Literature Review

Navy Telemedicine: A Review of Current and Emerging Research Models

CHERYL REED, Ph.D.,1 RALPH BURR, M.A.,2 and TED MELCER, Ph.D.2

ABSTRACT

With extended deployments of military telemedicine to remote, austere, and high-risk settings, there is a danger that implementation will outrun the research designed to assess it. A review of existing research evaluates current assessment models and indicates gaps around which future research may be designed. A review of models for evaluating telemedicine was conducted in September, 2001. Seven areas of assessment recurred that could be subsumed under the framework of "interoperability." A follow-up search was conducted during September, 2003, to ascertain the degree to which assessments of interoperability were being advocated. Although extensive deployments of telemedicine systems in high-risk settings make it imperative to know that human, organizational, and technical components are integrating productively (i.e., interoperaing), current research does not evaluate telemedicine as a system. The structure and functioning of telemedicine systems should be assessed empirically and systemically to avoid subjective or limited assessments.

INTRODUCTION

Currently, troops and the telemedicine systems that support their care are increasingly deployed to remote, austere, and high-risk environments. Clearly, all elements of a given system—human, organizational, and technical—must work together to enable the flow of information and to ensure quality medical care on demand. Hence, models for evaluating telemedicine must be adaptable to the evolving operational and funding climate so that implementation does not outrun current models for assessing efficacy and productivity. Navy telemedicine has been assessed using several models since the first implementation of telecommunications technologies. The Naval Health Research Center (NHRC), San Diego, California, has a long history of research evaluating the potential impact of advanced telecommunications for shipboard medical departments.1–6 The purpose of NHRC telemedicine studies to date has been to assess the need for these systems and their clinical impact.

2Naval Health Research Center, Field Medical Technologies, San Diego, California.

The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Department of the Navy, the Department of Defense, or the U.S. Government. This article is based on NHRC technical report number 02-35, approved for public release; distribution is unlimited.
ashore and afloat, and to evaluate whether telemedicine applications utilized in medical treatment facilities (MTFs) on shore might generalize to shipboard medicine.

Similarly, Carlos and Pangelinan described assessments of participant satisfaction with telemedicine technologies performed by the Telemedicine Working Group in Tricare Region 10. The group focused initial test and evaluation efforts on telepathology and teledermatology, but found little user buy-in for telepathology. The current report is a descriptive summary of tests on specific products used in teledermatology, and the establishment of working relationships among users and technicians. Based on their experience with testing and evaluating teledermatology technologies, the authors of this study strongly advocated getting physicians directly involved in planning for and implementing telemedicine in their facilities.

In this context, the variables examined have necessarily been limited so that the available data could provide clear answers to narrowly defined research questions. Thus, a complex system of health-care delivery has largely been studied as a series of separate applications operating as parallel platforms. A model has not yet been developed to account for Navy telemedicine as a dynamic system that relies on a key concept, interoperability, among hardware, software, personnel, and data.

The ways in which technical configurations and organizational structures affect the flow of information and resources in a given telemedicine system (i.e., interoperability) is an important research focus emerging for Navy telemedicine. Evaluating the degree of technical and structural interoperability can help ensure that each system is deployed to its maximum capacity. Assessing telemedicine as an integrated system simultaneously accounts for organizational and technological factors to evaluate how information gets distributed and processed. We propose to extend the concept of technical interoperability—the ability of software and hardware to exchange information and share tasks and resources—to human and organizational structures to provide a framework for assessing telemedicine as a system.

This report discusses current models for evaluating DoD telemedicine in light of two literature reviews, a preliminary review of models for assessing telemedicine and a follow-up review of recent articles concerning interoperability. The following section describes how the authors selected the articles for the preliminary review and offers comprehensive tables of both civilian and military studies that represent seven key criteria for assessing telemedicine systems. “Models in current military research” reviews representative models, program descriptions, and empirical telemedicine studies (if available) in military telemedicine to demonstrate the types and focus of existing telemedicine research in the DoD. “Follow-up review of interoperability literature” specifically reviews important recent research into the need for interoperability, an assessment category that has become particularly relevant to military telemedicine given the current alert status and funding environment. Possible structures for follow-up studies are suggested in “Recommendations for future studies,” whereas the concluding section summarizes emerging shifts in focus in telemedicine assessments.

