## CME ARTICLE

# Healthcare Technology Assessment: Methods, Framework, and Role in Policy Making

*Clifford S. Goodman, PhD*

*This activity is designed for healthcare organization managers and clinicians, particularly those involved in technology-related decisions, including coverage decisions, technology acquisition, practice guideline development, and evidence-based medicine.*

*GOAL*

To provide a basic understanding of the principles, methods, and systematic framework of healthcare technology assessment.

*OBJECTIVES*

1. Understand the role of healthcare technology assessment in policy making and the technical properties and impact assessed.
2. Become familiar with the categories and basic attributes of methods used in healthcare technology assessment.
3. Comprehend the ten-step framework for conducting a healthcare technology assessment.

**CONTINUING MEDICAL EDUCATION ACCREDITATION**

The Long Island Jewish Medical Center, the Long Island campus for the Albert Einstein College of Medicine, designates this continuing medical education activity for 1 credit hour in Category I of the Physician's Recognition Award of the American Medical Association

The Long Island Jewish Medical Center is accredited by the Accreditation Council for Continuing Medical Education (ACCME) to sponsor continuing medical education for physicians. This activity is cosponsored by the Johns Hopkins University School of Medicine.

This CME activity was planned and produced in accordance with the ACCME Essentials.

From The Lewin Group, Fairfax, VA.

Address correspondence to: Clifford S. Goodman, PhD, The Lewin Group, 9302 Lee Highway, Suite 500, Fairfax, VA 22031-1214. E-mail: cgoodman@lewin.com.

Managed care is in the thick of technology decision making. This decision making is under the increasing, and sometimes contentious, scrutiny of healthcare stakeholders, including technology companies, clinicians, patients, family members, the public, and the legal system. The growth and development of healthcare technology assessment in the private and public sectors reflect the demand for informed, well-founded, and accountable processes to help render decisions about the acquisition, coverage, and appropriate use of technology.

Healthcare technology assessment is the systematic evaluation of the properties, effects, or impact of medical technology. The main purpose of this assessment is to inform technology-related policy making in healthcare. Healthcare technology assessment is conducted by interdisciplinary groups using explicit analytic frameworks drawn from a variety of methods. Technology assessment methods are evolving and their applications are increasingly diverse. Managed care organizations with access to systematic, evidence-based, and adequately funded technology assessments are better positioned to harness the flow of technology for state-of-the-art, cost-effective healthcare.

## Role in Policy Making

Healthcare technology assessment is used in many ways to support policy making. Among its roles are to: (1) support the decisions of healthcare product companies regarding product development and marketing; (2) support regional and national allocation decisions for healthcare resources; (3) provide information for regulatory decisions by the Food and Drug Administration on market approval of a technology; (4) help healthcare payers and providers determine which technologies should be included in health benefits plans and help them formulate coverage (whether or not to pay) and reimbursement (how much to pay) policies; (5) help managers of hospitals, healthcare networks, and other healthcare organiza-

tions make decisions regarding technology acquisition or adoption; (6) inform clinicians, providers, and patients about the proper use of healthcare interventions for particular health problems (eg, practice guidelines and disease management programs); and (7) inform investors and technology companies concerning venture capital funding, acquisitions and divestitures, and other product and services transactions.

## Scope of Healthcare Technology

The analytic frameworks, methods, and data sources used in healthcare technology assessment must accommodate the diverse range of medical technology. For healthcare technology assessment, the three defining dimensions of medical technology are its form or material nature, its purpose or application, and its stage of development or diffusion.

The principal forms of healthcare technology are drugs (eg, antibiotics, antihypertensives, and chemotherapeutic agents); biologics (eg, vaccines and blood products); equipment, devices, and supplies (eg, computed tomographic scanners, cardiac pacemakers, diagnostic test kits, and surgical gloves); medical and surgical procedures (eg, coronary angiography, gall bladder removal, and psychotherapy); support systems (eg, electronic patient record systems, telemedicine systems, blood banks, and clinical laboratories); and organizational and managerial systems (eg, prospective payment systems, clinical pathways, and computer-based drug utilization review systems). The forms of technology affect assessment in many ways. For example, the typical hallmarks of clinical studies of the safety and efficacy of new drugs, including double-blinding, placebo controls, and randomization of several thousand patients, are impractical or impossible to implement for most medical devices.

The purpose or application of a healthcare technology can be categorized as prevention, screening, diagnosis, treatment, rehabilitation, or palliation.

When categorized by stage of development, technologies are considered future, experimental, investigational, established, or obsolete, outmoded, or abandoned. Future technologies are in a conceptual stage, anticipated, or in the earliest stages of development. Experimental technologies are undergoing bench or laboratory testing using animals or other models. Investigational technologies are undergoing initial clinical evaluation in humans. Established technologies are considered by providers to be a standard approach to a particular condition and are widely used. Obsolete, outmoded, or abandoned technologies have been superseded by other technologies or found to be ineffective or harmful.

Not all technologies fall neatly into single categories. Many tests and other technologies used for diagnosis are also used for screening. Certain "boundary-crossing" technologies that combine characteristics of drugs, devices, or other major categories of technology, such as implantable drug infusion pumps and bioartificial organs, may pose jurisdictional or administrative challenges to regulatory agencies, payers, standards-setting organizations, and other groups. Examples of technologies whose boundary-crossing attributes have complicated regulatory approval and coverage decisions in recent years are gallstone lithotriptors (used with stone-dissolving drugs) and positron emission tomography (used with radiopharmaceuticals). In addition, a technology might be considered established for certain patient applications but investigational for others. A technology once rejected or considered obsolete may return to established use for a better defined or entirely different clinical purpose; the use of thalidomide for treatment of leprosy is one such example. Many technologies, particularly medical devices, undergo multiple incremental innovations after their initial acceptance into general practice.

## Technical Properties and Impact Assessed

Healthcare technology assessment can involve investigating the technical properties, clinical safety, efficacy or effectiveness, economic attributes or impact, and social, legal, ethical, or political impact of a medical technology.

The technical properties of a technology include its performance characteristics and conformity with specifications for design, composition, manufacturing, tolerances, reliability, ease of use, and maintenance. Safety is a judgment of the acceptability of the risk (a measure of the probability of an adverse outcome and its severity) associated with using a technology in a given situation (eg, for a patient with a particular health problem) by a clinician with certain training or in a specified treatment setting.

Efficacy and effectiveness both refer to how well a technology works to improve patient health, usually based on changes in one or more pertinent health outcomes. In healthcare technology assessment, efficacy refers to the benefit of using a technology for a particular problem under ideal conditions (eg, within the protocol of a carefully managed randomized controlled trial) or at a "center of excellence" in patients who meet narrowly defined criteria. Effectiveness refers to the benefit of using a technology for a particular problem under general or routine conditions (eg, by a physician in a community hospital for various types of patients).

Healthcare technologies can have a wide range of microeconomic and macroeconomic attributes. Microeconomic concerns include the costs, prices, charges, and payment levels associated with a particular technology. Other microeconomic issues include comparisons of resource requirements and outcomes (or benefits) of a technology for particular applications through cost-effectiveness, cost-utility, or cost-benefit analyses. The macroeconomic impact of healthcare technologies includes the impact of new technology on national healthcare costs and the effect of technology on resource allocation among different health programs or among healthcare and other sectors.

