

3. Evaluating Health and Economic Effects

Ignorance never settles a question.

— *Disraeli*

Contents

	<i>Page</i>
Introduction	33
Evaluating Health Effects	33
Experimental Studies	34
Randomized Clinical Trials	34
Observational Studies.....	35
Cohort Design	35
Case-Control Design.....	37
Summary	38
Evaluating Economic Effects	39
Opportunity Cost	39
Marginal Valuation	39
Joint Production Considerations	40
R&D Costs	40
Overhead Costs	40
Cost v. Prices	40
Discounting	41
Conclusion	41

Evaluating Health and Economic Effects

INTRODUCTION

Evaluating health and economic effects of medical technologies is central to any assessment; indeed, some would argue that evaluations of health and economic effects are the essence of an assessment.

The first section of this chapter discusses health effects. The main technical decision to be made when testing for the health effects of a technology is which study design is most appropriate. This section describes the study designs available and

compares the designs presented in terms of their validity. Additional material on methods used to evaluate health effects is presented in appendix C.

The second section of this chapter concerns economic effects. It is primarily drawn from portions of a previous OTA report entitled *The Implications of Cost-Effectiveness Analysis of Medical Technology* (270). This section provides the reader with a brief discussion of the issues involved when evaluating economic effects.

EVALUATING HEALTH EFFECTS

Despite recent attention to the economic and social impacts of medical technology, the most critical aspect of the use of medical technologies remains their effect on health. An evaluation of health effects may examine efficacy (or effectiveness), safety, or both. *Efficacy** is the health benefit as measured under controlled conditions (such as those existing in a randomized clinical trial). *Effectiveness* is the benefit of technology under average conditions of use. Efficacy or effectiveness generally measure the intended effects of the use of a technology. *Safety* is a judgment of the acceptability of the risk** involved in using a technology.

There are many similarities between efficacy and safety —e. g., both are relative concepts and thus are discussed in terms of probabilities. Very importantly, however, their measurement may require different study methods. They differ in several key factors. In assessing efficacy, a study is usually oriented to a limited number of specific

benefits. The measurement of safety, however, usually involves a study design that is able to identify a broad range of risks; such risks are often unknown or unexpected, they may occur far in the future, and they may affect only a small percentage of individuals. These factors imply that efficacy and safety are not simply the plus and minus columns of a single measure. Each requires separate attention, although judgments of the importance of either a benefit or a risk should only be made in relation to the other.

There are methodologic principles that guide the design, conduct, and interpretation of any particular investigation. Specific methods for evaluating health effects of technologies are described below. Each method has its strengths, weaknesses, and limitations for detecting favorable or unfavorable outcomes associated with a technology.

Of particular concern in research design and analysis is the validity of the findings, which varies with the study design. Validity refers to the extent to which a situation (as observed or evaluated by other criteria) is reflective of the “true situation.” Four components of validity have *been* described (68). *Internal validity* refers to whether the observed effects of a medical technology, under the conditions of the study, are attributable to the technology and not to some other factors.

● OTA defines “efficacy” as “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” (266).

* Risk is a measure of the probability and severity of harm to the health of individuals in a defined population associated with use of a medical technology applied for a given medical problem under specified conditions of use (266).

Statistical conclusion validity, which is a subset of internal validity, refers to the appropriateness of statistical tests and their ability (or power) to determine whether observed effects could be explained by chance fluctuations and to detect true differences in performance of the technology under study. *External validity* refers to the generalizability of the observed effects to other patient populations, settings, or conditions. *Construct validity* refers to the adequacy of the theory that an investigator has about what makes the technology effective.

The appropriateness of a particular study design is dependent on the purpose of the study, the methods available, the effects to be measured, and the technology's pattern of use. The choice of method is also influenced by other factors such as ethical concerns, limits on the number of participants available for study, the need for timely results, and available budget.

The discussion that follows is focused on select study designs which are commonly used to assess the health outcomes of a technology.

