce long-term health care events and costs. Even if this is established, however, the manager of a hospital or health network must consider whether implementing such a program would be a good investment for that organization. That is, the placement of a new hub-and-spoke telepsychiatry program in a particular hospital or health network may need to demonstrate that it is technically feasible, is acceptable to clinicians and patients, and can attract enough users and realize enough revenue to meet its costs.

There is no single correct way to describe the scope or focus of a teleconsultation evaluation. In general, though, any such evaluation should specify at least the following elements, each of which can be broken down or described in greater detail:

- health care problem(s), e.g., diagnosis of dermatological problems, diabetes management, hypertension management, psychiatry, trauma, neurosurgical emergencies;
- patient population(s), e.g., children, non-elderly adults, elderly;
- technology(ies), e.g., particular videoconferencing system, teleradiology system, hub-based multi-site teleconsultation network;
- practitioners or users, including referring clinicians (e.g., general practitioners, mid-level practitioners) and consulting clinicians (e.g., radiologists, pathologists, dermatologists, surgeons);
- setting(s) of care: e.g., ambulances and emergency room, rural clinics and university-based teaching hospital; and
- properties (or impacts or health outcomes) to be assessed, e.g., efficacy or effectiveness, cost, cost-effectiveness, cost-utility, physician and patient satisfaction.

Describing the scope or focus of a telemedicine evaluation with elements such as these helps to strengthen claims for internal and external validity. For example, the internal validity of a controlled trial comparing a telemedicine intervention and usual care can be strengthened by specifying the particular conditions of a health care problem, patient population, setting of care, etc., thereby controlling for factors that might otherwise confound the causal effect of a telemedicine intervention on the endpoints of interest. External validity can be strengthened in that the generalizability of the findings of a particular study can be constrained or specified in terms of particular health care problems, patient populations, technologies, settings, etc.

Any evaluation should make explicit the purpose of the evaluation and the intended users or target audience of the evaluation. Knowledge of the intended users should affect the objectives and scope of the evaluation. Clinicians, health care managers, patients, payers, policy makers, and others have different interests and levels of expertise. As noted with regard to evaluation perspective, they tend to have different concerns about the effects or impacts of teleconsultation systems. They also have different needs regarding the scientific or technical level of reports, the presentation of evidence and findings, and the format of reports.

3. Perspective of Evaluation

The perspective of evaluation refers to the standpoint from which costs and benefits of a program or intervention are realized. For instance, the perspective of an evaluation of a teleconsultation program may be that of one or more of: society overall, third-party payers (e.g., Medicare, state Medicaid programs, or managed care organizations), a "hub" hospital in a teleconsultation network, referring primary care physicians, consulting physician specialists, patients, and even any funding agency that is supporting the program. Clearly, costs and benefits are not realized in the same way from each of these perspectives. Therefore, the findings of a telemedicine evaluation may be influenced or dictated by the evaluation perspective chosen. Many analysts favor using the broad perspective of society and identifying all costs and all benefits accordingly. However, what is cost effective from the perspective of a national agency (if it is assumed to have a societal perspective) may not be what is cost effective from the standpoint of a hospital manager or a patient.

The success of a teleconsultation program is likely to depend, at least in part, on the incentives and disincentives that prevail from the perspectives of stakeholders in the program. The flow of third-party payment, and therefore an incentive for participating in teleconsultations, is directly influenced by such factors as site of service, whether a physician is a referring or consulting physician, and the way in which a physician is compensated (e.g., based on salary or fee-for-service). For some physicians, such as already over-burdened referral physicians, the prospect of participating in telemedicine encounters (as opposed to simply referring patients to consulting specialists) may pose an increased workload. The time required to become proficient at practicing in a telemedicine environment may be a barrier to participation for some busy physicians. Persuading third-party payers of the cost-effectiveness of telemedicine may require demonstrating that current reimbursement for teleconsultations may diminish downstream adverse health events and utilization of services.

Multiple expert interviewees noted the significance of cost savings to patients and families, and that these savings should be accounted for in policy making pertaining to telemedicine services. Patients incur costs savings primarily as a result of the convenience of telemedicine, e.g., less time taken off work or school and lower costs of travel and accommodations, of particular importance for patients with conditions requiring regular visits, such as psychiatric treatment. Such costs and savings are of greater or lesser importance to a payer, depending on the payer's financial responsibility.

For example, while Medicare does not reimburse patients for travel time, virtually all state Medicaid programs do.

Adding to the complexity of identifying the perspectives of a telemedicine evaluation is that a given telemedicine program or network often has multiple applications, such as teleconsultations, continuing education, and administration. Further, it may be used by multiple clinical or administrative departments. Therefore, assessing the net value of a telemedicine program is likely to entail evaluating its costs and benefits for one or more perspectives for these multiple applications. In any case, evaluations of telemedicine should identify and describe its evaluation perspectives.

4. Comparator

Evaluating the impact of teleconsultations requires some basis of comparison. Consistent with remarks in the literature, several of our expert interviewees indicated that a recurrent weakness in telemedicine evaluations has been the lack of a clearly defined control group. In general, a comparator should be the standard or level of care that would be provided in the absence of the experimental intervention.

The design of a teleconsultation evaluation should specify, and justify, the comparator. For an evaluation of teleconsultations for patients in a local site that is remote from desired care (e.g., from a physician specialist), possible comparators include:

- no care;
- inadequate or underspecialized in-person care locally;
- in-person care remotely (requiring patient travel);
- delayed in-person care remotely (requiring patient travel); and
- delayed in-person care locally (requiring physician specialist travel).

Identifying an appropriate control group also depends on whether a telemedicine application substitutes for care provided by on-site personnel, or if it is additive to existing care.

In order to establish a realistic basis upon which to determine the true size of the effect of the teleconsultation, the selection of comparator should reflect as nearly as possible the usual care that would be available in the absence of the teleconsultations. For any given population, this may include any or all of the five possibilities listed above. Therefore, the way to achieve the most realistic comparison may be to randomize patients to usual care, which could include any care mode that they would seek in the absence of teleconsultations, and to teleconsultations.

Experimental designs with contemporaneous controls (i.e., current, parallel control groups) are generally stronger than those with historical controls. Using historical controls fails to account for the confounding effects of the passage of time, i.e., the

different prevailing conditions that may exist between the time of data collection for the historical control group and the time of data collection for the current intervention group. That is, changes may have occurred in the study population or aspects of health care delivery or administration during this time that would confound the study results concerning the causal effect between the type of care and the outcomes of interest. Historical controls can be sufficient if there is strong reason to assume that prevailing conditions have not changed over time, and that the relationship between usual (or no) care and the outcomes of interest has remained virtually constant. This confounding effect of time also pertains in instances where data are collected for a population prior to an intervention (e.g., the establishment of a teleconsultation program) and following the intervention. (This is sometimes referred to as "pre-test post-test" design.) Another methodological weakness of historical controls is the opportunity for selection bias to occur, i.e., in the selection of the basis for the historical control. Comparative studies of telemedicine have too often relied on historical controls.

The reliance on historical controls is due to a variety of reasons, including the practical difficulties of assignment of patients (randomly or not) to intervention groups and control groups. In some instances, once the telemedicine intervention was in place, it was impractical to keep patients from using it, thereby losing the basis of a contemporary control group. In other instances, the number of participating patients has been so small (e.g., in low-density rural areas) that dividing them into intervention and control groups would yield too little data upon which to base any statistically meaningful findings. In these instances where sample sizes are small, it may be desirable to conduct multicenter studies. Of course, this typically requires greater funding. Another approach is to use meta-analysis or similar statistical techniques to combine the results of multiple small studies (each of which may not have sufficient sample sizes to yield statistically significant findings) to yield a larger study that can achieve statistically significant findings. However, doing so requires making assumptions about the comparability of the populations and interventions used in the smaller individual studies. As noted above, a recent comprehensive attempt to conduct such a meta-analysis of research reports on telemedicine costs was unable to identify a sufficient number of studies to meet minimal criteria for combining the study findings (Whitten et al. 2000).

Another approach used in teleconsultation evaluations is to use matched populations served by different, yet similar health delivery sites. In these instances, one community retains usual care while the other gets the teleconsultation intervention. The validity of this type of design rests on assumptions about the similarity of the two populations, their respective health delivery sites, and other circumstances that might affect study results.

