

6.

**STATUS AND IMPLICATIONS  
OF EFFICACY AND SAFETY  
ASSESSMENT**

# 6.

## STATUS AND IMPLICATIONS OF EFFICACY AND SAFETY ASSESSMENT

Efficacy and safety are extremely important starting points in determining if technologies will be safe and effective in use. If a technology does not provide benefit with acceptable risk under optimal, controlled, research conditions, then it will not do so under average conditions of use. Simply stated, efficacy is essential to effectiveness. \*

Chapter 1 briefly mentioned the general importance of efficacy and safety data. That theme is further developed in this chapter, which presents information on the uses and users of such data. This chapter also describes a normative model of the generation, processing, and dissemination of efficacy and safety information, and contrast current programs and systems for assessment to the normative system. Finally, it examines the status of information on efficacy and safety.

### USES AND USERS OF EFFICACY AND SAFETY DATA

Any person or institution using or directly affecting the use of medical technologies is a user of efficacy and safety information. There are two basic types of users: "passive" and "active." Patients or consumers of medical care often can be viewed as "passive" users of efficacy and safety knowledge. Many Government and private sector programs, for example, several of the grant programs of the Department of Health, Education, and Welfare's (HEW) Health Services Administration (HSA), also are "passive" users. HSA, for example, may award a grant to a community for the establishment of certain specific health services. The agency does not require that technological services provided with these funds be of demonstrated efficacy and safety. This situation represents a passive use of efficacy and safety information, because the usefulness of the grant program depends in part on the effectiveness, and thus the efficacy, of the services purchased. "Active" users of efficacy and safety information include physicians, biomedical and health services researchers, nurses, and other health professionals, many public and private third-party payers, and personnel in Government regulatory programs and medical schools, and so on. Table 8 lists many of these users of information, the uses, and the sources of information.

Information from well-designed and valid studies of effectiveness can be of higher utility than studies of efficacy to most of the users listed, because many of them are concerned primarily with the benefit of a technology under actual or average conditions of use. Because of the difficulty of conducting evaluations of effectiveness, information on effectiveness is often lacking. Efficacy information, the next best source of guidance on

---

● However, even if a technology were safe, efficacious, and effective, it might lack social benefit if overriding ethical or other societal concerns were not addressed satisfactorily.