Experimental Studies

Experimental studies are characterized by the intentional application of a technology to a study population, and subsequent observation of effects. These studies must be carried out prospectively. They are frequently used prior to the dissemination of a technology, but can also be employed after the technology has diffused.

Randomized Clinical Trials

Randomized clinical trials (RCTs) are considered the most definitive experimental method for evaluating the efficacy or health benefits of a technology (60,148,187). An essential element of an RCT is randomization. Patients in an RCT are randomly assigned to one of at least two groups: one or more study groups, in which subjects are exposed to the experimental treatments, and a comparison group, in which the subjects are exposed to the control condition. The control condition can be either no treatment, the standard treatment (for comparison with a new treatment), or a variation (e.g., a different dosage) of the experimental treatment. The basic question to be

answered in an RCT is: Are the effects observed in the experimental group also observed in the comparison group? If the answer is essentially "no," the effects observed in the experimental group can be attributed, within the limits of probability, to the treatment technology.

RCTs are a family of designs that vary in size and complexity. The number of treatment conditions (e.g., dosage levels) can vary, as can the size of population tested and the statistical power* of the study. Small RCTs may be performed early in the development of a technology to demonstrate or test the efficacy of the technology's innovative elements. Large-scale, multicenter trials can be conducted at a later stage in the development of a technology to establish its efficacy and safety across a large population and in diverse settings (266), as well as to increase the statistical power that results from a larger sample size. A major goal of the multicenter trial is to improve external validity in regard to larger populations.

Sometimes a favorable or unfavorable outcome is observed (i.e., a participant gets better or worse) because the participant *believes* that the treatment will work or believes the treatment is harmful. This "placebo" effect,** psychologically related but nonetheless real, results in a change in the participant's condition. Further, the effect may be influenced by the investigator's expectations. To reduce potential bias from the placebo effect, treatment can be offered under conditions where the participant ("blinding") or both the participant and the health care provider ("double blinding") are not aware whether the participant is given the experimental or the control treatment. Another layer of "blinding" is added when the person analyzing the data is not told which group is the experimental and which is the comparison. That person may be a statistician, but frequently is a medical specialist, and also may be the provider.

The principal advantage of RCTs is that they have high internal validity, i.e., they permit relatively unambiguous conclusions as to whether

● The "power" of a study is the probability of its detecting an effect (of technology being tested) when one actually exists. The greater the power, the less likely one is to incorrectly reject an effective technology.

● *Although the placebo effect is discussed here under RCTS, it is not peculiar to such studies.

the observed effects of a treatment under the conditions of the study are due to the technology or some other factor(s). Randomization protects against potential selection bias in assignment of subjects to experimental and comparison groups. Within the limits of sampling error, the only difference between the groups is that the experimental group is given the treatment under study and the comparison group is not. Therefore, differences in outcome can be attributed to the differences in treatment, with a known probability of error due to chance.

Although well-designed RCTs are generally high in internal validity, they do not *necessary* resolve the problem of external validity (68). External validity is usually established only when large heterogeneous samples of participants are tested under a variety of circumstances, typically across a number of studies, or through large multicenter RCTs with carefully selected populations.

A disadvantage of RCTs is that they can be difficult to carry out in settings such as hospital clinics and physicians' offices and can be especially difficult for technologies that are already widely diffused and perceived as being effective (253, 401). In such situations, administrators and clinicians may be reluctant to make the changes in policies and procedures needed to conduct an RCT. Preexisting conclusions on the treatment being evaluated are a major obstacle to conducting RCTs (159). Such conclusions may subvert the randomization process. For example, the assessment of high-oxygen environments as a cause of blindness in premature infants was impeded by well-intentioned nurses (346). In one study nurses raised the oxygen level for the experimental group of infants in the belief that the low-oxygen environments were harmful. In another study, it was necessary to implement the treatment only partially, until evidence of the harmful effects of oxygen were more apparent.