A variety of technologies raise social, ethical, legal, and political concerns. Technologies such as genetic testing, fertility treatments, transplantation of scarce organs, and life-support systems for the critically ill challenge certain legal standards and societal norms. In dialysis and transplantation for patients with end-stage renal disease, ethical concerns arise from patient selection criteria, termination of treatment, and management of noncompliant patients. Ethical concerns continue to prompt improvement in informed consent procedures for patients enrolled in trials of investigative technologies.[1] Allocation of scarce resources to technologies that are expensive, inequitably used, or noncurative raises broad social concerns.[2] Political factors can prompt legislation mandating coverage for certain technologies or influence acquisition or placement of high-profile technologies in a given jurisdiction.

The terms *appropriate* and *necessary* often are used to describe whether a technology should be used in particular circumstances. For example, the appropriateness of a diagnostic test may depend on its safety and effectiveness compared with alternative available interventions for particular patient indications, clinical settings, and resource constraints. A technology may be considered necessary if withholding it would adversely affect the patient's health.[3]

Although the relationship between a preventive, therapeutic, or rehabilitative technology and its patient outcomes is typically direct (though not always easy to measure), the relationship between a technology used for diagnosis or screening and its patient outcomes is usually indirect. For example, the efficacy (or effectiveness) of a diagnostic technology can be determined along a chain of inquiry that leads from the technical capacity of the technology to changes in patient health outcomes to cost-effectiveness, as follows:

1. *Technical capacity* (does the technology perform reliably and deliver accurate information?)
2. *Diagnostic accuracy* (does the technology contribute to making an accurate diagnosis?)
3. *Diagnostic impact* (do the diagnostic results influence the use of other diagnostic technologies? For example, does it replace other diagnostic technologies?)
4. *Therapeutic impact* (do the diagnostic findings influence the selection and delivery of treatment?)
5. *Patient outcome* (does use of the diagnostic technology contribute to improved health of the patient?)
6. *Cost-effectiveness* (use of the diagnostic technology is cost-effective compared with alternative interventions?)

If a diagnostic technology is not effective at any step along this chain, it is not likely to be effective at any later step. In other words, efficacy at a given step does not imply efficacy at a later step.[4,5]

## Measuring Health Outcomes

Health outcome variables are used to measure the safety, efficacy, and effectiveness of healthcare technologies. Health outcomes have been measured primarily in terms of changes in mortality and morbidity. In a clinical trial comparing alternative treatments, the effect on health outcomes of one treatment relative to another (eg, a control treatment) can be expressed using various measures of treatment effect. These measures compare the probability of a given health outcome in the treatment group with the probability of the same outcome in a control group. Examples of such measures are absolute risk reduction, odds ratio, number needed to treat, and effect size.[6]

Although mortality and morbidity are usually the outcomes of greatest concern, they are not the only outcomes of importance to patients or others. Many technologies affect patients, family members, providers, and employers in ways not reflected in mortality or morbidity rates; this is particularly true for many chronic diseases. Health-related quality-of-life (HRQL) measures (or indexes) can provide a more complete picture of the ways in which healthcare affects patients. Health-related quality-of-life measures are designed to capture such dimensions as physical function, social function, cognitive function, anxiety and distress, bodily pain, sleep and rest, energy and fatigue, and general health perception. These general or disease-specific measures can be unidimensional or multidimensional. They may provide a single aggregate score or yield a set of scores, each for a particular dimension.[7-10]

Examples of general HRQL indexes are the Sickness Impact Profile,[11] Nottingham Health Profile,[12,13] Quality of Well-Being Scale,[14] Functional Inde-

pendence Measure,[15] Short Form (SF)-36,[16,17] EuroQol Descriptive System,[18] and Katz Activities of Daily Living.[19] Examples of disease-specific HRQL indexes are the New York Heart Association Functional Classification,[20] Arthritis Impact Measurement Scales,[21] and Visual Functioning (VF)-14 Index.[22]

Health technology companies are increasingly using HRQL measures to differentiate their products from those of competitors. Products that may have virtually indistinguishable effects on morbidity for particular diseases (eg, hypertension and depression) may have different profiles of adverse effects that affect patients' quality of life.

The quality-adjusted life year (QALY) is a unit of healthcare outcome that combines gains (or losses) in length of life with quality of life. Quality-adjusted life years represent years of life after a healthcare intervention that are weighted or adjusted for the quality of life experienced by the patient during those years. The QALY provides a common unit for comparing the cost-utility of different healthcare interventions and ranking priorities for allocating scarce healthcare resources. Healthy-years equivalents (HYEs), disability-adjusted life years (DALYs), and saved young-life equivalents (SAVEs) serve a similar function.

The scale of quality of life used for QALYs can be based on a method of eliciting patient utility for certain states of life. This dimension is typically standardized to a scale ranging from 0.0 (death) to 1.0 (perfect health). Three commonly used approaches for measuring utility are the standard gamble, time trade-off, and rating scale.[23,24] Quality-adjusted life years are useful for making comparisons among alternative technologies because they are generic units that can reflect changes brought about by different healthcare interventions for the same or different health problems. Quality-adjusted life years have also been used in "league tables" that rank diverse healthcare interventions according to their cost per QALY,[25] although the validity of such tables has been challenged because of inconsistencies in QALY construction and other methodologic variations.[25-27] Although the measurement of QALYs and their use in health policy remain controversial, interest remains high and methodologic work in this area continues.

## Timing of Assessment

There is no single correct time to conduct a healthcare technology assessment. Investors, regulators, medical specialty societies, managed care organizations, insurers, hospital managers, and others tend to make decisions about any given technology at particular junctures in its life cycle, and each may subsequently reassess the technology. Indeed, the determination of a technology's stage of diffusion may be the primary purpose of an assessment. For payers, technologies deemed "experimental" or "investigational" are usually excluded from coverage, but those that are established or generally accepted are typically eligible for coverage.

Tradeoffs are inherent in decisions regarding the timing for healthcare technology assessment. On one hand, the earlier a technology is assessed, the more likely its diffusion can be curtailed if it is found unsafe or ineffective. On the other hand, regarding the findings of an early assessment as definitive or final may be misleading. An investigational technology may not yet be perfected, its users may not yet be proficient, and its costs may not yet be stabilized. The technology also may not have been applied in enough circumstances to recognize its potential benefits, and its long-term outcomes may not yet be known. In addition, the "moving target problem" can complicate healthcare technology assessment.[28] By the time an assessment is conducted, reviewed, and disseminated, its findings may be outdated by greater data collection or changes in a technology, how the technology is used, or competing technologies for a given problem. As one technology assessor noted about the when-to-assess problem, "It's always too early until, unfortunately, it's suddenly too late!"[29]

## Ten Basic Steps of Healthcare Technology Assessment

The scope, selection of methods, and level of detail in the practice of healthcare technology assessment vary greatly. Nevertheless, most healthcare technology assessment activity involves some form of the steps shown in Table 1[30] and summarized below. Not every assessment includes all these steps, and the steps are not necessarily conducted in strict sequence. For example, many healthcare technology assessments do not entail primary data gathering (step 5), relying instead on synthesis of existing data (step 7). Meta-analytic approaches may involve concurrent or iterative evidence interpretation (step 6) and synthesis of existing data.

