
Chapter 4

Evaluation of Medical Technologies

Contents

	Page
Introduction	29
Methods of Evaluation	31
Preclinical . . 4	31
Informal	31
Epidemiological and Statistical	32
Controlled Clinical Trials	32
Formal Consensus Development	33
Cost-Effectiveness Analysis/Comprehensive Technology Assessment	34

Evaluation of Medical Technologies

INTRODUCTION

The process of biomedical research and development (R&D), from basic through applied to development research, produces new medical technologies. As noted in chapter 2, the pace of this process has been accelerating. Frequently, the benefits of new technologies have been clear and convincing. There are numerous examples of new equipment, drugs, and procedures which have significantly advanced the practice of medicine. Some diseases can now be effectively prevented, and medical innovations such as antibiotics have provided effective therapies for a number of other diseases. New diagnostic techniques have often made it possible to detect disease in time to apply an appropriate therapy. Even in cases of diseases for which no effective preventive or therapeutic measures are available, technologies have aided in relief of pain, amelioration of symptoms, and rehabilitation of individuals affected by chronic conditions (88). Finally, some new technology has increased access to health care, some has reduced the cost of care, and some has improved the outcome of care (45).

Yet advances in medical technology development have not occurred without concerns, particularly recently. They may be outlined as follows:

- A number of advances involve significant risks, some intrinsic and some which vary according to the setting in which and the skill with which they are applied. All invasive procedures, including the administration of drugs, surgery and the use of equipment, involve some finite risk to the patient. However, determination of the *safety* of new technologies is crucial, because some level of the risks that may be encountered must be judged acceptable in relation to the potential benefits.
- Many technologies have been widely diffused before their efficacy has been established. Concerns about efficacy are raised when a new technology is introduced without proof of its efficacy (e. g., electronic fetal monitoring (3)), when a widely used technology is later shown to be ineffective (e. g., oral anticoagulants in the treatment of myocardial infarction), or when the relative efficacy of alternative therapies is compared (e. g., the radical mastectomy) (88).
- Health care costs are escalating rapidly. The expanded use of medical technologies is an important factor in the rising costs, imposing *economic* burdens which cause problems for patients, for their families, and for society. Medical technologies contribute to medical care costs in various ways: Some have large capital investments, some require the use of costly supportive services, some present the possibility or requirement of costly followup care, some establish the need for continued use, some are overused after initial proof of reliability of efficacy (particularly diagnostic technologies), and some are used for inappropriate purposes.
- An increasing number of technologies raise *ethical* issues. The concerns may center on the use of the technology (e.g., as amniocentesis or renal dialysis), or on the use of human subjects during research on the technologies (e. g., as many cancer drugs).
- Medical technologies also raise other social issues. For example, with the advent of life-extending technologies such as artificial

● Efficacy refers to the probability of benefit to individuals in a defined population from a medical technology applied for a given medical problem under ideal conditions of use. Effectiveness, a term used interchangeably with efficacy by some, refers to the benefit of a technology under average conditions of use (85).

hearts and kidneys, modern technology has challenged society's traditional view of death and dying (88). Critics of the increased use of technologies charge that medicine is being dehumanized by the use of machines and scientific methods (85). Legal issues may arise in several areas, including allocation of resources, liability, and informed consent.

These concerns cover both technical and social issues. Although these sets of issues and the categories within them are often separated for discussion purposes, in fact they are inextricably linked. For example, ethical considerations, seemingly remote from technical matters, can hamper the determination of medical efficacy of some technologies. Although different methods are used to assess the technical and social impacts of new technologies, it must be recognized that problems (and their solutions) cannot truly be separated (88).

Because of these concerns, increasing attention is being focused on the process of medical technology development and use. * In principle, new technology should be introduced into the practice of health care when its benefits to society or individuals outweigh its costs. In practice, however, knowledge of either benefits or costs is often very limited at the time decisions (either formal or informal) regarding the dissemination of a new technology are made (45). Evaluation of technologies is the process in which the knowledge of benefits and costs is gathered and synthesized. It occurs, or should occur, after development and before diffusion and use. Thus, it is a vital component of formal technology transfer.