RCTs are generally considered more complex and expensive to conduct than other types of studies. The decision to initiate an RCT should be based on strong evidence that the hypothesis under consideration merits the possible expense and effort of conducting such a study.

Finally, RCTs maybe of limited utility in studying safety. As indicated above, safety is a measure of risk, and risks may occur after a considerable time, may occur infrequently, and may be unexpected. These types of effects maybe difficult to plan for and measure by an experimental study, thus necessitating the consideration of other forms of assessment.

Observational Studies

Observational studies may be valuable in generating or testing hypotheses about the health effects of a technology once the technology is widely diffused. They also may be considered in situations where experimental studies are inappropriate or impossible to conduct. The common element in all observational study designs is that the investigator does not control the application of the technology under study. The division of a population group into "cases" and "controls" or "exposed" and "unexposed" occurs through mechanisms unrelated to carrying out a study, such as the treatment preference of a physician (e.g., in the care of a stroke victim) or self-determination (e.g., in the choice of a method of contraception). Although the internal validity of observational study designs generally does not match that of experimental study designs, observational studies may allow evaluators to rule out competing explanations for the observed effects.

Because the investigator does not employ the deliberate or intentional modification of conditions between the study groups, steps must be taken to try to eliminate any potential bias in selecting the study groups. The investigator must try to control for bias, which may result when groups differ with respect to "confounding variables" (age, sex, health status, or any other characteristic which may account for observed outcomes). However, in nonrandomized studies the extent of selection bias cannot be known, and thus the effectiveness of the steps taken to minimize bias also cannot be known with certainty.

Cohort Design

Cohort studies begin with a "naturally occurring" population, or a sample thereof, chosen by

the investigator as defined by: 1) some criterion or combination of criteria such as specific age, location, time period, etc., and 2) exposure or nonexposure to a technology. The population is followed over time to observe the differences in health status between the exposed and unexposed groups. In a "prospective cohort study," the population is identified at the time of exposure and health status is assessed at a future time. If the population is identified after the exposure has occurred and the health status of the individuals is assessed at the present or a future time, it is termed a "retrospective cohort study."

A 1978 study by Roos and colleagues (319) employs a retrospective cohort study in assessing the effectiveness of tonsillectomy (with or without adenoidectomy) in preventing subsequent episodes of respiratory illness. This study illustrates many of the features that can be built into cohort designs to minimize the effects of confounding variables in an attempt to improve internal validity.

Roos and colleagues used medical claims and patient registration data provided by the Manitoba Health Services Commission to identify the population from which the cohorts were drawn. Two operated groups were created: one consisting of all patients operated on for tonsillectomy only during January 1973; the other consisting of all patients operated on for tonsillectomy only or tonsillectomy plus adenoidectomy for all of 1973. In addition, two comparison (nonoperated) groups were formed: children under the age of 14 whose records indicated evidence of tonsillar illness but no tonsillectomy operation during a 3-year period (1972-74); the other consisting of nonoperated siblings of operated patients.

Analysis of the data indicated that, on the average, the surgical procedures averted about one episode of respiratory illness per child over the 2 years following surgery. The greatest benefit accrued to the patients who had experienced the greatest number of episodes of respiratory illness in the year preceding surgery.

Had the investigators not taken measures to control for confounding variables, the observed results might have been explained by factors other than the surgical procedures. Specifically, maturation

(changes in health with age) was not believed to affect the results since the groups were similar in age and sex. The effects of "history," temporal events such as new health practices, was minimized by the use of concurrent controls. "Local history," the effect of family preference or physician factors, such as a predisposition toward surgery, was also minimized by including the sibling control group. These design features presumably minimized threats to the internal validity of the study.

