7
Critique of the Current System

It is one thing to show a man that he is in an error, and another to put him in possession of truth.

—John Locke

Contents

7
Critique of the Current System

INTRODUCTION

It has been well established that there is insufficient information regarding the costs, risks, and benefits of medical technologies. One purpose of assessing medical technologies is to produce information to help guide their appropriate use. Clearly, society wishes to promote the development and diffusion of safe and effective medical technologies. At the same time, society wishes to reduce ineffective and inefficient use of medical technologies. Finding a balance between these two goals is difficult. The complexity of our society and the mixed private/public nature of the health care system magnify the challenge of improving existing policies and processes for medical technology assessment.

Federal policies toward medical technology have developed in incremental fashion to meet rather specific goals. Only recently have the numerous policies that pertain to the development, diffusion, and use of medical technologies begun to be seen as elements of an overall system for guiding and promoting technological change, The collection of programs and activities that forms the current "system" for assessing medical technologies' has also been built-up in piecemeal fashion. Although many have realized that better information on the benefits, risks, costs, and social implications of medical technology is essential to guiding the development and use of technology without unnecessarily impeding innovation, progress toward developing a coherent system for assessing medical technologies has been slow.

The idea of a "strategy" for assessing medical technology is closely related to having a system. A strategy is in effect the underlying basis for the design and implementation of any coherent system for assessment. Any strategy will represent

a compromise among competing perspectives on the goals of medical technology and on the role and format of assessments. Strategies reflect desired policy directions and ideas on how best to move in those directions. Without a conceptually clear strategy, any system of assessment will suffer from inconsistencies and unclear objectives.

The basic objective of a strategy for medical technology assessment *is* to ensure that technologies of public policy importance are evaluated by appropriate methods in a timely fashion without unnecessarily harming innovation. This objective must be sought despite a number of formidable difficulties. One problem is the very limited amount of money available for original evaluations of medical technology. Costs of a single randomized clinical trial (RCT) can run as high as hundreds of thousand of dollars, yet the entire societal investment in original studies of medical technology probably does not exceed $200 million at any one time.

A sound strategy for assessment must take into account the stage of development of particular medical technologies-emerging, new, or existing. If emerging technologies are assessed too early in their development, innovation may be slowed; furthermore, the information gained by assessing the technologies may be valueless, because knowledge about their modifications and eventual uses will be limited. If new or existing technologies are assessed too late, the assessments will have little effect on the technologies' diffusion or use. Assessment information must be disseminated to appropriate parties in a timely manner.

A strategy for medical technology assessment must also deal with the universe of drugs, devices, and medical and surgical procedures—diagnostic, therapeutic, and preventive. Other important classes of medical technolog$_y$ to be dealt with are medical care delivery systems and organizational

---

● By the existing "system" of assessing medical technologies, OTA means the current set of programs and activities which are related to assessing medical technologies. In this context, the word "system" is used in the general sense and is not intended to imply that there currently is some overall design involved.

innovations, although, as noted earlier, they fall beyond the scope of this assessment.

A further objective of a strategy is to develop criteria to choose an appropriate method or methods for assessment. A strategy must permit one to determine when it is enough to know that a technology is efficacious, when it is desirable to have a formal cost-effectiveness study, and when a full-scale technology assessment with evaluation of social implications would be helpful and desirable.

A strategy should also address how the information gained from technology assessment is utilized and by whom. And finally, it is necessary to consider how technology assessment should affect health policy.

These are some of the challenges. This report cannot deal with all of them. The-section that-follows discusses the development of information on which to base decisions. The chapter ends with a critique of the current technology assessment system.

ASSESSMENT OF MEDICAL TECHNOLOGIES

Assessing medical technologies is a complex process, and no simple model can be devised to outline the steps that must be taken in all circumstances. Technologies are diverse, often lending themselves to be evaluated in diverse ways. The need for information about technologies by different people at different points in time varies, as well. Nevertheless, there are a finite number of technologies and a finite number of assessment methods. The information that is required to make rational and reasonable informed choices is also finite.

The previous chapters of this report have identified the needs for assessment information and the resources available to fill those needs. The needs are defined by those called on to make decisions. For example, the Food and Drug Administration (FDA) must decide whether to allow a drug or device to be marketed; it asks whether the technology works and whether it is sufficiently safe. The National Institutes of Health (NIH), in its basic research efforts, must set priorities regarding which technologies, including especially medical and surgical procedures, it will further investigate and develop. Having been called on by Congress to synthesize what is known about medical technologies in order to assist in their transfer, NIH also asks what the appropriate conditions and standards of use are. Similarly, Professional Standards Review Organizations (PSROs) ask whether local practice patterns conform to reasonable standards of care. Those who pay for care, whether they be the Health Care Financing Ad-

ministration (HCFA), Blue Cross, or individual patients, need to know whether the use of a medical technology is worth the cost. And, finally, the practicing clinicians, in consultation with individual patients, must make the final decision to use a technology. Throughout this process, the values and needs of society, the medical profession, and the patients themselves are interwoven.