Among our site visits and expert interviews, the single most often cited aspect of disparity between telemedicine interventions and usual care was third-party reimbursement. Several experts asserted that reimbursement drives utilization of telemedicine, and this theme was confirmed during our site visits. As several experts noted, the basis for comparison is may be undermined when reimbursement is available for usual care but not for teleconsultations. Reimbursement differences might not affect certain telemedicine evaluations, e.g., of the technical performance of a system, ease of use, or operating costs. However, reimbursement differences may confound findings

about clinician or hospital acceptance, access, utilization, health outcomes (if dependent on utilization), and other evaluation measures. Thus, for a valid evaluation of a teleconsultation program to be conducted, it may be necessary for the program to be conducted in the same payment environment as usual health care. Even to the extent that a teleconsultation program is shown to be effective and cost-effective, inadequate reimbursement could stand as a barrier to its use.

Reimbursement anomalies can have other unintended effects on telemedicine services. For example, the availability of reimbursement for telemedicine using video technology may prompt the medically unnecessary substitution of video-based encounters for simple telephone calls, which are not reimbursed.

The potential for differences in reimbursement status to confound comparative studies should be considered for demonstration projects where the results of such demonstrations are used to inform decisions about deploying or modifying telemedicine programs or establishing policies about telemedicine delivery or payment. In areas where reimbursement is not available for teleconsultations, a demonstration project should consider including funding for payment for teleconsultations that is comparable to payment for corresponding health care services. This funding could come from regular payers (e.g., Medicare, Medicaid, managed care organizations) on a special basis for the purposes of the demonstration, or it could be part of the demonstration budget itself. In either case, the process for providers to secure such reimbursement should not entail any different level or process of administrative than is entailed in reimbursement for usual care.

5. Randomization

In clinical trials or other comparative studies, randomization refers to the technique of assigning subjects (usually patients) to an experimental intervention (often a new treatment) group and a control groups based only on chance distribution. The purpose of randomization is to reduce the opportunity for selection bias when assigning patients to one group or the other. Proper randomization of patients is an indifferent yet objective technique that tends to neutralize the impact that any risk factors or other prognostic factors, known or unknown, may have on outcomes by spreading them evenly among the experimental and control groups. That is, randomization reduces the chances for any prognostic factor to be allocated unevenly between the experimental and control groups, which could thereby confound discerning the causal relationship between the experimental intervention and the outcomes of interest. For randomization to be successful, the number of patients (or other subjects) to be assigned must be large enough to achieve a high probability of evenly distributing any prognostic factors.

Given the need to minimize the influence of known as well as unknown sources of bias in comparative studies involving telemedicine, it is desirable to use random assignment whenever possible. Depending upon the investigation, it may be one or more of patients, physicians, or delivery sites that are randomized. For example, the randomization of patients to telemedicine intervention or standard of care minimizes the chances for

differences in such potential prognostic factors as age, disease severity, or socioeconomic status to confound results.

For telemedicine, randomization often is not a straightforward matter. As noted above, telemedicine interventions are not always discrete or self-contained technologies, and thereby present challenges to randomization. In the instances of a trial of a new drug, for example, the molecular entity itself is reasonably assumed to be the same from pill to pill. As long as physicians or patients follow the trial protocol for dispensing or taking the pills, there is presumed to be no interactive effect between the delivery of the medication and the molecular entity itself. However, in a trial of teleconsultations, the causal effects of the teleconsultation may be confounded by differences among the participating physicians, among the participating institutions (for a multicenter trial), or among other factors. For example, physicians include stronger and lesser proponents of telemedicine, have varying levels of confidence or satisfaction in teleconsultations, and have varying levels of skill in conducting them. Many telemedicine evaluations to date have involved physicians that have been self-selected as proponents, rather than also including physicians who may have tried telemedicine but rejected it, physicians that have been reluctant to participate in it, or physicians that have been interested but without access to telemedicine facilities. Clinics or other delivery settings vary in many ways that may affect the provision of teleconsultations. It may be difficult to standardize or control for these potentially confounding factors.

In principle, then, in addition to randomizing patients to either teleconsultation or standard care, it may be desirable to randomize participating referring physicians to teleconsultations or standard care, and to randomize participating clinics to providing teleconsultations or standard care. In practice, however, the numbers of participating physicians and delivery sites may be too small for randomization to distribute prognostic factors to teleconsultations or standard care evenly enough to neutralize their effects on the outcomes of interest. Although it is more preferable to randomize than not to randomize, even where small numbers prevail, the practical constraints may outweigh the benefits of doing so. Therefore, investigators need to make explicit assumptions about the similarities among physicians and among clinics, i.e., that they are not sufficiently different to have independent effects on the outcomes of interest. Clearly, these are important design considerations that can affect the validity of any findings of comparative studies of telemedicine evaluations.

Several of our expert interviewees called for large, multicenter RCTs of teleconsultations in which patients would be randomly assigned to teleconsultations or standard care, and health outcomes would be followed over time. To the extent that the centers involved and the nature of the teleconsultations can be assumed to be comparable, such RCTs could provide convincing evidence about the value of teleconsultations.

6. Time Horizon

The time horizon of a study refers to the study duration or length of follow-up for data collection. The time horizon for a comparative evaluation should be long enough to

capture the stream of relevant health and economic effects that are sufficient to detect any differences in these between the intervention and control groups. To not do so may yield misleading findings.

One of the challenges of telemedicine evaluations derives from the novelty of telemedicine applications. It is inevitable that provider institutions, physicians, patients, and other participants will require some time and practical experience to gain familiarity and skill with these applications. As such, some time will be required to "get up the learning curve" for these applications, with corresponding changes in efficiency and satisfaction. Study durations that begin during this ramping-up period may yield misleading results. Similarly, given the inefficiencies of resource use that arise when installing any new technology or program, the costs of operating a start-up telemedicine operation will not reflect the true, longer-term running costs of the program. In the case of rapidly evolving technologies, such as those used in telemedicine, the costs can change during the course of a study. The results of evaluations of cost or cost effectiveness of telemedicine operations can be very sensitive to the time spans for depreciation of capital costs and other accounting techniques for spreading costs over time.

The time horizon of a comparative evaluation of telemedicine should depend upon the endpoints or outcome measures of interest. Determining how teleconsultations change access to services may not require long follow-up periods, particularly if they are made available to large populations. However, it may take enough time to secure multiple visits for individual patients to gauge their satisfaction, and sufficient time for clinicians to become familiar with teleconsultations with a variety of types of patient indications to get a reliable measure of their respective levels of satisfaction.

Any evaluation of telemedicine that is intended to determine its effect on health outcomes must be long enough to capture the disease episode (for acute conditions) or normal course and fluctuations of disease (to capture changes in chronic conditions). Even longer follow-ups may be required to capture data on how the use of telemedicine can avert downstream progression of disease and adverse health events and their associated health care costs. Following up on longer-term health outcomes may require a more concerted tracking effort, including capturing patient data at multiple sites of service. Of course, increasing the time horizon of an evaluation generally increases its costs, and such evaluations are subject to cost constraints. Further, managers and policy makers usually seek study findings sooner rather than later, so there often is pressure to complete evaluations in as short a time as possible. Therefore, telemedicine evaluations should provide a rationale for how the time horizons correspond to the endpoints or outcomes of interest and any relevant constraints.

MAIN FINDINGS

Much of the present study serves to confirm and reinforce the 1996 IOM evaluation framework. The recent literature and interviews with telemedicine providers and other telemedicine experts provided examples of, and otherwise helped to elucidate many of the points raised by that framework.

In some cases, however, the present study provides greater depth or complexity, identifies supplemental issues, and calls into question the importance of ones included in the IOM framework. Some of these differences between this report and the IOM's derive from an additional four years of experience with telemedicine, including practical findings about the barriers to acceptance and use of telemedicine.