Postmarketing surveillance, the mechanism used to detect unsuspected adverse drug reactions *after* a drug is marketed, generally employs the prospective cohort design. * Typically, a user population of a particular drug is entered into a registry and followed over time for various "health events." Rates of such events are compared with rates in a nonuser population. Thus, unusual medical events may be associated with use of the drug. These studies, because they use relatively large populations, may detect associations between drug use and unusual adverse reactions which are generally not detected in small population studies such as those used in premarketing assessment of the drug.

Historical Controls.—Innovations in medicine often diffuse so rapidly and completely that new technologies or new treatment variations may become standard in a fairly short time. It may be impossible to assess the long-term outcome of the technology in a conventional prospective cohort study, for lack of a group of patients not given the new treatment. In these instances, if researchers are to conduct a study, they may use a variant of the cohort design which employs historical control groups, i.e., patients treated prior to the innovation. The use of historical controls, however, adds a serious limitation to the cohort study design: change, other than the change in the treatment being assessed, is constantly occurring in health care, and such change may affect the internal and construct validity of the study. Yet despite their limitations, historical control studies can be useful, particularly if the temporal gap between control and treated groups is small, since

*For a detailed discussion of postmarketing surveillance, see OTA's report entitled *Postmarketing Surveillance of Prescription Drugs* (281).

the likelihood of some validity problems is reduced. Great care, however, should be exercised in their use and interpretation.

One of the first studies assessing the efficacy of coronary care units (CCUs) (314) relied on historical controls. The first 200 patients with acute myocardial infarction (heart attack) admitted to the CCU at Royal Perth Hospital formed one cohort, and the last 200 patients treated for acute myocardial infarction prior to the opening of the CCU formed the comparison group. Although mortality rates by severity of infarction were somewhat better for patients treated in the CCU than for patients in the historical comparison group, the difference was not statistically significant. Because the patients were all treated at the same hospital (though at different times), the two groups were similar in many respects: the base population was similar, the hospital staff was basically the same, and hospital records, on which the study relied, were similarly kept. The validity of the study (i. e., did the CCU produce the effect), however, was compromised by the introduction at about the same time as the CCU of a number of other therapeutic measures (e.g., lidocaine and atropine to treat and prevent arrhythmias and the use of transvenous pacemakers to treat conduction blocks). This study of CCU efficacy was not definitive, and the value of CCUs, themselves not strictly defined entities, is still an open question. However, the study did raise enough questions to spawn further investigations.

Studies of the use of high-dose methotrexate chemotherapy for treating osteosarcoma, a form of bone cancer, * illustrate a case where the use of historical controls so compromised the study that erroneous results were obtained. Following the development of chemotherapy in the early 1970's, researchers began to experiment with ways to improve its apparent effectiveness. One approach was to treat patients with drugs before their cancer had spread. Studies using historical controls indicated that nearly half the patients treated in 1970 lived 2 years without a recurrence of the disease, compared to only 20 percent of a group of patients treated in 1960. However, the change in therapy from 1960 to 1970 was accom-

panied by other changes in diagnosis, treatment, and patients. For example, the patient mix undoubtedly changed over the 10-year span so that patients with the worst prognoses (i.e., metastatic cancer) no longer constituted the majority of those treated, rendering the cohorts noncomparable. One can have little confidence in the results of a study which seems to show the chemotherapy efficacious, when the confounding effects of the other secular changes that occurred between 1960 and 1970 could account for the effects of the treatment in analyzing the study data. In particular, the Mayo Clinic found that patients not treated with chemotherapy in the later time period also had higher survival rates.

In summary, cohort studies using historical controls serve a limited but sometimes helpful purpose. They may allow for an inexpensive preliminary inquiry as to the value of a technology, capitalizing on existing data. However, they seldom, if ever, provide definitive information on which to make decisions about the value of the technology.

Case-Control Design

Case-control studies compare a group of people with a disease (or other outcome event), cases, to another group without the disease, controls, and then determine whether they differ in their previous exposure to a presumed causal agent (e.g., a drug). These studies are retrospective in nature, the exposure having occurred prior to the identification of cases and controls.