MATERIALS AND METHODS

A preliminary literature review was conducted in September, 2001, in the PubMed database using the search term “telemedicine” in combination with “evaluation” or “model.” The search yielded 602 empirical and nonempirical evaluations of telemedicine, 69 (11.5%) of which evaluated telemedicine in military settings. Articles that discussed limitations that had already been resolved by subsequent technological advances (i.e., problems such as limited memory or slow processing speeds) were eliminated from consideration. The abstracts of the remaining articles were then reviewed and sorted by type of evaluation, such as cost/benefit analysis, provider satisfaction assessment, and so forth. NHRC technical reports and reports available through the Defense Technical Information Center (DTIC) were included in this analysis.

Because the articles reviewed in this search were largely written by MDs, a related literature search was conducted via the PubMed and the DTIC databases during September, 2001,
using the search term “telenursing” to get a preliminary indication of the perceptions of medical providers such as nurses and Independent Duty Corpsmen (IDCs). A total of 143 abstracts were read and sorted as previously described. These articles differed from the previous ones only in their specificity. They tended to focus on particular systems, equipment implementation, training needs, or treatment sites rather than models for assessing telemedicine as a system of care. These articles were subsequently omitted from this study in favor of a separate study to be conducted later.

We identified and reviewed an extensive number of evaluative models, descriptions of project implementations, and studies that attempted to assess the efficacy of specific components within larger telemedicine systems. Analysis of these revealed recurring proposals for or demonstrations of several specific types of evaluation. Seven areas of assessment recurred that could be subsumed under the concept of “interoperability,” a term that is being expanded to designate not only technical, but human and organizational compatibilities. Table 1 shows a composite of these criteria based on a comprehensive model suggested by Yawn and human factors criteria advocated in Yellowlees. The composite model combines technical, organizational, and human factors that are contained in the civilian literature, which may be adapted to specific Navy applications.

Proposed models, descriptions of actual programs, and empirical studies (if available) within each assessment category were reviewed further, and general criteria for assessing the military relevance of each type of assessment were developed. Studies were assessed to have military relevance on that basis of these criteria:

- Current and projected funding of NHRC Test and Evaluation research
- Projects currently funded within the Department of Defense (DoD)
- Projected DoD research, based on calls for proposals outlined in Broad Agency Announcements (BAAs)
- Feedback from Navy officers and researchers

See Appendix A for a summary table (Table 2) that lists the criteria for evaluating telemedicine systems that have been proposed in the literature. Table 2 summarizes representative studies in each of the seven categories noted in Table 1, to suggest their relevance for military telemedicine.

Finally, to test our conclusion that there is a gap in research in terms of the interoperability of technical, human, and organizational structures, the authors conducted a follow-up series of searches using PubMed on Sept 12, 2003. Three searches were conducted, using the terms “interoperability,” “interoperability and telemedicine,” and “interoperability, telemedicine, and military.” The term “interoperability” yielded 171 articles, with 129 published before 2002. “Interoperability and telemedicine” yielded 22 articles, with 18 published before 2002. “Interoperability, telemedicine, and military” yielded 2 articles, both published before 2002. These follow-up searches are discussed below.

# MODELS IN CURRENT MILITARY RESEARCH

A preliminary review of the literature on evaluating telemedicine revealed a mix of program descriptions, very short-term empirical studies focused on one or two components of a larger system, and conceptual models for evaluation. Several evaluative criteria have been proposed and analyzed in the literature; however, to date studies continue to address

<table>
<thead>
<tr>
<th>Table 1. Types of Assessment Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Task domain studies</strong></td>
</tr>
<tr>
<td>Assessment of tools and equipment</td>
</tr>
<tr>
<td>Evaluations of telemedicine</td>
</tr>
<tr>
<td>treatment settings</td>
</tr>
<tr>
<td><strong>Integration or outcome studies</strong></td>
</tr>
<tr>
<td>Cost analyses</td>
</tr>
<tr>
<td>Evaluations of participant satisfaction</td>
</tr>
<tr>
<td>Human factors evaluations</td>
</tr>
</tbody>
</table>

NAVY TELEMEDICINE

The composite model combines technical, organizational, and human factors that are contained in the civilian literature, which may be adapted to specific Navy applications.