For example, this report addresses greater depth or complexity in the matters of identifying appropriate comparators for telemedicine evaluations (particularly the crucial role of reimbursement inequities) and in the types of incentives and disincentives that arise from different economic perspectives of evaluation. Examples of supplemental issues identified in this report include the need to implement and interpret evaluations with due consideration of technological maturity and time horizons that do not produce misleading results, considerations for randomized design, the need to evaluate progress toward moving telemedicine programs into the health care mainstream, and the importance of independent financial viability as a prospective evaluation criterion. An aspect of the IOM framework that is viewed as being of lower priority is the need for continued emphasis on measuring patient satisfaction. The main findings described below incorporate and emphasize these issues of departure from the IOM framework.

As in evaluation of any health care technology, the evaluation of telemedicine can entail various combinations of properties or impacts (access, technical properties, safety, efficacy or effectiveness, cost, etc.) and methodological aspects (evaluation perspective, selection of comparator, time horizon, etc.). Among the telemedicine programs we examined and experts we interviewed, the properties or impacts viewed as being of highest evaluation priorities were patient access and "quality" (comprising some combination of technical properties, efficacy or effectiveness, and appropriateness of care). These were followed by clinician acceptance and cost and other economic impacts. Of lower priority was patient satisfaction, not because it is unimportant, but because it has so consistently been demonstrated to be high that continuing to emphasize it in evaluation would be redundant. Safety was generally regarded as not being at issue for telemedicine.

Among methodological issues, respondents emphasized the need to identify valid control groups to represent standard or usual care, as well as the challenges of doing so. Of the many factors that could confound a comparison of the impacts of a telemedicine program and standard or usual care, respondents most often pointed to differences in reimbursement, where usual care is reimbursed in a routine fashion and telemedicine services are not reimbursed at all, inadequately, or via non-routine or inefficient means.

The main findings of this report are as follows.

1. A fundamental consideration in evaluating a telemedicine application is specifying the purpose, target audience, and the scope or focus of evaluation. Although these often are not straightforward decisions, each evaluation should specify a minimum set of elements.

   - Telemedicine technology often is a means to facilitate or transmit care, or is used in conjunction with other technologies; therefore, it can be difficult to assess as an independent intervention. The technological scope of an evaluation may range from a particular store-and-forward technology or a two-way interactive television system to a full teleconsultation network.
   - In general an evaluation should specify at least such elements as: health care problem(s), patient population(s), technology(ies), practitioners or users, setting(s) of care, and properties (or impacts or health outcomes) to be assessed.
   - Evaluations should make explicit their purposes and intended users or target audiences. Knowledge of the intended users should affect the objectives, scope, and presentation of findings of the evaluation.

2. Patient satisfaction with telemedicine has consistently been demonstrated to be high. As such, resources for future evaluations may be better allocated to areas of higher priority.

   - The great majority of studies to date indicate very high levels of patient satisfaction, as patients have given virtually universal positive responses to receiving treatment to which they would otherwise not have access.
   - Patient satisfaction with telemedicine may now have been over-studied. Multi-question surveys of satisfaction can be a nuisance to patients in settings where clinicians, patients, and other participants in telemedicine are seeking to establish conditions that are as routine as possible.

3. Lack of reimbursement for telemedicine services has been a significant confounder in past evaluations of telemedicine. Future evaluation efforts (e.g., demonstration projects) should seek to establish comparable reimbursement environments for telemedicine and the usual care comparators whenever differences in reimbursement might affect study results.

   - Inequitable reimbursement conditions for telemedicine vs. usual care may confound findings about clinician or hospital acceptance, access, utilization, health outcomes (if dependent on utilization), and other evaluation measures. Reimbursement differences might not affect certain telemedicine evaluations, e.g., of the technical performance of a system, ease of use, or operating costs.
   - The administrative process for reimbursement should be the same as it is for usual care; that is, there should not be an added administrative burden or less convenience for securing reimbursement for telemedicine services. Non-existent

or separate billing procedures for telemedicine constitute further departure from the health care mainstream.

- Reimbursement inequities pose disincentives that contribute to underutilization of telemedicine services, including initial and follow-up encounters, which in turn affects determinations of their cost-effectiveness.
- Lack of conventional reimbursement procedures (e.g., capturing services data via CPT codes) can hinder data collection and evaluation.

4. The findings and utility of a telemedicine evaluation are likely to be influenced by the selection of economic perspective(s) of evaluation. To be of practical use, evaluations should account for one or more of multiple relevant economic perspectives, e.g., of clinicians, patients, hospitals, payers, or society-at-large.

   - Costs and outcomes or benefits of telemedicine programs accrue differently to multiple stakeholders. Accordingly, these stakeholders have different incentive structures for participating in or supporting telemedicine. Evaluations should account for perspectives of one or more of referring and consulting clinicians, patients, hospitals, managed care organizations, third-party payers, society-at-large, or others as appropriate.
   - Due in part to the relative difficulty of accounting for patients' direct non-medical and indirect costs, evaluations from their perspective have been insufficient. Given the central importance of patient participation in telemedicine applications, it is essential that evaluations from this perspective be properly undertaken.
   - Physician willingness to participate in, or satisfaction with, a telemedicine program may depend upon the physician's form of compensation (e.g., salary vs. fee-for-service). Persuading third-party payers of the cost-effectiveness of telemedicine may require demonstrating that current reimbursement for teleconsultations may diminish downstream adverse health events and utilization of services.
   - Determining appropriate evaluation perspectives should entail consideration of the multiple applications of many telemedicine programs, including for different health care departments and for educational and managerial purposes.

5. Telemedicine comprises an evolving portfolio of technologies and applications. As such, any prospective evaluation must allow for and be prepared to assess the impact (on efficacy or effectiveness, cost, cost-effectiveness, etc.) of applications that may not have been foreseen during the evaluation design.

   - Traditional evaluation methodology stresses prospective measurement of predetermined endpoints. This approach is generally appropriate for mature technologies that have reached steady-state applications. However, this approach does not account for evolving uses of technologies, such as those used in telemedicine, that change their utility in practice.
   - In many instances, the originally intended applications of telemedicine programs (e.g., teleconsultations or telepathology) have been overtaken or accompanied by

other applications (e.g., continuing education or management) that originally were unanticipated or considered to be of minor importance.

- Evaluations of telemedicine programs that maintain focus only on the performance of predetermined applications (some of whose utility may be diminished during pilot or demonstration stages) may ignore alternative applications, thereby yielding findings that underestimate the actual value of the telemedicine program.

6. Plans for evaluation of telemedicine programs should make explicit their assumptions regarding the relationship between the timing of evaluation and the maturity of the telemedicine program, and the evaluations should be designed accordingly.

   - Decision makers should recognize that studies of telemedicine applications that are prototypes or are not integrated into the health care mainstream provide only interim findings about the feasibility of such application, not how well they operate as mature applications.
   - The lack of technological or programmatic maturity has profound impacts on cost evaluations in particular. When start-up costs are high and initial utilization is low, the cost-effectiveness of a telemedicine application may appear to be unacceptably high if it is taken to represent cost-effectiveness at a more mature stage.
   - Further work is required to develop or adapt evaluation designs that take staged approaches commensurate with technological maturity. This may be analogous to stages or phases of evaluation used for other types of health care technology, although the evaluation models themselves that are used for pharmaceuticals and medical devices are largely inappropriate for telemedicine.

7. Given the need to minimize the influence of known as well as unknown sources of bias in comparative studies involving telemedicine, it is desirable to use randomized designs whenever possible. Depending upon the investigation, it may be appropriate to randomize one or more of patients, physicians, or delivery sites. However, randomization is often impractical or impossible for evaluating telemedicine applications.

   - Telemedicine presents challenges to randomized design. Telemedicine interventions are not always discrete or self-contained technologies (as in the instance of many pharmaceutical therapies). The causal effects of teleconsultations may be confounded by differences among the participating physicians, among the participating delivery sites, or among other factors.
   - In addition to randomizing patients to either teleconsultation or standard care, it may be desirable to randomize participating referring physicians to teleconsultations or standard care, and to randomize participating clinics to providing teleconsultations or standard care.
   - For randomization to be successful, the number of patients (or other subjects) to be assigned must be large enough to achieve a high probability of evenly

distributing any prognostic factors. This may not be practical for many telemedicine programs.

- If randomization of patients is not incorporated into an evaluation design, investigators should provide a rationale for this. If it is methodologically desirable but impractical to randomize physicians or delivery sites, investigators should provide a rationale for this, including their assumptions about the similarities among physicians or among the delivery sites.