### *Step 1: Identify Assessment Topics*

To a large extent, technology assessment topics are determined or bounded by an organization's mission or purpose. Most assessment programs have criteria for topic selection, although these criteria are not always explicit. Selection criteria used in setting healthcare technology assessment priorities for a managed care organization might include health

**Table 1.** Ten Steps of Healthcare Technology Assessment

| | |
|---|---|
| 1. | Identify assessment topics |
| 2. | Specify the assessment problem |
| 3. | Determine locus of assessment |
| 4. | Retrieve available evidence |
| 5. | Collect new primary data (as appropriate) |
| 6. | Interpret evidence |
| 7. | Synthesize and consolidate evidence |
| 8. | Formulate findings and recommendations |
| 9. | Disseminate findings and recommendations |
| 10. | Monitor impact |

problems with high morbidity or mortality rates, large number of enrollees affected, high unit or aggregate cost of a technology or health problem, substantial variations in practice within the network, sufficient research findings available on which to base assessment, and pressure by patients, clinicians, or employers to make payment decisions.

Of course, these criteria are for demand-side assessment priorities. On the supply side, assessment priority criteria for technology companies reflect their decision-making needs, such as potential market size, opportunity for market share, anticipated return on investment, need to demonstrate safety and efficacy for regulatory approval, and need to demonstrate value to providers and payers. Indeed, manufacturers' assessment priorities might bring assessment topics to the attention of managed care organizations. The processes for soliciting candidate assessment topics and ranking assessment priorities range from highly subjective to systematic and quantitative.[31,32]

### *Step 2: Specify the Assessment Problem*

One of the most important aspects of a healthcare technology assessment is to clearly specify the problems or questions to be addressed. This step affects all subsequent aspects of the assessment. A group conducting an assessment should have an explicit understanding of the purpose and intended users of the assessment. There is no single correct way to state an assessment problem. It may entail specifying at least the following elements: healthcare problem(s), patient population(s), intervention(s), practitioners or users, care settings, and health and economic endpoints of interest. Specifying inclusion and exclusion criteria for these elements links the assessment problem to evidence retrieval.

### *Step 3: Determine Locus of Assessment*

Healthcare decision makers can create or buy all or certain portions of healthcare technology assessments. Determining who is responsible for sponsoring or conducting an assessment depends on the nature of the problem, expertise of available personnel, time constraints, funding, and other factors. Even when an assessment report exists on a topic of interest, decision makers must determine whether it has a compatible perspective, whether the assessment problem is appropriate, how current the report is, whether the methodology is sufficiently credible, and whether the report is worth the price. Smaller healthcare provider and payer organizations often obtain reports from organizations that specialize in healthcare technology assessment; larger ones are more likely to have internal healthcare technology assessment programs. Some healthcare organizations commission selected components of an assessment, such as evidence retrieval and synthesis, and perform the other steps in-house.

Users of any healthcare technology assessment report should consider the potential for conflict of interest on the part of the organization that sponsored or conducted the assessment, whether it be a technology company, managed care organization, academic center, or government agency. Assessment reports should disclose sponsorship and other sources of support, authors and their organizational affiliations, and other documentation of methods and sources.

### *Step 4: Retrieve Available Evidence*

One of the challenges of healthcare technology assessment is to assemble the evidence—the data, literature, and other information—relevant to a particular assessment. For many technologies, the evidence is profuse, scattered, and of widely varying quality; for very new technologies, it may be sparse and difficult to find. Sources of evidence for healthcare technology assessment include computer databases of published literature; computer databases of clinical and administrative data; printed indexes and directories; government reports and monographs; reference lists in available studies, reviews, and meta-analyses; special inventories and registries of reports; health newsletters and newspapers; company reports and press releases; and colleagues and other investigators (Table 2).

Much valuable information is available beyond the traditional peer-reviewed published literature. This

"gray" or "fugitive" literature is found, for example, in industry and government monographs, regulatory documents, professional association reports, market research reports, special commission reports, conference abstracts, and on World Wide Web sites. The Internet is an extraordinarily broad and readily accessible medium that provides access to many of the information sources noted above. Although the gray literature can be timely and cover aspects of technologies not addressed in mainstream sources, it is usually not subject to peer review and must be scrutinized accordingly.

Searching for pertinent existing evidence generally is one of the first major tasks in conducting an assessment and should be planned accordingly.[33] The costs associated with evidence searches can be significant, coming in the form of staff time, computer access and acquisition of literature, data tapes, and other documentation. Although access to MEDLINE and other public-source databases is free or inexpensive, using some specialized scientific and business databases can be more costly. Database vendors offer a variety of database packages at varying prices. Organizations such as ECRI, the Blue Cross and Blue Shield Association, and the American Medical As-

**Table 2.** Information Resources for Healthcare Technology Assessment

**National Library of Medicine databases (MEDLARS)**
*MEDLINE:* citations for biomedical journal articles
*HealthSTAR:* citations for planning and administration health services literature on research and technology assessment
*HSTAT:* full text of clinical practice guidelines, consensus development reports, and technology assessment reports, primarily from federal agencies
*PREMEDLINE:* basic citation information and abstracts of articles before they are indexed and placed into MEDLINE
*PubMed:* World Wide Web access to MEDLINE, integrated molecular biology databases, links to participating online journals and related databases

*Also: AIDSDRUGS:* descriptions of substances used in trials on acquired immunodeficiency syndrome; *BIOETHICS:* citations for bioethics literature; *CANCERLIT:* citations for journal articles in cancer; *DIRLINE:* directory of organizations; *PDQ:* cancer treatment, supportive care, screening, prevention, and clinical trials; *TOXLINE:* citations on toxicologic and other effects of drugs and chemicals

**Evidence-Based Practice Center (EPC) Reports:** evidence reports generated by the 12 EPCs designated by the Agency for Health Care Policy and Research
**National Guidelines Clearinghouse (NGC):** sponsored by the Agency for Health Care Policy and Research, American Medical Association, and American Association of Health Plans; under development by ECRI
**National Comprehensive Cancer Network (NCCN):** evidence-based cancer care guidelines produced by an authoritative network or premier cancer centers
**National Institutes of Health Clinical Trials Registry:** database of government and privately funded clinical trials for serious and life-threatening conditions (under development; mandated by the FDA Modernization Act of 1997)
**Other guidelines:** produced by medical specialty societies and other authoritative organizations, such as the American Academy of Pediatrics, American Cancer Society, American College of Cardiology, American College of Obstetrics and Gynecology, and American College of Radiology.