The authors conducted a follow-up series of searches using PubMed on Sept 12, 2003.

Three searches were conducted, using the terms “interoperability,” “interoperability and telemedicine,” and “interoperability, telemedicine, and military.” The term “interoperability” yielded 171 articles, with 129 published before 2002. “Interoperability and telemedicine” yielded 22 articles, with 18 published before 2002. “Interoperability, telemedicine, and military” yielded 2 articles, both published before 2002. These follow-up searches are discussed below.

See Appendix A for a summary table (Table 2) that lists the criteria for evaluating telemedicine systems that have been proposed in the literature. Table 2 summarizes representative studies in each of the seven categories noted in Table 1, to suggest their relevance for military telemedicine.

Finally, to test our conclusion that there is a gap in research in terms of the interoperability of technical, human, and organizational structures, the authors conducted a follow-up series of searches using PubMed on Sept 12, 2003. Three searches were conducted, using the terms “interoperability,” “interoperability and telemedicine,” and “interoperability, telemedicine, and military.” The term “interoperability” yielded 171 articles, with 129 published before 2002. “Interoperability and telemedicine” yielded 22 articles, with 18 published before 2002. “Interoperability, telemedicine, and military” yielded 2 articles, both published before 2002. These follow-up searches are discussed below.
<table>
<thead>
<tr>
<th>Criteria for assessment</th>
<th>Study types</th>
<th>Representative references</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task domain</td>
<td>Evaluative model</td>
<td>Grigsby et al. (1999)1</td>
<td>List of health-care processes for evaluation (e.g., patient education, urgent evaluation)</td>
</tr>
<tr>
<td>Military relevance</td>
<td>Empirical study</td>
<td>Potter et al. (1997)2</td>
<td>Derm specialty most useful. Evac. decreased as telemedicine increased in data based projections. Patients treated via telmed rather than evac would receive more rapid access to urgent care specialist.</td>
</tr>
<tr>
<td>Tools and equipment</td>
<td>Program description</td>
<td>No relevant references</td>
<td></td>
</tr>
<tr>
<td>Military relevance</td>
<td>Evaluative model</td>
<td>Ganguly and Ray (2000)3</td>
<td>Interoperability of software systems at various levels is key to optimizing telemedical potential deployments and effective operations.</td>
</tr>
<tr>
<td>Tools and equipment</td>
<td>Program description</td>
<td>Birkmire-Peters et al. (1995)5</td>
<td></td>
</tr>
<tr>
<td>Telemedicine settings</td>
<td>Evaluative model</td>
<td>(1) Beach et al. (2001)6</td>
<td>The cost-effectiveness of telemedicine may be specific to the treatment setting. A telemedicine transaction model (TTM) can track the path of medical interventions from presentation to solution.</td>
</tr>
<tr>
<td>Telemedicine settings</td>
<td>Program description</td>
<td>Whitlock et al. (2000)8</td>
<td>Successful case studies describe telemonitoring laparoscopic surgeries on USS Abraham Lincoln.</td>
</tr>
</tbody>
</table>
### Table 2. Criteria for Assessing Telemedicine Systems (Cont’d)