8. A recurrent weakness in telemedicine evaluations has been the lack of clearly defined control groups. In general, a comparator should be the standard or level of care that would be provided in the absence of the telemedicine intervention.

   - The evaluation design should specify, and justify, the comparator. For teleconsultations, alternatives might include one or more of: no care, inadequate or underspecialized in-person care locally, in-person care remotely (requiring patient travel), delayed in-person care remotely (requiring patient travel), or delayed in-person care locally (requiring physician specialist travel).
   - Rather than methodologically preferable contemporaneous controls, telemedicine evaluations have too often relied on historical controls. Historical controls can be sufficient if there is strong reason to assume that prevailing conditions have not changed over time, and that the relationship between usual care and the outcomes of interest has remained virtually constant.
   - Similarly, designs other than contemporaneous, randomized controls, such as matched controls or pre- and post-test designs, may be more convenient, but have methodological weaknesses. Investigators should provide rationale for using these types of control groups, and address their implications for the validity of study findings.
   - As noted above, the existence of reimbursement for usual care and its absence for telemedicine services may undermine the validity of usual care as a comparator.

9. The time horizon for a telemedicine evaluation should be long enough to capture the stream of relevant health and economic effects that are sufficient to detect any differences in these between the intervention and control groups.

   - Given the novelty of telemedicine applications, it is inevitable that provider institutions, physicians, patients, and other participants will require some time and practical experience to reach a steady-state of operation.
   - Similarly, given the inefficiencies of resource use that arise when installing any new technology or program, the costs of operating a start-up telemedicine operation will not reflect the true, longer-term running costs of the program. The results of evaluations of cost or cost effectiveness of telemedicine operations can be very sensitive to the time span of the evaluation.
   - The time horizon of a comparative evaluation of telemedicine should depend upon the endpoints or outcome measures of interest. For example, the time

horizons required to determine differences in access, patient satisfaction, and the effect of early interventions on longer-term health outcomes and costs will differ.

10. In order to be successful, telemedicine must be integrated as smoothly as possible into existing, routine clinical and administrative functions, including facilities, scheduling and appointments, patient records, coding, and billing.

    - Unlike most new technologies that diffuse smoothly into health care delivery, implementing telemedicine programs often presents departures from standard means of health care delivery, administration, and financing.
    - Telemedicine evaluations should distinguish between any inefficiencies or lack of acceptance that are inherent in telemedicine applications themselves, as opposed to those that derive from an awkward fit between the telemedicine application and the clinical mainstream.
    - In order to assess the integration of a telemedicine program into the health care mainstream, one should consider the program's impact on practice patterns, patient flow, and revenue streams, along with any resulting physician or institutional resistance to the program. These findings can contribute to modifications toward achieving better integration of programs into delivery environments.

11. Independent financial viability of a telemedicine program will increase its prospects for integration into the health care mainstream and long-term success.

    - The ability of most programs to achieve financial viability will depend on their ability to secure appropriate coverage and adequate third-party payment for their services.
    - The single most important evaluation criterion for any telemedicine program may be the extent to which it achieves independent financial viability.
    - Proposed telemedicine programs should include multi-year business plans that describe how the program will progress toward financial viability as outside funding from grants or other temporary sources diminishes. Such plans should be considered in any grants review process, and financial status with respect to these plans should be regularly monitored.

APPENDIX A: HHS TELEMEDICINE EVALUATIONS

Within HHS, the bulk of telemedicine initiatives support and funding is provided by four agencies: the Agency for Healthcare Research and Quality (AHRQ), Health Care Financing Administration (HCFA), Health Resources and Services Administration (HRSA), and National Institutes of Health (NIH). Despite funding from multiple agencies, the utilization of teleconsultations remains low. The low level of utilization places limitations on conducting effective evaluations.

A. Agency for Healthcare Research and Quality

AHRQ's Center for Information Technology conducts and supports studies of health information systems, computerized patient record systems, and medical decision analysis, including data standards, automated medical records, and decision support systems. As a participant in the national High Performance Computing and Communications (HPCC) Program, AHRQ supports initiatives to promote increased speed and capacity of computers and electronic networks, as well as to make the transmission of data more secure.

The Evidence-based Practice Center program, coordinated by AHRQ's Center for Practice and Technology Assessment, is sponsoring a project by Oregon Health Sciences University titled "Medical Informatics and Telemedicine Coverage Under the Medicare Program."

B. Health Care Financing Administration

In October of 1996, HCFA initiated a demonstration project to allow reimbursement of teleconsulting services by Medicare beneficiaries at 57 Medicare-certified facilities. The objectives of this project are to assess the feasibility, acceptability, cost, quality, and access to services that could be made available through Medicare reimbursement for teleconsultation.

Prior to this demonstration project, HCFA contracted with the University of Colorado Health Sciences Center (UCHSC) to perform a preliminary study of telemedicine, conducted during 1993-1995. HCFA evaluated the program in terms of: 1) utilization, 2) access to care, 3) clinical efficacy and cost-effectiveness, 4) quality (both process and outcome), and 5) reasonableness of charges (considered to be a component of patient satisfaction). This effort involved a literature review, development of a conceptual framework for the analysis of studies examining effectiveness, selected case studies, a review of coverage policies of private third-party payers, and examination of utilization review and quality assurance/improvement models currently in operation as part of existing telemedicine systems.

The study concluded that few telemedicine services are actually being provided. Regarding clinical effectiveness of telemedicine, most scientific literature at the time pertained to teleradiology and telepathology. The report concluded that "very few papers had been

published concerning other applications of telemedicine, and the majority of articles that could be found were descriptions of applications rather than empirical research.

At the time of the preliminary study of telemedicine, there were no studies of cost effectiveness for any application of telemedicine, though some estimates were being made of "potential cost reductions" due to the availability of teleconferencing for remote consultations and continuing education, in a report published by Arthur D. Little, Inc., in 1992. The report concluded that "there were no data [to] test the validity of the model (outlined in the report) and thereby illuminate the matter of telemedicine's possible cost effectiveness."

A new HCFA demonstration project on the use of telemedicine for management of diabetes, titled "Informatics, Telemedicine, and Education Demonstration," is using specially modified home computers as "home telemedicine units" (HTU) linked to a clinical information system (CIS) maintained by Columbia Presbyterian Medical Center in New York City. The HTUs in patients' homes allow video conferencing, access to health information and access to medical data. Computerized devices read blood sugar levels, check blood pressure, take pictures of skin and feet for signs of infection, and screen for other factors that affect the management of diabetes. These data are fed electronically to the data system at Columbia. The CIS provides storage of clinical data for use in the development and application of patient care guidelines and clinical standards. Full-time nurse case-managers monitor the data and intervene if the data from a patient vary from guidelines. Patients receive feedback, including clinical data such as blood glucose levels, care reminders and suggestions on how to maintain good health.

The demonstration project is being conducted as an RCT. Half of the participants are receiving the intervention, consisting of an HTU and electronic services within a case-manager environment, and half continue to receive usual care for their diabetes. The demonstration consists of two components: an urban component in northern Manhattan, and a rural component in upstate New York.

The evaluation of this demonstration will consist of the following components:

- Physician profile. This is an analysis of differences between physicians participating in the demonstration and those not participating. Possible comparison groups include physicians approached for participation but who declined, and non-participating primary care physicians in the same geographic area. HCFA hopes to gain insight into whether specific characteristics of the physician might affect the probability of use of HTUs by the physician, which may have implications for the ultimate diffusion of the technology.
- Patterns of use analysis. This analysis will look at HTU utilization patterns over time by project participants, including separate descriptions for rural vs. urban, male vs. female, and younger vs. older participants.
- Analysis of access. This will compare the participants randomized to receive the intervention to the control group randomized to receive their usual diabetes care, using both an "intent to treat" and a "completer" analysis.

- Analysis of quality of life. Standard quality of life questionnaires will be used to compare baseline measures to follow-up measures for both the intervention and the control groups.
- Cost analysis. The primary analysis will be from the perspective of Medicare. The total Medicare expenditures per person for the control group will be compared to the total Medicare expenditures plus the intervention costs for the intervention group. A secondary analysis will consider non-Medicare covered expenses.