Like R&D, evaluation covers a broad spectrum of activities. These activities vary according to the nature of the technology being evaluated and according to the criteria being used. The historically most common, and perhaps most important, criteria used in the initial stages of evaluation of health-related technologies are safety, efficacy, technical feasibility, and tech-

nical performance. For commercial products (or potentially commercial products, even if developed with public or nonprofit funds), another basic criterion is potential profitability. Other evaluation criteria will then follow, including: effectiveness, reliability, suitability for the goals of its use, cost, cost effectiveness, affordability, potential or actual reimbursement status, repairability, convenience, esthetics, consumer satisfaction, social implications, legal impacts, patient protection, ethical concerns, and so on (93).

Clearly, some evaluation criteria pertain only to "product" technologies, such as devices or drugs. Other evaluation criteria pertain to the medical purpose for which the technology is used. Efficacy and safety, however, are the basic starting points in evaluating the overall utility of a technology. Other criteria, such as legal concerns, are rarely needed if the technology is shown to be inefficacious or unsafe. And, efficacy and safety information is often needed for evaluations of cost effectiveness or potential for reimbursement, for example (85).

The specific objectives of any evaluation depend on the specific criteria being used. In general, the purposes of evaluating medical technologies are:

- To ensure that technologies demonstrated to have potential benefits with acceptable risks are made available rapidly in the private and public sectors. Administrators of public regulatory and financing programs could make sounder and faster decisions regarding the use of the technologies with such information.
- To constrain the diffusion and use of technologies which either lack efficacy or cause excessive harm or whose total societal costs are judged greater than total societal benefits.
- To guide *appropriate* use of all technologies, because technologies are rarely completely inefficacious, unsafe, or undesirable to society.

Thus, the overall goal of evaluation is the production of information that can be used to affect the technology transfer process.

● Ch. 2 describes this process.

METHODS OF EVALUATION*

There are numerous methods available for evaluating medical technologies. The method used varies according to the criteria for evaluation (i. e., efficacy, safety, effectiveness, etc.) and according to the nature of the technology being evaluated (i.e., drug, device, procedure, etc). No technique is universally applicable for every technology. In many instances, less complex methods may be more appropriate than the sophisticated approaches. Frequently, combinations of techniques are used (85). And some methods, particularly those used for evaluating social impacts, are actually combinations of other methods.

This section describes six groups of methods: preclinical, informal, epidemiological and statistical, controlled clinical trials, formal consensus development, and cost-effectiveness analysis/comprehensive technology assessment. Generally, the first five groups* are used in evaluation concerned with technical issues relating to health effects (such as efficacy and safety), while the last two are used in evaluations concerned with social issues. It must be clearly stated that the categories are created to facilitate description; there are overlaps and combinations of methods in the categories used. In particular, the methods used for evaluating social impacts utilize some of the methods in the previous categories. Traditionally, clinical experience, based on informal estimation techniques, has been the most important. Other techniques, such as epidemiological studies, randomized controlled clinical trials, and formal consensus development are being used increasingly.

Preclinical

Many medical technologies are evaluated in biochemical and animal tests prior to human experimentation. There are two purposes for these tests: 1) to gather preliminary evidence to gain

the right to test with humans, and 2) to develop performance standard compliance to establish marketability. Biochemical tests include chemical analyses for purity, quantity, and quality of the active agents; analyses for potential pharmacological activity of filler and stabilizing substances; determinations of biocompatibility; and tests for long-term dissolution of body fluids and the possible presence of toxic residues in the production of plastic materials. Animal testing provides a guide to capacity to induce toxicity as well as potential therapeutic activity. Determining the degree of toxicity, or safety, is the major function of animal tests.

A controversial issue is the accuracy of animal models in determining the probable effects of drugs on people. Questions that arise include short-term high dose v. long-term low dose, animal species selection, population size, and controls. These questions are particularly pertinent with respect to carcinogenic agent evaluation. However, despite the inherent problems, an earlier OTA report concluded that animal tests are acceptable models for cancer studies and should probably be regarded as reasonable precursors to clinical studies (87).