<table>
<thead>
<tr>
<th>Criteria for assessment</th>
<th>Study types</th>
<th>Representative references</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(2) Walters (1996)11</td>
<td>(2) 70% of telemedicine consults impacted patient status by changing treatment (30%), preventing evac (10%) or significant improvement in care based on ratings by medical personnel (30%).</td>
</tr>
<tr>
<td>Military relevance</td>
<td></td>
<td>Burgess et al. (1999)12</td>
<td>Development of telemedicine programs should follow needs assessment from patient population.</td>
</tr>
<tr>
<td>moderate</td>
<td></td>
<td></td>
<td>Simulations used to project variables that lead to cost saving. Frequency of use was most important.</td>
</tr>
<tr>
<td>A model predicting the conditions where telemedicine would be most useful could measure potential impact in a deployed setting.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost analysis</td>
<td>Evaluative model</td>
<td>Cameron et al. (1998)13</td>
<td>Store-and-forward telemedicine projected to be cost effective for all Navy ship types. Teleradiology and video teleconferencing (VTC) for large ships only.</td>
</tr>
<tr>
<td>Military relevance</td>
<td>Empirical study</td>
<td>Stoloff et al. (1998)14</td>
<td>N/A Review of 32 empirical patient satisfaction studies found current research models lacking in rigor.</td>
</tr>
<tr>
<td>extensive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The military mission to provide care to troops anywhere supercedes cost factors. However, cost analysis is important in evaluating telemedicine as an alternative to MEDEVACS.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participation</td>
<td>Program description</td>
<td>None</td>
<td>Home monitoring encouraged help seeking by patients despite technical pitfalls.</td>
</tr>
<tr>
<td>satisfaction</td>
<td>Evaluative model</td>
<td>Sussmuth (2000)15</td>
<td>(continued)</td>
</tr>
<tr>
<td>Military relevance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>limited</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feedback on technology usability and provider/patient preferences may help predict telemedicine use.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
discrete components within telemedicine systems rather than how these components work together—i.e., the system as a whole. Much like the military telemedicine studies reviewed in this section, the civilian literature assesses local implementations or specialty applications of telemedicine.\textsuperscript{11–16} Although each of these current models evaluates telemedicine at some level, none of these frameworks allows a global or holistic assessment of telemedicine as a complex system in which human factors, organizational structures, and technical configurations interact.\textsuperscript{7} This section reviews representative models, program descriptions, and empirical studies that have made some attempts to assess how two or more components of a military telemedicine system work together. Such studies tend to focus primarily on either technical or “human” interoperability on a small scale; for example, describing human components as one factor in a technical setup or, conversely, technologies as tools to enable human interaction.

Assessments of technical configurations

At the Tripler Army Medical Center, Hawaii, the human and technical interface was evaluated in one treatment setting. Commercially available telemedicine technology was assessed using human factors usability criteria, such as technical acceptability, operational effectiveness, and clinical appropriateness, and technical specifications of each product using direct observation of provider behavior and interviews with medical users. Training needs were assessed by videotaping clinicians performing required tasks, and subsequently analyzing task time, error rate, and user preference. Interpretations of digital images captured with a video-otoscope were compared to examinations with a hand-held otoscope to assess the reliability, validity, and appropriateness of telemedicine equipment. Preliminary findings suggest that procedures are sensitive to testing instruments in terms of the types of errors generated and the time taken to complete the exam.

A technical and organizational analysis of telemedicine deployment was conducted in 1995 using data from the deployable teleradiology (DEPRAD) system in Bosnia and Hungary,\textsuperscript{19} including system configuration, inte-

<table>
<thead>
<tr>
<th>Criteria for assessment</th>
<th>Study types</th>
<th>Representative references</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human factors evaluation</td>
<td>Program description</td>
<td>May et al. (2003)\textsuperscript{16}</td>
<td>Providers who attempted to make telemedicine technologies fit their own standard practices failed to implement telemedicine successfully.</td>
</tr>
<tr>
<td>Military relevance extensive</td>
<td>Evaluator</td>
<td>Yellowlees (1997, 1998)\textsuperscript{17}</td>
<td>Provider investment in telemedicine technologies is critical for program success.</td>
</tr>
<tr>
<td>Important for understanding how new telemedicine initiatives can be successful and for ensuring effective use of technologies in various settings.</td>
<td>Empirical study</td>
<td>N/A</td>
<td>Human factors assessments are often combined with other types of evaluations.</td>
</tr>
<tr>
<td></td>
<td>Program description</td>
<td>N/A</td>
<td>Provider investment in telemedicine technologies is critical for program success.</td>
</tr>
</tbody>
</table>
gration and support, and telecommunications network technology. Over 10,000 radiological examinations were assessed. In most of these cases, reports from remote specialists were returned the next day.