Other HCFA telemedicine evaluations include demonstrations in Georgia and West Virginia.

C. Health Resources and Services Administration

The HRSA Office for the Advancement of Telehealth (OAT) promotes the use of modern telecommunications and information technologies to bring state-of-the-art health care and health information to every community, particularly medically underserved and isolated regions. OAT has worked to expand HRSA's evaluation activities and coordinate the development of evaluation tools with other agencies supporting telemedicine. Funding for telemedicine initiatives occurs through two grant programs: 1) the Rural Telemedicine Grant Program and 2) the Rural Health Services Outreach Program.

The goals of the Rural Telemedicine Grant Program are to improve access to quality health services for rural residents and reduce the isolation of rural practitioners through the use of telemedicine technologies. The program's objectives are to:

- demonstrate how telemedicine can be used as a tool in developing integrated systems of health care, thereby improving access to health services for rural residents; and
- evaluate the feasibility, costs, appropriateness, and acceptability of rural telemedicine services and technologies.

Eleven projects were funded in FY94 for a three-year project period, and 18 projects were funded in FY98 for a three-year period. The HRSA Office of Rural Health Policy (ORHP) originally administered this program. However, these activities were charged to the OAT when it was organized in August 1998.

The Rural Health Services Outreach Grant Program funds projects to support the direct delivery of health care and related services, to expand existing services, and to enhance health service delivery through education, promotion, and prevention programs. Outreach grants require the establishment of a network that is composed of three or more health care organizations, or a combination of health care, social service, and other organizations that support the delivery of health services. The grant program is administered by ORHP, however those projects concerning telemedicine are managed by OAT officers. In 1997, ORHP published *Exploratory Evaluation of Rural Applications*

of Telemedicine, which includes a nationwide survey of all rural hospitals, a follow-up survey of those indicating telemedicine capability, and site visits at four programs.

D. National Library of Medicine

The National Library of Medicine (NLM) offers two programs to encourage the use and development of telecommunications infrastructure. The High Performance Computing and Communications Program funds Internet access for health professions engaged in education, research, clinical care, and administration. The second program supports projects that develop and demonstrate the use of the National Information Infrastructure in health care, clinical research, and public health.

In October 1996, NLM awarded 19 multi-year telemedicine projects intended to serve as models for:

- evaluating the impact of telemedicine on cost, quality, and access to health care;
- assessing approaches to ensuring confidentiality of electronically transmitted health data; and
- testing emerging health data standards.

Two additional projects were awarded in September 1997. As appropriate, projects are to review and apply recommendations from the 1996 IOM report and the National Research Council publication, *For the Record: Protecting Electronic Health Information*.

E. Indian Health Service

The Indian Health Service (IHS) is funding 40-50 small telemedicine programs across the country. To date, the IHS has performed few formal evaluations. It performed a cost-benefit analysis on teleradiology, which included exclusively store-and-forward technology. The IHS also evaluated patient and clinician perceptions of a telepsychiatry program for Sioux children with Attention Deficit Hyperactivity Disorder (ADHD) in South Dakota. The focus of the evaluation was to compare measures of satisfaction between in-person and two-way interactive clinical encounters. Interactions were systematically rated by the evaluating physician, the child patient, and the child's parent or guardian. Provider reaction to and judgment about the interaction was evaluated by an 8-item questionnaire (containing a 5-point Likert scale) modified from Simonian et al., 1993. Child satisfaction with the interaction was measured using 7 of the 8 items of the Metro Assessment of Child Satisfaction (Simonian et al. 1993). Parent satisfaction with their child's health care visit was measured using the Pediatric Satisfaction Questionnaire (Finney et al., 1990), a modification of the Medical Interview Satisfaction Scale (Wolf et al., 1978). A total of 48 clinical interactions were included in the study, 20 of which were initial consultations. The results were as follows.

- The clinician was significantly less satisfied with the initial encounter in the audiovisual (AV) mode. This difference was not found for follow-ups, perhaps because the clinician felt the therapeutic relationship with the family was already established.
- The Child Satisfaction evaluation yielded no significant differences between onsite and AV interactions for either initial consultations or follow-ups.
- The parent questionnaire yielded no significant differences between the AV and the onsite conditions for either initial evaluations or follow-ups.

APPENDIX B: RELEVANT LITERATURE REVIEW FINDINGS

A review of the current literature was conducted to update, as appropriate, the IOM evaluation framework. The goal of the literature review is to update the IOM study with any relevant information having arisen within the last three years, and second, to supplement any gaps that may have been identified.

A. Literature search strategy

The literature review was conducted by a direct search of the MEDLINE database (citations of peer-reviewed journal literature), the primary bibliographic database maintained by the National Library of Medicine.

To identify pertinent articles, we applied selected Medical Subject Headings (MeSH) terms as shown in the table below. We excluded articles written prior to 1996 to avoid overlap with those articles cited in the IOM report. The search emphasized clinical encounters and consultations in telemedicine, as opposed to such areas as teleradiology, telepathology, and reviews of specific technologies or equipment.

The citations provided in the reference section of this document represent the selected list of pertinent articles chosen from the initial larger set of articles resulting from the literature search.

Literature Search Methods

Database Type	Database Name	Years	MeSH
NLM**	MEDLINE	1996-Present	Telemed* AND Cost-benefit Analysis AND Evaluat* AND Framework AND Evaluation Studies Evaluation AND Framework
	HealthSTAR	1996-Present	Telemedicine AND Evaluation AND Cost-benefit analysis AND Evaluat* AND Framework AND Evaluation studies Evaluation AND Framework

**All NLM searches were bounded by the subheadings: economics, education, legislation and jurisprudence, organization and administration, standards and statistics and numerical data.

B. Findings

The majority of telemedicine programs are in the earliest stages of usage. As asserted by Bashshur (1998), the absence of mechanisms for reimbursement and related funding for telemedicine programs will continue to constrain the maturation of such programs, in turn preventing appropriate evaluation.

The findings from the recent literature are broken into two broad categories:

1. Iterated points of the IOM evaluation framework.

2. Supplementary points of the IOM evaluation framework.

In general, little new information is offered by the recent literature. Most often, articles restate or echo the findings of the IOM framework. Given that the IOM framework is a comprehensive study, includes an extensive review of the literature, and was completed in 1995, this is not unexpected.

Several broad issues arose from the literature review that could add to the IOM framework, and which are incorporated into the present report. First, evaluations of telemedicine should take into account the maturity of the program being evaluated (e.g., pilot versus a "steady state" programs). Second, integrated into any evaluation should be a more substantial and specific cost-effectiveness analysis to adequately take into account the unique nature of telemedicine applications. Third, an appropriately rigorous methodology should be applied to the evaluative process to ensure that the data gathered are useful to the health care community and those that it serves, providing evidence-based findings that can be used to support coverage decisions as appropriate. Finally, a staged approach to evaluation, similar to that used for pharmaceuticals, is suggested. These four points are addressed in following sections.

Among the points arising in the recent literature that reinforce the IOM framework are: 1) the need for a sensitivity analysis to take into account potential changes in the applications, conditions of use or cost of a technology, and how these might affect outcomes or costs of interest, 2) the necessity of developing appropriate outcomes, and 3) the unique challenges to developing an evaluation of a telemedicine program. These points are addressed below.

1. Iterated Points of the IOM Evaluation Framework

Much of the literature on evaluation of telemedicine written since the IOM report has concurred with or further elucidated the information provided in the 1995 framework. Broad categories raised in the recent literature that enhance the IOM framework include:

- development of appropriate outcomes,
- the necessity of a sensitivity analysis, and
- challenges inherent in setting up an evaluation of a telemedicine program.

These points offer initial guidance in further developing a conceptual framework to supplement the IOM evaluative framework.

a) Development of appropriate outcomes

The issue of choosing the appropriate outcomes is addressed in the IOM framework, though the issue may be approached from several new perspectives not directly addressed by that report. The recent literature echoes outcomes noted in the IOM framework, such as clinical outcomes, health outcomes, patient and provider satisfaction, long-term versus intermediate outcomes, and others.