A strategy for medical technology assessment must take into account what information is known, what is not known, what is needed, what can be obtained, and what the cost of obtaining it will be. Information will never be perfect, and money and time will always be limited. Thus, it is important to make judicious use of evaluation methods. Fortunately, there are means to compensate for uncertainty when important information is lacking.

No clear-cut rules seem to be possible in devising a strategy for choosing a method to assess the safety, efficacy, and effectiveness of medical technologies. Often, the method of choice will be related to both the stage of diffusion of the technology and the extent of knowledge and belief as to its risks and benefits.

RCTs, for example, tend to be appropriate when a technology's risks and benefits are not well understood, when the technology is not yet in general use, and/or when costs of the technology are very high in relation to expected benefits, and when risks are expected to be low. Under these conditions, the purpose of an RCT is to establish

a cause-and-effect relationship. Reasonable candidates for RCTs would thus be new drugs, new invasive devices, new expensive equipment, and new elaborate services requiring capital expenditures (e.g., neonatal intensive care units). When risks and benefits are not well known or are not believed, randomization can be used without violating some of the ethical principles noted in chapter 3. If RCTs are used early in a technology's diffusion, nonrandomized designs, especially case-control studies, can later be used to establish the effectiveness of the technology as it diffuses into diverse settings.

However, when a technology is in widespread use, risks and benefits are either already known or are widely believed to exist, and randomization may be neither possible nor appropriate. In this case, nonrandomized designs" can be used to establish relationships which can later be tested, if desired, by more rigorous methods, including randomization.

Economic analyses are similarly varied, and no one technique is applicable in all cases. However, economic information may be worthless without good safety and effectiveness information. For the user of the technology, the price is the cost. That price must somehow be compared to the perceived value of the use of technology. But for more general decisions, especially at the societal level, economic analyses are very complex, requiring both technical expertise and good judgment. A cost in one instance may be ignored or even counted as a benefit in another.

As discussed in chapter 4, decisions concerning the development and use of certain medical technologies often have profound social and ethical implications. Especially at the Federal policymaking level, these implications are important to consider, even though they cannot be precisely quantified.

Finally, informed decisions rest on the analysis of all available information. Chapter 5 discussed a number of techniques that can be used to synthesize information from research studies in a systematic manner. Additionally, group process techniques such as Delphi and nominal group process are available to assist policymakers and technically expert groups in making decisions. None of these methods for synthesizing information is perfect, but each has potential value in the development of more orderly processes for setting policies regarding medical technologies.

CRITIQUE OF THE CURRENT SYSTEM

The present system of medical technology assessment, like the medical delivery system, is pluralistic, and many of the public and private sector activities reviewed in this report were undertaken for purposes other than medical technology assessment. The diversity of activities is not necessarily a weakness. Such diversity capitalizes on the wealth of ideas and interest of many different people and organizations. Nevertheless, it makes the job of fashioning a more coherent system of assessment more difficult.

Perhaps the principal reason for the difficulties with the present system is that the main parts were developed separately over a long period *of time* with specific, sometimes inconsistent, goals. The existing programs and activities were not devised as elements of an overall system of technology assessment. In the case of the Food, Drug, and Cosmetic Act, amendments over the past three quarters of a century have been internally consistent, tending to build on and complement previous legislation. However, most other legislative and nongovernmental efforts that affect medical technologies have not been so well coordinated. Thus, the country has a system of physicians, hospitals, planning agencies, PSROs, health survey activities, research activities, and insurance claims networks, all of which use, dispense, regulate, evaluate, collect information on, or otherwise affect medical technologies, but which often do not complement one another's needs for technology assessment. *

_____
*This criticism does not necessarily apply to the Veterans Administration's system, which OTA did not study to any appreciable degree. Nevertheless, it is known that the Veterans Administration is developing a system whereby potential investigators are informed of program needs, a research agenda is developed to satisfy those needs, and useful information generated by research is made available to those who need it (168).

To examine the extent to which the needs of an overall system for medical technology assessment are met, the programs and activities comprising the present system are discussed in the remainder of this chapter with reference to the four phases of the technology assessment process mentioned earlier: 1) the identification of technologies needing assessment, 2) the testing of technologies to develop information concerning their health and economic effects, 3) the synthesis of information, and 4) the dissemination of the information that is available.