**Cochrane Library**
*Cochrane Database of Systematic Reviews:* systematic reviews of controlled trials on hundreds of topics
*Database of Abstracts of Reviews of Effectiveness (DARE):* structured abstracts of systematic reviews from around the world, critically appraised by the National Health Service Centre for Reviews and Dissemination in the United Kingdom
*Cochrane Controlled Trials Register:* bibliography of controlled trials, including sources outside peer-reviewed journal literature
*Cochrane Review Methodology Database:* bibliography of sources on research synthesis

**Other Resources**
*EMBASE (Excerpta Medica database):* citations for biomedical journal articles (Elsevier)
*American College of Physicians Journal Club on Disk (ACPJCOD)*
*Bandolier:* evidence-based healthcare journal
*Canadian Clinical Practice Guidelines Infobase*
*TRIP Gwent Database:* single searchable index to 14 EBM resources
*ESToC (Elsevier Science Table of Contents):* tables of contents of journal articles in science (Elsevier)
*SCISEARCH:* citations for scientific journal articles (Institute for Scientific Information)
*Current Contents:* tables of contents of journals in science, social sciences, and other fields (Institute for Scientific Information)

sociation sell their healthcare technology assessment reports on a subscription basis. Some market research monographs and other reports produced for health product companies, investors, and other business interests are priced in the thousands of dollars.

### *Step 5: Collect New Primary Data (As Appropriate)*

Although many healthcare technology assessments are based on available evidence, some entail collection of new primary data. Healthcare technology assessments by managed care organizations and other payers to make coverage determinations typically do not entail new data collection but rely on existing data. It is beyond the scope of this article to describe the planning, design, and conduct of clinical trials, observational studies, and other investigations for collecting new primary data.[34] Nevertheless, describing how the methodologic validity of these primary data sources are interpreted in technology assessment is important. Certain attributes of primary studies produce better evidence than others. In general, prospective studies are superior to retrospective ones; controlled studies are superior to uncontrolled ones, randomized studies are superior to nonrandomized ones; large studies (ie, involving enough patients to detect with acceptable confidence levels any true treatment effects) are superior to small ones; contemporaneous controls are superior to historical ones; and blinded studies (patients, clinicians, analysts) are superior to unblinded ones.

New data on the effects of healthcare technology in humans can be generated through the following (methods are listed in rough order of most to least scientifically rigorous for internal validity—that is, for accurately representing the causal relationship between an intervention and an outcome in the particular circumstances of a study):

- large or small randomized controlled trials
- nonrandomized trials with contemporaneous or historical controls
- prospective cohort studies
- retrospective case-control studies
- cross-sectional studies
- surveillance studies (eg, using registers or surveys)
- consecutive case series
- single case reports.

The principles underlying the relative scientific rigor of methods form the basis of evidence interpretation. There are tradeoffs between studies with high internal validity, such as large randomized controlled trials, and those with high external validity (or generalizability), such as various observational studies or "natural experiments." Investigators have made progress in combining some of the desirable attributes of both. For example, while retaining the methodologic strengths of a prospective, randomized design, "large, simple trials" use large numbers of patients, more flexible patient entry criteria, and multiple study sites to improve external validity.[35] Assessors must consider resource requirements and time constraints when deciding to undertake a new primary data collection. Except for large organizations and for high priority areas, health plans and other payers do not undertake large randomized controlled trials that may cost tens of millions of dollars and take several years to conduct. Assessors should weigh the marginal investment required to conduct primary data collection with the expected marginal value of the data (eg, for making a coverage decision).

### *Step 6: Interpret Evidence*

A challenge in conducting any healthcare technology assessment is to derive credible findings from evidence drawn from different types of studies of varying quality. Technology decision makers increasingly demand stronger scientific evidence regarding the effectiveness and cost-effectiveness of technologies. (The recently coined term *evidence-based medicine* refers to the incorporation of this demand into clinical decision making.) Strength of evidence refers primarily to the scientific validity underlying findings about causal relationships between healthcare interventions and health outcomes. Grading evidence according to its methodologic rigor is increasingly becoming a standard part of healthcare technology assessment. Such grading can take various forms, each of which involves structured, critical appraisal of the evidence against formal criteria.[36,37] Evidence tables that summarize attributes of study design, patient characteristics, patient outcomes, and derived summary statistics are useful for displaying important qualities about available studies. In any case, the authors of technology assessment reports should document the criteria or procedures by which they use study data.

Some analysts consider the results of studies that do not have randomized controls subject to such great bias that they should not be included when determining the effects of an intervention. Others say that studies with weaker designs should be used, but given less weight or adjusted for their biases.[38] Evidence standards should be commensurate with practical differences among healthcare technologies. Certain

methodologic hallmarks of pharmaceutical assessment, such as sample sizes of thousands of subjects, placebo controls, and double-blinding of patients and clinicians, are often impractical or impossible for medical devices, surgical procedures, diagnostic procedures, and other technologies. Methodologic requirements for establishing with a high degree of certainty the safety and effectiveness of a technology for a given indication depend on the inherent risks and potential health improvements posed by the technology compared with standard care. Thus, for example, methodologic requirements for heart valve prostheses, left-ventricular assist devices, and neurologic implants differ from those of surgical instruments and from those of hearing aids. Still, for any type of technology, the trend is to demand stronger and better documented evidence. By raising the bar on the evidence hierarchy, decision makers increase the burden of proof for approval, coverage, use, or acquisition of technologies.

### *Step 7: Synthesize and Consolidate Evidence*

After the merits of individual studies are considered, healthcare technology assessment involves some form of analysis that pulls together available findings. Methods used to combine or synthesize data include nonquantitative literature reviews, group judgment or "consensus development," meta-analysis or other systematic literature syntheses, decision analysis, and economic analyses.

Because of biases inherent in traditional means of consolidating literature, such as nonquantitative literature reviews and editorials, healthcare technology assessment emphasizes more structured, quantified, and better-documented methods.[39]

Virtually all healthcare technology assessment efforts involve some form of group judgment at one or more junctures, particularly to formulate findings and recommendations. Group judgment may be unstructured and informal or involve formal group methods, such as the nominal group and Delphi techniques. Although these processes typically involve face-to-face interaction, some group judgment efforts combine remote, iterative interaction of panelists with face-to-face meetings.[40,41] However, the opinion of an expert committee is not in itself primary scientific evidence.

Meta-analysis refers to a group of statistical techniques for combining results of multiple studies to obtain a quantitative estimate of the overall effect of a particular technology (or variable) on a defined outcome. This combination may produce a stronger conclusion than can be provided by any individual study. Meta-analysis is generally used when there are no definitive studies on a topic and the non-definitive studies are in some disagreement. Like traditional methods for consolidating literature, meta-analysis can be limited by biased selection of studies, poor quality data, insufficiently comparable studies, and biased interpretation of findings. However, the systematic selection and organization of evidence used in meta-analysis can minimize these shortcomings. Advances in the conduct and interpretation of meta-analysis continue, particularly in strengthening the methodologic links between clinical trials and meta-analysis.[42-44] The basic steps of meta-analysis are listed in Table 3.