A point raised in the recent literature addresses more specifically the level at which the outcomes are assessed. Several articles call for a move away from assessment of individual technologies toward assessment of how a telemedicine program would work at the health care system level, or societal level. Bashshur (1998) argues that "a number of these [technology specific] issues are no longer of concern .... the question of clinical safety should be put to rest." Siwicki (1997) concurs, arguing that the technology behind the medicine has been adequately demonstrated. What is needed now is "a vast number of legitimate, in-depth studies that spell out that telemedicine delivers quality health care that is cost-effective."

Taylor (1998) addresses at length the issue of appropriate outcomes. Levels of assessment may include an improvement in the well-being of a population, a reduction in the costs of providing a service, an increase in the knowledge of general practitioners, an improvement in the quality of information received, or increased patient compliance. Taylor's general argument is that evaluations of specific technologies and pilots are anecdotal and do not greatly increase the level of knowledge with regard to system evaluation. More useful, Taylor argues, is an assessment of the effects of telemedicine systems, rather than the more narrowly focused assessments of individual technologies.

b) The necessity of sensitivity analysis

As mentioned in the IOM discussion and restated several times in the literature, telemedicine is a dynamic field. Technology is constantly improving, and new, sometimes unintended applications are continually arising. However, as in the evaluation of any technology, a static evaluation may be obsolete by the time it is completed. Given this situation, it is essential to integrate into any evaluation a sensitivity analysis that would attempt to account for such potential changes in the applications, conditions, use, or costs of telemedicine technology, and how these might affect outcomes or costs of interest, as well as other unintended uses and consequences (Bashshur 1998, Crowe 1998, Sisk and Sanders 1998). Further, an ongoing evaluation built into a telemedicine program may be most effective in assessing the true success or lack of success of a maturing program. Such a sensitivity analysis reflects the essential principles for an evaluation framework.

c) Challenges to developing an evaluation of a telemedicine program

Mintzer (1997) proposes a number of lessons learned from analysis of a program involving thirteen telemedicine networks funded in 1994 by the HRSA Office of Rural Health Policy's Rural Telemedicine Program. While many of these points do not have direct implications for evaluation of telemedicine, they collectively represent the necessary backdrop for conducting an effective evaluation. Knowledge of these challenges integrated into an evaluative framework would allow those conducting evaluations to account for, and possibly avoid, similar pitfalls. These include the following, although some of these points were also made in the IOM framework.

- Expect to expend considerable effort in training and convincing practitioners to try telemedicine.
- Utilization is as likely to be initiated by specialists as by rural practitioners.
- Look for non-conventional clinical applications.
- Conduct a thorough needs assessment and have regularly scheduled telemedicine clinics.
- Transmission costs are high and need to be factored into long-term plans for sustaining a telemedicine network.
- Confidentiality and privacy may be bigger concerns in theory than in practice.

This same article contains a comparison of start-up challenges versus operational challenges. While start-up issues centered on delays in obtaining equipment, phone connections, and properly working hardware and software, operational issues included problems such as equipment residing in inaccessible areas (e.g., far away from emergency room staff), or off-hours inaccessibility (i.e., equipment located in a room that is normally locked during the night shift).

Other start-up issues cited by Mintzer included provider reluctance to use telemedicine and lack of provider comfort with equipment. An article written as an interim report of a telehome health project evaluation cited as additional barriers the concerns among staff that the technology would replace the nurses, and that the physical distance between patient and provider would threaten their professional relationship with patients (Johnston 1997). To alleviate some of these concerns, those conducting the study implemented a communications plan to keep staff involved in project development, and attempt to preempt misperceptions regarding the program.

As mentioned, these lessons learned should be viewed as precursors to carrying out an evaluation of telemedicine that will both be implemented and executed in an efficient fashion and yield worthwhile results.

2. Supplementary Points to the IOM Evaluation Framework

Several main issues were raised in the recent literature that supplement the IOM framework in important ways. These points are outlined below.

a) Pilot Versus "Steady State" Evaluation

A telemedicine program in the very early stages of maturity will have very different costs and results than a program that has matured to a steady state. Evaluation of a telemedicine program past the initial pilot phase and into the steady state phase of implementation is necessary to assess fully and accurately the viability of such a program. As one measure, the costs associated with a pilot program are different, and often greater, than those associated with a more mature program. In some cases, this is due to the lack of economies of scale, or early cost burdens associated with extensive training, staff familiarization, and equipment set-up. As Crowe (1998) states regarding communication costs, "The collection of data on communication costs, often the major part of system costs, in a pilot telemedicine project may not necessarily reflect the costs likely to be incurred in a mature telemedicine system." Mintzer (1997) cites other challenges of evaluating a pilot telemedicine program versus a mature program. These include:

- delays in obtaining equipment;
- delays in getting telephone connections made;
- equipment and software technical difficulties;
- training of new or inexperienced staff;
- provider reluctance to use telemedicine; and
- developing comfort among staff who will be using the telemedicine.

The point at which a program matures into a steady state varies for each program. Further, it may be difficult to place a time limit on the pilot phase, as it has been shown that even two years may not be enough for a program to reach a steady state (Mintzer 1997). As a supplement to the steps for evaluation planning, this aspect of telemedicine evaluation should be taken into account to ensure reliable outcomes.

b) Cost-effectiveness Evaluation

Because cost structures and expenditures change over time, the issue of how to best carry out a cost-effectiveness evaluation of telemedicine is closely related to evaluation of programs at the pilot versus mature stage. Crowe provides extensive consideration of this topic in a cost-effectiveness analysis of telemedicine published in 1998. Specifically addressing the issue of evaluation of pilot versus mature programs, Crowe states, "There is a problem that a health-related telemedicine service may be evaluated in isolation as a pilot project, but, as a mature service, may be integrated with other services such as tele-

education and telebanking for a rural community." Exhibit 1 provides a breakdown of cost types, as according to the article by Crowe (1998).

Exhibit 1: Cost Types for Telemedicine Evaluation

Cost Type	Cost Elements
Project establishment costs	▪ Preparation of submissions for funding approval ▪ Selection processes to decide which projects are to proceed ▪ Recruitment of staff ▪ Feasibility studies ▪ Preparation of tenders for equipment ▪ Selection and installation of equipment ▪ Revision of organizational arrangements ▪ consultation with staff ▪ Training of staff in new systems and procedures and in use of equipment ▪ Establishment of an evaluation framework involving procedures for the collection and analysis of data for both the status quo and the new initiative and often involve computer staff
Equipment costs	▪ Computers and associated hardware (modems and video boards) ▪ Videoconferencing and document display software
Maintenance costs	▪ Suggested that maintenance charges be calculated at 10-15% per year of the capital cost of the equipment ▪ Travel times and costs ▪ Downtime loss
Communication costs	▪ Because of economics of scale, communications costs should decrease substantially in a mature program
Staffing costs	▪ A successful telemedicine program (in a steady state) should make demands on staff time less, and should therefore cost less ▪ Suggested that an hourly rate is used for staff specialists and an appropriate fee for visiting consultants

Source: Crowe (1998)

Sisk and Sanders (1998) also address the issue of cost-effectiveness analysis of telemedicine programs, citing the need to specify the full range of actual alternatives and the unique barriers to cost-effectiveness analysis: "Multiple uses of a telemedicine system may have joint costs that are difficult to apportion to one service, the existence of a system may lead to expanded indications of use, and technological change may make an evaluation outdated." Sisk and Sanders outline some of the potential cost implications of a telemedicine program, briefly summarized here as follows:

- any savings and expenditures incurred in treating a patient earlier in the course of the condition;
- changes in the productivity of health professionals;
- patient time saved; and
- changes in transportation costs.

The above costs and benefits of a telemedicine program accrue both to society, in general, and to the party responsible for payment of the relevant health care services, in particular. However, the costs and benefits to payers of telemedicine are particularly dynamic, as changing times in the health insurance marketplace illuminate opportunities for savings among payers. Historically, a lack of insurance coverage for telemedicine services has been an impediment to adoption with fee-for-service payment. Under capitation payment and fixed budgets, however, providers have financial incentives to use the most efficient method to deliver services. With the expansion of integrated health care delivery systems and such capitated payment arrangements, plans and providers are likely to weigh a broader range of costs against potential benefits in deciding about investments in telemedicine. While some of these points are made within the IOM framework, the articles by Sisk and Crowe add value to the framework by considering these points from an economic analysis standpoint.

c) Rigorous Methodology

In the context of higher standards for evidence-based health care, relatively few studies have been conducted that apply a rigorous methodology to the study of telemedicine. This is a necessary first step in developing a framework for evaluating telemedicine programs.