Substantial biases are possible in case-control studies. The most serious result from the selection of an inappropriate comparison (control) group. Because it is not possible to achieve complete comparability between the comparison group and the case group, controversies about the interpretation of case-control studies generally revolve around the question of whether or not the controls are an appropriate representation of the population that gave rise to the cases. Other problems also exist: the retrospective nature of the method implies no control over the treatment, forces reliance on individuals accurately recalling past events, and forces reliance on records that were kept for reasons other than those of carrying out a study.

*G. B. Kolata, "Dilemma in Cancer Treatment," *Science* 209:792, 1980.

The studies of a possible association of estrogen therapy with endometrial cancer illustrate the problems encountered in using case-control designs. The major dispute among researchers (191, 196) concerns the appropriateness of the control group. The traditional approach, to compare patients with endometrial cancer to control patients with other genitourinary cancers, has found a consistently high association between the use of estrogen and endometrial cancer. Critics of this approach note that because estrogen use may provoke uterine bleeding, and because a woman with bleeding is very likely to seek medical attention, there may be a higher percentage of women carefully examined and tested in the group taking estrogens than in the population of women not taking estrogens. This would lead to a higher rate of detection of endometrial cancer in the estrogen group than in the nonestrogen group.

Horwitz and Feinstein (191) contend that because of this increased surveillance, cancers are detected in the estrogen group that otherwise would not come to clinical attention during the lifetime of the women, and that if the nonestrogen group were tested as carefully, more cancers would be detected in that group. To counteract this potential selection bias, these investigators recommended selecting controls from among women surgically treated for noncancerous uterine diseases. The use of such a population to create the control group should adjust for the bias resulting from increased surveillance and diagnosis. Horwitz and Feinstein showed that when this selection procedure was employed, the likelihood of estrogen being linked to cancer was significantly lower than under previous study approaches. As Cole (61) has stated, however, patients undergoing the same diagnostic procedures as the cases can be "an inappropriate control group," since the same causal agent may be responsible for their illnesses.

These studies have not resolved the issue, however, and proponents of traditional control selection procedures claim that there is little detection bias in their method, since most cases of endometrial cancer are eventually diagnosed (196). These critics maintain that the controls used by Horwitz and Feinstein are biased, because they do not give an appropriate picture of estrogen use

in the underlying population. However, recent evidence from autopsy studies has shown that many cases of endometrial cancer indeed *are* unsuspected during life and are first detected, if at all, at autopsy (192).

Is there more or less bias in Horwitz's and Feinstein's control group selection than in the traditional approach? The two approaches might be viewed as providing a range of estimates for the relationship being examined. Because of the internal validity problems associated with this method, the use of different control groups to estimate the range of relative risk estimates might be considered.

In summary, case-control studies are relatively inexpensive, can be carried out in a relatively short time and usually employ smaller sample sizes than other study designs. The case-control design lends itself to ascertaining the associations between known rare events or outcomes and suspected causal agents when the events occur only years after the exposure. For example, this design might be used to investigate the relationship between a commonly used drug and a rare adverse effect. Case-control studies can be used to explore a hypothesis without disrupting medical practice. However, case-control studies are not useful for discovering previously unsuspected effects or discovering adverse effects of rarely used drugs.

Summary

Observational study designs used to assess the outcome of a medical technology are those in which the investigator does not control the application of the technology to the study population and applied in essentially the same manner as observational designs would be applied to examine other risk factors for disease. These designs are most applicable for detecting or ruling out specified but unforeseen adverse consequences of a technology after the technology has been diffused. Experimental designs, those in which the investigator controls the application of the technology according to specific criteria, are in theory and often in practice more useful, especially in determining efficacy.