A projected structural model by Chimiak et al. examined the organizational concepts and technologies needed to establish a telemedicine system for the entire Navy Fleet. Gomez et al. described telemedicine support for humanitarian missions in Somalia, Croatia, Macedonia, Germany, Italy, Kuwait, Ivory Coast, Egypt, Panama, Virgin Islands, Kenya, and Haiti from the Water Reed Army Medical Center (WRAMC) in Washington, DC. The study also analyzed data from clinical consults from February, 1993, through February, 1996, at WRAMC to assess the responsiveness of the telemedicine service. They reported that off-the-shelf equipment and low-bandwidth transmission could support multiple remote treatment sites simultaneously.

Calcagni et al. examined human factors as well as the structural model of Phase I of the Primetime III system, which provided telemedicine support to the U.S. Army during Operation Joint Endeavor in Bosnia. Case reports showed the feasibility and effectiveness of telemedicine in shipboard and U.S.-based medical departments. The majority of the consults assessed were in radiology; however, video-teleconferencing (VTC) ENT and dermatological problems were also examined. The authors emphasized human factors in the overall success or failure of the system.

Assessments of human or organizational structures

The effect of videoconferencing in the Inpatient Psychiatry Unit at the Tripler Army Medical Center, Hawaii, on human interaction in a telemedicine system was examined. Family members were linked with patients and their therapists via the videoconferencing system. Patients were able to see and receive support from geographically remote family members.

Clement et al. examined patient outcomes in a telemedicine neuropsychology clinic linking Brooke Army Medical Center, Fort Sam Houston, Texas, with Army community hospitals. Initial assessments of patients with neurological disorders or brain injuries were completed at the medical center, whereas follow-up visits occurred at local Army hospitals via videoconferencing. A total of 32 patients were seen in 87 videoconferencing sessions. Telemedicine made it possible to determine whether soldiers with a recent brain injury were fit for return to duty or should be separated from deployment.

Hunter et al. described the training and collaboration opportunities promoted by two teleoncology systems. The Pacific Oncology Outreach Project, an Internet-based system at Tripler, used mostly still pictures and live audio. Its goal was to avoid evacuations from outlying islands for healthcare. The Region 10 Integrated Cancer Network primarily used videoconferencing, with the goal of promoting distance learning and collaboration among medical personnel. Both systems were successful and well accepted by providers, while each one had distinct advantages. The Internet-based system at Tripler allowed users greater access. The Cancer Network’s Integrated Services Digital Network (ISDN) minimized administrative tasks.

Cubano et al. examined outcomes of laparoscopic procedures telementored aboard the USS Abraham Lincoln, the first demonstration of this kind aboard a combat ship, to assess telemedicine’s effect on the mentoring relationship. The Battlegroup Telemedicine system is configured to link the USS Abraham Lincoln, the Johns Hopkins Applied Physics Lab, and the Naval Medical Center at San Diego (NMCSD) via intraship, ship-to-ship, and ship-to-shore modalities. The authors discussed five cases in which remote specialist surgeons successfully mentored generalist ship’s surgeons in performing laparoscopic hernioplasties. The study results suggested that successful telementoring depends on the working relationship among operating surgeons even more than on the hardware-software configuration.

Norton et al. described early experiences in the telemedicine network established among the Tripler Army Medical Center, nongovernment agencies such as the University of Hawaii, and remote civilian treatment populations in Micronesia. They reported the establishment of a successful clinical consultation...
and health education network using low-bandwidth equipment and primarily preexisting communications systems.

A program description by Cook et al. briefly described the Tripler Army Medical Center’s implementation of digital communications in telediagnosis in the Pacific region using a “hub-and-spokes” model. Operation “Shooting Star” airlifts a Deployable Telepresence Unit to remote locations and connects primary care providers and patients with the Tripler’s medical specialists via satellite or ground communications. Digital radiology simplifies logistics and allows remote care, even in harsh environments.