**Table 8.—Users of Efficacy and Safety Information**

User	Actions taken on the basis of efficacy and safety information	Major sources of information
<b>Non-Federal public or private programs:</b>		
Physicians (and nurses, other health professionals)	<ul style="list-style-type: none"> <li>• Clinical decision making relative to diagnosis, treatment, and prevention of health problems</li> <li>• Decisions to adopt new technologies</li> <li>• Publishing, communicating to professional associations, colleagues, etc.</li> </ul>	<ul style="list-style-type: none"> <li>• Own experience</li> <li>• Colleagues</li> <li>• Professional meetings</li> <li>• Professional literature</li> <li>• Detail men, other manufacturers' representatives</li> </ul>
Professional associations	<ul style="list-style-type: none"> <li>• Set standards for use of technologies</li> <li>• Assess competence for certification, etc.</li> <li>• Communication to membership, etc.</li> </ul>	<ul style="list-style-type: none"> <li>• Professional literature</li> <li>• Experience of members and other health professionals</li> </ul>
Schools of medicine or public health	<ul style="list-style-type: none"> <li>• Instruction</li> <li>• Set agendas for future research</li> </ul>	<ul style="list-style-type: none"> <li>• Knowledge and experience of faculties</li> <li>• Professional literature</li> </ul>
Private sector third-party payers	<ul style="list-style-type: none"> <li>• Decisions to place a technology on the coverage schedule</li> <li>• Decisions to reimburse for specific uses of a technology</li> </ul>	<ul style="list-style-type: none"> <li>• Professional opinion</li> <li>• Professional literature</li> <li>• Associations</li> </ul>
<b>Federal Government programs:</b>		
Food and Drug Administration, PHS	<ul style="list-style-type: none"> <li>• Decisions to allow investigational use of drugs or devices</li> <li>• Decisions to allow marketing of drugs or devices</li> <li>• Decisions to allow products to stay on market</li> </ul>	<ul style="list-style-type: none"> <li>• Manufacturer or sponsor</li> <li>• Professional literature</li> <li>• Staff knowledge</li> <li>• Outside professional advisors</li> </ul>
Medicare program, HCFA	<ul style="list-style-type: none"> <li>• See private third-party payers</li> </ul>	<ul style="list-style-type: none"> <li>• Office of Health Practice Assessment, PHS</li> <li>• NIH, and other Federal programs</li> <li>• See private third-party payers</li> </ul>
Medicaid program, HCFA	<ul style="list-style-type: none"> <li>• See private third-party payers (HCFA recommends such decisions but the States have the decision authority)</li> </ul>	<ul style="list-style-type: none"> <li>• Medicare decisions</li> <li>• See private third-party payers</li> </ul>
National Institutes of Health, PHS	<ul style="list-style-type: none"> <li>• Decisions on research agendas</li> <li>• Decisions on demonstration and control programs</li> <li>• Dissemination of information</li> </ul>	<ul style="list-style-type: none"> <li>• Research conducted at or supported by NIH</li> <li>• Professional literature</li> <li>• Outside advisors</li> <li>• Staff knowledge</li> </ul>
Health Resources Administration, PHS	<ul style="list-style-type: none"> <li>• Set national guidelines for health planning</li> <li>• Develop planning guidance for certificate-of-need determinations</li> </ul>	<ul style="list-style-type: none"> <li>• Other Federal agencies</li> <li>• Contracts with private organizations</li> <li>• Professional literature</li> </ul>
Office of Professional Standards Review Organizations, HCFA	<ul style="list-style-type: none"> <li>• Set guidelines for medical care reviews</li> <li>• Set guidelines for reviews of institutional and length-of-stay admissions</li> </ul>	<ul style="list-style-type: none"> <li>• See Health Resources Administration</li> </ul>

•And State and private sector programs linked to HRA, such as health systems agencies,

••And private sector, local PSROs

appropriate use of technology, is therefore utilized more frequently. For example, regarding new technologies, there is usually little or no experience with them under average conditions of use for the development of even informal professional consensus of effectiveness. For these reasons, it is important to develop and disseminate the most valid and comprehensive efficacy and safety information possible, within resource and methodological constraints.

## A SYSTEM FOR ASSESSING EFFICACY AND SAFETY

The adoption and use of medical technologies by health care professionals should be based on well-validated information regarding their benefits and risks. This statement does not imply that every aspect of every technology must or can be subjected to randomized, controlled clinical trials. That would be an impossible task for several reasons, including financial and human resource limitations, the excessive time requirements, philosophical and political considerations, the complexity of medical technologies and their uses, etc. However, it does imply both the existence of accurate and relevant information, which is developed to the extent desired and practical, regarding the effects of technologies and the dissemination of such information to the individuals and groups in need of it. Also, this information should pertain to the benefits and risks of a technology under the conditions in which it will actually be used. Because of the difficulty of obtaining effectiveness and safety data, decisionmakers substitute efficacy and safety data as a somewhat equivalent measure of the technical effects of technology.

This section presents a model of the process of generating, processing, and disseminating information on efficacy and safety. This model is then compared to the current systems and programs in order to examine whether shortcomings exist in the current systems.