Problems mentioned in the literature include small sample sizes, flawed study design, and inaccurate and imprecise measurement (Bashshur 1998). Suggestions to improve current methodology include pooling of data across programs, using randomized controlled trials (RCTs), and case control studies with relevant meta-analyses (Yellowlees 1998). Nitzen et al. (1997) attempted to ensure methodological rigor by establishing a gold standard, requiring that each patient be examined by multiple physicians, conducting the in-person visits and teleconsults within a very short time span, and conducting matched-pair analyses on all study data. Finally, the researchers calculated kappa coefficients, both for comparison of their findings with other studies and as a check on their success in reducing bias in the study design (Nitzen et al. 1997).

In the first of a two-part series, Taylor (1998) proposes a comprehensive set-up of telemedicine evaluation. The article broadly outlines an evaluation of telemedicine consisting of three phases:

1. identification of the technical specification of equipment required for the particular telemedicine application;

2. tests to ensure that the evaluation is being conducted in the appropriate settings; and

3. establishment of a set of standards and guidelines to ensure that the telemedicine system is used to the best advantage.

By considering a specific study (which is generalized here), the evaluative process specifics are broken into four elements, each of which has key issues associated with it, as summarized in Exhibit 2.

Exhibit 2: Elements and Key Issues of a Sample Evaluation

Element	Key Issues
Select a set of cases to provide a suitable basis for answering the questions of interest.	▪ An adequate number of cases must be used. ▪ Awkward or difficult cases must be included. ▪ The range of cases should reflect the specific questions addressed in the study.
Interpret cases both 1) using telemedicine (the study condition) and 2) not using telemedicine (the control condition).	▪ The roles of the study and control groups must be clearly distinguished. ▪ The situations in the study and control groups should be comparable. ▪ Any possibility of confounding or transfer between the conditions should be minimized. ▪ Subjects should be given clear instructions and, if appropriate, training in the use of the new technology.
Interpret cases to develop a "gold standard."	▪ If it is not possible to establish a gold standard, then a design, which does not require a gold standard, may be better. ▪ Any effect whereby determining the gold standard systematically excludes cases should be minimized. ▪ If a gold standard is required, it should be established independently of the control and the study conditions.
Compare the conclusions of interpreters in the study and the control conditions to the gold standard and indices of diagnostic accuracy.	▪ The statistical analysis used should be appropriate to the question being answered. ▪ The conclusions drawn should be clearly warranted by the analysis. ▪ Statistics should not be used unnecessarily.

Source: Taylor (1998)

d) Staged Approach to Evaluation

As mentioned in the IOM framework and the recent literature, a sensitivity analysis is an essential aspect of any telemedicine evaluation. Taking this need into account, one possible approach to evaluating telemedicine programs at the technology level may be a staged approach similar to that currently in practice in the pharmaceutical industry (i.e., preclinical testing, Phases I, II, III, and IV). This is presented by DeChant et al. (1996) in an article titled "Health systems evaluation of telemedicine: a staged approach," in which comparisons are made to the method by which pharmaceuticals are developed.

DeChant et al. propose an analogous set of stages for telemedicine, and that "in each stage of the analysis, the evaluation is tailored to the technology's state of development." These results would then be used to improve the technology before dissemination occurs. The method entails addressing to varying degrees the three primary concerns outlined in the IOM framework: quality, access, and cost. DeChant et al. argue that not all three of these components would play a role in each stage of the process, but should be considered only as appropriate. The intent is not only to adopt the evaluation to the maturity of the technology, as appropriate, but to integrate into this methodology aspects from the evaluation of pharmaceuticals in order to better "capture telemedicine's potential to produce system-wide change." Such a staged approach may provide a

method for taking into account the maturation of telemedicine technologies over time, and the resulting effects on important outcomes.

e) Other Issues

Other issues that were included in the recent literature, but not expanded upon extensively include the following.

- Sisk and Sanders (1998) raised the issue of economic discounting when conducting evaluations. This "reflects the fact that people place a higher value on events in the present than in the future, and that funds (or effort) invested in the present can reap interest over time." While this is a somewhat less critical point in conducting an evaluation and is not specific to telemedicine, evaluators of telemedicine should be aware of and take into account this added factor.
- Bashshur (1998) and others point out that a "triage system" may be necessary to avoid potential over-utilization of telemedicine technology. This should include establishment of specific telemedicine-related protocols to reduce arbitrary or frivolous use of the technology. While this is primarily a program development issue, an implication for evaluation is the determination of appropriateness and necessity of technology utilization. That is, one aspect of an evaluation should be whether telemedicine is being used in an appropriate fashion, and when necessary. Appropriateness evaluation may be done retrospectively (e.g., through medical record review) or prospectively, as a method to supplement clinical decision making.
- Farand et al. (1997) conducted a study designed to examine, in part, the "clinical problem-solving processes in the context of a telemedical consultation, in order to verify to what extent the technological environment preserves the characteristics of medical reasoning that are known to occur in more traditional clinical settings." This represents a somewhat unique approach to evaluating telemedicine, focusing on the actual interaction and subsequent consequences of the interaction rather than the effectiveness or utility of a technology. They concluded that an evaluation should take into account the interacting problem-solving modalities that may be encountered in the context of telemedicine consultation, that is, the reasoning that a health professional may use to make a clinical determination, and the changes, if any, in the interaction between the physician and the patient.

C. Conclusions

Elements drawn from the recent literature, including those concurring with the IOM framework and those that may augment the IOM framework, provide an initial understanding of the current state of knowledge regarding evaluation of telemedicine. Based on the recent literature, the IOM framework may be augmented in the following ways. (These are incorporated into the present report.)

- Take account of the maturity of the program (pilot vs. steady state) in evaluating any telemedicine program.

- Integrate into any telemedicine evaluation a specific cost-effectiveness analysis.
- Apply an appropriately rigorous methodology to the evaluative process to ensure that the data gathered are useful to the health care community and those that it serves, and to provide evidence-based findings that can be used to support coverage decisions as appropriate.
- Take a staged approach to evaluation of telemedicine programs to account for the maturation of telemedicine technologies over time and the resulting effects on outcomes.

There has not been a substantial amount of new information on telemedicine evaluation since publication of the IOM report. This outcome was not unexpected, however, given that the IOM framework is rather extensive, includes an in-depth review of the pertinent literature, and was developed relatively recently. In designing future evaluations of telemedicine activities, evaluators should consider carefully the IOM framework and the supplemental evaluation aspects identified in the present report.

APPENDIX C: SITE-VISIT INTERVIEW PROTOCOL

(This protocol was used as a general guide in conducting site visits to ensure that discussions covered a broad range of issues.)

On what area of evaluation do you think evaluators of telemedicine should be focusing? (Please rank in order of most to least important.)

Quality of care and health outcomes

- What types of health outcomes measures are appropriate for telemedicine evaluation?
- What questions can evaluators ask to determine the effects of telemedicine on health (clinical) outcomes compared to the alternative(s)?

Access to care

- What types of access measures are appropriate for telemedicine evaluation?
- How can evaluators determine the effect of telemedicine on the use of services (utilization) or the level or appropriateness of care compared to the alternative(s)?
- How can evaluators determine if and how telemedicine affects the timeliness of care or the burden of obtaining care compared to the alternative(s)?
- What questions can evaluators ask to determine the barriers to utilization of telemedicine?

Costs and Cost-effectiveness

- What questions can evaluators ask to determine the costs of a telemedicine application for participating health care providers or health plans compared to the alternative(s)?
- What questions can evaluators ask to determine the costs of telemedicine for patients and families compared to the alternative(s)?
- What questions can evaluators ask to determine the costs for society overall compared to the alternative(s)?
- What questions can evaluators ask to determine how the costs of telemedicine relate to its benefits, compared to the alternative(s)?
- What types of cost evaluation measures are appropriate for telemedicine evaluation (e.g., cost benefit, cost-effectiveness, and cost minimization)?