Delaplain et al. described a telemedicine outreach initiative between the Tripler Army Medical Center, Hawaii, and outlying small islands in the Kwajalein Atoll. The patient population is approximately 5,000 American workers under contract with the DoD. The Tripler Army Medical Center provided remote consults for 59 cases in various specialties during the period of the study. Fifteen evacuations were avoided, saving an estimated $2,000 per trip. The authors conclude that cost savings to date justify extending telemedicine services to other remote sites in the Pacific Basin.

FOLLOW-UP REVIEW OF INTEROPERABILITY LITERATURE

Our initial literature review revealed the need for a framework to assess several components of a telemedicine system simultaneously. Hence, we propose an extended definition of interoperability to include the manner in which technical, human, and organizational structures affect the flow of information and resources in a given telemedicine system. Many of the evaluative criteria proposed in the literature may be synthesized into an assessment model that encompasses technical compatibility and the human and organizational factors. Research using this framework would assess individual components as they relate to each other and to the system as a whole.

A follow-up review of the literature on interoperability suggested the following. First, interoperability—including technical, organizational, and human aspects—has been a frequent topic and continues to be emphasized in the telemedicine literature; empirical studies assessing particular systems are rare. Second, the telemedicine literature contains relatively few papers that focus on the interoperability of telemedicine systems, particularly military telemedicine. Most papers present arguments on the need for interoperability without specific solutions or evidence-based tests of these solutions. Exceptions include proposed solutions to semantic and technical interoperability but no evidence-based evaluations. The specifics of how interoperability may advance military telemedicine remain relatively unexplored, except for isolated demonstrations of technical setups without evaluation data. In brief, there is a gap in systematic, empirical research in this area, and more work is needed to integrate health care within the U.S. civilian population, the DoD, and the international community. Recent “preparedness” exercises in the United States suggest present limitations in interoperability that exacerbate vulnerability to the impact of chemical and biological attacks. Some regional interoperability initiatives show that ensuring all components in a system work together has great potential for cost savings and for improving the quality of clinical outcomes for patients. However, systematic evaluation data are very rare. Most published work is descriptive (specific programs) or theoretical (advocating a set of principles or guidelines that should be implemented).

RECOMMENDATIONS FOR FUTURE STUDIES

Bashshur has raised several challenges for evaluating practices based on rapidly changing technologies:

- Variable costs and capabilities of telemedicine systems.
- Limitations in current technologies (rather than in the practice of telemedicine itself).
- The tendency of studies to over generalize results from specific technologies to other, unrelated applications.
NAVY TELEMEDICINE

• Midcourse changes made in operating parameters during the period of a study.

He concluded that a financially stable environment, regional networks, and the ability to redistribute resources in response to change should all be present if evaluation of a system is to be useful. Clearly, military health care is well situated to fulfill all of these capabilities. Military telemedicine sidesteps some of the main concerns of civilian applications, such as medical licensing or legal issues; state regulatory policies or political support; untrustworthy sources of information; the corporate or commercial structure of health care; community funding; unequal access to care for individual patients; and the effect of telemedicine on general practice. However, most military evaluations of telemedicine have focused on cost/benefit projections for developing a system of telemedicine, or on the implementation of telemedicine in a particular specialty or treatment setting. Tripler Army Medical Center in Hawaii has been especially prolific in producing telemedicine studies that deal with different treatment populations.

The need for a framework within which to utilize and optimize the Navy’s own ongoing assessment is readily apparent. This section suggests two types of studies to assess system integration while keeping data collection manageable and also providing the necessary empirical measures for assessment. Both types represent the first necessary steps in shifting research designs toward a comprehensive model that assesses telemedicine systems as an integrated whole.

A technology-based approach for classifying telemedicine systems

Military telemedicine systems vary widely, both in treatment facilities and in the clusters of technologies that are used to deliver care. Levels of telemedicine capabilities do not easily correlate to levels of care. Several different telemedicine suites may be implemented in a single level of care. The continuum of treatment sites ranges from highly mobile to wholly fixed sites, with different combinations and levels of technologies available for implementation within different levels of care. Often the success of the implementation of telemedicine is not based on the setting or level of care, but rather the specialized training and investment of the personnel using it (Fig. 1).