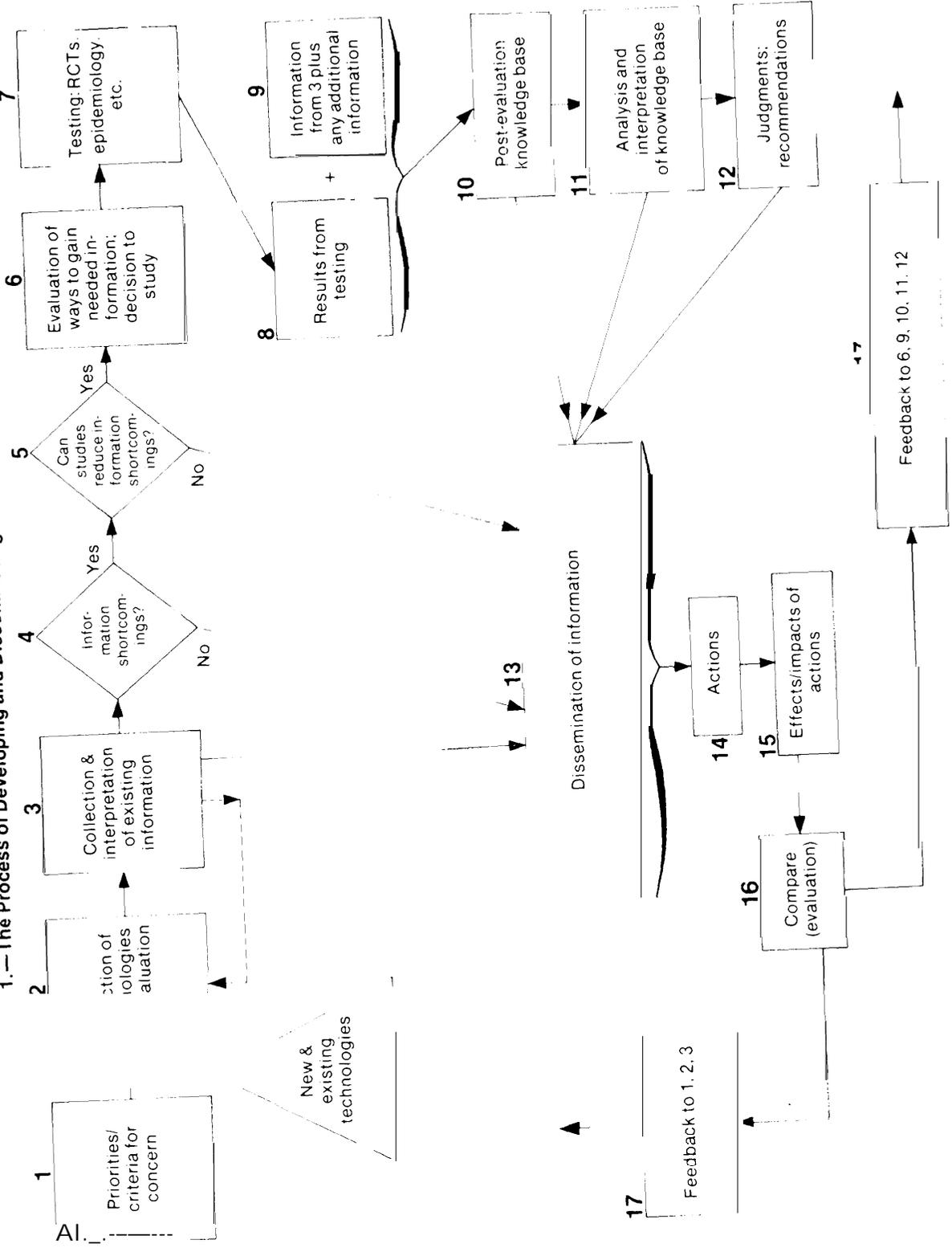
Developing and disseminating information on efficacy and safety is a tremendously complex process. Although many of the intricate details of the process are not germane for the purposes of this report, the complexity of this process should not be forgotten. To illustrate some of this complexity, figure 1 depicts many of the elements involved in assessing efficacy and safety. Even that relatively complicated process described in figure 1 represents a simplified abstraction of the reality.