Decision analysis uses available quantitative estimates to model the sequences of alternative healthcare strategies (eg, of diagnosis or treatment). Each healthcare strategy is assigned a probability that certain events and outcomes will occur, along with the values of the outcomes for patient populations and associated costs. Decision models can be used to predict the distribution of outcomes for patient populations and the associated costs of care. They also can be used to help develop clinical practice guidelines for specific health problems. For individual patients, decision models can be used to relate the likelihood of potential outcomes of alternative clinical strategies or to identify the clinical strategy that has the greatest

**Table 3.** Basic Steps in Meta-Analysis

| | |
|---|---|
| 1 | Specify the problem of interest |
| 2 | Specify the criteria for inclusion of studies (eg, type and quality) |
| 3 | Identify all studies that meet inclusion criteria |
| 4 | Classify study characteristics and findings<br>■ study characteristics (eg, contexts, participants)<br>■ methodologic characteristics (eg, sample sizes, measurement process)<br>■ primary results<br>■ derived summary statistics |
| 5 | Statistically combine study findings (eg, by averaging effect sizes), relate these to study characteristics, and perform sensitivity analysis |
| 6 | Present results |

personal utility.[45,46] The main steps of decision analysis are summarized in Table 4.

Studies of costs and related economic implications can involve attributes of primary data collection or synthetic methods, or both. Although economic data often are collected in conjunction with clinical trials and other primary data collection methods, economic analyses typically are conducted using data from multiple existing sources. Greater emphasis on demonstrating the value of spending on healthcare technology has increased the demand for cost-effectiveness analyses and related studies of the economic impact of healthcare technology.

There is a variety of approaches to economic analysis, the suitability of which depends on the purpose of the technology assessment and the availability of data and other resources. Identifying and quantifying all costs and all benefits (or outcomes) is rarely possible or necessary, and the units used to quantify these may differ.[47] The main types of economic analysis are:

1. *Cost-of-illness analysis* (determination of the economic impact of an illness or condition, such as smoking, arthritis, or bedsores, including its associated treatment costs)
2. *Cost-minimization analysis* (determination of the least costly among alternative interventions assumed to produce equivalent outcomes)
3. *Cost-effectiveness analysis* (comparison of costs in monetary units with outcomes in quantitative nonmonetary units, such as decreased mortality or morbidity)
4. *Cost-consequence analysis* (form of cost-effectiveness analysis that presents health outcomes and consequences, such as mortality, morbidity, and adverse effects, and costs associated with an intervention, without aggregating or weighting across outcomes)
5. *Cost-utility analysis* (form of cost-effectiveness analysis that compares costs in monetary units with outcomes in terms of their utility, usually to the patient, measured, for example, in QALYs)
6. *Cost-benefit analysis* (comparison of costs and benefits, both of which are quantified in common monetary units)

Cost-minimization analysis and cost-effectiveness analysis, including cost-consequence analysis and cost-utility analysis, necessarily involve comparisons of alternative interventions (including no intervention). Although cost-benefit analysis typically involves comparisons of alternative technologies, this is not necessary. In principle, a technology is cost beneficial as long as its costs are outweighed by the monetary value of its benefits or outcomes.

Because it measures costs and outcomes in monetary terms, cost-benefit analysis enables comparison of disparate technologies, such as coronary artery bypass graft surgery and breast cancer screening. A controversial drawback of cost-benefit analysis is the difficulty of assigning monetary values to all pertinent outcomes, including changes in the length or quality of human life. Cost-effectiveness analysis avoids this limitation by using more direct or natural units of outcomes, such as lives saved or strokes averted. As such, cost-effectiveness analysis can only compare technologies whose outcomes are measured in the same units. In cost-utility analysis, estimates of utility are assigned to health outcomes, enabling comparisons of disparate healthcare technologies.

In practice, wide variations exist in economic study methodologies.[48] Although some variation is unavoidable, differences in perspective, accounting for direct

**Table 4.** Basic Steps in Decision Analysis

| | |
|---|---|
| 1 | Develop a model (eg, a decision tree) that depicts the set of important choices (or decisions) and potential outcomes of these choices. For treatment choices, the outcomes might be health outcomes (health states); for diagnostic choices, the outcomes might be test results (eg, positive or negative). |
| 2 | Assign estimates (based on available literature) of the probabilities (or magnitudes) of each potential outcome given its antecedent choices |
| 3 | Assign estimates of the value of each outcome to reflect its utility or desirability (eg, using a health-related quality-of-life measure) |
| 4 | Calculate the expected value of the outcomes associated with the particular choice(s) leading to those outcomes. This calculation typically is made by multiplying the set of outcome probabilities by the value of each outcome. |
| 5 | Identify the choice(s) associated with the greatest expected value. Based on the assumptions of the decision model, this is the most desirable choice, because it provides the highest expected value given the probability and value of its outcomes. |
| 6 | Conduct a sensitivity analysis of the model to determine if plausible variations in the estimates of probabilities of outcomes or utilities change the relative desirability of the choices. (Sensitivity analysis is used because the estimates of key variables in the model may be based on limited data or expert conjecture.) |

and indirect costs, time frames, discounting, and other aspects result from lack of expertise and may reflect biases on the part of the investigators or sponsors. This diminishes the comparability and transferability of study results and the credibility of findings. In recent years, national and international groups have been developing and revising standards for conducting and reporting health economic studies.[24,49,50] These efforts should improve the practice of economic analysis, particularly if payers and providers require that economic dossiers presented by technology companies use such standardized methods.

Documents outlining standards for economic analyses provide checklists and other guidance on how to conduct these analyses. Although reviewing these guidelines is beyond the scope of this article, there are some well-recognized principles of economic analysis. For example, the economic perspective of the analysis (ie, society, payer, provider, or patient) should be made explicit. Both direct costs (medical and nonmedical) and indirect costs (such as loss of productivity) should be included. Using actual costs is preferred to using charges or prices, and opportunity costs are the best gauge of true costs. The time horizon of analysis should be specified and should cover a clinically justifiable span of accrual of costs and outcomes. Marginal cost analysis, which compares marginal changes in costs to marginal changes in outcomes, is preferred to average cost analysis. Costs and outcomes should be discounted over time, and discount rates should be specified and justified. Sensitivity analysis should be conducted to determine how reasonable changes in the estimates of input variables affect results.

Methodologists and other stakeholders in economic analysis differ on the preferred choice of perspec-

**Table 5.** Evidence Grading for AHCPR Practice Guidelines

| | |
|---|---|
| **Type of Evidence** | |
| I | Meta-analysis of multiple, well-designed controlled studies |
| II | At least one well-designed experimental study |
| III | Well-designed, quasi-experimental studies, such as nonrandomized controlled, single-group pre-post, cohort, time-series, or matched case-controlled studies |
| IV | Well-designed nonexperimental studies, such as comparative and correlational descriptive and case studies |
| V | Case reports and clinical examples |
| **Strength and Consistency of Evidence** | |
| A | There is evidence of type I or consistent findings from multiple studies of types II, III, or IV. |
| B | There is evidence of types II, III, or IV, and findings are generally consistent. |
| C | There is evidence of types II, III, or IV, but findings are inconsistent. |
| D | There is little or no evidence, or there is type V evidence only. |

From the Agency for Health Care Policy and Research (AHCPR), US Department of Health and Human Services, 1994