The degree of validity, particularly internal and external validity, of the findings varies with the

study design chosen. Observational studies' (e.g., case-control and cohort) lack the high degree of internal validity found in the design of choice for experimental studies, the RCT. That is, it is usually more difficult in observational as opposed to experimental studies to determine whether the observed differences can be attributed to the technology under study. Because observational studies can more accurately reflect the conditions of use of medical technologies in the population, they may, in some cases, have a higher degree of external validity than experimental studies.

The study design ultimately selected depends on several factors, including the developmental

stage of the technology, the purpose of the study, ethical considerations, the population available, and budget constraints. Seldom is assessment a one-time event. Associations of cause and effect can *rarely* be established through a single study. In theory, judicious decisions in study design selection should be based on a review of previous and ongoing studies so that each new study becomes a building block toward the total assessment, leading to sound policy decisions. In practice, sometimes they are not.

EVALUATING ECONOMIC EFFECTS

The economic effects of medical technology have been assessed through a variety of methods, most notably cost-benefit analysis, efficiency studies, cost-effectiveness analysis (CEA), cost-impact (total costs associated with a technology), or private sector-oriented techniques such as return-on-investment analysis.

Currently, the most visible and potentially the most useful of these techniques is CEA. CEA is not simply an economic technique; it is a blend of economics and clinical information. As such, it will be described in chapter 5 with synthesis.

No matter which form of analysis is chosen, certain methodologic considerations need to be taken into account. These considerations were identified and examined in previous OTA reports (270,271), and the discussion presented here is based on that earlier work.

Opportunity Cost

The principal concept when evaluating the economic effects of a medical technology (or any activity) is opportunity cost. The opportunity cost of a resource is its value in its next best use. Thus, the true cost of a resource is not necessarily its market price tag. Rather, it is what one must give up elsewhere in order to use that resource.

An illustration should help to clarify the difference between a market price tag and a re-

source's true opportunity cost. From the perspective of a hospital accountant, volunteers' time is free; it is not found on the hospital's wage bill and the accountant would ignore it. But is volunteer labor not a true cost of running the hospital? Volunteers definitely contribute to the output of the hospital. And from a social perspective, if the volunteers' labor would have been donated elsewhere had the individuals not worked at the hospital, such labor clearly has value. In essence, the opportunity of using the labor productively in other activities has been foregone. From a social perspective, therefore, the volunteers' time should be included in an assessment of costs. Although determining an appropriate dollar value maybe difficult, the social value of volunteer time should not be ignored in an analysis.

Furthermore, as stated in chapter 2, both direct costs—resources purchased directly—and indirect costs—the value of the lost “production” time seeking care or being sick—should be included in an analysis.

Marginal Valuation

The worth of a technology should be assessed at, what economists term, “the margin.” That is, an analysis should seek to compare the added, or marginal, cost of producing the next unit of benefit.

In an evaluation of computed tomography (CT) scanning, the issue is no longer whether the technology itself is cost effective, but, rather, whether the various applications of the technology are cost effective. Should CT be used for confirming suspected brain disease/trauma, or for ruling out brain disease/trauma when persistent headaches are presented? In what instances are body scans indicated—or cost effective?

In general, the relevant inputs or costs which must be considered in the case of a medical technology will be tied to whether the technology is already in place or whether it has yet to be adopted/purchased.

Joint Production Considerations

Many technologies have multiple applications, and the technological process being studied is seldom applied in isolation. These two considerations can have enormous effects on cost calculations.

For instance, since a single blood test can be and is often used as a source of information for numerous diseases and bodily functions, analyzing the cost of drawing blood for only one purpose is inadequate if the total cost is used; it either overstates the associated costs, understates the potential benefits, or both. Likewise, a CEA of a Pap smear program should be done in recognition of the fact that many other health evaluations are not only possible but are ordinarily performed during the examination, whether formally or informally. That is, a woman who is given a Pap test may be screened for other pelvic disorders, high blood pressure, fever, skin rashes, weight problems, and many other conditions. All of these procedures carry certain potential benefits and all of them should be assigned some of the cost (or, conversely, less cost should be assigned to the Pap test); or the analysis should be evaluating the complete examination rather than just the Pap test.