In this report, the process is viewed as an interdependent and nondiscrete flow of four types of actions:

- Identification: Monitoring technologies, selecting those in need of study, and deciding which to study. (Steps 1-6 of figure 1)
- Testing: Conducting the appropriate analyses or trials. (Step 7)
- Synthesis: Collecting and interpreting existing information and the results of the testing step, and, usually, making recommendations or judgments of efficacy and safety. (Steps 8-12, and often 3)
- Dissemination: Providing the synthesized information, or any other relevant information, to the appropriate parties who use or make decisions concerning the use of medical technologies. (Step 13)

The action steps represented in figure 1 are not within the scope of this report. For a description of some of the possible actions, see table 8. Also, an HEW report on medical

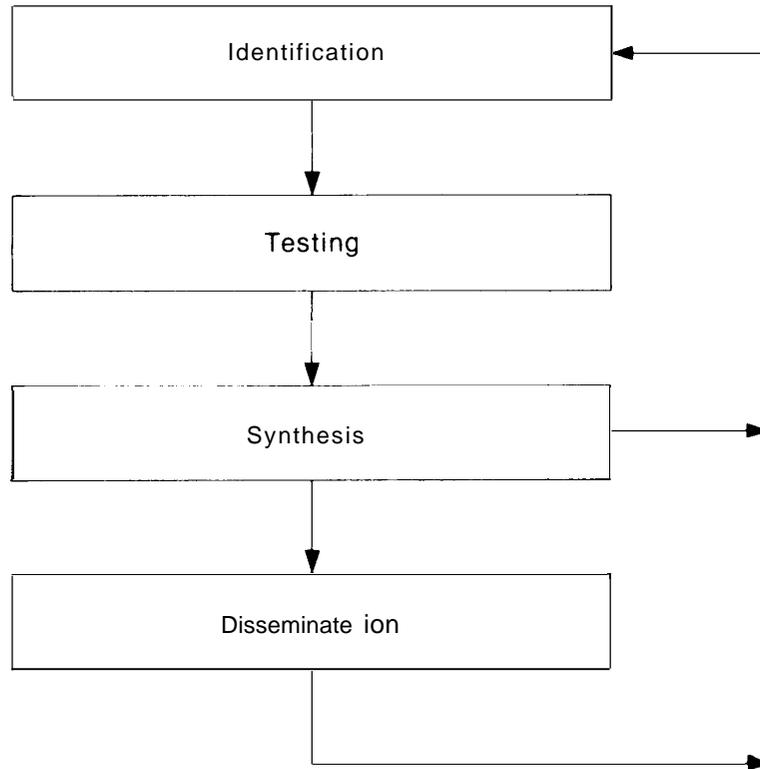
1.—The Process of Developing and Disseminating Information on Efficacy and Safety



technology management (369) describes in greater detail the potential actions and relates them to a similar model.

The four elements of a normative system for developing and disseminating efficacy and safety information are depicted in figure 2.

**Figure 2.—Simplified Process for Developing and Disseminating Efficacy and Safety Information**



This model represents only one possible method of viewing the process of assessing medical technologies. It is designed to serve as a logical standard against which existing assessment programs may be evaluated.

## SHORTCOMINGS OF CURRENT SYSTEMS AND PROGRAMS

The primary shortcoming in current assessment methods is the lack of a formal *or well-coordinated* “system” for developing and disseminating safety and efficacy data (53,250,357,369). Some elements of the process are operating and performing well. However, the elements are not linked together and do not follow each other logically. The Assistant Secretary for Health of HEW has stated (357):

There are, of course, informal mechanisms for the assessment of the health-care technology. It is probably true that such informal approaches served us reasonably well in the past. But for a variety of reasons, we can no longer rely on such informality.

HEW recognized the lack of a “strategy for managing medical technology. . . and . . . an analytical paradigm upon which to develop such a strategy” (369). A report to the Secretary in December 1977, outlined the components of “such a strategy.” Responding to that study, the Secretary of HEW established an Office of Health Technology in January 1978. The Office was designed to include these functions: testing and demonstrating the strategy developed in the 1977 study, serving as a focal point for health technology policy development in the Department, and providing recommendations to the Health Care Financing Administration (HCFA) on the advisability of reimbursement for specific medical technologies (287). As of September 1978, however, insufficient implementation of the proposed HEW system had taken place. Consequently, the Office of Technology Assessment (OTA) was unable to analyze the actual functions being fulfilled.