Patient perceptions

- How can evaluators determine if patients are satisfied with the telemedicine service compared to the alternative(s)?

Clinician perceptions

- How can evaluators determine if attending and consulting clinicians are satisfied with the telemedicine application compared to the alternative(s)?
- Below is a list of questions that we have collected from the IOM framework. Which of these questions is more important? Less important? Irrelevant?

  - What is the perspective of the evaluation? Society? Network? Site? Physician?
  - What is the (1) setting of care (e.g., hospital or physician's office); (2) condition being treated (e.g., dermatology or psychiatry); (3) technology used (e.g., store-and-forward)?
  - What is the alternative to which telemedicine is being compared? That is, what would happen in the absence of telemedicine, or similarly, what is the control group in the evaluation?
  - How is success measured?
  - How is the telemedicine implemented? That is, is it well-integrated into patient care? Does it effectively meet population needs?
  - How is the program dealing with issues such as confidentiality, privacy, equipment and protocols, if at all?

- At what point in a program's development (i.e., its life cycle) should an evaluation begin?
- What types of research design/evaluation methods do you feel are most appropriate for telemedicine?
  - large randomized controlled trial (RCT);
  - small RCT;
  - nonrandomized trial with contemporaneous controls;
  - nonrandomized trial with historical controls;
  - cohort study;
  - case-control study;
  - cross-sectional study;
  - surveillance (e.g., using registers or surveys);
  - series of consecutive cases;
  - single case report.

- Based on your experience/knowledge of telemedicine evaluation, do you believe that sufficient scientific rigor has been applied to current and past evaluation of telemedicine?

- Please comment on how telemedicine evaluation is best conducted with respect to the following methodological factors.
  - perspective of analysis (e.g., society, payer, provider patient)
  - accounting of direct costs (medical and non-medical)
  - accounting of indirect costs (e.g., loss of productivity, patient time)
  - use of charges or prices versus actual costs
  - choice of time horizon for analysis (i.e., short-term versus long-term)
  - use of sensitivity analysis

- To what degree do you feel reimbursement drives and/or directs use of telemedicine services, and subsequent evaluations of such programs?
- What do you see as the emerging issues (in terms of policy and evaluation) within the field of telemedicine?

APPENDIX D: TELEMEDICINE EXPERTS INTERVIEWED

Thelma Armstrong – Eastern Montana Telemedicine Network
Susan Capalbo – Associate Professor, Dept. of Agricultural Economics, Montana State University, Trace Research Center
Catherine Finley – Health and Human Services Policy Analyst, Southern Governor's Association
Bill Grigsby, Ph.D. – Senior Research Associate, Telemedicine Research Center
Susan Gustke – Executive Director, Eastern Area Health Education Center
Michael Hillman, M.D. – Marshfield Clinic Telehealth Network
Douglas Perednia, M.D. – Director, Advanced Telemedicine Research, Telemedicine Research Center, Oregon Health Sciences University
Curtis Rooney, J.D. – American Hospital Association
Jay Sanders, M.D. – Global Telemedicine Group
Bill Siwicki – Senior Editor, *Health Data Management*, Faulkner & Gray
Dennis Vidmar, M.D. – Captain MC, U.S. Navy, Department of Dermatology; Walter Reed Hospital, Department of Dermatology
Margaret VanAmringe – Vice President, External Relations, JCAHO
Robert Waters, MPA, J.D. – Partner, Arent Fox; Center for Telemedicine Law
William Weissert, Ph.D. – University of Michigan, School of Public Health
Pamela Whitten, Ph.D. – Michigan State University, Telemedicine Program

REFERENCES

Ackerman, Michael. Personal interview. 3 May 2000.

Association of Telehealth Service Providers. October 2000. Telemedicine FAQ. Available at: http://www.atsp.org.

Bashshur RL. Rethinking the evaluation and priorities in telemedicine [editorial]. Telemed J. 1998; 4(1):1-4.

Crowe BL. Cost-effectiveness analysis of telemedicine. J Telemed Telecare. 1998; 4 Suppl. 1:14-7.

DeChant HK, Tohme WG, Mun SK, Hayes WS, Schulman KA. Health systems evaluation of telemedicine: a staged approach. Telemed J. 1996; 2(4):303-12.

Farand L, Lafrance JP, Arocha JF. Collaborative problem-solving in telemedicine and evidence interpretation in a complex clinical case. Int J Med Inf. 1998; 51(2-3): 153-67.

Finney JW, Brophy CJ, Friman PC, et al. Promoting parent-provider interaction during young children's health-supervision visits. J Appl Behav Analysis. 1990; 23:207-213.

Goodman C. The moving target problem and other lessons from percutaneous transluminal coronary angioplasty. In: Szczepura A, Kankaanpaa J, eds. Assessment of Health Care Technologies: Case Studies, Key Concepts and Strategic Issues. New York, NY: John Wiley & Sons; 1996:29-65.

Grigsby J, Schlenker RE, Kaehny MM, et al. Analysis of expansion of access to care through use of telemedicine. Report 4: Study summary and recommendations for further research. Denver: Center for Health Policy Research. 1994.

Grigsby J, Kaehny MM, Sandberg EJ, Schlenker RE, Shaughnessy PW. Effects and effectiveness of telemedicine. Health Care Financ Review 1995;17(1):115-31.

Grigsby J, Schlenker RE, Kaehny MM, Shaughnessy PW, Sandberg EJ. Analytic framework for evaluation of telemedicine. Telemed J. 1995 (1); 1: 31-39.

Gustke SS, Balch DC, West VL, Rogers LO. Patient satisfaction with telemedicine. Telemed J. 2000; 6(1):5-13.

Institute for the Future. Health & Health Care 2010. The Forecast, the Challlenge. Princeton, NJ: Jossey-Bass, 2000.

Institute of Medicine. Telemedicine: A Guide to Assessing Telecommunications in Health Care. Field MJ, ed. Washington D.C.: National Academy Press, 1996.

Kennedy M. Telemedicine: promising tool or Pandora's box? WMJ. 1998; 97(1): 26-8.

The Lewin Group. Outlook for Medical Technology Innovation. Report 3: Technology Assessment by Public and Private Payers. Washington, D.C: AdvaMed, 2000.

Mintzer CL, Wasem CJ, Puskin DS. Program activity in the second year of the Rural Telemedicine Grant Program Part 2. Telemed Today. 1997; 5(6):32-3.

Nitzkin JL, Zhu N, Marier RL. Reliability of telemedicine examination. Telemed J. 1997; 3(2):141-57.

Ohinmaa A, Hailey D, Roine R. The Assessment of Telemedicine: General Principles and a Systematic Review. INAHTA Project on Telemedicine. Finnish Office for Health Care Technology Assessment and Alberta Heritage Foundation for Medical Research, 1999.

Rendina MC. A net present value analysis of neonatal telecardioldogy. Telemed Today 2000;April:23-25.

Sisk JE, Sanders JH. A proposed framework for economic evaluation of telemedicine. Telemed J. 1998; 4(1):31-7.

Simonian SJ, Tarnowski KJ, Park A, Bekeny P. Child, parent, and physician perceived satisfaction with pediatric outpatient visits. Devel Behav Ped. 1993; 14:6-12.

Siwicki B. Measuring the benefits of telemedicine. Health Data Manag. 1997; 5(11): 79-80, 82, 84.

Taylor P. A survey of research in telemedicine 1: Telemedicine systems. J Telemed Telecare. 1998; 4(1): 1-17.

Taylor P. A survey of research in telemedicine 2: Telemedicine services. J Telemed Telecare. 1998; 4(2): 63-71.

Telemedicine 2000 Buyer's Guide & Directory. Telemed Today 2000;(Special Issue).

Whitten P, Kingsley C, Grigsby J. Results of a meta-analysis of cost-benefit research is this a question worth asking? J Telemed Telecare 2000;6(Suppl 1):S4-6.

Wolf MH, Putnam SM, James AS, Stiles WB. The medical interview satisfaction scale: Development of a scale to measure patient perceptions of physical behavior. J Behav Med. 1978; 1:391-401.

Yellowlees P. Practical evaluation of telemedicine systems in the real world. J Telemed Telecare. 1998; 4 Suppl 1:56-57.