The primary focus of research is evaluating how information and resources flow within given levels of care and technological configurations. This could be accomplished in a series of short-term studies that build on data collected at several levels. For example, initial studies might classify military clinics by their telemedicine capabilities. A continuum of capabilities would be established, with fully equipped telemedicine facilities at one end and clinics with minimal telecommunications resources at the other. Each class of clinic would be assessed based on outcome measures such as clinical impact, operational efficiency, and integration of various components of the system with overall clinic functioning.

Next, telemedicine capabilities would be ranked from basic to advanced technologies and correlated with outcome studies conducted at various levels of care. For example, the basic telemedicine suite for a clinic might consist of telephone and FAX for remote consultations. More advanced classes of telemedicine suites would consist of added technologies such as e-mail, Internet, computed radiography, various other diagnostic videoscoping procedures (e.g., ENT otoscope), and live videoconferencing. Important variables such as medical staffing at each clinic, patient population, and geographic location could also be assessed and isolated statistically to permit a controlled comparison of clinics on outcome measures.

Alternatively, telemedicine systems could be reviewed based on correlating clinical requirements to technical configuration and availability at different levels of care. The ultimate goal of such studies would be to develop a method for determining which levels of telemedicine should be deployed at different levels of care.

Assessing a pilot telemedicine system or site

A more conventional research model would evaluate a single pilot system or site systemi-
cally. In other words, human, organizational, and technical interoperability could be tested across levels of care in a typical small, manageable system. A systemic evaluation would allow researchers to analyze how the various areas of assessment function in relation to each other and to the larger system.

For example, the Medical Data Surveillance System (MDSS), a Web-based system currently being developed to track epidemiological trends, could be evaluated as a closed “telemedicine” system. MDSS uses telecommunications technologies to enable early medical threat detection in a near-real time health surveillance system. Although this database is used for record keeping and preventive medicine rather than treatment per se, its implementation in the field incorporates many of the characteristics of telemedicine systems. It connects experts with generalists; it enables the flow of data in virtual real time from one treatment facility to another; and it employs many of the same communications technologies traditionally thought of as telemedicine. Medical personnel have interactive access to MDSS with a standard Web browser interface, which enables them to view outpatient visits over time, compute estimated outbreak curves, and generate standard reports. Additionally, on a daily basis, MDSS automatically performs a series of Dynamic Change-point Detection (DCD) analysis to alert users of any significant epidemiological events within the populations served. In field tests, the system detected an upswing in gastrointestinal complaints that was traced to a tainted water supply. Early detection of such
trends can facilitate appropriate diagnosis and treatment of naturally occurring illnesses as well as those resulting from the use of biochemical weapons.

Thus, this system operates within a more complex system of technology-based health care delivery, and it contains several elements of the larger system. Hence, it constitutes a microcosm of the larger system. Although its function is not directly related to patient care, its deployment uses communications technologies to connect seemingly disparate events and alert caregivers to trends that might not be evident to them. Tracking trends in virtual real time has the potential to prevent large-scale health-care crises, such as those that would result from an undetected biochemical attack. Lessons learned in one structure can then be tested in other parts of the larger system, mitigating the logistical challenges of assessing a complex, multitiered system at every level all at once.

Each of the seven components for evaluation proposed in the telemedicine literature can be assessed within the MDSS system: task domains, tools, settings, integration, costs, participant satisfaction, and human factors. These components can be assessed in two ways. First, each component can be evaluated as part of the MDSS system to determine how effectively individual factors operate within the larger system. Second, the interoperability of these seven components can be assessed to determine the cumulative or systemic effect of all seven components.