Development and dissemination of information on the efficacy and safety of drugs and devices more closely approximates a coherent system than does the assessment of medical and surgical procedures. Beginning in 1906 with the passage of the Federal Pure Food and Drugs Act, various laws have been enacted to regulate the safety and/or efficacy of both drugs and medical devices. Surgical and other procedures that depend primarily on providers’ techniques have not been subject to similar Federal controls. Assessment of safety and efficacy for these procedures has remained primarily in the hands of the profession.

There are a number of factors which help explain the differences in the safety and efficacy evaluations for products and procedures. One of these is the physical nature of products. Investigators can learn much about products before they are tested clinically (394). For procedures, however, clinical testing is the essence of their development. In addition, procedures are complex, and therefore, their evaluations are correspondingly complex.

Source of sponsorship also distinguishes products and procedures. Drugs and devices usually are developed for marketing by profitmaking firms. Mechanisms have been created to regulate industries. Procedures, however, are usually developed by an individual physician or medical team. Given the history of relative autonomy the medical profession has enjoyed in our society, it is not surprising that the profession has been given the responsibility for regulating its own members and their use of technology (125,332,334). It appears, therefore, that one major problem in assessing efficacy and safety centers on procedures which develop without control or planning in the private sector of medical practice.

## **Identification**

Presently, there is no complete list or catalog of either existing medical technologies or those that particularly require assessment for efficacy and safety. Partial lists do exist. The Food and Drug Administration (FDA), for example, has lists of approved drugs and devices. The fact remains, however, that many medical procedures, which are not on reimbursement schedules, but are important to assess (bed rest for certain diseases, for example) are not cataloged in one source.

No existing system completely identifies developing technologies that will need evaluation for safety and efficacy. The National Institutes of Health (NIH) does a yearly study of its clinical trials and publishes a catalog of those trials it supports. Other agencies, such as the Veterans Administration (VA), have similar catalogs or lists. Through

its premarket approval process, FDA gathers information on drugs and devices that are being developed. If medical and surgical procedures were to be evaluated before they came into widespread use, however, some comprehensive system for recognizing them in a timely fashion would be necessary. A variety of sources could produce such a catalog. Professional literature is one source. Another is institutional committees that review research for adherence to ethical standards. Complete lists of clinical trials would provide the beginning of an "early warning system."

Even if funds for, and numbers of, clinical trials were greatly expanded, setting priorities for study would still be necessary, because it is neither possible, nor desirable, to study every efficacy- or safety-related aspect of medical technology. Such priorities might help to ensure that all areas of medicine, such as prevention, are considered. Priorities for assessment might include beneficial technologies that are neglected or technologies that are suspected to be useless or dangerous. Technologies that are, or are expected to be, either expensive or widely used also could be given priority. For new technologies, potentially important advances could be assessed rapidly.

In sum, there is no formal process for selecting which technologies are to be studied; indeed, there is not even a set of priorities for such selection. New drugs and new devices are, however, subject to the FDA market approval process and thus are automatically identified for study, at least in regard to the efficacy and safety claims of the manufacturers.

## **Testing**

The testing phase includes stimulating, requiring, funding, or conducting studies. Shortcomings related to the testing phase center around four issues: 1) the state of the methodologies for conducting controlled trials, consensus activities, and other tests; 2) the level of financial support, particularly for controlled clinical trials; 3) the relative appropriateness of the questions and technologies being studied; and 4) the number of personnel qualified to conduct such research.

Although the state of clinical trial methodologies has improved dramatically in the past 30 years, there are still uncertainties involved in the design of each trial. This report is not directly concerned with the technical methodologies for testing, but it should be noted that "there is no standard textbook on clinical-trial methodology" (147), and that the further development and dissemination of methodological information would complement efforts to assess efficacy and safety.