**Table 6.** Evidence Grading for Clinical Preventive Services

| | |
|---|---|
| **Quality of Evidence** | |
| I | Evidence obtained from at least one properly designed randomized controlled trial |
| II-1 | Evidence obtained from well-designed controlled trials without randomization |
| II-2 | Evidence obtained from well-designed cohort or case-controlled analytic studies, preferable from more than one center or research group |
| II-3 | Evidence obtained from multiple time series with or without the intervention. Dramatic results in uncontrolled experiments (such as the results of the introduction of penicillin treatment in the 1940s) could also be regarded as this type of evidence |
| III | Opinions of respected authorities, based on clinical experience, descriptive studies, or reports of expert committees |
| **Strength of Recommendations** | |
| A | There is good evidence to support the recommendation that the condition be specifically considered in a periodic health examination. |
| B | There is fair evidence to support the recommendation that the condition be specifically considered in a periodic health examination |
| C | There is insufficient evidence to recommend for or against the inclusion of the condition in a periodic health examination, but recommendations may be made on other grounds. |
| D | There is fair evidence to support the recommendation that the condition be excluded from consideration in a periodic health examination. |
| E | There is good evidence to support the recommendation that the condition be excluded from consideration in a periodic health examination. |

From the US Preventive Services Task Force. Guide to Clinical Preventive Services. 2nd ed. Baltimore, MD: Williams & Wilkins; 1996.

tive of analysis. Some contend that the societal perspective accounting for all direct and indirect costs is the only appropriate one, while others argue that, as a practical matter, decision makers are concerned only with the streams of costs and outcomes that accrue to them. Certainly, any decision regarding the acquisition or coverage of a technology has multiple stakeholders with different economic perspectives. The strategic approach for any one stakeholder conducting an economic analysis may be to account for the economic perspectives of the relevant parties and to use this information to forge effective coverage and reimbursement policies, marketing efforts, pricing, or other actions flowing from an assessment.

The validity of a cost-related study depends on the sources of the data for costs and outcomes. Increased attention is being given to collecting cost data in more rigorous, prospective studies, including randomized controlled trials. However, the closer integration of economic and clinical studies raises important methodologic issues.[51] To promote more rational diffusion of new technologies, reliable cost and outcomes data should be generated during the early part of a technology's life cycle, such as during any randomized controlled trials required for marketing approval. However, such trials might have methodologic limitations with regard to analysis of costs and outcomes. As noted previously, randomized controlled trials are expected to yield information on efficacy rather than effectiveness, and the costs of care in the careful and specific protocols of these trials might not represent the costs of care in more general settings. In deciding how to use economic data from clinical trials, assessors should weigh these tradeoffs in internal and external validity in the context of the policy decision at hand.

For example, in support of a coverage decision, assessors should not only examine the methodological rigor of the source data, but also should compare the populations and care settings in which available data were collected to the care settings and populations to be affected by the coverage decision.

### *Step 8: Formulate Findings and Recommendations*

Findings are the results or conclusions of an assessment; recommendations are the suggestions, advice, or counsel that follow from the findings. Recommendations can be made in various forms, such as a coverage policy, practice guideline, or directive. These should be evidence based, relating the validity of individual studies to the overall quality of the evidence for the findings and the strength of the recommendations. Even for those aspects of an assessment problem for which there is little useful evidence, an assessment group may have to provide some findings or recommendations. In any case, authors of healthcare technology assessment reports should make explicit the analyses and reasoning used in making their recommendations and with what level of confidence the recommendations were made. Authors of technology assessment reports increasingly annotate their recommendations with different levels of strength that reflect the grades of the evidence and direction of findings pertaining to each recommendation.

Examples of evidence grading hierarchies linked to recommendations are shown in Tables 5 and 6. The Agency for Health Care Policy and Research (AHCPR) uses the framework shown in Table 5 for linking the quality of evidence to the strength and consistency of recommendations in its practice guidelines. The AHCPR approach incorporates quantitative overviews of evidence (eg, meta-analyses), where available, in classifying levels of evidence and grading recommendations. A similar approach is used by the US Preventive Services Task Force, which assigns levels of strength of recommendations for various preventive services to the quality of the available evidence, as shown in Table 6. Of course, as shown in Table 7, the quality of evidence and strength of recommendations can vary among different forms of a technology.

**Table 7.** Evidence Grading for Screening Tests for Colorectal Cancer

| Intervention | Level of Evidence* | Strength of Recommendation* |
|---|---|---|
| Annual fecal occult blood testing of persons aged 50 and older | I, II-1, II-2 | B |
| Routine sigmoidoscopy in persons aged 50 and older | II-2, II-3 | B |
| Routine digital rectal examination | III | C |
| Routine barium enema | III | C |
| Routine colonoscopy | III | C |

*Level of evidence and strength of recommendation as defined in Table 6.
From the US Preventive Services Task Force. *Guide to Clinical Preventive Services*. 2nd ed Baltimore, MD: Williams & Wilkins; 1996.

Explicit evidence interpretation in healthcare technology assessment enables users to understand the reasoning behind the assessment findings and recommendations and to challenge them when appropriate. It also helps assessment programs and policy makers determine if a reassessment is needed as relevant new evidence becomes available.

### *Step 9: Disseminate Findings and Recommendations*

Dissemination efforts must compete with the burgeoning flow of health-related information being transmitted across diverse channels using increasingly sophisticated means. Worthy findings and recommendations can be lost because of misidentified and misunderstood target audiences, poor packaging, wrong transmission media, bad timing, and other factors.[52,53] Dissemination should be planned and budgeted at the outset of a technology assessment along with other assessment phases or activities. However, dissemination plans should not be rigid; the nature of the findings and recommendations themselves may affect the choice of target groups and type of messages to be delivered. The results of the same healthcare technology assessment may be packaged for dissemination in different formats (eg, for patients, clinicians, or policy analysts) and delivered accordingly via different media.

Healthcare technology assessment organizations and potential users of assessments should be aware of conditions or limitations on the dissemination of reports. Of particular note, especially since passage of the FDA Modernization Act of 1997, is the ability of the Food and Drug Administration to regulate the dissemination by technology companies of published studies of off-label uses of their products (ie, unapproved indications for products already approved for certain indications).

### *Step 10: Monitor Impact*

The impact of a healthcare technology assessment should be measured primarily in terms of its ability to achieve its purposes, such as influencing a regulatory decision, an acquisition or adoption decision, a coverage decision, or clinician behavior, as appropriate. Although some assessments are translated directly into policies with clear and quantifiable impacts (eg, Food and Drug Administration approval for marketing a new medical device), others are for voluntary use. For example, assessments conducted by the Blue Cross and Blue Shield Association Technology Evaluation Center are provided as nonbinding guidance for coverage decision making by the association's 55 independent member plans. Measuring the impact of a healthcare technology assessment can range from elementary to infeasible. The impact of an assessment depends not only on the qualities of the report itself, but also on how it is disseminated; the target groups' professional, legal, contractual, or administrative obligation to comply with it; and ongoing changes in the healthcare environment.

## Conclusion

Healthcare technology assessments are evolving and their applications are increasingly diverse. The heightened demand for technology assessment, particularly from the private sector and public institutions whose technology-related policy making are under scrutiny, is pushing the field to evolve keener processes and user-specific products.

Managed care has great potential to strengthen the evidence base of healthcare technology assessment. Advances in health informatics and maturation of integrated health systems, particularly in the form of computer-based patient records, can provide a substantial stream of data that can be used to assess the short- and long-term impact of interventions on outcomes and cost. More broadly, managed care organizations can incorporate, adapt, and emphasize systematic and transparent technology assessment programs.