Informal

The increasing need to formally evaluate medical technologies, particularly for efficacy and safety, has been described. However, the majority of such evaluations are still based on informal approaches. They may take place during medical school and specialty training and through personal or peer experience. Physicians and other health personnel are constantly exposed to medical technologies throughout medical school, residency, and special courses. Usually, students assume that these technologies are efficacious and safe. Some of them have undergone formal assessments, but most are recommended based on previous experiences or training received by the instructor. Personal experience, the actual use of technologies, is a common qualitative method used to assess both efficacy and safety (and other evaluation criteria). Although it has limited statistical value and

*This section is based largely on OTA's previous report, *Assessing the Efficacy and Safety of Medical Technologies* (85). The section "Cost-Effectiveness Analysis/Comprehensive Technology Assessment" is derived from the 1980 and 1976 OTA reports (89,114).

*Formal consensus development belongs in both categories.

lacks control over scientific quality, it may be advantageous in some cases. Personal knowledge of a patient, for example, may promote beneficial adjustments to the type and level of treatment. Peer experience is more explicit than personal experience; information may be exchanged by items such as journal articles, pamphlets, and personal communication.

Informal techniques are based on the clinical approach of qualitative, artful decisions as compared to the scientific approach of quantitative, mathematical decisions. In any comprehensive system of evaluation, there is a place for both approaches, since each extreme may be appropriate in certain situations. In addition, many assessments require combinations of techniques. Furthermore, cooperation between clinicians and statisticians must exist to attain appropriate decisions when the more rigorous techniques are used.

Epidemiological and Statistical

Epidemiology is the study of the determinants and the distribution of diseases and injuries in human populations. It also incorporates the study of the impact of medical interventions on diseases and injuries. There are a number of epidemiological methods useful in evaluating the effects of medical technologies. Each of these methods involves the collection of data, for at least two groups, on disease manifestations, on changes after the medical intervention (or lack of it), and on certain factors which may be associated with the determinants or distribution of the disease or injury under study. Once the data are collected, statistical analyses are performed to compare the two groups.

The methods differ in the types of data collected, the way the groups for study are selected, and the time frame studied. Retrospective studies compare groups of people who have a disease with those that do not. These studies are designed to determine whether the two populations differ in terms of percentage exposed to certain critical factors. The relationship between oral contraceptives and thromboembolism was established this way. Most information used in retrospective studies is derived directly from the

patients, their relatives and friends, and their medical and other records. Thus, there may be doubt about the uniformity, accuracy, and completeness of information (especially on death certificates). In addition to incomplete or biased data, the selection of appropriate comparison groups presents a major problem with this method. There are advantages with the method, however, especially utility, low cost, and quick results.

Prospective studies follow the histories of persons both exposed and unexposed to a critical factor under study. The incidence of deleterious effect (or improvement) resulting from such exposure is then determined for persons in the two groups. A major advantage of prospective studies is the relatively clear designation and selection of both the study and the comparison groups by means of matching characteristics with minimum bias before the disease develops. Disadvantages of these studies include their high cost and long latent periods before results are obtained, the possible occurrence of changes in patients and methods over the duration of the test.

Computer modeling and simulation are methods used most effectively in evaluation when mechanisms of a technology are understood. By simulating physiologic conditions on the computer, the evaluator can apply the technology and obtain information about its effects in different clinical situations without ever involving patients. A major drawback to these methods is that the means of applying them is not yet adequately developed. In addition, they require a fair amount of knowledge about the effects of the technology in order to apply them. They may be particularly useful in evaluating effectiveness of certain technologies, however, and provide information at an accelerated rate at less risk to patients (45).

Controlled Clinical Trials

Controlled clinical trials are a powerful tool in evaluating the impact of technologies on individuals, because they involve the actual controlled application of the technology and objective observation of the results. Perhaps the most

important type of controlled clinical trial is the randomized clinical trial (RCT). * In an RCT, patients who agree to participate are randomly assigned to one of two (or more) groups: one which is exposed to the experimental treatment, and one which is exposed to the standard treatment which may be no therapy such as a placebo (for comparison with a new treatment) or a variation (e. g., a different dosage) of the experimental treatment. Clinical tests and examination of the members of each group are used for evaluations of the relative benefits and risks of the technology.

The principal advantage of RCTS is that they have high internal validity, i.e., they permit relatively unambiguous conclusions as to whether the observed effects of a treatment are due to the technology or some other factor(s). RCTs are the most useful when: 1) the benefit of a new technology is uncertain, or 2) the relative benefits of existing therapies are disputed. There is much statistical theory that supports the scientific utility of the randomization procedures in these trials. And, if a large sample of patients and conditions are tested, external validity (the generalizability of the observed effects to other patient populations, settings, or conditions) may be high.