Including the effects of joint production adds greatly to the problems of measurement and valuation, but these difficulties in no way diminish the conceptual importance of fully considering these effects in a complete CEA. Sometimes, for instance, a very small incremental (or marginal),

increase in cost to an existing production process can have large benefits spread over multiple applications. However, some large cost increases may produce fewer benefits than existing production processes when their contributions to all the applications are taken into account.

R&D Costs

R&D costs may pose a problem when evaluating a technology's worth. In general, where R&D is an integral part of the immediate program in question (e.g., when analyzing the costs and benefits of a new technology in a medical research center), the R&D resources should be included along with the program's operating inputs. When the R&D has preceded the program being evaluated—that is, its existence is independent of the immediate policy decision—R&D resources should be excluded from consideration.

Overhead Costs

Determining how to allocate overhead costs is particularly difficult. If the use of the technology at issue is truly marginal to the overall enterprise, one might be tempted to ignore overhead, to look only at the marginal resource needs associated with the program. However, if the existence of some of the overhead depends on the program in question, clearly it must be identified and included. The general principle of seeking the marginal inputs still holds, but often in practice one may have to attribute to the program a share of overhead proportional to the program's share of the total enterprise.

Costs v. Prices

Uncritical use of market prices can lead to large gaps between cost estimates and true costs. Illustrative of this problem is the use of hospital charge data to reflect the costs of hospital care. A common practice, this form of "pricing" ignores the known idiosyncrasies of hospital accounting in which hospitals charge well above true marginal costs for certain services and use the profits to subsidize other services for which charges do not cover marginal costs. If the deviations from marginal costs were small, one might reconcile accept-

ance of imperfect hospital data as a readily available source of information providing a qualitative valid picture. However, studies of the discrepancies between true costs and charges show dramatic differences. For example, hospital pharmacy charges can vary from 10 to 1,000 percent of the true costs of drugs depending on the frequency of their use, their level of cost, purpose, etc.

Discounting

Costs and benefits seldom occur at the same point in time. Through the application of a method termed discounting, however, they can be treated as if they all occurred in the present.

The rationale for discounting future costs and benefits stems from the fact that resources can be productively invested for future gains, as well as from the observation that people expect to be rewarded for postponing gratification. For instance, in order to induce individuals to save, in-

terest must be paid, even in the absence of inflation. The rate of interest determines the future value of the amount invested. Thus, for example, \$100 invested at 5-percent interest this year will become \$105 next year. Discounting is the reverse process: \$105 next year has a “present value” of \$100 when the discount rate is 5 percent.

Although there is general agreement among economists and policymakers that discounting future moneys is conceptually correct, there is no consensus concerning what discount rate should be used, and there is still some confusion as to the proper method of valuing future nonmonetary benefits/effectiveness. However, when benefits are long delayed, almost any discount rate will reduce benefits substantially (to near zero in extreme cases), making them less important to the outcome of the analysis (270). Thus, this phenomenon results in making the rate used and the uncertainty of future events less important than they otherwise would be.

CONCLUSION

Choosing a research method for assessing health effects depends on various factors. In general, one should opt for the study design that produces results. But constraints such as economic and social/ethical factors limit one’s choice. There is a role in medical technology assessment for each of the methods discussed in this chapter. The important point to remember is that each method has its inherent strengths and weaknesses, and one must always exercise caution in accepting the results of a study without carefully taking note of the study’s limitations.

Evaluating economic effects requires careful attention to the principles outlined in the latter portion of this chapter. An important point to always keep in mind is that costs are usually not what they appear to be, especially in health care.

The evaluation of the social and ethical implications of medical technologies is discussed in the next chapter. The following chapter reviews methods for synthesizing results from research studies, CEA, and group decisionmaking techniques.