Various factors undermine evidence-based approaches and contribute to inappropriate use, including overuse and underuse, of technology. Demand for the latest technology, whether its safety and efficacy are proved or not, is driven by dazzling news reports, physicians' desire to specialize and use new technology, incentives and habits remaining from fee-for-service healthcare, and persuasive direct-to-consumer marketing. The underuse of many proven technologies, including beta-blockers for heart attack survivors, cochlear implants, glucose monitoring for diabetes, management of cancer pain, and many more, results from professional turf battles, poor understanding of the relative risks and benefits of technologies, and sociopolitical stances, among others. Technology demand is also driven by awareness of potential and real denial of proven interventions and information about care options. Reactions to these concerns, such as legislative mandates for minimum hospital stays after childbirth or mastectomies and for access to investigational procedures outside of clinical trial protocols, can circumvent or impede the evidence process.

## ··· REFERENCES ···


1. Passamani E. Clinical trials: Are they ethical? *N Engl J Med* 1991;324:1589-1592.
2. Ferguson JH, Dubinsky M, Kirsch PJ. Court-ordered reimbursement for unproven medical technology. *JAMA* 1993;269:2116-2121.
3. Kahan JP, Bernstein SJ, Leape LL, et al. Measuring the necessity of medical procedures. *Med Care* 1994;32:357-365.
4. Fineberg HV, Bauman R, Sosman M. Computerized cranial tomography: Effect on diagnostic and therapeutic plans. *JAMA* 1977;238:224-230.
5. Institute of Medicine. *Assessing Medical Technologies*. Washington, DC: National Academy Press; 1985.
6. Laupacis A, Naylor CD, Sackett DL. How should the results of clinical trials be presented to clinicians? *Ann Intern Med* 1992;116(suppl 3):A12-A14.
7. Fletcher A, Gore S, Jones D, et al. Quality of life measures in health care. II: Design, analysis, and interpretation. *Br Med J* 1992;305:1145-1148.
8. Gill TM, Feinstein AR. A critical appraisal of the quality of quality-of-life measurements. *JAMA* 1994;272:619-626.
9. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993;118:622-629.
10. Testa MA, Simonson DC. Assessment of quality-of-life outcomes. *N Engl J Med* 1996;334:835-840.
11. de Bruin AF, de Witte LP, Stevens F, Diederiks JP. Sickness Impact Profile: The state of the art of a generic functional status measure. *Soc Sci Med* 1992;35:1003-1014.
12. Doll HA, Black NA, Flood AB, McPherson K. Criterion validation of the Nottingham Health Profile: Patient views of surgery for benign prostatic hypertrophy. *Soc Sci Med* 1993;37:115-122.
13. Jenkinson C, Fitzpatrick R, Argyle M. The Nottingham Health Profile: An analysis of its sensitivity to differentiating illness groups. *Soc Sci Med* 1988;27:1411-1414.
14. Kaplan RM, Anderson JP, Wu AW, et al. The Quality of Well-Being Scale: Applications in AIDS, cystic fibrosis, and arthritis. *Med Care* 1989;27(suppl 3):S27-S43.
15. Linacre JM, Heinemann AW, Wright BD, Granger CV, Hamilton BB. The structure and stability of the Functional Independence Measure. *Arch Phys Med Rehab* 1994;75:127-132.
16. McHorney CA, Ware JE Jr, Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36). III: Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care* 1994;32:40-66.
17. Ware JE Jr, Sherbourne CD. The MOS 36-item Short Form Health Survey (SF-36). I: Conceptual framework and item selection. *Med Care* 1992;30:473-483.
18. Essink-Bot ML, Stouthard ME, Bonsel GJ. Generalizability of valuations on health states collected with the Euro-Qol© questionnaire. *Health Econ* 1993;2:237-246.
19. Lazaridis EN, Rudberg MA, Furner SE, Cassel CK. Do activities of daily living have a hierarchical structure? An analysis using the longitudinal study of aging. *J Gerontol* 1994;49:M47-M51.
20. O'Brien BJ, Buxton MJ, Patterson DL. Relationship between functional status and health-related quality-of-life after myocardial infarction. *Med Care* 1993;31:950-955.
21. Meenan RF, Mason JH, Anderson JJ, Guccione AA, Kazis LE. AIMS2: The content and properties of revised and expanded Arthritis Impact Measurement Scales health status questionnaire. *Arth Rheum* 1992;35:1-10.
22. Steinberg EP, Tielsch JM, Schein OD, et al. The VF-14: An index of functional impairment in patients with cataract. *Arch Ophthal* 1994;12:630-638.
23. Neumann PJ, Zinner DE, Wright JC. Are methods for estimating QALYs in cost-effectiveness analyses improving? *Med Decis Making* 1997;17:402-408.
24. Gold MR, Siegel JE, Russell LB, Weinstein MC. *Cost-Effectiveness in Health and Medicine*. New York, NY: Oxford University Press; 1996.
25. Maynard A. Developing the health care market. *Econ J* 1991;101:1277-1286.
26. Drummond M, Mason J, Torrance G. Cost-effectiveness league tables: Think of the fans. *Health Policy* 1995;31:231-238.
27. Mason JM. Cost-per QALY league tables: Their roles in pharmacoeconomic analysis. *PharmacoEconomics* 1994;5:472-481.
28. Goodman C. The moving target problem and other lessons from percutaneous transluminal coronary angioplasty. In: Szczepura A, Kankaanpää J, eds. *Assessment of Health Care Technologies: Case Studies, Key Concepts and Strategic Issues*. New York, NY: John Wiley & Sons; 1996:29-65.
29. Buxton MJ. Problems in the economic appraisal of new health technology: The evaluation of heart transplants in the UK. In: Drummond MF, ed. *Economic Appraisal of Health Technology in the European Community*. Oxford Medical Publications; 1987.
30. Goodman C, Snider G, Flynn K. *Health Care Technology Assessment in VA*. Boston, MA: US Veterans Health Administration, Health Services Research & Development Service; 1996.
31. Donaldson MS, Sox HC Jr, eds. *Setting Priorities for Health Technology Assessment: A Model Process*. Washington, DC: National Academy Press; 1992.
32. Phelps CE, Mooney C. Correction and update on 'priority setting in medical technology assessment.' *Med Care* 1992;30:744-751.
33. Goodman C. *Literature Searching and Evidence Interpretation for Assessing Health Practices*. Stockholm, Sweden: Swedish Council on Technology Assessment in Health Care; 1993.
34. Spilker B. *Guide to Clinical Trials*. New York, NY: Raven Press; 1991.
35. Peto R, Collins R, Gray R. Large-scale randomized evidence: Large, simple trials and overviews of trials. *J Clin Epidemiol* 1995;48:23-40.
36. Chalmers TC, Smith H Jr, Blackburn B, et al. A method for assessing the quality of a randomized controlled trial. *Control Clin Trials* 1981;2:31-49.
37. Evidence-Based Medicine Working Group. Evidence-based medicine: A new approach to teaching the practice of medicine. *JAMA* 1992;268:2420-2425.
38. Chalmers TC, Hewett P, Reitman D, Sacks HS. Selection and evaluation of empirical research in technology assessment. *Int J Technol Assess Health Care* 1989;5:521-536.
39. Chalmers I. The Cochrane Collaboration: Preparing, maintaining, and disseminating systematic reviews of the effects of health care. *Ann NY Acad Sci* 1993;703:156-163.
40. Fink A, Kosecoff J, Chassin M, Brook RH. Consensus methods: Characteristics and guidelines for use. *Am J Pub Health* 1984;74:979-983.