OTA finds that the current system for evaluating medical technologies exhibits major deficiencies in each of the four phases of the assessment process. For technologies at different stages of development (i.e., emerging technologies, new technologies, existing technologies, and new applications of existing technologies), as well as for technologies classified as either drugs, devices, or surgical or medical procedures, the adequacy of the present system differs.

The existing system for identifying technologies to be assessed, except for FDA's system of identifying new drugs and devices, is unnecessarily poor. (Among the many reasons for this is the inadequacy of the synthesis of research information.) In the testing of medical technologies, many studies generate evaluative information, but the quality of such information varies widely. FDA's research requirements for new drugs and devices seem adequate for the premarket approval process, and much NIH-sponsored research has resulted in significant information for society. In other areas, however, high-quality studies are few, and most of them are not helpful in setting policy. High-quality, objective syntheses of research findings-a prerequisite for developing policy or setting medical practice standards—are rare. Many syntheses are informal, overly subjective, group-generated norms and are not based on a rigorous assessment of the scientific evidence. Although there are increasing efforts to disseminate technology assessment information, much of the information has questionable value. The excessive adoption, diffusion, and use of some medical technologies indicate a need for improved dissemination efforts.

In the expanded critique that follows, special attention is paid to the identification of medical technologies to be assessed, since OTA finds that this is the critical phase of any overall assessment strategy.

Identification

Any system for medical technology assessment must have mechanisms to identify technologies to be assessed and to set priorities among candidates for assessment. Clearly, no single mechanism is appropriate for all occasions and all technologies. What works for drugs may not be suitable for surgical procedures, and what is appropriate for identifying emerging technologies may not be adequate for established ones.

Methods of identifying technologies for assessment can be thought of as falling into one of three generic categories: 1) routine mechanisms, 2) priority-setting mechanisms, and 3) mechanisms of opportunity. Routine *mechanisms* systematically identify a class of technologies and are usually connected with a particular event with which all technologies in the class are associated. (Examples are FDA's requirement that all drugs and devices be registered with it prior to marketing or testing in human beings and, if taken advantage of, HCFA's reimbursement coverage determinations.) *Priority-setting mechanisms* are not routine and are often mechanisms or processes used by some group to determine priority technologies for assessment based on some implicit or explicit criteria. (Examples are the processes the institutes of NIH and HCFA use to establish their research agendas, the processes the Office for Medical Applications of Research (OMAR) of NIH uses to set priorities for consensus development conferences, and the process the National Center for Health Care Technology (NCHCT) formerly used to establish priorities for technology assessments.) *Mechanisms of opportunity* are means for identifying technologies for assessment as opportunities happen to occur. These are less well defined than mechanisms in the previous two categories, but are not necessarily less important, because technologies that suddenly become important to assess often do so for safety or ethical reasons. (Examples

of mechanisms of opportunity are FDA's spontaneous drug reporting system, PSROs' medical care evaluation studies, and the systematic generation of patient outcome data which researchers can analyze or the public can challenge.)

The specific purpose of identifying technologies for assessment will vary according to the stage of development of the technology in question. Emerging technologies, especially drugs and certain devices, need to be assessed for safety and efficacy. In the case of drugs and certain devices, assessment in the emerging stage is necessary because of the well-accepted social responsibility, as expressed in FDA law, to protect the public. There may be other social and ethical reasons to assess emerging technologies as well. New and existing technologies need to be assessed for safety and effectiveness, sometimes for cost effectiveness, and at times for social and ethical consequences. Physicians and their patients need to know what works and what the benefit/risk balance is. Patients and insurers, including HCFA, need to understand the economic implications of technologies, especially when there are alternatives. Information is also needed by PSROs. PSROs are charged with assessing whether HCFA funds are being used on "needed" services, and these include both new and existing technologies. Additional information is needed by health planning agencies, which are charged with determining whether major new technologies should be purchased by hospitals.

A critique of the current system's record in identifying medical technologies for assessment yields mixed results. For the purposes of the discussion below, it is helpful to think of the identification of technologies at the four typical stages of development defined in chapter 1: 1) emerging technologies, 2) new technologies, 3) existing technologies, and 4) new applications of existing technologies.

Identification of Emerging Technologies

Certainly, the most thorough system for identifying emerging medical technologies is FDA's premarketing approval process for drugs and Class III medical devices (see ch, 6). This process is clearly a routine identification mechanism, as defined above.

The only other notable systems for identifying emerging medical technologies are priority-setting mechanisms. These include the processes for establishing the research agendas of NIH, the National Center for Health Services Research, and while it was funded, NCHCT. Although each institute and research agency has its own internal systems, the process of establishing priorities for intramural research and extramural contracts is essentially determined by institute or agency staff. Research priorities tend to be set by informed professional staff who know their particular field well and thus know which questions are important to address. Grants can be either solicited or unsolicited. In either case, the projects are generally selected on the basis of technical merit and judgments about their importance. The research agenda priority-setting processes of NIH and other Federal research agencies generate information for a base of knowledge which can lead to unpredictable but substantial future dividends that may be difficult or impossible to measure. Often, however, the processes do not address the immediate policy priorities of operating agencies such as HCFA and other social priorities such as Congress may have for the health care system.

Two other priority-setting mechanisms for identifying emerging technologies for assessment are also deserving of mention. HCFA's research arm, the Office of Research and Demonstrations, is charged with assessing technologies of interest to its operations. Some of these technologies may be classified as emerging, although most probably would not be. Seldom, however, are the technologies clinically related (e.g., some are concerned with information systems). The other mechanism was the NCHCT's "emerging technology list. " This was a systematic and broad-based approach for identifying health-related technologies under development which were expected to be used in the practice of clinical medicine within 5 years. However, critics from industry charged that the compilation of such a list threatened the innovation process, and the 1981 reauthorization of NCHCT specifically withdrew the center's authority to compile this list.

The third type of identification mechanism, taking advantage of unforeseen "opportunities," is generally not relevant for emerging technologies,

except in rare special cases such as the artificial heart.

Discussion: FDA's routine identification of all emerging drugs and emerging Class III devices seems adequate and appropriate. Emerging medical and surgical procedures do not seem to lend themselves to being identified through routine mechanisms. The most appropriate identification method for emerging procedures would seem to be the subjective priority-setting mechanisms such as those being used by the institutes of NIH and other Federal research agencies.

NCHCT's "emerging technology list" had the advantage of cutting across categorically related programs and also forced each program of the Department of Health and Human Services to explicitly identify technologies which were emerging and of importance. The 1980 format for submission of an entry included the name and identification of the technology, a technical description, a statement of importance or potential impact, an evaluation of the technology's present status and data needs, and any special considerations. NIH staff have commented that the exercise of compiling such a list was useful in taking stock of what was happening in their respective fields. If one objective of the technology assessment system is to more actively manage technologies, compiling such a list would seem to be quite useful in that it allows one to make predictions and to plan for the future. The charge of industry that the list inhibited innovation is not supported by any data that OTA could find, but the issue could be further examined.

OTA concludes that emerging drugs and devices are adequately and appropriately identified, but that emerging medical and surgical procedures could be better identified. Overall, the identification of emerging technologies is not a critical weakness of the present system.

Identification of New Technologies

As a group, new technologies, those in the adoption phase, are the most easily identified. In particular, such technologies are the most obvious candidates for identification through routine mechanisms. Most new medical devices (i.e., Class I and 11 devices) are routinely identified as

required by FDA law (see ch. 6). * New medical and surgical procedures, including the use of new drugs and/or devices, are potentially identifiable routinely through the reimbursement process, since the question of coverage should arise when a new procedure is identified.

All new medical technologies are also potential candidates for being identified through the priority-setting mechanisms discussed in the preceding section on emerging technologies. And, in fact, new technologies are identified for assessment through the priority-setting processes of the institutes of NIH and the other research agencies of the Public Health Service (PHS) and HCFA.

All new technologies are also logical candidates for being identified through mechanisms of opportunity. The recent maternal serum alpha-fetoprotein (MSAFP) controversy illustrates the use of such a mechanism (see app. E). FDA was on the verge of approving widespread use of the MSAFP screening test when a special interest consumer group (the parents of children with spina bifida) questioned the validity of FDA's data. A major assessment of MSAFP was subsequently carried out, and new regulations were issued. This case illustrates a mechanism of opportunity (i.e., publishing data and making decisions under public scrutiny).

Discussion: As a class, new technologies are the most easily identified as candidates for assessment, especially by routine mechanisms. In the case of new Class I and Class II medical devices, FDA's routine identification process seems adequate and appropriate. In the case of new medical and surgical procedures, however, there is currently no systematic mechanism for identification. To some extent, new procedures are identified through the reimbursement system; however, in contrast to the structured identification process of FDA, identification through reimbursement decisions of HCFA and other public and private insurers is much more haphazard. While there is considerable potential for the reimbursement system to be used routinely as a primary means of identifying new procedures for assessment, prob-

---

'More invasive devices (i.e., Class 111 devices) are identified in the "emerging" phase, because they must receive FDA's approval before being tested in clinical trials.

lems persist. Even the process of identifying which procedures are new seems to be unsatisfactory: terminologies and codes on claims forms are often not accurately labeled or are not standardized; new procedures often do not have a procedure code number. However, to the extent that there is increased scrutiny by third-party payers of bills submitted for new procedures and more than occasional denial of payment for such bills, the provider has a strong incentive to request payment for an already existing standard procedure, rather than a new one, thus complicating the identification process.