There is no "correct" level of financial support for clinical trials; no one can set an exact figure for the amount that should be invested in trials and other forms of testing. Does the current level of funding, then, represent a shortcoming? This question must be answered positively because important areas of health care are not receiving adequate investigation, according to the evidence gathered by OTA. New or developing immunization and screening technologies and new procedures are studied relatively infrequently, as are existing technologies of all types. This discussion applies to both the second and the third shortcomings listed above.

Often, the decision to investigate a certain question (for example, what specific effects of a technology are being examined?) has been influenced by such factors as investigator curiosity, research needs, and so on. \* The concerns and information needs, for

---

\*Many of the shortcomings of the testing phase are intimately related to the inadequate identification phase.

example, of health planning agencies or Professional Standards Review Organizations (PSROs) are much less frequently considered in these decisions (369). Changes in the level or direction of the Nation's activities in assessment of efficacy and safety would highlight the limited number of personnel presently qualified to conduct such research. Biostatistics and epidemiology have been less affluent areas of health research (57). Consequently, the number of epidemiologists, statisticians, and others essential to efficacy and safety assessments may be inadequate for future needs.

In short, the country has the potential to develop a good capability for testing efficacy and safety, but the actual effort could perhaps be expanded or at least organized according to somewhat different priorities. Such an effort may require an expanded base of qualified research personnel.

## Synthesis

Synthesis involves a critical analysis of the results of testing (available data from preclinical to clinical experience, epidemiological studies, and controlled trials) and all other available and relevant information. This analysis involves a "putting together" of the data into a summary of the efficacy and safety of the technology in question. It usually takes the form of judgments or recommendations regarding the appropriate indications for use of the technology. Consensus development, which is described in chapter 4, also can be considered a synthesis activity. Syntheses are most commonly found as review articles in the medical literature. However, this literature varies in quality and is usually not directed toward the needs of practitioners. Williamson notes that "many, if not most, health sciences publications are detailed, highly technical research reports directed by the investigator to his fellow researchers, " and that "interpretation of many. . . requires an understanding of technical terminology, research design, and analytical statistics that is beyond the scope of the average professional. . . ." (428).

The validity of published information also has been questioned. Two studies of research reports in leading medical journals found nearly 75 percent of the publications analyzed to have invalid or unsupported conclusions as a result of statistical problems alone (**115,300**). Other studies that focused on research design, data collection, and analysis in specific areas of medicine found that none of the articles studied yielded valid or supportable results (137, 189). When Juhl and his coworkers examined the literature in gastrointestinal diseases, they found that few well-designed trials were conducted. Additionally, they observed a preponderance of positive trials, indicating a bias toward positive results (186). Furthermore, 80 percent of the trials dealt with new treatments; few were concerned with evaluating "established treatments. "

Federal Government synthesis activities are expanding. The consensus development activities of NIH are too new for evaluation of their effects. The hypertension synthesis (see chapter 5) seems to have had positive impact. Many of the consensus exercises planned for 1978 by the Institutes of NIH, however, appear to be modifications of seminars and conferences planned previously. How well these activities fulfill the synthesis function remains to be seen, but there is great potential. The Alcohol, Drug Abuse, and Mental Health Administration (ADAMHA) has used a technique related to consensus development in the area of psychosurgery. However, that agency contends that a more formal and quantitative technique should be developed. The process of recommending coverage decisions to Medicare (chapter 5) by the Office of Health Practice Assessment represents another synthesis activity. That Office has stated, however, that because of the ad hoc nature of the process there is "no assurance that the best and most reliable data are utilized in a given case" (369).

Despite the recent expansion in synthesis activities, they still represent a modest level of activity that have suffered, at least in part, from lack of quality in both content and process. Furthermore, synthesis activities are hampered by the lack of well-validated information on efficacy and safety.