**41.** Veatch RM. Consensus of expertise: The role of consensus of experts in formulating public policy and estimating facts. *J Med Philos* 1991;16:427-445.
**42.** Dickersin K, Berlin JA. Meta-analysis: State-of-the-science. *Epidemiol Rev* 1992;14:154-176.
**43.** Laird NM, Mosteller F. Some statistical methods for combining experimental results. *Int J Technol Assess Health Care* 1990;6:5-30.
**44.** Yusuf S. Meta-analysis of randomized trials: Looking back and looking ahead. *Control Clin Trials* 1997;18:594-601.
**45.** Pauker SG, Kassirer JP. Medical progress: Decision analysis. *N Engl J Med* 1987;316:250-258.
**46.** Thornton JG, Lilford RJ, Johnston N. Decision analysis in medicine. *Br Med J* 1992;304:1099-1103.
**47.** Eisenberg JM. Clinical economics: A guide to the economic analysis of clinical practices. *JAMA* 1989;262:2879-2886.
**48.** Elixhauser A, Halpern M, Schmier J, Luce BR. Health care CBA and CEA from 1991 to 1996: An updated bibliography. *Med Care* 1998;36(5 suppl):MS1-9, MS18-147.
**49.** Canadian Coordinating Office for Health Technology Assessment. *Guidelines for Economic Evaluation of Pharmaceuticals: Canada*. 2nd ed. Ottawa, Canada: Canadian Coordinating Office for Health Technology Assessment; November 1997.
**50.** Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *Br Med J* 1996;313:275-283.
**51.** Coyle D, Davies L, Drummond MF. Trials and tribulations. Emerging issues in designing economic evaluations alongside clinical trials. *Int J Technol Assess Health Care* 1998;14:135-144
**52.** Goldberg HI, Cummings MA, Steinberg EP, et al. Deliberations on the dissemination of PORT products: Translating research findings into improved patient outcomes. *Med Care* 1994;32(suppl 7):JS90-JS110.
**53.** Mittman BS, Siu AL. Changing provider behavior: Applying research on outcomes and effectiveness in health care. In: Shortell SM, Reinhardt UE, eds. *Improving Health Policy and Management: Nine Critical Research Issues for the 1990s*. Ann Arbor, MI.: Health Administration Press; 1992.

***See following page for Suggested Reading List.***

## ··· SUGGESTED READING ···


- Antczak-Bouckoms A, Burdick E, Klawansky S, Mosteller F. Using medical registries and data sets for technology assessment. *Int J Technol Assess Health Care* 1991;7:123-128.
- Banta HD, Luce BR. *Health Care Technology and Its Assessment: An International Perspective*. New York, NY: Oxford University Press; 1993.
- Banta HD, Werk L, Cranovsky R, et al. Special section: Report from the EUR-ASSESS Project. *Int J Technol Assess Health Care* 1997;13:133-340.
- Begg CB. Biases in the assessment of diagnostic tests. *Stat Med* 1987;6:411-423.
- Brook RH. Using scientific information to improve quality of health care. *Ann NY Acad Sci* 1993;703:74-84.
- Cook TD, Cooper H, Cordray DS, et al. *Meta-Analysis for Explanation: A Casebook*. New York, NY: Russell Sage Foundation; 1992.
- Drummond MF, O'Brien B, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*. 2nd ed. New York, NY: Oxford University Press; 1997.
- Eddy DM. *A Manual for Assessing Health Practices & Designing Practice Policies: The Explicit Approach*. Philadelphia, PA: American College of Physicians; 1992.
- Eddy DM. Clinical decision making: From theory to practice. Cost-effectiveness analysis. Is it up to the task? *JAMA* 1992;267:3342-3348.
- Fowkes FG, Fulton PM. Critical appraisal of published research: Introductory guidelines. *Br Med J* 1991;302:1136-1140.
- Franklin C. Basic concepts and fundamental issues in technology assessment. *Intensive Care Med* 1993;19:117-121.
- Frazier HS, Mosteller F, eds. *Medicine Worth Paying for: Assessing Medical Innovations*. Cambridge, MA: Harvard University Press; 1995.
- Gold MR, Siegel JE, Russell LB, Weinstein MC. *Cost-Effectiveness in Health and Medicine*. New York, NY: Oxford University Press; 1996.
- Guyatt GH, Sackett DL, Sinclair JC, et al. Users' guides to the medical literature: IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group. *JAMA* 1995;274:1800-1804. (See others in this series.)
- Hasselblad V, McCrory DC. Meta-analytic tools for medical decision making: A practical guide. *Med Decis Making* 1995;15:81-96.
- Institute of Medicine. *Assessing Medical Technologies*. Washington, DC: National Academy Press; 1985.
- Littenberg B. Technology assessment in medicine. *Acad Med* 1992;67:424-428.
- Petitti DB. *Meta-Analysis, Decision Analysis, and Cost-Effectiveness Analysis. Methods for Quantitative Synthesis in Medicine*. New York, NY: Oxford University Press; 1994.
- Peto R, Collins R, Gray R. Large-scale randomized evidence: Large, simple trials and overviews of trials. *J Clin Epidemiol* 1995;48:23-40.
- Rettig RA. *Health Care in Transition: Technology Assessment in the Private Sector*. Santa Monica, CA.: RAND; 1997.
- Sackett DL, Richardson WS, Rosenberg W, Haynes RB. *Evidence-Based Medicine*. New York: Churchill Livingstone; 1997.
- Spilker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*. 2nd ed. Philadelphia, PA: Lippincott-Raven; 1996.
- Szczepura A, Kankaanpää J, eds. *Assessment of Health Care Technologies: Case Studies, Key Concepts and Strategic Issues*. New York, NY: John Wiley & Sons; 1996.
- Temple R. Problems in the use of large data sets to assess effectiveness. *Int J Technol Assess Health Care* 1990;6:211-219.
- Thornbury JR. Clinical efficacy of diagnostic imaging: Love it or leave it. *Am J Roentengol* 1994;162:1-8.
- US Congress, Office of Technology Assessment. *Identifying Health Technologies That Work: Searching for Evidence*. OTA-H-608. Washington, DC: US Government Printing Office; 1994.
- US Congress, Office of Technology Assessment. *Tools for Evaluating Health Technologies: Five Background Papers*. BP-H-142. Washington, DC: US Government Printing Office; 1995.
- US Preventive Services Task Force. *Guide to Clinical Preventive Services*. 2nd ed. Baltimore, MD: Williams & Wilkins; 1996.
- Ware JE Jr. The status of health assessment 1994. *Ann Rev Pub Health* 1995;15:327-354.
- Yusuf S, Held P, Teo KK, Toretsky ER. Selection of patients for randomized controlled trials: Implications of wide or narrow eligibility criteria. *Stat Med* 1990;9:73-83.