The process of identifying new technologies through the priority-setting processes of the institutes of NIH and agencies of PHS has essentially the same strengths and weaknesses as were discussed in the connection with the identification of emerging technologies. From an academic point of view, the system seems appropriate. The weakness stems from the lack of an adequate system to identify priority candidates for the operating agencies, especially HCFA, PSROs, and planning agencies. Theoretically, HCFA has its own research arm, the Office of Research and Demonstrations, to accomplish this. As stated previously, however, that office has not been very involved to date with either identifying technologies for assessment or assessing them.

Whether the mechanisms of opportunity for *identifying new* technologies are adequate is difficult to assess. Since standardized, high-quality data on technology use and health outcome are not generally available, it is likely that they are not.

OTA concludes that new drugs and devices are adequately identified for assessment, but that new medical and surgical procedures are not. The most pressing need is for a routine mechanism to identify new procedures before they are widely adopted. The reimbursement system, because coverage and payment decisions are critical points in the diffusion of many technologies, might be given primary consideration. In addition, the priority-setting systems of the institutes of NIH and of other Federal research agencies (e.g., NCHSR) are adequate **and appropriate for their respective mandates, but there is not an adequate similar**

system to fulfill the needs of operating agencies (e.g., HCFA, pIanning agencies). Finally, sufficient mechanisms of opportunity for identifying new technologies do not exist but could be developed. Medical specialty societies could be helpful in this area.

Identification of Existing Technologies

As a group, existing medical technologies tend to be the least likely candidates for routine identification, primarily because there is no natural triggering mechanism such as introduction. Consequently, the timely identification of existing technologies must depend largely on priority-setting procedures or mechanisms of opportunity.

Theoretically, if emerging and new technologies had been adequately identified (and assessed) as they developed, there would be less need to identify (and assess) them after their adoption and general diffusion. But, as indicated in previous OTA reports (e.g., 266,270,279), most existing medical technologies have not been adequately assessed. At a minimum, existing medical technologies should be monitored for risks that may not have been previously apparent. A review of the activity in this area reveals a very poor record, with a few exceptions and a few encouraging signs.

One encouraging sign is the interest in postmarketing surveillance systems for drugs. * Postmarketing surveillance systems are noteworthy, because there is increasing concern that FDA's premarket approval process is not sufficient to protect the public after a drug or device is marketed and in use (281). Although often regarded as testing techniques, postmarketing surveillance systems can also be thought of as sophisticated systems for identifying technologies needing further investigation. Such systems represent a hybridization of a routine mechanism, a priority-setting mechanism, and a mechanism of opportunity, Although data may be collected routinely under postmarketing surveillance systems, not all drugs would automatically be screened. FDA can set its own research agenda, and independent research investigators, at their own initiative, can be ex-

---

'This topic is considered at greater length in a separate OTA report entitled *Postmarketing Surveillance of Prescription Drugs* (281).

pected to use the data to identify fertile areas for future study.

Another encouraging sign for the identification of existing technologies for assessment are private sector initiatives using the priority-setting method. These include Blue Cross/Blue Shield's Medical Necessity Project and the American College of Physicians' new Clinical Efficacy Assessment Project (see ch. 6). In the Federal sector, discussed previously, priority-setting processes (including NCHCT's and OMAR's) are also used to identify existing technologies.

The most glaring omission in the system for identification of existing technologies for assessment is the lack of identification by operating agencies, especially HCFA and State Medicaid agencies. Even with its PSRO arm, HCFA does not have an adequate system to question technologies that are already in widespread use.

One mechanism of opportunity that can be used to trigger identification of an existing technology in need of assessment is the identification of a competing technology. To some degree, this mechanism is used implicitly. For instance, computed tomography scanning was likely to have been compared with existing technologies such as skull X-rays. Whether such opportunities for identification are always, or even frequently, taken advantage of is not clear.

It was stated earlier that mechanisms of opportunity are particularly useful for identifying technologies that are currently in use. FDA has a spontaneous reporting system for adverse drug reactions which illustrates how technology assessment opportunities surface "spontaneously. " Similar systems could be used by other agencies such as HCFA. A functioning identification system of opportunity requires a method by which a technology assessment issue can be reported and a means to act on that information.

OTA concludes that the system for identifying existing technologies in need of assessment is inadequate. One promising possibility is postmarketing surveillance techniques. As was true with emerging and new technologies, the priority-setting procedures of Federal research agencies may be adequate for those agencies' respective

needs, but not for the needs of operating agencies such as HCFA. And the operating agencies themselves do not adequately identify existing technologies for assessment. Medical specialty societies could be helpful in this area. Finally, NCHCT's activities of identifying nationally important priority technologies for assessment were valuable but are not now funded. Thus, no body is currently undertaking this important task.

Identification of New Applications of Existing Technologies

The consideration of new applications of existing technologies is important for two reasons. First, a new application of a technolog$_y$ means that previous information about the technology may no longer be applicable; and second, a technology's new use may provide an opportunity to identify it through a routine mechanism. At present, OTA is unaware of any systematic method of identifying new applications of existing technologies as candidates for assessment.

These technologies can be identified through priority-setting procedures and mechanisms of opportunity, as can existing technologies. It would seem, though, that the most rewarding approach for identifying new applications of existing technologies would be through a routine mechanism, namely, the reimbursement system.

OTA concludes that new applications of existing technologies are not adequately identified for assessment. To facilitate the identification of such technologies, the most promising approach may be the use of the reimbursement system to link the diagnosis with the use of the technology. Medical specialty societies could be helpful in this area.

Testing

Many of the deficiencies of the testing phase of the current system for medical technology assessment are intimately related to the inadequacies of the identification phase. In order to know what to test for, one must have identified the appropriate technology for assessment, the relevant policy concern (e.g., safety, efficacy, or cost effectiveness), and the information which is lacking. Thus, an adequate testing phase requires an ade-

quate identification phase. It is not surprising, therefore, that the strengths and weaknesses of current testing activities closely parallel those of the identification phase.

FDA adequately identifies emerging and new drugs and devices that need assessment and also determines what information it needs. Furthermore, FDA carefully reviews the research protocols of the industries it regulates and requires that the protocols be used. Resulting testing by industry seems adequate. As suggested previously, however, FDA does not have an adequate means to identify which drugs and devices need further testing once they are released into the market. Thus, FDA cannot develop protocols for further testing of products in new settings or under different applications. As indicated in chapter 3, adequate protocols can be developed (see also ref. 281).

Chapter 6 discussed the testing of medical and surgical procedures through the funding activities of NIH. Since the individual institutes of NIH subject all research protocols to an intensive peer review process (270), the quality of the research is generally good. Any problems with such activities center around either the need for additional funding or the agenda-setting process* (the latter is essentially an issue of identification). It is important to note that NIH does not have the mission to ensure that all medical and surgical procedures are proven to be safe and effective. (Nor does any other agency or organization.)

Currently, the overriding weakness of the testing phase is in the testing of new and existing med-***ical*** and surgical procedures. Since procedures tend to be developed within the practice of medicine, they are generally adopted and accepted by the medical community without a routine, formal examination of their merits. A good deal of the problem in this area stems from a lack of research funding. Another problem concerns the development and use of research methods, since RCTs are not appropriate for all clinical inquiries.

---

*During the current period of fiscal restraint, substantially increased Federal research budgets seem unlikely; this, it may be worthwhile to explore the possibility of joint private/public efforts. The theme is explored in ch. 8.

Data systems can be linked and then used to identify technologies for assessment, and such systems can also be used, though to a lesser extent, to evaluate safety, effectiveness, and cost effectiveness. Prospective studies could be initiated to link technology use to health outcome and cost. One model which could be further examined is the Clinical Data Acquisition Plan which was being developed by NCHCT (see ch. 6). Data systems may be adequate in some cases to provide sufficient evidence of safety and effectiveness of technologies, especially if they are used to complement more rigorous testing methods.

FDA's postmarketing surveillance activities for drugs, mentioned earlier, are being developed to monitor adverse reactions to drugs (281). Such systems may be adaptable for other technologies as well.

The Federal Government has not used its potential leverage to test technologies through the reimbursement system. For instance, if HCFA could use its system to study whether new procedures were safe, effective, and cost effective, or needed further testing before final reimbursement decisions were made, many ineffective technologies might be identified and discarded well before they were accepted by the medical community. The ambiguous "reasonable and necessary" clause of HCFA's statutory language has been an obvious impediment to such activity.

Although the private sector has been actively involved in testing medical technologies, its direct support for well-controlled clinical trials has not been very extensive (except for the research monitored by FDA). Research protocols tend to be of a nonrandomized design and often rest on the information derived from available data bases and recordkeeping systems (e.g., 209). Nevertheless, there is evidence of increasing private sector interest in research on technologies. Much of the interest seems to stem from the belief and concern that resources are not used efficiently.

Finally, it should be noted that currentl, no public or private body has responsibility for determining either the cost effectiveness or social/ethical implications of medical technologies. FDA and NIH are both primarily oriented towards safety and efficacy issues. It ***is*** true that NCHSR and to

a lesser extent HCFA do selectively fund some cost-effectiveness studies, but no one body is charged with systematically examining the larger social issues.

OTA concludes that, in general, drugs and devices are adequately tested for safety and efficacy prior to being marketed. Medical and surgical procedures, however, are not well tested for either safety or effectiveness. No class of technologies is adequately evaluated for either cost-effectiveness or social and ethical implications. Finally, there is no organization whose mission it is to ensure that medical and surgical procedures are assessed for safety and efficacy or to evaluate any class of technologies for cost effectiveness and for social/ethical implications.

Synthesis

Synthesis activities in the area of medical technology assessment are generally of two major types: 1) synthesis of the results of individual research studies; and 2) synthesis of a body of research findings with other concerns such as risk, social, ethical, or cost factors. The former, which is more focused and technical than the latter, seeks to answer questions such as those concerning the safety, efficacy, or effectiveness of a given technology. The second, which is more policy directed, often seeks to develop guidelines or standards for medical practice or reimbursement policy. The value of the latter depends, in large part, on the adequacy of the former. That is, one cannot consistently set good policy regarding medical technologies without knowing what the collective research says about a given set of issues.

The challenge for synthesizing research evidence concerning a technology is to make sense out of a growing body of information-some bad, some good. Techniques available to do this were described in chapter 5.

Synthesis activities are inherently a part of conferences sponsored by individual institutes of NIH and other Federal agencies (and numerous other organizations). Among the more formal synthesis-type activities within the Federal Government are the consensus development conferences sponsored by OMAR of NIH.

The goal of consensus development is to synthesize the scientific literature on safety and efficacy/effectiveness and to recommend to physicians the appropriate use of technologies. In many respects, consensus development conferences are well done and important activities. As discussed in chapter 5, however, the NIH consensus conferences have demonstrated weaknesses in terms of objectively synthesizing scientific information and in recommending guidelines for the appropriate use of the technologies they consider.

For instance, although the NIH panels are generally composed of eminent physicians, a methodologist (i.e., a biostatistician or an epidemiologist) is not always included, and the validity of evidence from scientific research is not always explicitly examined (see app. C). Thus, the methodological limitations of a given study may be overlooked. Another limitation of NIH's format is the process itself. For instance, the use of adversary groups and task forces has been almost entirely abandoned, and the questions that have been posed are strictly limited to issues on which there is enough factual evidence to reach agreement. For the purpose of synthesizing available knowledge, this approach may be adequate (assuming that the available knowledge is all included and understood), but for the purposes of identifying gaps in knowledge and needs for future research, this approach is weak. Of equal importance is that consensus development conferences tend to examine in depth only two aspects of medical technology assessment: safety and efficacy. This limits the usefulness of the conferences and calls into question the appropriateness of their setting guidelines for clinical use of a medical technology (e.g., frequency of Pap smears or the use of mammography).

Setting medical standards (e.g., indications for using respiratory therapy) by professional organizations and governmental agencies, though not customarily characterized as a synthesis activity, does depend on the integration of available information. Ideally, these organizations and the individuals within them should first systematically and objectively review the clinical research evidence. A knowledge base (see ch.5), such as the National Library of Medicine's (NLM's) Hepatitis

Knowledge Base, may be useful in this regard. An important output should be the identification of fertile areas for further research. However, the common pattern is for standards to be set, whether by PSROs, HCFA, professional organizations, or NIH, that are based on the group's belief of good medical practice, much of which is unsupported by scientific evidence. Thus, not only are important opportunities lost for further research, but perhaps more important, current practice patterns tend to be validated when they should not be. Finally, not only is the research evidence generally not reviewed systematically and objectively, neither is the standard-setting process. Formal decision-assisting techniques such as Delphi and nominal group techniques are seldom applied.

OTA concludes that the synthesis phase of the present system of technology assessment is unnecessarily weak within both the private and public sectors. Research evidence regarding the safety, efficacy, and effectiveness of medical technologies is seldom examined systematically and objectively. Federal agencies and private insurers and organizations set policies, guidelines, and regulations, and/or make reimbursement coverage determinations, many of which profoundly affect the adoption and level of use of medical technologies. Yet, their decisions are usually based on informal, subjective, group-generated norms which tend to support the status quo. Formal, more objective techniques both for evaluating research evidence and for making decisions and setting policy could be used more often to aid in better decisionmaking.

Dissemination

The issues associated with making sure that the right people have access to technology assessment information transcends technology class (i. e., drug, device, procedure). However, the dissemination issue is particularly important for the decisionmaker at the point of a technology's adoption. At that point, the insurers, hospitals, physicians, or patients need to assimilate safety, efficacy, and cost information in order to make a rational decision based on their individual conditions, values, and objectives.

This report does not deal with the entire scope of information transfer. It does, however, brief-ly examine the ability of the Federal Government to make available research findings and the activities of NLM in indexing and providing access to the biomedical and other health-related literature that may be useful for medical technology assessment. These issues are addressed in greater detail in a separate technical memorandum entitled ***MEDLARS and Health Information Policy (276).*** That document also discusses the relationship between NLM and private sector organizations that index and provide computerized access to the biomedical and other health-related literature.

Specific problems associated with communicating information about medical technologies appear to be similar to those in other fields of science and technology. Paradoxically, the amount of information available is at once too much and too little. The "publish or perish" syndrome has led to an explosion in the quantity of literature without an accompanying improvement in quality. One way to ameliorate the problem of an overabundance of primary literature has been to rely more on secondary sources, particularly bibliographic data bases that can be read by a computer.

NLM has excelled in collecting, indexing, and making accessible biomedical literature by a computerized bibliographic system. An earlier OTA staff paper of this assessment indicated that about *76* to **98 percent** of the relevant biomedical journals are covered by NLM's major biomedical data base MEDLINE *(278).* But subject coverage of the health care delivery field by contrast is poor: less than *40* percent of all relevant citations were contained in MEDLINE. Many of the missing citations were in economic, business, and sociological journals. The coverage of citations in the health care delivery field is limited, not only because many of the citations are in economics and business journals, but because a large number are also older than the NLM health file. The percentage of relevant citations held in MEDLINE will be significantly greater for articles citing the more recent literature.

References to other sources of information on medical technology assessments such as ***mono*-** graphs, reports, conferences, and Government documents are not nearly as accessible in other bibliographic data bases as references to the journal literature. Thus, many useful Government re-

search documents may not be used or may have to await for the authors to publish the results in a refereed professional journal.

Along with the growth in literature in the biomedical field has come confusion on the part of many users about obtaining information. The large number of primary publications and even secondary publications (e.g., bibliographies) makes it difficult for the occasional user to find information efficiently. Users with access to a well-trained and competent information specialist or librarian find their search simplified. However, the quality of libraries or information centers and the quality of the staff vary. Furthermore, there is no comprehensive single source where information about existing federally generated data bases in a field can be obtained. This complicates even an informed user's search and has resulted in the unnecessary duplication of information.

Two important issues related to NLM's usefulness in the dissemination of technology assessment information are: 1) whether NLM should include more Government reports and other nonserial literature (especially in the area of health services) in its data bases, and 2) whether NLM should modify its indexing process to indicate more useful information as to articles dealing with research findings. With regard to the first issue, it should be noted that the National Technical Information Service (NTIS) has major responsibility for Government reports and perhaps NLM should not duplicate the collection, although NLM is expand-

ing its data base somewhat in this direction. An effort could be made to link existing data bases so that a single search could access both NLM and NTIS data bases as well as any other sources relating to health questions. With regard to the second issue, one possibility would be for NLM to carry a code within its citations that is related to the methodological and statistical nature of the article. The editors of research journals could be asked to supply the necessary information (276).

Finally, the potential impact of the widespread distribution of microcomputers in physicians' offices in the future could be significant. For instance, NLM's data bases could be immediately accessible, and if knowledge bases such as NLM's Hepatitis Knowledge Base were available, the dissemination of technology assessment information could be much enhanced.

OTA concludes that better methods need to be found to communicate information about medical technologies to physicians, researchers, and policymakers. OTA also concludes that Government-generated reports, many of which maybe important to technology assessment, are not as accessible as they could be. There is no mechanism through which all health-related Government reports can be identified or obtained. Finally, NLM's mission and capabilities should be examined to determine whether more Government reports should be included in its data base, and whether NLM should index articles to indicate their methodological and statistical nature.

CONCLUSION

Thus, OTA finds that there are major problems with each of the four components of the present system of medical technology assessment. The last chapter of this report provides Congress with options to address what appear to be some of the most striking weaknesses.