Yet RCTs have a number of problems. The most controversial problems are ethical; they are based on a concern for both patient and physician rights and responsibilities. Critics of randomization point out that: physicians must make clinical judgments and act according to their consciences (which is precluded by acting according to a protocol); personal physicians must influence whether their patients enter a trial and what treatment is administered; patients must be given the best possible information in consent forms; and patients should be able to choose which treatment is delivered. Other criticisms do not focus on randomization, but instead on the processes used in the trials. Questions about the rights of patients are raised, particularly for children. For example, when can informed consent be given by a child? at what

age? with what medical conditions or illnesses? And, who, if not the child, will guard those rights? The long-term effects of treatments or other medical technology interventions can be serious and long in evidencing themselves, particularly in children.

There are those who defend the ethics of using controlled clinical trials. One reason is that physicians can not do just what they "believe" best, since their practice should be based upon sound scientific evidence. Further, if each patient is so unique as to be ineligible for statistical randomization, how can the individual physicians use clinical judgments based on past experience as the optimal guideline for determining the treatment of the next patient? Another defense of RCTS states that the rights of patients are protected in their ability to refuse participation in the trial.

There are also more practical problems involved in the use of RCTs. One is that many trials require a long period of time and large commitments of money, resources, and subjects. In addition, they can be difficult to conduct in settings such as hospital clinics and physicians' office. RCTs can also be especially difficult to conduct for technologies that are already widely diffused. In these situations, administrators and clinicians may be reluctant to make the changes in policies and procedures necessary to conduct the trials. Finally, a priori conclusions on the treatment being evaluated are a major obstacle to conducting RCTs, since such conclusions may subvert the randomization process.

Overall, there are no unequivocal answers to the concerns raised. In general, many articles note the problems, but recommend cautious use of the technique.

Formal Consensus Development

Formal consensus development is an evaluation method which synthesizes evaluation results from earlier, more specific studies. It is generally employed when evidence from previous studies does not lead to an unequivocal decision on the effectiveness, safety, etc. of the technology under consideration. A consensus group is a panel of experts formed both to

*Controlled clinical trials which do not involve randomization are also important, particularly if the treatment may have the potential of curing the disease but killing the patient,

evaluate all pertinent available information, which may range from informal to detailed statistical studies, and to recommend its findings to the medical community.

There are two types of consensus groups relevant here. One type of group evaluates the current state of efficacy and safety knowledge regarding either a particular medical technology or technologies that relate to a specific medical condition. This type of group is found at NIH and will be discussed in depth in chapter 5. A second type of group both analyzes a medical technology, particularly devices, and recommends possible standards to be used in the conduct of future assessments.

Cost-Effectiveness Analysis/ Comprehensive Technology Assessment

This category of methods represents evaluation techniques whose primary feature is that they are actually formal processes. As such, they incorporate other methods of evaluation. Both require basic information on the technical impacts of the technology being assessed and are used when the evaluation criteria are "social" in nature. Another characteristic of these methods is that they are intended to be decision-assisting ones.

Cost-effectiveness analysis (CEA) can be thought of as a synthesis of both the health effects and the economic effects of a technology. In an earlier OTA study, *The implications of Cost-Effectiveness Analysis of Medical Technology* (89), this method was studied in depth. OTA found that the value of CEA lies more in the process of performing the analysis than in any numerical results which are derived from it. In addition, there is no one "correct" way to do an analysis. The most appropriate approach to CEA and similar methods is to perform it in an open forum such that assumptions and underlying values can be challenged; to identify, measure, and, to the extent possible, value all relevant benefits and costs; and to present the results of the analysis in an "array" of effects rather than forcing them into some aggregate single measure.

Comprehensive technology assessment is a form of policy research that evaluates the short- and long-term social consequences (e. g., societal, economic, political, ethical, legal) of the application or use of technology. Like CEA, comprehensive technology assessment was the focus of an earlier OTA report, *Development of Medical Technology* (88). The principles that apply to CEA also apply here; the major difference is that comprehensive technology assessment covers a broader range of factors, especially those of a social nature.