### **Dissemination**

Many of the comments relating to synthesis also apply here. Federal agencies have not assigned a high priority to disseminating information. FDA sometimes sends letters to all physicians as one mechanism for distributing important information. The National Center for Health Services Research (NCHSR) frequently disseminates information to a wide audience by issuing a series of NCHSR Research Reports that describe the results of projects funded or conducted by that agency. Also, NIH has provided information primarily to the professional community through its demonstration and control projects, through the National Library of Medicine, and through other activities, including a regular feature in the *Journal of the American Medical Association*.

As described in chapter 5, the private sector also has multiple channels which encourage the flow of information. Professional societies are expanding their activities in this area.

The Federal Government provides little information for such public agency activities as health planning programs. In the case of the computed tomography (CT) scanner, for example, the Bureau of Health Planning and Resources Development, the Federal agency which administers health planning activities, contracted with a private firm to produce planning guidelines for such devices. Likewise, third-party reimbursers, such as the Medicare program, seldom receive assistance from such agencies as NIH in deciding benefits.

## **STATUS OF EFFICACY AND SAFETY INFORMATION**

The shortcomings described above would be much less deleterious if the state of knowledge about the efficacy and safety of medical technologies were adequate. Conversely, if the state of information were inadequate and there were no shortcomings in the processes and systems of assessment, perhaps little could be done to improve the information base. The data inadequacies, and the corresponding difficulties in using technologies, might then be the inevitable result of the inherent complexities in the field of medicine. However, there *are* shortcomings in the current ways in which efficacy and safety information are developed and disseminated. Therefore, data inadequacies and their effects— inappropriate diffusion and use of technologies—need examination.

Many technologies have been shown to lack efficacy or be unsafe only after enjoying widespread use. A psychosurgical procedure called leucotomy or lobotomy, for example, was widely adopted in the early 1950's and was subsequently abandoned when its efficacy and safety were seriously challenged. The Wassermann test for diagnosing syphilis was used for over 40 years until it was discovered that only half of the patients with positive test results actually had the disease (223). More recent examples include internal mammary artery ligation (see chapter 3, case 8), colectomy (surgical removal of the large intestine) for epilepsy (162), carotid-jugular shunts for mental retardation, lumbo-dorsal sympathectomy, uterine suspension, and gastric freezing.

Questions of efficacy have been raised recently regarding a number of medical technologies currently in use (72,124,162,179,223). As mentioned earlier, White has stated that only 10 to 20 percent of all procedures used in present medical practice have been shown to be of benefit by controlled clinical trials; many of the other procedures may not be efficacious (426). In fact, many technologies in use have had their efficacy and safety questioned, including oral drug treatment for diabetes (64,236), respiratory therapy (19,24), oral decongestants (207), thermography for diagnosing breast cancer (248), ergotamine for migraine headache (410), immune serum globulin for preventing hepatitis (303), intensive care for pulmonary edema (152), coronary care units (233), and radical mastectomy (228).

Such widely used technologies as tonsillectomy, appendectomy, and the Pap smear have not been completely assessed for efficacy (see chapter 3, cases 1, 9, and 10). Others, such as electronic fetal monitoring (EFM) and coronary bypass surgery, have been diffused rapidly before careful evaluation (see chapter 3, cases 7 and 8). Concern about risks has led to questions regarding the use of mammography and skull X-ray (see chapter 3, cases 4 and 6).

The above are only examples. Others could be listed. The systems for assessing efficacy and safety have made the compilation of such a list possible. However, the same systems were not able to provide early and adequate information in order to prevent or delay the spread of technologies until their effects had been predicted more clearly. Further, since these examples can be cited, there are probably many others. Although perfect information on efficacy and safety can never be attained, shortcomings in assessment systems may be impeding a closer approximation of that goal. The status of efficacy and safety information cannot be exactly determined, but the combination of long lists of examples of technologies inadequately assessed and shortcomings in assessment procedure processes may indicate that improvement is possible.