Following are some questions to be considered:

- **How efficient is the overall system?** Is electronic tracking of trends more productive than current methods for detection of a disease outbreak? Does the answer to this question change in deployed versus nondeployed settings? What does electronic medical surveillance allow us to do that pencil and paper or other conventional methods such as telephone don’t?
- **How does each component interact with and affect each of the others?** Does the ability to conduct disease surveillance affect treatment outcomes? Costs? The actual tasks performed? Equipment needed? Human factors? Does medical surveillance translate into patient care at some point?
- **If one or more components is not functioning at full capacity, how does that affect the efficiency of the overall system?** If data are not available as often as expected, does this affect outcomes? Does lag from “real-time” data input to MDSS reduce its effectiveness relative to conventional methods of surveillance? If so, how much? Does potential lag time affect costs, or human factors such as provider investment in the surveillance system? Does patient satisfaction affect the military relevance of tracking medical trends?
- **How does each component fit into the system as a whole?** How does setting (primarily, varying degrees of access to health databases) affect the function of MDSS? How do differences in the frequency of electronic updates across treatment settings affect the way the data can be used? Which treatment settings and patient outcomes are most impacted by electronic surveillance of medical trends?
- **Are any of these elements working against each other?** Is “real-time” data available in settings where bandwidth allocation for medical tasks is limited? Does the system operate differently depending on the frequency of update (in different treatment settings, from one to several times a day)? How does provider satisfaction affect the implementation of the data received?

Some components to consider in each of the seven evaluation categories are as follows:

- Assessment of current research/efficacy of telemedicine at the test site.
- Criteria for judging proposed research or areas to implement telemedicine.
- Criteria for assessing which projects should get funding and approval to proceed.
- Structures and procedures for assessing work in progress (projects as they develop).
- Structures for generating written reports of work in progress.
- Structures for promoting completed or successful projects to the larger telemedicine community.
- Ongoing assessment of the overall health of the telemedicine project.
CONCLUSION

An assessment of telemedicine as a complex system is needed to demonstrate how current processes, usage, training, equipment, and expenditures measure up to the emergent needs of the system as a whole. Assessing telemedicine as an integrated system of medical surveillance and/or health-care delivery would simultaneously account for organizational and technological factors to evaluate how information gets distributed and processed. This type of analysis would extend the concept of technical interoperability—the ability of software and hardware to exchange information and share tasks and resources—to human and organizational structures. Evaluating the degree of technical interoperability and the organizational ability to pool knowledge and resources in a system can help ensure that that each system is deployed to its maximum capacity.

Measuring the interoperability in a telemedicine system is important for military studies in three key areas. First, telemedicine’s self-diagnostic and surveillance capabilities increase the likelihood of preventive care and help-seeking behavior in deployed settings where autonomy and competence are highly valued and individuals are tempted to ignore their own health needs in the interests of the mission. Second, the successful implementation of a technology depends in large part on the perception of users that it is useful and accessible. Any strategy for estimating the level of telemedicine needed at a particular level of care should take into account how the technology will change, or clash with practices already in place, especially at the level of the individual user. Even seemingly minor non-standard implementations can erode the efficiency of the system and the overall readiness of the troops. Finally, as deployments increase in settings where the level of care and the type of technology deployed are limited by the surrounding environment, the efficient sharing of information and resources that telemedicine promises is literally a matter of life and death. Although descriptive evaluations from the battlefield are vital, we need a way to assess the overall structure and functioning of these deployed systems empirically.

The logistics of systemic, empirical assessments of telemedicine are daunting, even in less challenging environments, as evidenced in recurring, but unrealized, proposals for systemic models for evaluation in both civilian and military literature. A shift in research focus is needed to realize such assessments. Reconceptualizing interoperability to encompass human, organizational, and technical structures may provide a framework for evaluating multiple components of a telemedicine system simultaneously.

ACKNOWLEDGMENTS

Report No. 02-34 was supported by the Office of Naval Research, Arlington, VA, under Work Unit No. 637006/N M2332.001-60004. The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, or the U.S. Government. Approved for public release; distribution unlimited.

This research has been conducted in compliance with all applicable Federal Regulations governing the protection of human subjects in research.

REFERENCES

APPENDIX A: TABLE 2 REFERENCES

This article has been cited by: