TNO Prevention and Health

Nederlandse Organisatie voor
toegepast-natuurwetenschappelijk
onderzoek/Netherlands Organisation
for Applied Scientific Research

B 20

TNO report

2002.019

Standard of health services purchased in the national health care system in Poland

Appendix on HTA methodology

Date	January 2002
Author(s)	H.D. Banta
Project name	Standard of health services purchased in the national health insurance system (contract nr. 3.4.1.41) - Poland
Project number	011.41053

Summary

This text is the basic methodological reference for the health technology assessment framework of the Basic Benefit Package in Poland. There are two aims of this report: 1) to define a general scientific framework which will minimise the problems of interpreting and realising the methodological aspects of HTA in discussion and reports in Poland, and 2) to explain and define useful concepts and conventions so that large parts of explanatory texts can be avoided in other reports, especially in the tasks 2, 3 and 4 of this project.

The text refers to the main concepts and definitions in health technology assessment (HTA). The reader of this text will, at the end, have a reasonable grasp of the methods used in HTA.

The methodological framework is based on international literature in the HTA area. Furthermore, it closely connects to the ongoing European Union work in standardising HTA. It also is consistent with already existing Polish pharmaco-economic guidelines.

The definition of HTA, "a comprehensive form of policy research that examines the short-and long-term social consequences of the application or use of technology" goes beyond the single meeting between the doctor and his patient and also beyond ordinary clinical treatment development in the health care. HTA is used to make decisions about use and regulation of health care resources, and how to protect and empower patients.

Health technology here is defined as the "drugs, equipment and devices, and the medical and surgical procedures used in health care, and the organisational and supportive systems in which they are delivered." Thus, the definition of health technology is a broad one, encompassing all knowledge-based aspects of health care.

HTA takes a social perspective. The social perspective differs depending on the place of the organisation or individual in the health care system. For a formulary committee in a hospital, the issue might be the population of the hospital catchment area. For the insurance company it is the current and (especially) future policyholders. The basic interest of the Polish government is the Polish population. Also, the scientific perspective has to be explicit. HTA is a normative science and it is a question about not only whose values should count and whose not, but what should be valued and what not?

The process of technology development, diffusion and prioritisation in health care is a complex system. Ideally the process goes through the steps of development, assessment, policy, use and control. Decisions how to direct and control the process is likewise a complex and sometimes informal process. The process emerges from patient/consumer needs, by a professional/industry, HTA and price processes, the government, insurance funds, health care providers and financial control. All involved parties need as accurate information as possible.

The methods and techniques of evaluations of health technologies are comparably well standardised. Efforts are made all over the world to agree on how to do these evaluations in a manner that make the results comparable. This appendix will review some of the most important concepts and definitions and will also give an indication of

the underlying rationale. It will concentrate on the social parts of the HTA, how to determine if health benefits exist, if the intervention is relatively safe, how to balance the health effects to the use of resources, and how to grapple with complex social and ethical issues.

The first concept is "efficacy". Efficacy refers to health benefits under ideal or experimental conditions. Before any technology is used outside of experimental situations, it should be shown to be efficacious. Otherwise, patients are given false hope and they are exposed to potentially dangerous procedures without the assurance of a chance to have health benefit. In the best of circumstances, with enthusiastic physicians, appropriate monitoring, good equipment, etc., clinical practice may approach the standards of efficacy. There are a number of ways to determine efficacy; the "gold standard" for efficacy is the randomised clinical trial (RCT).

Another key concept is "effectiveness". Effectiveness refers to health benefits under the usual conditions of practice after the technology has diffused into widespread use. Because community conditions are generally not up to the standard found in the best medical centers, effectiveness is essentially less than expected from evidence of efficacy. There are a number of reasons that this is true, but training and experience of the clinicians, motivation of clinicians, and budget constraints are important possibilities.

Safety or risk must be carefully balanced against efficacy. No technology is totally safe. In each individual patient, the chance of being helped must be put in relation to the chance of being harmed. In some cases, a technology (a new drug, for example) may have such harmful side effects that it is not approved for use even though it may be efficacious.

Generally speaking, those involved in HTA do not collect original data on efficacy, effectiveness and safety. Instead, they rely on the scientific literature. While the scientific literature is the underpinning of all good practice, it has been imperfectly used in most cases. Reviews of the literature, for example, generally are partial and biased and confirm the views and wishes of the person(s) doing the review. Therefore, the gold standard for using the scientific evidence is to carry out a "systematic review" of all scientific literature. This requires expertise in electronic searching, including knowledge about available databases for HTA and systematic reviews.

If efficacy/effectiveness can be assured, the next concept of interest is cost in relation to benefit. *Cost-effectiveness* is routine effectiveness related to the cost difference of the compared alternatives. However, data on effectiveness are rarely available, so cost-effectiveness must generally be estimated.

Cost should mean resources, and not only money. A resource is an asset that can be used as an input to produce a result. Staff, equipment, energy, food, diagnostic tests, assistive devices and many other things and services are resources. Also, time and simple human efforts are resources. A distinction is made between the financial cost and the resource cost. A difference is found if, for instance, subsidies, taxes or insurance premiums are involved.

There are two steps in the both clinical practice and policy-making. The first question deals with efficacy and effectiveness and answers to the question "What treatment is the

best choice, not considering resources?" The second one considers if available resources are sufficient. Finally, financial cost is conditional on these two questions and answers the question "How can this chosen treatment be financed?"

Cost data are mainly available in accounts and are collected from an accountancy perspective. But HTA studies can only use such data after careful scrutiny. Often the cost data has to be adapted and supplemented with data especially collected for the purpose of the specific HTA study. One aid for doing such primary estimations is the opportunity cost principle, which answers to the alternative use of resources. If a new and costly treatment method is used for a certain diagnosis the opportunity cost measures the resource use in the perspective of how much should have been used in the best alternative case, for example, the existing routine treatment method.

One of the consequences of the opportunity cost principle is the focus *on marginal costs: the cost it takes to treatment an additional patient (or an additional day, hour or minute for one patient). The average cost is the total cost divided by the number of treated patients.*

In some cases the desired marginal cost may be hard to assess or even define. If the average cost is used instead of the marginal cost it implies an overestimate of the true cost. It can not be an underestimate.

Finally, the time perspective plays a role for the marginal cost. In a short term perspective (days) only the time use of staff, certain materials and e.g. hospital bed use is variable (and then have an opportunity cost) and thus can be included in the marginal cost. In a longer-term perspective (years) almost all facilities are variable. Even buildings can be demolished or rebuilt, and the marginal cost will be larger since the variable cost will be a larger part of the total cost.

One of the most widespread misunderstandings of economic appraisal, and also an important cause of erroneous policy decisions, is the perspective that economic evaluations are limited to financial costs and effects. This is known as a "cost offset" estimate. Following this perspective, the effects of health care and the social costs of it are not explicitly recognised.

Standing in contrast to "cost offset" estimates is the "welfarist" perspective, which is the main one used by HTA. The welfarist looks for costs and the value for money in terms of health effects. There is also a perspective known as "extra-welfarist", which on the one hand recognises that someone may want to pay for someone else's health care, and on the other hand also takes the view that some aspects of health care need to be left to the decision-maker. This differs from the welfarist view in terms of the way that quality-of-life issues are measured.

Costs and effects must be valued differently, depending on the time perspective in which they appear. All individuals are more or less impatient. In economic terminology it has been described as if we all have an intrinsic "impatience rate". We would like to receive a benefit, money or otherwise, earlier than later if we can choose. We would also like to postpone sacrifices/costs if possible. This has also consequences for health care. It is perceived better to save a life today than to wait and save one tomorrow. The individual time preference is aggregated in HTA to a social discount rate, in which the effects of waiting of patients on costs and health effects is included in assessments.

There is general agreement that the social discount rate should be approximately five percent.

The concept of *efficiency* is often used in contrast to the concept of *productivity*. The latter responds to the question whether something can be done faster than before or with fewer people - in other words it is a method of cost containment. But the concept of productivity does not say anything about the desirability of treatment effects. For that the concept of efficiency has to be introduced. Efficiency answers the question whether resources are used for the right purpose and whether the goal fulfilment is appropriate.

A common opinion is that health care resources should go to people with the greatest need. HTA results give recommendations about how to put resources to the best use, which instead means to give treatment to the patients that can utilise the treatment, but also not to give treatment to those who can not utilise it, even if they have great needs, because that would mean that some resources would be used without giving any or low effect.

Early cost-benefit analyses concentrated on economic issues such as public and private values, pollution problems, etc. There was little interest in such things as quality of life or utilities. However, these broader effects are important to health care activities and there has been continuous discussion about how to handle them and to make them explicit, for instance, when handicap or rehabilitation activities are assessed. This early "consumerist" approach was criticised in public debates. In the late 1930s and later in the 1970s, economic analysts often took the "decision-maker's" perspective. In this approach, such immaterial values could be included under the plea that in a democratic society the decision-maker (e.g. the Ministry of Health) is selected to make such non-consumerist choices and to make them deliberately.

As a community, we pay money and other resources to health care and we require that the health care system use these resources to prolong the length of life and improve the quality of life - compared to a situation without health care. Yet, what is the value of life? While such a question is relevant to discussions of human values, purpose of life and human dignity, it is not the main question of HTA studies. Instead, HTA questions concern, for example, how treatments can be compared where, in one case, 100 people die per year and in the other 110. In some way or another it is necessary to value saving lives to be able to choose between health technologies.

Such choices are obviously influenced by social and cultural issues, and can never be entirely "rational". HTA has paid insufficient attention to this context for care, and certainly does not have any clear methodology that has been validated to carry out such a task. Nonetheless, it seems increasingly important to pay explicit attention to social consequences and ethics in making clinical and policy choices in dealing with health.

Relevant and appropriate methodology is the first requirement for doing a trustworthy assessment of health care procedures, and relevant and appropriate data is the second. Data about the health effects suffer from both validity and reliability problems. As stated already, in general, the preferred method of getting reliable effect data is to use a randomised controlled trial (RCT). This study design attempts to measure the outcomes in an efficacy type situation, where all possible factors affecting the outcome need to be controlled and the effect of the change in technology can be certified. In many cases it is not possible to do formal RCTs. In that case, several second-best choices can be

made: non-randomised controlled trial, cohort prospective studies with control, cohort retrospective study with control or case/controlled or "Before/After" studies.

For an HTA study, however, other needs also should be met. Effectiveness, as it is defined in this appendix, focuses on the routine use of the technology. In most cases that is not the focus of a RCT, which studies a technology under ideal circumstances. In some cases, therefore, separate HTA studies have to be done to check effectiveness.

Experience shows that cost data are not a great problem from a methodological point of view, but that knowledge about the right way of collecting and use these cost data could be improved. Ever since the end of the sixties, the HTA community has categorized costs as being direct and indirect. This distinction is only a convention and stems from an accountancy habit, describing direct costs as those directly involved in the production process. Indirect costs are not involved in the treatment, but may be a secondary consequence. Often the lost income is an indirect cost of the treatment. More and more HTA analysts try to move away from this direct/indirect convention.

Often issues of co-payment are linked to HTA. Co-payments in any health care system will reduce the volume of care. Unfortunately, part of this reduction will affect efficacious and cost-effective care. It will thus have effect in the sense that it will not supply efficient health care to the appropriate number of people.

Contents

1 Introduction

This appendix has been developed to support other activities in the work with the Polish Ministry of Health concerning health technology assessment and the basic benefit package. The main outcome for the project is to propose standards for the determination of this package, as well as assuring cost-effectiveness health care, through, for example, contracts.

The adoption and use of health technology should be based on well-validated information on its benefits and risks first and foremost. Assuming that benefits and risks are acceptable, information on financial implications is also necessary. In some cases, analysis of the social and ethical dimensions of a particular technological area may be helpful to decision-makers.

Such information can be produced by a process that has come to be known as technology assessment, developed because of general societal concerns about technology and its effects. By the mid-1970s, such concerns had become very visible in the health field, and medical technology assessment (later renamed health technology assessment) began to be developed.

Technology assessment is seen as a comprehensive form of policy research that examines short- and long-term social consequences (for example, societal, economic, ethical, legal) of the application of technology (Arnstein, 1977; Coates, 1977; Office of Technology Assessment, 1976). Thus, technology assessment is an analysis of primarily social rather than technical issues, and it is particularly concerned with unintended, indirect, or delayed social impacts. The goal of technology assessment is to provide policy makers with information on policy alternatives, such as allocation of research and development funds, formulation of regulations, or development of legislation.

Technology assessments have been completed in virtually every area of technological change, from examining global warming to the placement of nuclear missiles. Since 1975, the field of health technology assessment (HTA) has been developed.

HTA deals with the implications of health care above and beyond the single meeting between the doctor and his patient. It is also of a wider concern than ordinary clinical treatment development in health care itself (i.e., the internal product development), but it uses the documentation of clinical developments as a basis for analysis. HTA should be used especially concerning consequences outside health care, of how to make decisions about using the resources of health care, how to regulate its applications and how to protect patients and enhance and empower their own choice of care.

As a consequence of the "social" aspects, it is of great importance to define and express the perspective of the HTA analysis. It may be different in different situations. For a formulary committee in a hospital it might be a matter of the population of the hospital reception area. For the insurance company it is the current and (especially) future policyholders. The basic interest of the Polish government is the Polish population (rather than the voters or even the budget of the ministry). *Example: The costs to the ministry budget of developing a capacity for a certain massive prevention program may initially be very high, but for the nation the value in terms of saved future illness can be well worth the efforts.*

Finally, the perspective in terms of social consequences needs to be supplied by awareness of the scientific perspective taken. It is a question about not only who's values should count and who's not, but what should be valued and what not? HTA is a conglomerate of different methodological tools from several disciplines, and every discipline has its own internal scientific discussion. One issue that has been subject to discussion over the years is the welfarist or extra-welfarist approach to socio-economic evaluation and consequently concerning the basis of HTA. The focus of the discussion is about what constitutes the society. Is it just a sum of individuals or are there merit values also beside the total health effects of these individuals (Brouwer et al. 2000). *Examples: Has Poland got a better health care system if it can provide equal access to health care? Is it a value to society that individuals can work and provide production to the nation?*

If such effects should count they should also be recognized in the HTA value system. Today many HTA studies include productivity changes in the studies but very few include equity issues.

To assure that assessment information is available to the appropriate decision-makers at the time it is needed, many believe that a policy concerning technology assessment is needed at the national level. Chapter Three presents the structure of such a policy framework. Chapter Four sets forth the issues of efficacy and safety and their measurement and estimation. Chapter Five presents briefly the evaluation of costs, including cost-effectiveness, cost-benefit and cost-utility analysis. For those who wish to go into more depth on health economics, Chapter Six gives a more "expert" perspective. Lastly, Chapter Seven introduces social consequences of health technology and their evaluation.

2 Decision making in health technology development

What are the health technology problems that Poland has?

1 Dependence on other countries for its health technology;
2 Irrational and wasteful diffusion;
3 Physician freedom and demands;
4 Patient demands;
5 Lack of awareness of problems with efficacy and cost-effectiveness;
6 Probably great problems with indications of use, maldistribution, overuse, and underuse of health technology;
7 Dominance of industry in adoption decisions;
8 And others.

To overcome some of the difficulties, a HTA based process has been proposed leading to a straightened and formalised supervision of the proposed Basic Benefit health insurance Package (BBP).

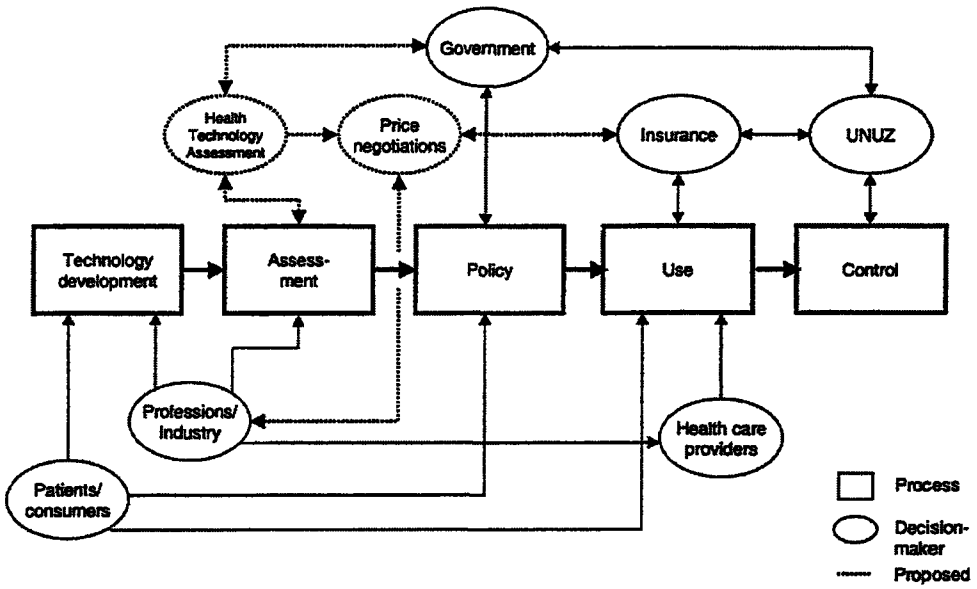
The process of technology development, diffusion and prioritisation in health care is a complex system, as depicted in Figure 1 below.

Ideally the process starts in *development*: innovations, discoveries and refinement of earlier technologies. The technologies are then *assessed*, either explicitly by formal analyses, or by implicit judgements by professionals, managers or policymakers. The assessment results in a *policy* decision, which in turn is operationalised into actual *use* by health care providers. Finally, management and financial audits *control* the use and feedback information to policymakers.

Decisions how to direct and control the process is likewise a complex and sometimes informal process. Ideally the initiation of a development process emerges from *patient/consumer* needs (demand-pull) or, probably more common, by a *professional/industry* development collaboration (technology push). Typically assessment of the new technology is provided by the developers themselves, and if not restricted by *HTA and price agencies*, will be connected close to industry marketing and professional career making.

The government sets out policy, but is influenced by the democratic control of citizens, industry and professional lobbying, and financial constraints from providing the health care capacity. The government also sets policy *for insurance funds and health care providers*, and the latter two execute the demand and supply market of what technology should be paid for. Finally, in Poland the audit agency *UNUZ* executes the financial control of the health care use.

Figure 1: Decision making in health technology development



Arrows in Figure 1 represents information flows.

Each decision-maker in the system can profit from the use of HTA. Hopefully, developing HTA will reduce the influence of political factors in policy decisions.

In summary, the users of HTA methods are:

Patients	Knowledge of HTA principles will help in judging new treatment development and to lobby patients' case.
Industry	Many producers of pharmaceuticals and a growing number of devices use HTA arguments to promote the acceptance into the market. HTA arguments are also used in pricing negotiations with insurance or government. Transparent and defined methods are needed.
Professions	HTA is the accepted professional tool to check health care resource use related to health effects. HTA work has become a professional merit for management positions in health care.
HTA process	A national HTA co-ordinating centre is the excellence centre of methods and holds the responsibility for information, research, updating and adaptation to specific Polish circumstances. Regional or hospital HTA offices execute the national policy.
Price negotiating process	A national (or insurance) centre for price negotiations between producers, consumers and financers to balance costs and usefulness. Uses information from the HTA agency for setting priorities. Regional or hospital price setting offices execute the national policy if necessary.

Government	Responsible for policy, judicial framework, investment in health care facilities and efficient use of health care resources, based on solid HTA principles.
Health care providers	Responsible for efficient use of the budget, but also of participating in clinical development. HTA principles give opportunity to take a resource perspective instead of a limited accountancy perspective.
Insurance	Balance between health care insurance costs and the legitimate claims of value for money. Use of actuarial principles in long term HTA assessment.
UNUZ	Financial supervision of insurance, but also of effective use of insurance money. HTA is a valuable tool in this work.

All involved parties need as accurate information as possible about the parts of the HTA methods applicable to them and the same methods should be used by all parties. In general, the international literature, European Union guidelines, Polish pharmacoeconomic guidelines and this methods appendix can serve as basic references.

3 A System for Health Technology Assessment

Effective planning requires information on which technologies are beneficial and how beneficial each is. It requires that technologies be applied to important health needs. It requires an idea of how much each technology will cost. It requires some idea of the social context of the technology: will it be acceptable to patients, does it challenge ethical systems, will physicians and other providers use it? Ultimately, the consequences of technological change must be well understood. The goal is to identify important developments and to assess their implications. Ultimately, such information could be used in making policy decisions.

Decisions about today's technologies should be made today. Decisions about tomorrow's technologies should be made as early in their respective life cycles as is reasonably possible. Both existing and future technology can be dealt with in a system of health technology assessment (HTA).

3.1 What Is Health Technology Assessment?

Technology assessment arose because of general social concerns about technology. The field began formally about 1965 in the Committee on Science and Astronautics of the House of Representatives of the United States Congress (Daddario, 1967). One result of the reports and discussions of the Committee was the formation of the U.S. Office of Technology Assessment (OTA) in the Congress in 1972. Thus, from it beginnings, technology assessment had the goal of providing policy makers with information on policy alternatives.

As health technology became an important policy issue, primarily because of the rising costs of health, it was natural to attempt to apply these developing ideas to it. The U.S. National Academy of Sciences published a 1973 report that examined the implications of four technologies: in vitro fertilisation, choosing the sex of children, retardation of ageing, and modifying human behaviour (Committee on the Life Sciences, 1975). The National Institutes of Health carried out an assessment of the totally implantable artificial heart in 1973 (National Heart, Lung and Blood Institute, 1973). In 1975, the Office of Technology Assessment established a health program. The establishment of the health program may be said to be the birth of the formal field of health technology assessment, although studies had begun simultaneously in Spri in Stockholm.

Within the broader context of technology assessment, differences with health technology were recognised. The first was the goal of health: to improve health. In addition, health technologies are often small and discrete, as compared to huge investments in industry, defence, or energy. The impact of health technology depends on its pervasiveness: it touches virtually everyone's life. Therefore, as it has developed, health benefits of technology have been the main focus of health technology assessment, with increasing attention to financial costs over time. The main application of health technology assessment has been to make specific decisions within specific policy areas, to be discussed in the next section.

It is important to realise, however, that the field of technology assessment encompasses much broader studies. Understanding diffusion of health technology is important in effective HTA, and this issue needs more research. The role of technology in society

and the health system should be better understood. Research on health technology inevitably moves into health policy-making and suggestions for improving policy. The public role and the public's understanding of technological issues need to be taken into account. Technology assessment should develop in such a way as to shed light on these broader questions.

3.2 Using Technology Assessment in Public Policy Mechanisms

There are a variety of public policy mechanisms dealing with health technology. Governments fund health-related research and development. Drugs are regulated for efficacy and safety. Medical devices are similarly regulated in some societies. Physicians are licensed to practice medicine and, in the United States, they may become board-certified in medical specialties; hospitals and other facilities are licensed and accredited; health-planning bodies sometimes regulate technology. And payment policy may be the most important determinant of technology use. Payment for health services by sick funds and insurance companies means that patients and physicians use technologies either free or at a small cost to themselves. Hospital, regional, provincial and even national budgeting systems severely constrain technology acquisition.

Using technology assessment as an aid to decision making in these programs is a relatively new idea. Only in the area of drug regulation is evidence concerning the technology systematically collected and used as an important determinant in decision making. And in this case, it is mainly used for the initial market introduction and labelling, but not for the ultimate use of the drug.

Each policy mechanism, each regulation, each decision concerning payment and payment levels, could be improved by using the best available information on a technology or a group of technologies. The actual information needed varies from program to program. In the area of drug regulation, scientific evidence on safety and efficacy is the main need; economic and social implications play relatively little part in the decisions. In regulation of facilities, efficacy and safety are important, but costs and efficiency are also key variables; costs and cost-effectiveness studies can be very helpful at this stage. A number of countries engage in price regulation of new technologies. In other countries, prices are determined by the manufacturer or the market. In a truly competitive market, prices will ultimately approximate the costs of production, but the medical market is far from a truly competitive market. For manufactured technologies, patents are awarded to encourage investment in research and development, which may permit monopolistic pricing except where pricing is regulated. Nevertheless, in all cases, prices should be reasonable and commensurate with the benefits provided. HTA is also used increasingly in defining health benefits, or making coverage determinations for health insurance or health systems; efficacy and effectiveness are the first concern in these programs, but other dimensions may also be important.

Today, a serious problem in using health technology assessment as an aid to decision-making is that information from assessments is often not available when it is needed. Assessments are generally not done with strategic purposes in mind. Ideally, assessments should be done in phases concurrent to the life cycle of a particular technology.

Thinking of the life cycle of a technology in five stages may help clarify this point:
- Future technology - technology not yet developed;
- Emerging technology - technology prior to adoption;
- New technology - technology in the phase of adoption;
- Accepted technology - technology in general use;
- Obsolete technology - technology that should be taken out of use.

This scheme can make some of the decision points obvious. If a future technology is a drug or device, then industry must decide whether or not to commit resources to develop it. Industry must then decide if it wants to market and promote it and it must ultimately decide whether to maintain, alter, or discard it.

If a new drug is to be marketed, then the drug regulatory system must decide if it is to be allowed on the market. If the new technology is to be used in practice, someone must decide whether to provide it, purchase it, or pay for it. Decisions must also be made to stop providing obsolete technology. Hospitals must decide whether to purchase, and practitioners and their patients must decide whether to use, the technology. Lastly, accepted or existing technology needs to be reviewed for new uses and modifications or the possibility of obsolescence.

If the technology is a medical or surgical procedure, then there are fewer decision points (although complex procedures often involve drugs and devices). Procedures are not regulated and industry is not directly involved. The main decisions, then, are institutional and practitioner decisions to provide the technology and payment decisions. The policy mechanism that can affect procedures is then the payment or budget mechanism.

To affect the diffusion of technology in a constructive way requires attention to the needs for information. Information needs differ for different technologies, but they also differ according to such factors as the stage in the life cycle and the importance of the technology for society.

One important stage in the life cycle of a technology is before it has actually been developed or evaluated: when it is an issue for the future. Until recently, there had been relatively little research concerning future health technology. However, this is now an active area of research in government-funded HTA programs, and international collaborative efforts to identify and assess new and future technology have been established, especially in Europe (Carlsson and Jorgensesn, 1998; Robert et al, 1999; Stevens et al, 1998)

3.3 A System for Assessing Health Technology

A complete system for technology assessment of all new health care technology would monitor technological change at all different stages of technological development and diffusion. The following assessments would be part of its activities:

1 Prospective assessments. Certain technologies expected to be important - in terms of costs, impact on health, or impact on the health system - could be assessed before they are developed. This is a speculative type of assessment that is concerned primarily with social effects of the technology. The attempt is to identify technology that is likely to develop. A project in the Netherlands examined five technological

areas prospectively: neurosciences, lasers, vaccines, genetic screening, and home care technology. One conclusion was that expanding opportunities in human genetics raises social issues such as discriminating against people with certain genetic susceptibilities for jobs or insurance (Banta and Gelijns, 1988). The suggestion was to make the use of genetic tests for this purpose unlawful. Another conclusion was that technology can help make home care more possible and more cost-effective (Banta and van Beekum 1988). Strategic investments in this area of research and development could give great returns to society.

2 Assessments for efficacy and safety. Traditionally, these assessments are early in the life cycle. At present, only drugs and biologics are systematically examined in this way. Many technologies come into widespread use without such assessments. Policies could be used to slow technological diffusion until such evidence was available. In addition, the system for identifying future technology would alert policy makers when a it was nearing readiness for diffusion, and assessments could be required or funded at that point. It is also necessary to assess efficacy and safety as the technology diffuses. The technology continues to improve, and it is used for ever-broadened indications. Therefore, an iterative process of assessment is necessary (Banta and Thacker, 1990).

3 Assessments for cost-effectiveness. When data are available on efficacy, cost-effectiveness calculations can be made. In many cases, cost data can be collected during early clinical trials. Cost-effectiveness studies can be done at any stage in the life cycle, but are probably most useful before widespread diffusion (Culyer and Horisberger, 1983; Drummond and Stoddard, 1984).

4 Assessments after diffusion. When a technology has diffused widely, generally little attention is paid to it. However, there are a number of reasons for examining a technology at this stage after it has been widely accepted. The costs of the technology tend to come down over time. Medical devices may become easier to handle because of manufacturer's modifications. The usefulness of the technology may be quite different in the community than it was in the university hospital. A different group of people may receive the technology. Indications may be broader: the technology may be used with less severe cases of the disease or with older or younger people. The providers may be less (or more) skilful. Patients may be less prone to follow physicians' advice.

In some cases, important technologies would be examined prospectively and then examined iteratively (Banta and Thacker, 1990). In other cases, only a few studies to demonstrate efficacy and safety might be necessary. The number and type of studies would be determined by the importance of the technological decisions to be made and by the resources available for assessment. Such assessments should be done without preconceptions. The technology might be encouraged; the assessment might show that it is valuable enough to be diffused very rapidly. The technology might be discouraged. Or it might be left alone to diffuse without active policy intervention.

An alternative approach has been called "needs-based technology assessment" (Bergevin and Tugwell, 1995; Tugwell et al, 1995). The planning process starts with the burden of illness, considering how much is modifiable by technology. The technology is assessed in controlled circumstances, those in need are identified, and community effectiveness is estimated. The cost of a program is put in relation to its benefits. All

information is then combined synthetically to arrive at recommendations on facilities planning and diffusion of technology. In its early years, HTA dealt almost entirely with specific technology. More recently, health needs and problems have become an increasing focus for HTA. For example, the Swedish Agency for Technology Assessment in Health has carried out a large assessment on back pain, and mental health services have been assessed in relation to their ability to deal with mental illness in a number of countries.

Such a complicated set of activities requires a systematic approach. A system or process of technology assessment may be viewed as an interdependent and nondiscrete flow of actions:

1 Identification: Monitoring technologies and selecting those in need of study and deciding which to study.

2 Priority-setting: Resources for evaluation are insufficient to meet the needs for assessment, so priorities must be set between candidates for evaluation.

3 Synthesis: Available scientific literature is identified and characterised. A question often asked is "Is the literature adequate for reaching judgements on the technology?" If the answer is "no", prospective research such as clinical trials may be necessary for assessment can proceed.

4 Data collection: This step involves collecting and analysing new data. A frequent situation is that cost information is not available in the literature, or is not applicable to the situation being assessed. Some cost information must then be collected.

5 Final synthesis: The information from the literature review is synthesised with the information collected prospectively to make conclusions concerning the technology. Often, recommendations or judgements about appropriate use are a result of HTA.

6 Dissemination: The synthesised information, or any other relevant information, is provided to the appropriate persons who use or make decisions concerning the use of health technologies.

7 Implementation: Since changing behaviour of clinicians and administrators is difficult with information, strategies for implementation are generally necessary. For example, linking HTA with coverage decisions will help assure implementation of HTA results in practice.

A system for assessing health technology should include all of these actions. This is not to say that one institution must carry out all of these actions. The system could be made up of a number of sub-parts. However, the system should be co-ordinated and integrated.

Identification

A decision to conduct technology assessments must be preceded by the identification of technologies that are candidates for assessment. Some identification of technologies is carried out routinely. For example, the process of regulating drugs as carried out in Poland and other countries requires that all drugs be registered before they can be

marketed. Thus, a complete list of present drugs is theoretically available, and some information on future drugs is available through this mechanism. A similar process exists for the introduction of major devices and diagnostic technologies in the United States and Europe, especially for invasive devices.

A technology that is a discrete medical procedure is unregulated. Examples include renal dialysis; gastrointestinal endoscopy; heart transplant. Some information is available through health planning mechanisms and special studies. However, this information covers only a small part of the universe of new and existing health technology. In addition, new or expanded uses of existing drugs, devices, and diagnostic technologies are often not identified. Therefore, special efforts must be made to identify procedures. One possibility, more and more used, is to examine lists of reimbursed procedures in different countries.

Priority-Setting

The setting of priorities among candidate technologies is hampered without a rather complete list of technologies, since priorities can only be set among technologies that are already known. Thus, although priorities may be set for carrying out health technology assessments, these would be improved with more information about technologies needing assessment.

Priorities for assessment might include beneficial technologies that are neglected or technologies that are suspected to be useless or dangerous. Technologies that are, or are expected to be, either expensive or widely used could also be given priority. For new technologies, potentially important advances should be given priority so that they can be assessed rapidly. For old technologies, syntheses of available knowledge can indicate if the technology has been well-tested or that it might be obsolete.

Priority setting has received increasing attention (Donaldson and Sox, 1992;208; Henshall et al, 1997; Oortwijn, 2000). The most extensive attempt to develop an analytic model was done by a committee of the Institute of Medicine (Donaldson and Sox, 1992). The model incorporates seven criteria with which to judge a topic's importance. It combines scores and weights for each criterion to produce a priority ranking for each candidate topic.

Using the model also requires judgements by a panel, data gathering by program staff, and review of a national advisory council. The model was tested during 1991. Beginning about 1995, it was used in the HTA program of the Basque Country in Spain (Henshall et al, 1997). However, most HTA programs have found the process too unwieldy and too dependent on data that in fact are not available.

The development of quantitative models seems a helpful step in the field of health technology assessment. However, all priority-setting programs also include human judgement. This is important both because of the inherent limitations of the models and because health technology assessment is itself part of human and political processes (Banta and Buch Andreasen, 1990).

Oortwijn (2000) has argued for the incorporation of social indicators in priority setting, and she has shown the feasibility of such a step in studies in the Netherlands. The

chapter in this report on social implications of health technology may indicate the importance of such a development.

Synthesis

Synthesis is a critical part of technology assessment (Marcus et al, 1987; Wortman and Yeaton, 1987), involving a critical analysis of available data from pre-clinical to clinical experience, epidemiological studies, and experiments and all other available and relevant information.

The advantage of synthesis is that it provides focused, user-oriented information at a relatively low cost (Marcus et al, 1987). If done carefully, with attention to limitations of knowledge, then synthesis can both guide technological decision-making and research to answer important questions (Woolf et al, 1990).

Until recently, reviews of health research have been carried out in an unsystematic fashion, often simply reinforcing the prejudices of the author of reflecting the relative ease of access of literature (Liberati et al, 1997). Therefore, traditional reviews have usually been based on examination of only some of the evidence. In addition, the scientific reliability of studies has not been taken into account when weighing the evidence.

This method has been replaced a systematic review of the literature. The method has been pioneered by the Cochrane Collaboration and others (NHS Centre for Reviews and Dissemination, 1996; Oxman, 1994). The systematic review has been defined as " . . . the applications of scientific strategies that limit bias to the systematic assembly, critical appraisal and synthesis of all relevant studies on a specific topic" (Cook et al, 1995).

A systematic review should be based on a clear written protocol that defines the following elements (Liberati et al, 1997):

1 **The question(s) to be addressed**, the way in which the literature will be searched, the criteria that will be used to decide which types of studies are to be included, what outcomes will be measured, and how they may be differently treated because of variations in validity are the characteristic features of a systematic review. In addition, the protocol will outline the methods to be used for extracting data from the identified studies and the analysis that will be used. By stating the protocol before conducting the main body of the review, opportunities for post-hoc alterations in the review process are reduced, leading to a more honest and unbiased assessment of the state of research knowledge.

2 **A defined and systematic search strategy** for ensuring the identification of the maximum amount of relevant research evidence increases the information available and so increases the precision of overall estimates of effect and reduces likelihood of bias due to selective identification of research results.

3 **Clear criteria for assessing the validity of studies and deciding which studies are included** in any qualitative or quantitative synthesis. This ensures that the results of the review best reflect the evidence rather than the subjective beliefs of the authors.

4 **Clear process for extracting data** from the primary study reports and checking accuracy.

5 **Strategies for analysis** is important because authors have to decide whether the review will include quantitative (statistical) pooling of data or whether this latter step is not warranted. If data are too sparse, too low quality, or too heterogeneous to proceed with their statistical aggregation, a qualitative summary of evidence is appropriate. If quantitative pooling is appropriate, and possible, then authors should specify the statistical method(s), the approaches to evaluate heterogeneity, possibly declaring the expected sources of it (on clinical/public health grounds), the kind of subgroup analyses they intend to perform, and the type of sensitivity analyses needed to explore the robustness of data. Here it should be decided, for example, whether to analyse groups of studies separately according to their quality.

If all relevant literature is not identified and assembled, the assessment will not be scientifically defensible and may be biased. The literature search is best carried out in collaboration with librarians/documentalists who are experienced at searching and database retrieval. An explicit search strategy is important, and exclusion and inclusion criteria for studies are needed. In addition, database searches should be supplemented by scanning reference lists of retrieved literature and hand-searching (where possible) of key journals. In addition, it may be helpful to consult subject experts and search databases of gray literature and conference abstracts. Furthermore, data in different languages may be important for the final conclusion. Happily, the Cochrane reviews follow these guidelines closely, and if a Cochrane review is available, it is probably reasonably reliable. Reports from HTA agencies are not so rigorous, but do generally follow these guidelines.

The question of quality of the evidence will be further discussed in the next chapter.

There are situations in which there is less reliable evidence available on the effect on final health outcomes or when that evidence is difficult to synthesise. Because the methods of HTA used in these situations are less well developed and standardised and involve more judgement, they often produce more controversial results. Methods that deserve particular attention in this respect include, among others, the role of opinion, use of modeling, qualitative synthesis, and surrogate endpoints, all to be discussed below.

Data Collection

This step includes stimulating, requiring, funding, or conducting studies and collecting data. Testing can and often should be done at various stages in the life cycle of the technology. **Obviously, any original data collection, including clinical trials, should be guided by what is already known.**

The assessment of future technology is limited (Carlsson and Jorgensen, 1998). The technology does not yet exist, so its effects cannot be directly evaluated. Nor is its costs known. Thus, only limited data collection can be done. One example would be to collect data on existing health practice that the technology is destined to change and then model the difference that a particular technology might make. Or perhaps the most important data collection done at this stage would deal with social effects. If a technology were identified as being potentially very important for the societal or the

health delivery system, it might be desirable to monitor its progress systematically or periodically. The first stage might be a thorough analysis of what could be projected concerning its potential costs, legal, economic and other relevant implications, and so forth.

Under an ideal model, all technologies would be tested for safety and efficacy before they come into widespread use. In practice, that goal is only achieved in part in the area of drug regulation. Otherwise, the testing of technologies for benefits by well-designed studies is done less often than desirable (Office of Technology Assessment, 1978).

It would be desirable to collect information on technology costs and economic consequences of the use of specific technologies. As already noted, benefit is only one part of what the policy-maker needs to know. "Benefit at what cost?" is the most important question. Once the technology exists, cost data can be collected. For example, early clinical trials can include collection of cost data; this practice is growing (Drummond and Stoddard, 1984; Coyle et al, 1998)

Social evaluations can also be done early in the diffusion of a technology. As the technology begins to be used, the early results can be very important. Ethical implications that could only be speculated upon may become clear. The technology may challenge certain important societal beliefs. It might be that early in the diffusion of a technology is the most important time to collect data on its social effects. For example, in organ transplantation, there may be an unanticipated socio-economic imbalance whereby upper-class patients tend to be the donor recipients while lower socio-economic class individuals or minorities may tend to be the donors. However, there is no clear-cut and accepted method to evaluate social effects, so several studies by different investigators would often be necessary.

Once the technology is in the clinical inventory, studies are less frequently done. One critical question concerns how the technology is actually used in practice. Efficacy information has usually been collected in well-controlled studies in academic situations -- what might be called "ideal" circumstances, with good facilities, well-trained health providers, and highly motivated staff. The results in this kind of study may have little relation to the benefit at the community level. More studies of the real-life situation with technology are needed. This has been called the "**effectiveness**" of technology (Office of Technology Assessment, 1978). Increasingly, the importance of trying to assess effectiveness is being recognised (Liberati et al, 1997).

Overall, insufficient data collection on health technology is done for many technologies. Clinical trials of technology are funded by different sources as part of applied research; however, the amounts of money are small relative to the economic and health consequences of their use. Few cost-effectiveness analyses or other economic studies are funded, although their numbers are increasing.

Synthesis

Synthesis of the information generated by the prospective studies and during the testing phase of the assessment process is necessary to provide a responsible basis for decisions made on the technology. Since policy makers and others are not trained in study design and interpretation, presenting the raw data may be of little use. Nevertheless, the data needs to be used to develop informed decisions. The main purpose of synthesis, then, is

to make knowledge relevant to policy. Synthesis reports can be used in such decisions as whether or not to pay for a given technology and can be used to examine quality of care (Chassin et al, 1987).

A cost-effectiveness analysis that does not involve the prospective collection of data is essentially a form of synthesis. So is a consensus meeting that is based on a systematic review. For policy makers such as health planners, insurers and health care providers, synthesis as almost synonymous with assessment, and this is generally appropriate. To them, assessing a technology means examining available information, summarising and analysing it, and arriving at a judgement. Without this step, there is no assessment. Synthesis might be considered as the most important step in the assessment process.

Synthesis varies quite a lot from one part of the life cycle of a technology to another. For future technology, the only information readily available may be that generated by scientific studies. Early in the diffusion of a technology, only limited information on benefits and costs is available. Later in the life cycle, much more information is available. This is one reason, perhaps, that most assessments that have been done have dealt with older technology. It is surely true that information collected later in the life cycle of a technology will be more accurate (Yeaton and Wortman, 1985). This is not an argument for waiting, however. Decisions must be made during the critical stages of the diffusion of technology, and they should be based on the best information that can be developed. If mistakes are made in early assessments, they can be corrected if assessment is developed as an iterative process.

The most widespread method of synthesis continues to be the literature review, as already noted. A report from the United Kingdom states, "Basic scientific principles are frequently disregarded, and, as a result, reviewers and their readers may be misled by biases or random errors that are similar in size to the likely effects of the technologies being reviewed" (Advisory Group on Health Technology, 1992). Increasingly, good assessments make use of **meta-analysis** techniques, a quantitative process that combines results of scientific studies, weighting studies by design characteristics (Glass, 1976; Office of Technology Assessment, 1982; Sachs, 1987; Petitti, 1994). Thus, double blind studies with large samples are given more weight than unblinded smaller studies. Meta-analysis has little possible applications in Poland at this time, and will not be discussed further.

Cost-effectiveness analysis can also be very useful in synthesising data (Office of Technology Assessment, 1980). The number of such studies has been increasing steadily in the clinical literature since the 1960s.

Group decision-making methods are very popular synthesis techniques. They seek to evoke expert opinion to aid decision-making. The National Institutes of Health Consensus Development Program, described later in this Chapter, is an example (Institute of Medicine, 1990).

Decision-analysis is a method of analysing medical decisions (Weinstein and Fineberg, 1980). The method begins from a set of symptoms or a definite health problem and develops a decision tree to show alternatives. Each decision is then assigned a probability, based on such data as expected outcomes following treatment. The entire process of diagnosis and therapy, or part of it, can then be evaluated mathematically.

Another possibility that is in widespread use in HTA is a method **of non-quantitative synthesis** involving a literature review, expert reviews of draft material, revision(s) of the draft, and development of policy options or recommendations. This method can be combined with meta-analysis, if time and data are available. Not only is sufficient information for a meta-analysis often not available, but the questions asked in HTA are such that statistical pooling of results generally does not answer the questions asked, such as effectiveness or social consequences. However, the methodological aspects of qualitative synthesis are less developed than the quantitative ones (Liberati, 1997). Until such methods have been developed, the emphasis must be on making the process explicit and transparent. To the extent possible, quantitative synthesis should follow the guidelines for any systematic review.

Another type of synthesis leading to standards for appropriate use has been developed by the RAND Corporation in California (Brook et al, 1990;Chassin et al, 1987; Fink et al, 1991; Hilborne et al, 1991; Liberati, 1997). A detailed literature review is done and presented to a panel of experts. After examining the literature, the experts rate different indications for a given procedure. A procedure is deemed appropriate for a given conditions when the benefits of performing the procedure outweigh the risks by a sufficient margin that it is worth doing and preferable to alternatives.

A type of synthesis that is gaining great visibility in the United States and Europe is developing practice guidelines, also sometimes referred to as practice policies or practice parameters (Field and Lohr, 1992; Office of Technology Assessment, 1994). Guidelines are being actively developed by physician organisations, insurance companies, government agencies, and others. Eddy (1990) defined practice guidelines as "recommendations issued for the purpose of influencing decisions about health interventions." Guidelines can be voluntary or mandatory, they can be flexible or rigid, and they can be coupled with rewards and penalties, such as financial penalties. They should be based on a systematic review or an HTA. Murphy et al (1998) have discussed the use of consensus development methods in making clinical guidelines. Eddy (1991) is sceptical about the value of guidelines as a control procedure because of the need for information on outcomes and on patient preferences, both of which are lacking. Practice guidelines might improve quality of care and avoid malpractice, but evidence that they do in fact lead to such improvements is generally lacking.

The general lack of data means that conclusions will often only be supported by uncontrolled clinical data. Such tentative conclusions, reached with the help of experts, should be better than policy decisions reached after advice from one or two specialists. It is important to emphasize that technology assessment is an iterative process, often giving tentative conclusions that may need alteration later (Banta and Thacker, 1990).

In the long run, the most important issue in synthesis is the validity of the results (Institute of Medicine, 1985). Does any synthesis method reach the right answer? More research is needed on methods to assure that the results of a synthesis are unbiased and based on the best available data.

This raises the issue of **expert opinion.** An extreme position would be that health practice should be based on scientific evidence, and not clinical opinion. However, this position cannot be supported for a number of reasons. First, clinical opinion does have value and often points out problems in scientifically-derived standards. Second, evidence from randomised trials and other studies will seldom if ever be adequate for

answering important policy questions, such as appropriate indications of use or most appropriate siting of technology. Third, the scientific questions concerning efficacy and safety may be settled by studies, but questions concerning the diffusion and use of the technology will inevitable rest on individual and social values (Liberati et al, 1997).

The RAND method, discussed briefly above, is an example of the use of expert opinion. Consensus conferences, discussed in the next section, are another. The most important question concerning expert opinion is if it in fact helps arrive at the right answer; unfortunately, although many feel that it does, there is no scientific validation of this position. Another question concerns how much the opinion of the experts is swayed by panel choice and composition; at least this points out the need for a structured and explicit approach to the use of expert opinion. It is also interesting to consider whether expert opinion helps in the dissemination process.

Consensus conferences. Consensus conferences are part synthesis, part group decision-making, using expert opinion. Since this model has been used by number of countries, it will be discussed briefly. However, consensus conferences seem to have fallen out of favour, and now seem to be used primarily as a method of dissemination of HTA results.

The National Institutes of Health (NIH) in the United States inaugurated its consensus development program in 1977, and has conducted more than 100 conferences. The goal is to bring together various concerned parties - physicians, consumers, ethicists, etc. - to seek consensus on the safety, efficacy, and appropriate conditions for use of various health technologies (Jacoby, 1985). Judgements are intended to be based on the available scientific evidence. The consensus development process is designed to produce a written document, called a "consensus statement," that can be accepted by clinicians, researchers, and the public. The statement is supposed to identify both what is known and not known about the technology.

A panel of experts is selected by NIH to hear presentations by the leading medical researchers addressing a specified set of questions. After a two-day meeting, the panel spends the evening of the second day drafting a statement intended to address the questions. The next morning, the consensus statement is read to the audience, which is invited to comment. The conference concludes with a press conference. A few final changes may be made in the statement, and then it is released to thousands of organisations and individuals and is published in leading medical journals.

The NIH consensus conference model has inspired a number of European countries to hold their own conferences. The Netherlands, Sweden, the United Kingdom, Denmark, and France have held conferences. Each country, however, developed a model specific to its own needs. This section will describe some of those differences.

In the NIH model, the focus is efficacy and safety of the technology. The audience is primarily physicians, although implicitly it is specialised physicians and researchers. NIH does not want to become involved in policy issues such as regulation, planning, economics, or reimbursement (Institute of Medicine, 1985, p. 132).

In other countries, this focus is considered too narrow. In Denmark, for example, a health technology is only socially acceptable when it is available to all persons who need the procedure (Buch Andreasen, 1988). Therefore, access to the technology is as

important as its theoretical benefits. Inevitably, then, financial costs and other considerations are a central issue.

In Denmark experts were purposely excluded from the panel, although a panels did include distinguished physicians, as well as administrators, economists, and public representatives. Reports were disseminated widely, especially to the policy makers, including Parliament. Consensus conferences in Denmark, then, were seen as part of a democratic process of dealing with technology.

The Swedish model followed closely that developed by the NIH (Calltorp, 1988). Still, in Sweden the panels are broader than in the United States, resulting in an examination of other impacts than efficacy and safety, including societal, organisational, ethical, and economic issues. The target group for the statements includes physicians, politicians, administrators, and the public. Statements were published in professional journals and are extensively covered by the lay press, radio, and television.

In the Netherlands, consensus conferences are organised by the medical profession itself. The meetings are used for the development of criteria and standards for good medical practice. Subsequently, these standards can be used as the basis for audit procedures in hospitals.

In the United Kingdom, conferences also followed the United States model closely, but panels were not experts. They represent a range of interests, including consumers. The hope is to foster a societal discussion. The questions put include cost and other aspects, but the panels do not receive sufficient information to consider "competing demands for services".

Although the NIH reports are publicly available, the focus is on physicians. In other countries, the public has played a larger part.

Some evaluative studies of consensus development are reviewed later in this Chapter. The U.S. Institute of Medicine (1990; Goodman and Baratz, 1990) has made suggestions for improving the NIH program, and has also developed standards for any program of consensus development. The standards include one important in the context of this book: "Prior to a consensus development conference, programs should provide an ordered and categorised compilation or synthesis or research reports and related evidence . . ." (Goodman and Baratz, 1990, p. 150).

There are other consensus-type activities to assist in technology decision-making. For example, the State of Oregon in the United States has employed population-based consensus meetings to set priorities for decision on public preferences for priority setting (Eddy, 1990; Office of Technology Assessment, 1992; Welsh and Fisher, 1992). In addition, physician organisations in the United States poll their members to reach expert consensus on the appropriate use of technologies.

One of the main reasons for the loss of favour of the consensus model is probably the rise in HTA. In its early years, consensus conferences were seen as a relatively inexpensive way to develop information about a technology or a class of technologies. Later, they came under criticism for not carrying out a thorough and systematic review of the available scientific literature. As HTA became more prominent, and it was clear that HTA results were in fact based on such a review, the rationale for consensus

conferences was weakened. Today, as mentioned above, the consensus conference seems mainly for the purpose of gaining professional input or informing the profession and the general public about the results of an HTA.

Dissemination

Information from HTAs are spread in several ways. **Diffusion** is a passive process by which information is spread to an audience. **Dissemination** is an active process of spreading a message to defined target groups (Granados et al, 1997).

Diffusion of information is not very effective in implementing behavior change. Research has shown that even the most active dissemination of information generally has only a modest impact on behavior (Granados et al, 1997; Office of Technology Assessment, 1994).

Dissemination of health technology assessment information is generally done in a rather passive way. The main vehicles are the medical scientific literature and meetings (Advisory Group on Health Technology, 1992). This presents a number of problems:

1 The scientific literature is oriented to research studies, not to synthesis;

2 There are large time lags in publication of studies;

3 Physicians and others do not necessarily carefully keep up on the literature;

4 The medical literature often presents studies that purport to give data on efficacy and safety but which are not based on rigorously controlled studies; and

5 The journals have little interest in cost or other social effects.

A number of studies have examined the quality of the biomedical literature (see discussion in Banta and Luce, 1993). These studies, in sum, have shown that reports on research in the literature, even in the best journals, are not totally reliable and often do not conform to generally accepted statistical methods. This problem could be made worse by increasing industry support for studies, if care is not taken.

This quality problem contributes to the practitioner attitude of scepticism toward the scientific medical literature. As a technique of scientific communication, the biomedical literature has proved its utility many times over. The issue here, however, is the dissemination of information to policy makers and practising physicians, who are not able, for reasons of time and training, to analyse the results of studies.

This problem is compounded by the explosion in the number of articles and journals. The "publish or perish" syndrome seems still to spreading through the world, resulting in a continual expansion in numbers of articles printed, with no apparent improvement in the quality of information available. The large number of journals has made the development of computerised indexing systems such as the United States National Library of Medicine's MEDLINE system necessary (Office of Technology Assessment, 1982).

These systems, however, give limited information on the usefulness of particular articles for defined purposes. One major problem is that journals require such brevity that one cannot assess the methods and weaknesses of studies to critically analyse them.

Little attention has been paid to effective distribution. That would require making contact with the intended audience and convincing the audience of the importance and validity of the information. Dissemination as part of a technology assessment process should be designed to influence behaviour. An important aspect of dissemination is convincing physicians to modify their behaviour. Professional organisations and continuing education courses can play an important role.

Better methods are needed to communication information about health technologies to physicians, researchers, and policy makers (Office of Technology Assessment, 1982). There has been little research on methods for engaging in effective dissemination or for evaluating the success of dissemination (Granados et al, 1997). What is clear is that available good information is not well-used, for example, in determining good medical practice.

Implementation is a very active process of dissemination, including interventions to reduce or eliminate barriers to behaviour change and/or activities to promote behaviour change. It may involve attempting to influence health policy or the general public, including informing the media.

Obviously, changing policy and practice is a complex process. Unfortunately, there is little empirical research on the effectiveness of dissemination and implementation strategies on changing the behaviour of policy makers or the media (Liberati, 1997). There is even less examining the cost-effectiveness of different potential strategies.

HTA has years of experience in influencing policy. In general, the experience can be frustrating. One must always have the perspective that HTA is only one input into a policy decision, which inevitably is made based on many political, social, and cultural factors.

Successful implementation requires attention to barriers to behaviour change. Barriers to behaviour change can be categorised using the following framework (Liberati et al, 1997):

Environmental barriers (e.g., political climate, lobbying by special interest groups, professional practice characteristics, financial disincentives, cultural characteristics);

Personal characteristic barriers (e.g., perception of risk, clinical uncertainty, information overload); and

Prevailing opinion barriers (e.g., difficulty dealing with acceptance of uncertainty or risk, standards of professional practice, opinion leaders, social standards).

Since barriers to change vary by target group, barriers need to be assessed for each target group. Implementation activities should be designed to overcome these barriers. Some suggestions for effective dissemination and implementation are given in Granados et al (1997).

3.4 Impact of Technology Assessment Results

It is often asked, quite reasonably, what is the effect of carrying out health technology assessments? Do patterns of care change? Are new policies developed? Because of its importance for the field of HTA, this subject will be briefly discussed.

The prior condition for technology assessment having an impact on either health policy or health practice is, obviously, that its results must be used by decision-makers. What is not so simple is identifying and measuring such use. Furthermore, many assessment programs are not able to require change; they may only reach conclusions or give recommendations, which tends to weaken their impact.

Possible positive results of technology assessment would be to allocate health resources more efficiently; to affect the adoption and use of medical innovations; to hasten abandonment of ineffective therapies; and to resolve controversies about competing treatments. Technology assessment may be most useful when either the relative benefit of a new treatment is uncertain or the relative benefits of exiting therapies are disputed.

There are two basic types of uses of technology assessment: 1) direct, program-oriented uses, such as in the use of RCT results by the US Food and Drug Administration (FDA) in its regulation of drug approvals; and 2) indirect, non-mandated ones that arise from technology assessment as part of an information strategy. In the latter case, the information generated affects the behaviour of decision-makers independent of any mandated use or sanctions for not using it.

There is little debate about the effects of direct uses of technology assessment information. In the FDA program, for example, the approval of a new drug depends on the availability of assessment information. The FDA approves a research plan; the industry conducts clinical trials and submits the results to FDA;

FDA carries out an assessment of the results of the trials; and FDA approves and disapproves the new drug. There is little doubt that the process of assessment has a great impact.

Increasingly, technology assessment is used as part of decision-making in insurance programs. In the Netherlands, as already mentioned, cost-effectiveness analyses of heart and liver transplant and in vitro fertilisation were performed with support from the Dutch Sick Funds Council to determine if coverage should be granted (Bonsel, 1991). The resulting analyses affected the policy debate, although they could not be said to have determined the outcome (Van Rossum, 1990;1990). In recent years, the use of HTA in coverage policy has spread to a number of countries (Cranovsky et al, 1997).

A specific example of technology assessment and its impacts concerns preventive services, particularly vaccines, in the U.S. Medicare program. Since its inception, the Medicare program has not included coverage for prevention. Congress was mainly concerned about providing coverage for hospital care and was apparently worried that preventive activities such as health education could not be limited. They were perhaps also sceptical about the benefits of prevention.

In 1979 the Office of Technology Assessment analysed the cost-effectiveness of the (then new) vaccine against pneumococcal pneumonia (Office of Technology

Assessment, 1979). One result found was a cost-effectiveness ratio in those 65 years of age and older of $1,000 per healthy life year gained, generally considered to be a good return in health CEAs. Congress responded and amended the Medicare law to cover pneumococcal vaccine in 1981. OTA also studied influenza vaccine in 1981 and found it to be a net medical benefit and money saving with use of the vaccine in the population over the age of 65 (Office of Technology Assessment, 1981). In 1987 the Congress amended the Medicare law to cover influenza vaccine as well. Congressional decision-makers found the results of these analyses useful, and both House and Senate committees requested OTA to examine other preventive measures applicable to Medicare.

After a report on preoperative routine testing was published by the Swedish Council for Technology Assessment in Health (SBU) in 1989, a questionnaire was used to evaluate the situation before the report was completed and was repeated in the same hospitals both in 1990 and 1991 in order to evaluate the impact of the report and the activities that have followed. The valuation in 1991 showed a significant decrease in routine preoperative testing which continued in the 1991 measurement. The savings in economic terms, besides the increase in quality of care, was 50 million crowns per year, or five times the yearly budget of SBU.

The Quebec Conseil d'evaluation des technologies de la sante (CETS, the Council for the Evaluation of Health Technologies) commissioned an evaluation of its reports in 1991. Ten reports were selected, and a wide range of policy makers, clinicians, and administrators were asked about the impact of the reports in semi-structured interviews. Eight of the ten reports had a noticeable impact on policy, six on the organisation of care and two on clinical practice. The financial impact of these reports was found to be considerable. The evaluation found that the political choice of the agenda was important, in part because available resources allowed assessment of only 5-10 percent of technologies needing assessment.

In the mid-1990, a study of the influence of HTA on a number of technologies was carried out in eight countries (Battista et al, 1994). The influence was found to vary from great influence to almost no influence. Overall, HTA carried out in Sweden was found to make an important contribution to policy decisions. The influence of HTA was also considerable in other countries, including the Netherlands. However, in Germany, which had no public approach to HTA at that time, the influence of HTA was very small.

A similar study of the influence of HTAs on prevention in members of the European Union was carried out in 2001 (the results will be published). The study found that in countries that had an active publicly-funded HTA program, assessments were used in making policy decisions. The most successful countries included Sweden, France, the United Kingdom, and the Netherlands. In other countries, with less activity in HTA, or with newer programs, the links between assessment and policy-making were often lacking.

The issue of influence of HTA on management ("meso") decisions and clinical (micro) decisions is obviously important. While the literature on effect of scientific studies on practice is extensive, and generally shows only modest impact, there are few studies of the direct effects of HTA on these levels.

The consensus conference model has been extensively evaluated, especially in the United States. Wortman and Vinokur (1982) examined the process of consensus development as carried out by the NIH, raising a number of criticisms, including the use of evidence in the consensus process (756). Wortman and Yeaton (1987) have suggested that all consensus conferences could be based on a formal synthesis document, for example, since the NIH did not consistently include experts on methodology in its panels, and it generally did not do a complete literature review. Apparently, none of the consensus programs that have recently been initiated in Europe carry out an independent literature review (Buch Andreasen, 1988; Carlsson et al, 2000). In some cases, recommendations of NIH consensus panels have not been based on the best experimental evidence (Institute of Medicine, 1985, p. 133).

The RAND Corporation carried out an evaluation of the effects of consensus conferences, but also analysed the content of consensus statements, pointing out, for example, that they tended to have an academic and scholarly tone that might not be ideal to appeal to practising physicians (Kahan et al, 1988). Overall, the statements were found to have a modest effect on physician practice (Kanouse et al, 1989; Kosokoff et al, 1987).

In Sweden, the effects of the first five consensus conferences on health policy and administration were evaluated by a questionnaire and interviews with leading politicians and administrators. More than half of the respondents indicated that they had found the statements from one or more conferences to be of practical value. In some cases, statements had a direct effect on political decisions (Calltorp, 1988). The effects of the first four conferences on physicians were also evaluated. The main target group was defined as hospital-based physicians in supervisory positions within relevant clinics. A mail survey was sent to 1,668 physicians, and 86 percent responded. Awareness of a particularly consensus conference was high. According to 7-10 percent of the respondents, a consensus statement had changed clinical practice. However, most physicians said that there had been no change, because the consensus statement reflected clinical practice prior to the conference (Johnsson, 1988).

In the Netherlands, too, the effects of consensus conferences have been evaluated (van Everdinger, 1988). Conferences on blood transfusion policy in hospitals, melanoma of the skin, prevention of bedsores, and diagnosis of deep vein thrombosis were assessed by documents, interviews, and questionnaires. The evaluation showed that most physicians were aware of the findings of the conferences and applied them in practice. For example, in the case of bedsores, the consensus document advised against many traditional forms of treatment, including ice, warm air, bandaging, local ointments, and soaps. Two years later, a survey showed that the majority of the centres had adapted and modified their practices.

The central point seems to be that information itself is not enough. Any program based on this strategy must also pay attention to such factors as the context for the use of the technology, as discussed in this chapter.

In summary, based on limited literature, one can conclude that technology assessment has had considerable impact on policy making. Consensus conferences seem to have had substantial impact in some areas. More information is needed concerning impact at the meso and micro-levels of the health system.

3.5 Conclusions

Technology assessment of health practice seems a necessary policy tool in this modern age. Still, existing programs are relatively few in numbers and their resources small. This indicates the necessity that they plan their own activities carefully. A systematic approach to such activities is thus necessary.

The advantages of a system for identifying and assessing health care technology seem obvious. Policy decisions would be improved with better information. Society's resources would be better used. Are there dangers in such a process? The main danger is that innovation could be slowed. If mistakes were made, this slowing could be inappropriate. This is one advantage of the iterative approach proposed here: mistakes could be corrected. In any case, a cautious approach is called for. Nonetheless, it seems clear that developing a system such as that sketched in this chapter offers important advantages.

4 Evaluation of Efficacy and Safety

Efficacy (benefits) and safety are the starting points for evaluating the overall utility of a health technology. If a technology is not efficacious, it should not be used, and if its efficacy is unknown, statements about its overall value cannot be made. In addition, efficacy and safety data are needed to evaluate the cost-effectiveness of a technology. Neither the need for a technology nor its appropriate use can be established without good information on efficacy and safety.

The health system and the policy-makers are, or should be, concerned about efficacy and safety because of their role as protector of the public health. In addition, the policy-makers have specific roles in the development and use of health technology. Because public funds pay for health, policy-makers have a responsibility to make good decisions in health.

Efficacy and safety have been keystones of medical practice since its beginning. However, the standard has been an intuitive one, or based on logical deduction from scientific knowledge of that particular time. The scientific evaluation of safety and efficacy is a relatively recent development, since the first randomised trials of health technologies were only carried out in the late 1940s.

4.1 Problems with Efficacy and Safety

Many technologies are not adequately assessed before they come into widespread use. It has been estimated that only 10 to 20 percent of all procedures used in medical practice have been shown to be of benefit by controlled clinical trials (Office of Technology Assessment, 1978). With the accelerating rate of technological change in health, the present situation might be worse. However, the increasing attention to evaluation and the large number of randomised trials may also mean that new technologies are better evaluated today than previously, but that medical practice is still largely based on older technologies that have seldom been evaluated.

A number of technologies have come into widespread use and then been shown to be without benefit. Gastric freezing for peptic ulcer (Fineberg, 1979) and diethylstilbesterol (DES) for pregnancy complications (Apfel and Fisher, 1987) are two examples. A number of operations as a treatment for coronary artery disease had been in fairly widespread use without evidence of benefit before coronary artery bypass grafting was shown to be efficacious (Preston, 1977).

A commonly used technology without evidence of benefit is electronic foetal monitoring (EFM). EFM enables evaluation of the foetal heart rate during labour in relation to uterine contractions and facilitates detection of certain types of abnormal patterns (Banta and Thacker, 1979). EFM was marketed beginning about 1968 and quickly spread into use in delivery rooms around the world. It presently is beginning to diffuse into less developed countries such as China.

In 1976, the first controlled clinical trials of EFM were published, showing no benefit to either mother or baby. These trials have largely been ignored by the obstetrical community. By 1988, 9 trials had been completed, including two very large ones, one in Dublin and one in Texas, still without clear evidence of benefit (Thacker, 1987). In

1988, the U.S. Preventive Services Task Force made the following recommendation: "Foetal heart rate should be measured by auscultation . . . in . . labour. Electronic foetal monitoring should not be performed routinely on all women in labour. It should be reserved for pregnancies at increased risk for foetal distress" (US Preventive Services Task Force, 1989). A review by the Cochrane Collaboration in 2000 finds no clear evidence of benefit from EFM (Thacker and Stroup, 2000) Nevertheless, EFM seems to continue to be standard practice.

This case illustrates a common problem, probably more common than total lack of benefits. A technology is developed and seems to have certain benefits, especially for those with a severe form of a particular disease. When the benefits seem relatively clear, the technology diffuses into use, but is used with more and more indications and also with different conditions, without further evidence of benefits. The result is a widespread use of technology in situations where it is probably not efficacious, or where it has a very low level of benefit. The computed tomography (CT) scanner is a good example. It is commonly used in people with headache and no other problem, although it is known that it is very rare to find abnormalities in this case (Office of Technology Assessment, 1978). Some common technologies that are often overused in a similar way include laboratory tests; many x-rays; gastroendoscopy; many surgical procedures, including tonsillectomy, hysterectomy, and appendectomy; intensive care units; many drugs (antibiotics are a dramatic example); and many procedures in obstetrics, such as Caesarean section. It is particularly difficult to see that newer diagnostic methods replace the old ones (Eisenberg, 1989). Obviously, over-use of diagnostic and therapeutic procedures can have profound implications for effectiveness, safety, and financial costs.

Balancing efficacy and safety can be a difficult problem. The example of retrolental fibroplasia (RLF), a cause of blindness in babies, is interesting (Silverman, 1985). RLF was first recognised in Boston in 1941. By the late 1940s, an epidemic of RLF was recognised in the United States, and to a lesser extent in other countries. The disease was seen in premature infants who had been placed in intensive care after birth. Gradually, evidence accumulated that oxygen supplements were responsible for the problem and attempts were made to limit oxygen. The rates of onset of RLF fell. With the growth of neonatal intensive care in the 1970s, which included oxygen supplementation, the survival rate of very small premature infants increased, but the rate of RLF also increased dramatically. However, in the 1980s, the oxygen hypothesis appeared to be too sim the role of oxygen in the genesis of RLF and in the levels of oxygen necessary to sustain life in premature new-borns. This question is still not settled (Silverman, 1985, p. 178-179).

Safety problems seem much more common than is recognised in the clinical situation. More than 3 million infections are acquired each year in U.S. hospitals, afflicting 1 in 18 people admitted to an acute care institution (Haley, 1986). In one study of 815 consecutive patients on a general medical service in a teaching hospital, 36 percent experienced an iatrogenic illness, most arising from drugs or invasive procedures (Steel et al, 1981).

4.2 The Concepts of Efficacy and Safety

Efficacy and safety have been defined many times (Office of Technology Assessment, 1978). In these definitions, four factors can be identified: benefit to be achieved,

medical problem giving rise to the use of the technology, population affected, and conditions of use under which the technology is applied. Efficacy may be defined as the probability of benefit to individuals in a defined population from a health technology applied for a given medical problem under ideal conditions of use (Office of Technology Assessment, 1978).

The question of benefits is perhaps the most important in this area. Traditionally, outcome criteria have been largely restricted to measurement of mortality and morbidity. Less consideration has been given to life expectancy or to psychosocial and functional factors. However, quality of life is becoming more and more of an important factor. In fact, measuring quality of life has become a central part of cost-effectiveness analysis. Furthermore, most valued outcomes have been defined by the health system itself, but it seems clear that the patient's preferences and values must play a part. For example, some patients will choose higher quality of life, but shortened life expectancy.

The range of outcomes can be varying a great deal when considering a particular technology. Diagnostic technologies are particularly difficult, where benefit may be examined at 5 levels (Fineberg et al, 1977):

1 Technical capability - does the technology perform reliably and deliver accurate information?

2 Diagnostic accuracy - does use of the device permit accurate diagnoses?

3 Diagnostic impact - Does use of the device replace other diagnostic procedures, including surgical exploration and biopsy?

4 Therapeutic impact - Do results obtained from the technology affect planning and delivery of therapy?

5 Patient outcomes - Does use of the device contribute to improved health for the patient?

In fact, few studies of diagnostic technologies have gone beyond the level of diagnostic impact. Most studies in the medical literature either give diagnostic accuracy or merely report that diagnoses can be made with the technology. If one assumes that the purpose of diagnosis is to improve health, one can readily conclude that this situation is not satisfactory.

The specification of benefit is often difficult for other classes of technologies as well. For example, should coronary artery surgery be evaluated for its impact on life expectancy or for its ability to relief pain? In the case of electronic foetal monitoring, earliest statements were that it would prevent deaths in the foetus and new-born. As that assertion seemed less and less likely, it was then stated that it would prevent neurological damage.

A technology's efficacy can only be evaluated in relation to a particular medical problem. It would be meaningless merely to state that a technology is efficacious. Still, the specification of medical problems can be complex. Many technologies are efficacious for one indication but not for another. Many technologies are beneficial for

dramatic cases, but benefit for less dramatic cases of the same condition have not been demonstrated.

The effect of a technology varies, depending on the population affected or the individual to be treated. A common practice is to carry out clinical trials in severely ill patients, and then to extrapolate the benefits to less severely ill people. Another common practice is to test drugs in adult males and to extrapolate the results to pregnant women and children, where testing is ethically difficult. While such extrapolations may produce correct results, strictly speaking the technology can only be considered to have proven benefit in the specific group in which it was tested.

Conditions of use are critical in a technology's efficacy. The outcome of use of a technology depends on the skills and knowledge of physicians, nurses, and other health personnel (Feeny, 1995); by the quality of the drugs, equipment, and institutional settings; and by support systems used by those personnel. In general, for example, institutions that carry out complex procedures more frequently have better outcomes (Office of Technology Assessment, 1988). Efficacy is generally estimated from studies done in excellent settings, such as university hospitals. The benefits found from use in average settings will usually be less. For this reason, effectiveness is considered to be benefits from a technology under average or usual conditions of use. Unfortunately, little information is available on the effectiveness of technology.

Safety, like efficacy is a relative concept relating to several dimensions. No technology is safe. A statement of safety is a judgement of the acceptability of a risk. Risk can be defined as the probability of an adverse or untoward outcome's occurring and the severity of the resultant harm to health of individuals in a defined population, associated with use of a health technology applied for a given medical problems under specified conditions of use (Office of Technology Assessment, 1978).

Efficacy and safety can only be evaluated fully in terms of each other. A technology may provide benefits, but the value of those benefits depends on part on the risk involved in using the technology. Thus, any use of a health technology involves a compromise between potential benefit and risk.

A difficulty in this balancing is that benefits tend to be rather clear, while risks occur in low rates and harm is often found much later. When a technology is used, some will benefit and some will be harmed. These effects do not divide equally. Some people will be harmed without benefit. Since benefit and risk are both probabilistic, the provider must be careful to be sure that undue risk is not associated with the use of the technology. In general, physicians seem to pay insufficient attention to risk.

4.3 Estimating Efficacy and Safety

Techniques used for estimating efficacy and safety range from the informal judgements of individual physicians to large complex randomised clinical trials. No technique is universally applicable. In many instances, less complex methods may be more appropriate than more sophisticated approaches. It is not possible to do multiple randomised trials on every technology. A strategy is then needed to determine which method is appropriate in a given case.

Preclinical. Many health technologies are evaluated in laboratories and in animal tests prior to human use. Such tests are usually required by regulatory agencies in the case of drugs, and often in the case of devices. The major function of animal studies is to give an initial idea of toxicity or risk.

The importance of technical studies is often overlooked by clinicians and policy-makers. However, the clinical efficacy of a drug depends on its chemical composition, purity, and dosage. The clinical efficacy of a device depends on such factors as its raggedness and its reliability.

Informal clinical assessment. Despite the increasing visibility of formal studies, informal evaluation is the norm (Office of Technology Assessment, 1978). Personal experience by a provider is key to physician attitudes about efficacy. In fact, physician experience is a key determinant of diffusion of technology.

Peer experience and consensus is more explicit than personal experience. Much of the medical literature is made up of such informal experience, without controls. Peers also interact in such form as medical society meetings. A community consensus gradually develops concerning any technology (Greer, 1988). Consensus development programs have the goal of drawing on the experience of clinical providers, as well as the scientific literature, in arriving at conclusions on efficacy and safety.

Informal assessment can seldom give more than preliminary indications on efficacy and safety. The development and course of disease is not understood sufficiently to give much credibility to observations made without control groups. At the same time, the gradual evolution of medicine has probably depended more on individual impressions than on scientific studies.

Epidemiology. Epidemiology is the study of the determinants and distribution of diseases and injuries in human populations. The methods of epidemiology are often used to study the impact of medical interventions. Descriptive studies have often established links between interventions and outcomes. The classic epidemiological investigation was that of Snow in 1855, in which he established that contaminated drinking water was responsible for an outbreak of cholera. Removal of the handle from the pump of the offending water supply ended the outbreak.

While descriptive epidemiology can give evidence of efficacy, it is more often used in determining safety. Controlled studies usually do not include sufficient subjects to find risks or to estimate their frequency. Large descriptive studies, for example of a group taking a certain drug, can reveal such risks.

In recent years, there has been increasing attention to the use of large data sets for assessing technology, especially effectiveness. Methods to use computerised data sets for this purpose are still in development and their results have not been validated (Sisk, 1990). Nonetheless, the use of large databases offers attractive possibilities for the future (Moses, 1990; Safram, 1991). Tierney and McDonald (1992) have done a useful review of eight large practice databases.

Computer modelling. Epidemiological and statistical methods are now used frequently in computer modelling to examine the efficacy of medical interventions. Computer models are also developed in cost-effectiveness analyses. Modelling is a way to

simulate the future, including complicated features of a real life process to reveal what variables have the greatest effects. Modelling is most useful when based on strong empirical investigations.

Models can be useful to:

- Deal with situations where there are gaps in or a lack of available scientific information; using a model can also be a structured way to include expert opinion.
- Generalise results to other settings (especially relevant in economic analyses.
- Extrapolate beyond the data observed in a trial or apply trial information to specific populations.
- Make a link between intermediate clinical endpoints and final outcomes.
- Bring together evidence from multiple independent sources.

The principles of developing a model are simple (Institute of Medicine, 1985). First, the important factors or variables that influence the benefits of the technology must be identified. Then the relationships between the factors must be specified. A mathematical model uses mathematics to define the relationships between variables. Mathematical models vary greatly in their level of detail and complexity, and therefore in how much time and money is needed to develop them.

However, it is most important to realise that the model is not a replacement for other methods such as controlled clinical trials (Liberati et al, 1997). The quality of the model depends on the quality of the available data. The model can be flawed or biased. Since it usually depends on a number of assumptions, it is relatively easy to manipulate the results, depending on a desired outcome.

A model should be:

- Explicit and clear. The structure of the problem needs to be clear and all relevant variables should be included. The decision-maker needs to consider explicitly the possible outcomes whenever more than two final outcomes are possible.
- Transparent. The rationale of each pathway should be explicitly described.
- Documented. It should be explicitly described where the probability and cost estimates come from. Probability or utility assessments and cost estimates may be derived from the medical/economic literature, clinical experience, expert opinion, or some combination of these.
- A comprehensive sensitivity analysis (see next chapter on cost-effectiveness) always needs to be performed to assess the degree to which the results are influenced by particular variables and assumptions.
- Any model should, where possible, be validated and comparison made with other study or model results.

Perhaps the most important point about modelling is that it is not a substitute for primary research data. They complement each other. The evidence from clinical trials provides the basis for better models. Modelling may point out key questions that a clinical trial might help to answer (Liberati et al, 1997)

Controlled clinical trials. All subjects who agree to participate in a controlled clinical trial are normally assigned to experimental and control groups. A "cohort" study follows a group of people who have received an intervention. Such information has limited value without a control group, however, because it is not known what would have happened to the group without the intervention. Commonly, historical controls are used, that is, the outcome for the group is assumed based on information from the medical literature or other sources. However, since many factors others than the intervention may be different, this method is not very reliable. Other possibilities include making up a control group from people who received no intervention or another interventions, sometimes in another hospital or clinical. Such groups can be matched to each other by parameters such as age and sex. This method is generally more reliable than historical controls, but is still not highly reliable, because it is not fully known what differences between the two groups might be of importance in the outcome.

In a randomised controlled trial (RCT), considered the "gold standard" in such evaluations, the determination of group is done by random assignment. In a controlled clinical trial intended to assess efficacy, the experimental group would be treated or diagnosed by the technology under examination; usually the control group would be treated by an established standard technology, although sometimes a placebo control is used.

Randomised controlled trials are most useful when, 1) the benefit of a new technology is uncertain, and 2) the relative benefits of existing therapies are disputed. There are three major advantages to randomisation. First, bias may be eliminated from the assignment of treatment. Secondly, randomisation prevents bias with respect to variables that exist in the experiment but are not directly considered in the design (in other words, the control group and the experimental group are comparable in all respects). Third, statistical tests of significance that are used to compare treatments are only valid in such a design.

Ethical objections to randomised trials are often raised. This issue is discussed later in this appendix.

The method of carrying out a controlled clinical trial and many of its implications have been extensively discussed in books and articles (Bailar and Mostellar, 1984; Byar et al, 1975; Institute of Medicine, 1985; Schwartz, 1992; Liberati, 1997).

From the standpoint of technology assessment, the main issue in analysing RCTs is the validity of the evidence presented. All RCTs are not the same. The method must be rigorously applied if the findings are to approximate "truth." Methods can be critically assessed by analysing reports and articles presenting trial results. A number of rating scales and standards have been devised to judge RCTs. Meinert (1980) has provided one comprehensive set of criteria:

- The source of funding for the trial and an indication of whether the reported results are a subgroup of a larger data set;

- A list of the treatment groups and the rationale for the choice of treatments;
- A description of the method to allocate patients to treatment groups, including reference to the blinding used in each group (i.e., none, single, or double blinded);
- The safeguards used in the trial to protect patients' informed consent and privacy;
- The criteria used to include patients in the trial;
- The rationale for the number of patients studied, including a statement of assumptions used in calculating the sample size;
- A statement of the length of time required to complete patient enrollment;
- A description of the population from which patients were selected;
- A description of the baseline and follow-up examination schedule;
- A specification of the key outcome variable(s);
- The descriptive information on the baseline comparability of the treatment groups;
- The number of patients assigned to each treatment group;
- The number of patients followed to the end of the study or to death;
- The number of deceased patients;
- The number of patients unable or unwilling to return for followup examinations, including a count of the number who could not be located at the end of the study;
- A description of quality control procedures used in collecting data;
- A description of the methods of analysis, including an indication whether the reported p values resulted from a single or repeated evaluation of the data; and
- A discussion of the power of the study.

An important issue in clinical trials is the role of **surrogate outcomes**. It is often difficult to obtain information on health outcomes, especially when the final results of an intervention may occur at some time in the future. Trials are sometimes funded with time limitations, such as three years. Such factors as these have led to an over-dependence on surrogate outcomes that are not necessarily related to important clinical outcomes such as mortality or morbidity.

Surrogate outcomes are typically physiological or biochemical markers that are taken to be predictive of important clinical outcomes (Liberati et al, 1997). The point is that experts assume a relationship between the intermediate and the eventual (health) outcome, even when such a link has not been demonstrated.

Examples are when the outcome of a treatment for cardiac rhythm problems is assessed in terms of its effect on cardiac rhythm rather than on cardiac morbidity or mortality and when treatments for enlarged prostate glands are assessed by urine stream rather than health outcomes such as urgency and frequency of urination.

Gotzsche et al (1996) have provided a useful set of examples of surrogate outcomes being used to assess the effectiveness of medical interventions. The authors describe three categories of possible outcomes:

- Situations where the hints provided by surrogate outcomes have eventually proved to be right;
- Cases where a rapid diffusion of a not yet proven technology has occurred; and
- Instances where results are still controversial, but the use of surrogate endpoints has stopped research on more relevant endpoints.

There are also examples where use of surrogate endpoints covered up the net harm caused by the technology. The negative and controversial cases outnumber the positive ones. This points out the need for caution.

The lesson seems clear: the use of surrogate outcomes should be avoided unless a clear link has been demonstrated between the outcome.

Formal synthesis. Synthesis of existing scientific information with clinical experience may result in valid estimations of efficacy. Synthesis may be done with the assistance of mathematical models, briefly discussed above.

In synthesising evidence of efficacy, the quality of the evidence is extremely important. It may be necessary to reach conclusions in the face of poor evidence, and often clinical evidence may be considered to be valid, yet the actual scientific evidence should ideally still be summarised.

One effective method of dealing with this problem is to make explicit the quality of the evidence behind any recommendation. The Canadian Task Force on the Periodic Health Examination graded the quality of the evidence and classified its recommendations accordingly. Quality of evidence was graded as follows (Spitzer, 1979):

I: Evidence obtained from at least one properly randomized controlled trial.

II-1: Evidence obtained from well-designed cohort or case-control analytic studies, preferably from more than one centre or research group.

II-2: Evidence obtained from comparisons between times or places with or without intervention. Dramatic results in uncontrolled experiments (such as the results of the introduction of penicillin 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.

Recommendations were classified as follows:

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 poor evidence regarding the inclusion of the condition in a periodic health examination, and 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.

These "decision rules" were accepted with little modification by the United States Preventive Services Task Force (Woolf et al, 1990).

4.4 Experiences with Assessing Efficacy

As indicated in other parts of this report, the assessment of efficacy is not new. What is relatively new is the application of statistical principles, instead of individual judgement, in such assessment. The first randomised trials were carried out in the late 1940, showing that using such methods is relatively recent.

Science is not enough in assessing a technology. The scientific base may be perfectly sound. The test, though, is does the technology work? This can only be tested in practice. Learning from controlled experimentation is central to progress in health care (Institute of Medicine, 1985).

One problem is that physicians often value their own experience more than experimental results (Geertsman et al, 1982). Physicians need more training in the importance of prospective evidence on efficacy.

Unfortunately, there are not enough experts in controlled experimentation, nor enough money, to support the careful assessment of all technologies. The present problem for health care is that most technology has never been adequately assessed. The randomised controlled clinical trial began to gain visibility in the 1960s and 1970s. Since that time, more controlled experiments have been done, and the quality of the evidence for many newer technologies is good. An important remaining problem is the lack of evidence on older technologies.

4.5 Conclusions

Limited assessment resources means that optimal studies cannot be done in many cases. This means that decisions must often be made without data from controlled studies. While the methods mentioned above, especially the application of epidemiological principles, can give suggestive evidence, one must always question its validity (Wortman and Save, 1982). The lack of good evidence is a serious problem. OTA

estimated that only 10-20 percent of commonly used medical and surgical procedures had been subjected to controlled clinical trials (Office of Technology Assessment, 1978). An examination of the syntheses done by the Office of Health Technology Assessment, which presently gives advice to the Medicare program, found that evidence of efficacy was insufficient in from 63 to 76 percent of the technologies reviewed, depending on the year (Lasch et al, 1987).

One important implication is that more primary research is needed. More research physicians with training in quantitative methods are needed. More epidemiologists and statisticians are needed. More research teams are needed. Synthesis is of limited value without data, and may be dangerous if its conclusions are not firmly based.

5 Introduction to Costs and Their Evaluation

Increasingly, health cost is a policy issue. As the percentage of the gross national product spent on health reaches and exceeds 10 percent, societies are questioning the value of such care. Economists have long warned that resources always are limited. Now, the health system is truly faced with those limited resources.

It is this background that has given health technology assessment its primary impetus. During the 1970s, most policy makers expected technology assessment to be helpful in containing costs. Gradually, they have realised that costs must be contained by financial means, such as budgets and prospective payment. The issue for the 1980s has become choice: how does a society determine how to spend its limited resources? Technology assessment seems definitely useful in aiding such choices.

Thus, the basic issue with health technology is not the absolute amount of resources spent on health, but how much benefit is obtained from such expenditures in comparison with other possible expenditures.

The methods of economic appraisal can help determine the answer to this question. A number of sources have been instrumental in helping to standardise the concepts and methods of economic appraisal as it is or may be used in HTA (Drummond et al, 1997; Gold et al, 1996; Office of Technology Assessment, 1980). Drummond et al (1997) have provided a source that is not only clear, but gives useful advice on how to actually carry out certain operations, such as using a discount rate or performing a sensitivity analysis, only touched on in this chapter. For a more technical discussion on the economic analysis see Chapter 5.

5.1 The Role of Technology in Health Expenditures

A number of studies have tried to quantify the impact of health care technology on costs. These analyses have been somewhat controversial, in part because of the complexities of the subject. The first complexity is that the definition of health care technology is broad. Using a broad definition, it seems rather obvious that technology must account for a major part of the expenditures. The second is that studies analyse the issue of costs from a number of different perspectives.

This issue has been reviewed by Banta and Luce (1993). Only a brief discussion will be provided here, focusing on the aggregate impact of technology and the impact of particular technologies or classes of technology.

Total expenditures for health reflect 1) changes in population and health needs; 2) overall wage and price inflation; 3) wage and price inflation in health in excess of general inflation; and 4) changes in service intensity, that is, the quantity of inputs per unit of health (499). The last two factors are technology-related and provide a general indication of changes in health technology and its effects on costs. Using such methods, analysts have concluded that new technology accounted for 21-24 percent of the rise in hospital costs in the United States during the 1970s and 1980s (Banta and Luce, 1993.)
Another starting point is to consider particular technologies or classes of technology. For example, Redisch (1992) looked at the cost and operating data of 7 types of ancillary services (pathology, pharmacy, nuclear medicine, laboratory, diagnostic X-

ray, therapeutic X-ray, and blood banks) in approximately 1,500 hospitals. He found that about 40 percent of the increase in operating costs was due to a higher use of these facilities per admission.

The costs of equipment are a relatively modest part of the overall expenditures for health. A typical figure might be that pharmaceuticals account for around 10 percent of health expenditures in industrialised countries, while medical equipment accounts for half of that amount or less.

An active debate has gone on for years concerning whether so-called "big ticket" (expensive) and "small ticket" (inexpensive) technology contributes more to costs. Big-ticket technology includes CT scanners, open-heart surgery, and megavolt therapy. An example of small-ticket technology is the auto-analyser in the clinical laboratory. Groot (1986) estimated that big ticket technology in the Netherlands accounted for 14.3 percent of total costs. He concluded that small ticket technologies might raise costs much more. However, both Showstack et al (1982; 1985) and Scitovsky (1985) found that between 1971 and 1982 small ticket procedures appear to contribute less than before. Laboratory tests did not contribute to rising costs and new imaging techniques were mostly substituted for older, more invasive procedures. Both studies found that some new and relatively expensive technologies, however, did increase costs. Some studies conclude that increased costs may result more from diagnostic than from therapeutic procedures (Banta and Luce, 1993).

Some studies have examined the cost of specific technological programs, such as treatment of end stage renal disease by dialysis, a major contributor to expenditures in many countries. Some other major contributors to rising expenditures include surgery for people admitted for myocardial infarction; delivery of a baby by Caesarean section; respiratory distress syndrome of the new-born; and the provision of other intensive treatments for the critically ill (of course, the clinical laboratory test is an important part of intensive care).

One problem with the literature cited is that these are studies of past technology. New and future technology can, in many cases, be cost-reducing. For example, new methods of surgery make shorter lengths of stay in hospitals, and even outpatient surgery, possible. Miniaturisation means that complicated diagnostic technology, heretofore only available in large hospitals, can be made available in peripheral hospitals and clinics. Drug delivery devices make it possible to control drug therapy much more closely, making expensive hospitalisation often unnecessary. A variety of technologies make home care more feasible for both medical and functional problems. Future analyses will need to examine such technologies.

Overall, there seems little doubt that new technology raises health expenditures. However, as will be discussed further below, this fact is not very relevant. The question is: what has been obtained from these expenditures?

5.2 Evaluating Specific Technologies

The formal techniques of comparing costs and benefits have been applied to health programs only since the late 1950s. Cost-benefit analysis (CBA) and cost-effectiveness analysis (CEA) are designed for programs that receive an insufficient evaluation in the marketplace. The marketplace fails to reflect the full effects of using a health

technology for a number of reasons. Unintended side effects occur. Knowledge is incomplete. Technologies such as immunisation have broader consequences than those for the individual because of herd immunity. But the most significant reason that market evaluation is insufficient in health is because of the nature of the health system. Consumers do not pay directly for health, or pay only a small part of the cost, in many industrialised countries. They are then unable to make rational choices that take into account all anticipated costs and benefits. Furthermore, many decisions are now made by the policy structure and by the health system. These decisions also need guidance. Increasingly, cost-benefit analysis, and especially, cost-effectiveness analysis, is the tool used to aid these decisions.

The terms cost-effectiveness and cost-benefit analysis refer to formal analytical techniques for comparing the negative and positive consequences of alternative projects. These terms are used in different ways by different people, and they may encompass a range of techniques from large prospective experimental studies that include collection of cost and effectiveness data to partially intuitive, best-guess estimates of costs and benefits done to assist a pressing policy decision. The main point is that some technique of this nature must be part of any rational decision. People use such techniques in their own personal decisions. For public policy purposes, the techniques may be more formal, but the philosophy is similar.

Both cost-effectiveness analysis and cost-benefit analysis analysts to identify, measure, and compare all of the relevant costs and consequences of alternative ways of addressing a given problem. Costs are the monetary valuations of resource inputs required to produce a health outcome. Consequences are the monetary and non-monetary results of applying a particular intervention. It is vital to recognise that socio-economic evaluations are grounded in the clinical efficacy of the interventions, which they evaluate; the clinical effects of a medication must be clearly understood before the relevant socio-economic hypotheses can be generated. The objective of these evaluations is to structure the information so that it is helpful to the decision- or policy-maker.

Through socio-economic evaluations, it is possible to make better-informed decisions regarding the use, distribution, and financing of health services and technologies. Although socio-economic evaluations are not a panacea for rising health costs, they provide information which health decision-makers can use to optimise their resource allocation decisions.

An ideal socio-economic evaluation must deal with the clinical situation. The clinical intervention normally is intended to change the health status of a patient, which is roughly analogous to the safety and efficacy or effectiveness, including quality of life impact of the technology. A change in health status has an intrinsic value; that is, people are willing to pay, and do pay, for better health. However, the intervention can directly change health services resource use, and also for instance labour force utilisation, without impacting health status. For instance, a new laser technique may permit outpatient instead of inpatient cataract surgery without affecting efficacy, or an oral formula of a medication may replace an intravenous formula of that same medication. That is, different sets of health care and other resources can create different sets of health effects. Chapter Six will discuss the importance of separating resources from effects in more detail.

5.3 A Brief Review of the Literature

The volume of cost-benefit (CBA) and cost-effectiveness (CEA) studies in the scientific literature has been increasing at a rapid rate since the mid-1960s (Office of Technology Assessment, 1980) and the pace shows signs of picking up. Overall, there has been an historical trend toward more CEAs (Elixhauser, 1992). Whereas the majority of studies from 1979 to 1990 were of U.S. origin (66 percent), that gap has been decreasing rapidly since 1986, probably mainly due to an increase in the volume of the European literature.

It has commonly been said that CBA/CEA studies are primarily for administrative decision-making, such as decisions to regulate, cover (for insurance purposes), reimburse or set price and purchase. Also, it is common wisdom that physicians are not particularly interested in costs. Yet, most studies are published in clinical journal rather than in the non-clinical health journals. Why this publication pattern? Physicians are still critical to the acceptance of a technology. Perhaps the clinical literature may be considered to include more influential journals. Finally, physicians themselves are often the authors or co-authors of CEA/CBA studies.

The literature was also analysed by class of technology. Between 1979 and 1990, 20 percent of the studies concerned preventive interventions, 36 percent concerned diagnostic technologies, and 45 percent concerned medical or surgical procedures. Thus, it is clear from a gross look at the health literature that there is a growing demand and interest in cost-benefit and cost-effectiveness studies internationally. However, the overall volume and quality of this literature is disappointing. For example, the US Prevention Task Force (1996) commented on the small number of studies and their poor quality in the area of prevention. A report from the United Kingdom concerning interventions to reduce the burden of disease made a similar remark (NHS Centre, 1995). A review of reports from the International Network of Agencies for Health Technology Assessment (INAHTA) carried out by the author of this appendix found that direct analyses of cost-effectiveness were rare, presumably in part due to lack of good literature, and that comments on costs were generally rather superficial.

5.4 Types of Socio-economic Evaluation

There are a number of evaluations that can estimate the socio-economic effects of medical technologies. They are distinguished primarily by whether and how costs and consequences are measured. Three types of studies assess both the costs and consequences of medical interventions: cost-benefit, cost-effectiveness, and cost-utility analysis (Drummond et al, 1987).

In cost-benefit analysis, the costs and the consequences of a technology are expressed in monetary terms. This often entails placing a dollar value on health outcomes. Because all consequences are valued in the same metric, cost-benefit analyses allow comparisons of disparate types of interventions with widely divergent outcomes. For example, it would be theoretically possible to compare the construction of a hospital to the construction of a dam. One drawback of CBA, however, is that not all consequences are easily estimated in monetary terms, and placing dollar values on human lives can be problematic. Furthermore, consequences that are not easily expressed in monetary terms may be ignored in a CBA, thus potentially mis-estimating the true consequences of an

intervention. In addition, policy-makers and health professionals may find it distasteful to explicitly place a monetary value on life and limb (Drummond et al, 1987).

Cost-effectiveness analysis was developed to address this limitation of CBA, and it has been used extensively in medical care. In CEA, the value of all resources consumed are measured in monetary units, but health outcomes or consequences are measured in their natural units, such as number of lives saved, years of life saved, cases diagnosed, or cases prevented. The strength of CEA lies in the fact that no dollar value is placed on human life or health outcomes. Instead the value of the technology is compared to other technologies with the same kind of outcome and the cost per achieved outcome can be assessed. Although CEA allows researchers to examine the costs per unit of health outcomes, only interventions whose outcomes are measured in equivalent units can be directly compared. For example, in a CEA, it is assumed that all years of life are equivalent: adding ten years to one person's life has the same value as adding one year of life for ten people. A year in the life of a debilitated cancer victim is considered the same as a year in the life of a patient with simple high blood pressure.

Cost-utility analysis (CUA) addresses this limitation of CEA by measuring the "utility" or value of years of life rather than just enumerating them. Outcomes are measured in terms of their quality or states of consumer preference. One drawback of this method is that the field of utility analysis is relatively young and the methodology is still developing. Unlike CEA, which can examine a number of intermediate health outcomes such as cases found or cases prevented, there is only one outcome measure in CUA: quality-adjusted life years (QALYs), healthy years equivalents (HYEs) - or a variant, disability-adjusted life years (DALYs). For the choice of what technology to apply this gives a simpler answer, but fails also to recognise possible multidimensional aspects of the technology. The uncritical use of QALYs has come under increasing criticism in recent years (Nord, 1999). Although increasingly used in HTA oriented and economic studies, CUA has not found much application in policy decisions.

Together with the three main techniques above another method termed cost-minimisation analysis (CMA) concentrates solely on netting the direct health costs associated with interventions. This technique is generally useful only when efficacy is similar or identical for each of the interventions being compared, and that is in many cases difficult to show unless a proper CEA is done (Briggs and O'Brien 2000). However, this method is relatively simple and has been widely applied in HTA.

Collectively, these types of socio-economic evaluations attempt to measure the social and economic effect of diseases and their treatments. They step beyond an assessment of the clinical efficacy of a treatment and attempt to understand the effects of treatments in a much broader sense. The specific formulation used in a study depends on the problem to be analysed, the data that are available, and the valuation methods that will be used. One type of socio-economic evaluation is not inherently better or worse than another. However in any particular case, one method is likely to be more appropriate than another based on the types of questions asked and the types of data, which are available.

Although there is no single "correct" method of carrying out a CEA, there seems to be an agreement on general principles (Office of Technology Assessment, 1980; Drummond, 1997; Warner and Luce, 1982). OTA (1980) has summarised the principles as follows:

1 Define the problem. The problem should be clearly and explicitly defined and the relationship to health outcome or status should be stated.

2 State the objectives. The objectives of the technology being assessed should be explicitly stated, and the analysis should address the degree to which the objective are (expected to be) met.

3 Identify alternatives. Alternative means (technologies) to accomplish the objectives should be identified and subjected to analysis. When slightly different outcomes are involved, the effect this difference will have on the analysis should be examined.

4 Analyse the benefits/effects. All foreseeable benefits/effects (positive and negative outcomes) should be identified, and when possible, should be measured. Also, when possible, and if agreement can be reached, it may be helpful to value all benefits in common terms in order to make comparisons easier.

5 Analyse costs. All expected costs should be identified, and when possible, should be measured and valued in currency units (dollars, Zlotys).

6 Differentiate the perspective of the analysis. When private or program benefits and costs different from social benefits and costs (and if a private or program perspective is appropriate for the analysis), the differences should be identified.

7 Perform discounting. All future costs and benefits should be discounted to their present value.

8 Analyse uncertainties. Sensitivity analysis should be conducted. Key variables should be analysed to determine the importance of their uncertainty to the results of the analysis. A range of possible values for each variable should be examined for effects on results.

9 Address ethical issues. Ethical issues should be identified, discussed, and placed in appropriate perspective relative to the rest of the analysis and the objectives of the technology.

10 Discuss the results. The results of the analysis should be discussed in terms of validity, sensitivity to changes in assumptions, and implications for policy or decision-making.

This list should guide anyone carrying out an economic analysis. It is also a handy list to use in evaluating an economic analysis carried out by other investigators or HTA programs.

As mentioned above, cost-effectiveness analysis can be carried out prospectively or retrospectively. Until the last decade or so, studies have been predominantly retrospective, based on clinical data and estimates of economic effect. However, there is a strong trend toward carrying out such studies prospectively, so that economic data is collected as part of a clinical trial, making direct calculation of cost-effectiveness possible.

5.5 The Research Question

A socio-economic evaluation begins with a clinically-based hypothesis and a research question which states the objective of the study. The research question outlines the alternatives that will be examined, the perspective taken, and the pathway of clinical management that will be considered.

To develop relevant **socio-economic hypotheses**, one must first understand the clinical effects of the technology. For example, hypotheses that are designed to test the costs and consequences of a medication are derived from clinical information on the efficacy and safety of the drug. The clinical efficacy and safety profile is the basis for any socio-economic advantages or disadvantage. A technology with socio-economic advantages, when applied appropriately, may lead to lower costs for outpatient care (fewer concomitant medications, fewer outpatient visits), lower inpatient costs (fewer hospitalisations, shorter lengths of stay), lower costs in the non-medical sector of the economy (fewer days absent from work, longer productive life), and greater quality of life (improved social and emotional functioning, greater sense of well-being). Conversely, a poor safety profile can lower quality of life and increasing health utilisation. When applied appropriately, a technology can induct costs, lower quality of life due to side effects while producing few health and economic benefits.

Most socio-economic **evaluations compare one treatment alternative with another**, even if the comparison is "no treatment." Ideally, socio-economic evaluations should examine those alternatives that are actually available or that would be realistic options in the clinical setting. It is important that the choice of alternatives is not too narrow. For example, there may be a tendency to compare one drug against another drug when a feasible and possibly illustrative option is to compare a drug to a non-drug therapy such as surgery or behavioural/educational approaches.

The choice of alternatives can introduce unacknowledged biases into the study: the cost-effectiveness of a particular technology depends to a large extent on the alternative analysed. Thus, the analyst should justify why specific alternatives were chosen and why others were not.

Perspective refers to the viewpoint from which the study is performed; in other words, whose interests are considered in the evaluation? Perspective largely defines the types of costs and consequences that will be assessed and is also a powerful determinant of the conclusions that will be drawn from the study results. The most common perspectives are society, the third party payer (insurance companies or national sickness funds), and the health provider. Studies are rarely performed from the perspective of patients alone. Generally, the societal perspective is preferred because it considers the general social welfare rather than only the well-being of a specific player in the health arena. Because studies conducted from the point of view of a specific player will tend to examine only those costs and consequences that are relevant to their budgets, the solutions from these narrower perspectives are almost always suboptimal and may lead to wasteful decisions when examined in the context of the general social welfare.

For example, a government agency in charge of decisions about reimbursement for a new imaging technology may wish to examine the impact of that device on its own health expenditures. Such a narrow analysis would ignore a vast array of costs and consequences that should be considered in making a judgement regarding the adoption

of imaging technologies. Such costs might include non-reimbursed costs borne by the patient, the costs covered by other third party payers, and the costs absorbed by health providers.

Inclusion of these costs and consequences provides a much richer base of information for decision-makers, allowing them to understand the varied effects of a technology, those costs that are often hidden, and those consequences which fall outside the limited scope of the agency's perspective. Ignoring these costs and consequences can result in decisions, which are not optimal for society in general. One agency's budget may benefit, but overall costs may increase. It is often erroneously assumed that the government's perspective is necessarily societal. Although in some instances this may be the case, when a specific component of the government is concerned with the impact of a new program or technology on its budget to the exclusion of all other budgets, a definite non-societal perspective has been taken.

The research question also sets the parameters for the pathways **of clinical management and epidemiology of disease** that will be assessed in the evaluation. These pathways define the clinical stream and epidemiologic of events which form the basis for estimating the socio-economic stream. It is at this point that the potential areas of resource use are defined. The results of the study can be altered considerably by the addition or deletion of a particular category of costs or consequences, for example, by including one pathway of clinical management, rather than another. Clinical knowledge is absolutely vital at this stage of describing the pathways of clinical management because without it important inputs and outcomes may be neglected. The same is true of the epidemiology of chronic disease such as diabetes or osteoporosis.

5.6 Fundamental Concepts of Economic Analysis

Two fundamental concepts lie at the heart of all economic analyses: opportunity cost and marginal analysis.

Opportunity Cost: The true economic cost of an intervention is the value of the benefits that would be derived from using the resources required for that intervention in their next best use. This is termed the opportunity cost. The concept of opportunity cost is the basis of the valuation of medical resource use.

In a truly competitive market, prices will equal economic or opportunity costs. The medical care market, however, is not a truly competitive one because of a number of influences such as lack of consumer sensitivity to prices (due to health insurance coverage) and providers' control of demand for services. As a result, the prices or charges affixed to goods and services do not generally reflect their true economic value. In the case of resources in global budgeted systems such as those found in hospitals outside the United States, neither costs nor charges for resources are routinely encountered at all. Because of the difficulty of estimating opportunity cost, many socio-economic evaluations use prices, charges or per diem values as a proxy for costs. To the extent that the purpose of the analysis is to assist in making choices between alternatives rather than to determine true societal value these substituted values may be an acceptable proxy for costs.

Marginal Analysis: The second fundamental concept of economic analysis is marginal analysis, which is a consequence of the opportunity cost principle above. Generally, the

basic question assessed in socio-economic evaluations is not whether to employ a particular intervention, rather, it is when to employ it, how often, under what clinical conditions, and in what specific circumstances. Thus, the question is "what is the additional cost of producing one more unit of that good or service and how much additional benefit will be derived from that level of investment?" This question deals with the marginal or incremental costs of an intervention and the marginal benefits that are expected.

Although socio-economic evaluations should ideally examine marginal costs and consequences, a major limitation faced by these studies is that generally only data on average costs and consequences are available. By using average rather than marginal costs and consequences, it is much more difficult to determine the optimal use of a technology.

5.7 Measuring Costs and Consequences

Ideally, socio-economic evaluations should include all potentially relevant costs and consequences over all time. However, practical compromises must inevitably be made. Because of limitations in money, time, and data availability, many studies examine only those particular costs and consequences that are expected to be most salient. Researchers must use judgement in simplifying a broadly stated research question into a workable study plan. Treatment costs include the value of actual changes in resource use that are attributable to the technology and also the patients own costs for participating in the treatment. It is important to examine not only those direct medical costs that are associated with the technology, but to look beyond technical-related costs to other sectors of the health market. Other nonmedical costs include resource expenditures outside the medical care market, such as costs borne by patients in getting health related information and seeking care.

Costs are often subdivided into fixed and variable costs. Variable costs vary with the volume of services rendered while fixed costs remain constant across the entire range of service volumes. By definition, variable costs such as medical supplies or medications are gained or lost depending on whether they are used and in what volumes. On the other hand, fixed costs are "fixed," that is, they cannot be "saved," at least not in the short run, regardless of whether or not they are used. From a policy- or decision-maker's standpoint, it is critical that costs be examined carefully to estimate, which are truly variable to the program that is affected by the medical technology under study.

The consequences of medical interventions are the clinical, psychosocial, and economic outcomes of employing that intervention. Clinical consequences include measures of death, disability, and illness. Psychosocial consequences include personal outcomes such as pain and anxiety, and social effects such as work days lost and job changes resulting from illness. Many of these outcomes were considered intangible until methodological advancements such as utility analysis and quality of life assessment were made. Of course, very important economic consequences of a medical intervention are the changes in health costs themselves: savings or additional expenditures that result from the use of particular interventions in different points in time. But these consequences should be regarded as different ways of using the health care and other resources in order to achieve the health effects. Although the goal of socio-economic evaluation is to assess the full range of consequences that result from the use of a

technology, such a broad approach is often not feasible, practical or sometimes even necessary.

Measuring consequences is a challenging issue. Many important outcomes are not quantifiable in monetary terms, and thus cost-benefit analysis is not always possible. Clinical consequences such as morbidity and mortality may be appropriate for a given research question. However, these measures often do not capture the full range of outcomes that result from treatment.

Quality of life assessment and utility analysis are means of measuring those effects that have previously been considered intangible. In quality of life studies, patients are asked to respond to a series of questions about their social, psychological, physical, and intellectual well-being in order to identify precisely what aspects of a patient's life are affected by a disease or intervention. Such studies can result in detailed accounts of the impact of interventions and can provide the basis for choosing among interventions which have very similar costs but which can have different effects on patients' personal well-being. These consequences of therapy are not assessed using typical clinical measures. Some of the quality of life assessment issues will be discussed in detail in Chapter Six.

Utility analysis seeks to compress these psychological effects into one summary measure in order to compare vastly divergent outcomes in terms of a common metric, usually expressed on a scale that ranges from 1 (perfect health) to 0 (a situation equal to death). Life-years weighted by utility values form the denominator of a cost-utility ratio, a measure that is easily grasped and which effectively summarises health effects information. As already mentioned, this method is widely used in literature but has rarely been used in decision-making and is frequently criticised.

In general, researchers should begin a study by enumerating the full range of consequences that are based on the clinical evidence that is available. They should evaluate which consequences are most relevant and which are most likely to differ between the alternatives compared. They should examine these consequences to appraise which are measurable in monetary terms, which are measurable in natural units and which can be valued using quality of life or utility measures. Finally, in reporting results, researchers should make explicit which potentially important outcomes are not included in the final results and what the impact of these omissions might be on the interpretation of results. A qualitative description of the consequences that have been excluded from the analysis is often helpful.

5.8 Sensitivity Analysis

Researchers face significant challenges in accurately estimating the impact of interventions on costs and on patients' well-being. These challenges often mean that relatively gross estimates must be made and that these estimates are subject to considerable uncertainty. Assumptions must often be made regarding resource use, the costs of resources, or the health effects of an intervention. When such assumptions are made, researchers cannot be certain that their conclusions are tenable, that is, whether the conclusions are "sensitive" to changes in these assumptions. Sensitivity analysis is an important analytical tool, which can be used to test whether the conclusions of a study change as assumptions are altered. It should always be performed whenever there is uncertainty about key variables.

When faced with uncertainty about the true value for some costs or consequences, researchers will generally begin by choosing a value that is a "best estimate." They will then vary that estimate, usually suggesting a high value and a low value and then repeat the calculations to see whether their results change under the high and low assumptions. If the conclusions drawn from the study do not change as these values are altered, the results of the analysis are "insensitive" to changes in this variable and one can be comfortable with the conclusions. However, if the conclusions change with the sensitivity analysis, then one's faith in the analysis is less certain.

5.9 Discounting

When costs or consequences do not occur within a relatively short time frame, or when costs and consequences do not occur at the same time, the results of the socio-economic evaluation should be adjusted to reflect the positive value of time preference. This means that costs and consequences incurred in the present have greater importance than those that occur in the future, even if inflation and bank interest would not exist. Basically, discounting reduces future costs and consequences to their present value by a discount rate, commonly between three to five percent annually.

5.10 Other Study Design Issues

As mentioned above, socio-economic evaluations can be conducted by using prospective or, more commonly, retrospective data. Prospective data might be collected in conjunction with a randomised controlled trial (RCT) of a medication whereas a retrospective study uses data from existing sources, such as the clinical and economic literature or a claims data base from a third party payer (Drummond and Stoddart, 1984).

Inclusion of economic measures and quality of life surveys within a clinical trial is a relatively recent phenomenon, but this practice can be expected to increase in the future. Such prospective studies can supply highly valid information because of the strength of the RCT design. However, there may be problems in generalising the results to other patient groups. Because RCTs are expensive, they are often performed on relatively small and specialised groups of patients, therefore it is not clear that the results can be generalised to other populations (Coyle et al 1998).

On the other hand, retrospective studies often make use of large population-based data sets, or they use "best estimate" values in conjunction with analytic modelling techniques to simulate general population values. These studies suffer from a different kind of uncertainty. Because the data are seldom derived from well-controlled studies, the precision of values may suffer. Sensitivity analysis as discussed above, is particularly helpful in analysing the results and improving the "believability" of retrospective studies.

5.11 Conclusion

One of the principal reasons that health technology assessment is an issue today is due to its impact on the cost of medical care. Health technologies tend not only to be expensive themselves, especially new technologies, but their use is apt to have significant economic and health consequences.

Socioeconomic evaluations can be important tools for decision-makers in the health field. Health decision-makers face increasingly difficult choices between competing alternatives, such as choices between imaging modality or medications with apparently equivalent efficacy. If these choices are based simply on a comparison of the monetary costs per scan or dose of the drugs, other important costs and consequences may be disregarded: these include costs such as those associated with the utilisation of medical resources to treat significant adverse reactions, or consequences such as differences in quality of life.

Socioeconomic evaluations can provide balanced and impartial appraisals of the relative costs and efficacy of interventions and are increasingly essential tools for decision-making.

However, health economic evaluation is still developing. In the last 10 years, analyses of costs and cost-effectiveness have been increasingly used as an aid to public policy making. This trend is probably due to the increasingly severe budget constraints and high costs of care in all countries. The analyst has a heavy responsibility to present all relevant information concerning the economic consequences. Policy-makers often seek the easy decision. They have more interest in the "bottom-line" than in the actual framework of the analysis and in simple cases that might be enough. But the methods described in this chapter are crude, and to base decisions on them without the input of human judgement would not be wise.

Although socio-economic evaluations attempt to measure and evaluate the costs and consequences of a medical treatment as accurately as possible, they face restrictions of having to gather data in the real world and having to draw conclusions on the basis of incomplete information. As a result, the costs measured in economic analyses are sometimes not true opportunity costs. They represent either an imperfect price set in a non-competitive marketplace or an arbitrary price set by a government commission or other proxy. If this imperfect price deviates markedly from the true price is in some cases not known. And even though marginal analysis is a critical goal of an economic evaluation for optimal decision-making, average costs and consequences are usually the only costs available. The same lack of knowledge prevails: are there a large difference between the marginal and the average prices? In many cases maybe not, in other - possibly. The important thing is if it would make a difference for the decision to be taken.

Furthermore, socio-economic evaluations generally collect data on only the most salient costs and consequences on a few alternatives, and do so over a relatively restricted time frame. As a result, some important costs and consequences may be neglected and some important alternatives ignored if not the analyst can correctly identify and discriminate between economic essentials and non-essentials. Once again a great responsibility lies on the analyst to do a professional job.

6 An Expert Presentation of Concepts and Methods of Economic Appraisal

The previous chapter has attempted to provide a basic introduction to financial costs and economic analysis. In this chapter, many of the same concepts are discussed, but in more depth and with examples. Those curious to know more about economic appraisal are encouraged to read this chapter.

The methods and techniques of economic appraisal are comparably well standardised, as mentioned in the previous chapter. Efforts are made all over the world to agree on how to do these evaluations in a manner that make the results comparable. This chapter will review some of the most important concepts and definitions and will also give an indication of the underlying rationale. It will concentrate on the social parts of HTA, and how to balance the health effects to the use of resources.

6.1 Costs Versus Resources

One of the most important concepts of HTA studies and the central concept in all economic studies is the cost: *the value of the amount of one kind of resource used to produce a health result or effect.*

A resource is an asset that can be used as an input to produce a result. That means that not only money is a resource. Staff, equipment, energy, food, diagnostic tests, assistive devices and many other things and services are resources. Also time and simple human efforts are resources, sometimes not exchangeable to money. "I would like to go to the doctor. It doesn't matter what it costs but I don't have the time."

The *price* is the cost for one unit of the resource. That means that the price is dependent on how the amount is measured (i.e. in this case price per mg, ml or DDD). Example: If 140 mg of a certain chemotherapeutic is used to treat cancer in the induction phase, and the *price* of the drug is 5.10 PLN per mg, the cost of this treatment resource will be 714 PLN.

Very often costs are not what they seem to be. In many countries health insurance makes immediate health care cost less, or even disappearing to the patient. It is often not recognised that the patient on average already has paid for the care by his insurance premiums. Thus, the *financial cost* to the patient of a doctor's visit is lower than the actual *resource cost* of the doctor to the hospital. Also, if the government subsidizes certain activities, like highly qualified procedures, the cost to the hospital will be lower than the actual resource cost.

Furthermore, if the tax system uses income taxes and/or pension schemes there is a difference also between the hospital resource cost and the salary to the doctor. In this case the financial cost is higher than the resource cost. That is also the case for cost where VAT taxes are laid upon the selling price. If a hospital purchase a set of patient beds, the cost for the hospital is greater than the benefit to the bed producer, because the government will have a certain fraction of the price per bed as a tax. Often the evaluator will need to consider if tax or subsidy issues play a significant role or if they can be neglected. If taxes or subsidies shall be included or not in the evaluation is a matter for

experienced evaluators. Some guidance can also be found in for instance Sugden and Williams (1978).

To sum up: *charges are not cost*!

In the following discussion, the difference between the two perspectives will be called the *resource perspective* and the *financial perspective* respectively. This has consequences for the correct assessment of costs, and it also leads to two consecutive kinds of decision-making. The first one (resource cost) answers the question "What treatment is the best choice, considering benefits and resource use?" The second one (financial cost) is conditional on the first and answers to the question "How can this chosen treatment be financed?" If the second question can not be answered, for instance because there is no budget available, the decision has to go back to the first question and find a second-best choice.

The primary goal for the economic analysis is to be helpful to the decision-maker. It is there obviously important, both from the efficacy point of view, but also from the resource point of view. Cost data are mainly available in accounts and are collected from an accountancy perspective. It follows from above that economic studies can only use such data after careful scrutiny and often cost data has to be adapted and supplemented with data especially collected for the purpose of the economic study. This should be done under supervision from a professional with economic expertise.

6.1.1 The Opportunity Cost Principle

One of the most important issues in HTA is the concept of opportunity cost. The opportunity cost principle answers the question concerning *the alternative use of resources*. If a new and costly treatment method is used for a certain diagnosis the opportunity cost measures the resource use in the perspective of how much should have been used in the best alternative case (the control group), most probably the existing general accepted routine treatment method. In the case of a new treatment for a previously non-treated diagnosis only, the alternative of a placebo treatment could be considered. Should there be treatments of another kinds, for instance a new primary care drug treatment aims to replace an old hospital based expensive machinery treatment, the latter is the proper comparison.

The opportunity cost principle can also be applied to other alternatives than variations in treatment. Consider a program for screening of children for having celiac disease (gluten intolerance). The illness can be silent for years but still lead to serious secondary problems like infertility, osteoporosis and some forms of cancer. By finding silent cases early, recommending strict non-gluten diet, the secondary consequences may be reduced. The two alternatives are consequently on the one hand the screening cost plus the cost of the diet, and on the other hand the cost of treating secondary consequences later in life.

6.1.2 Cost Checklist

Many HTA analysts have problems of identifying resources involved in a health care process. The resource concept itself is confusing, the relation to alternative actions and the distinction between the financial and the resource perspective does not make identification easier.

Example:
Costs and health effects of relatives and friends are frequently neglected as a resource use of "informal care" resources. The reason is that such services are seldom paid, they do not appear in any registers, and the opportunity cost, that is what these informal caregivers should have done otherwise, is not very clear even to themselves. Besides, time is in many cases not seen as a resource. As a consequence such resources are undervalued, for instance in home care programs.

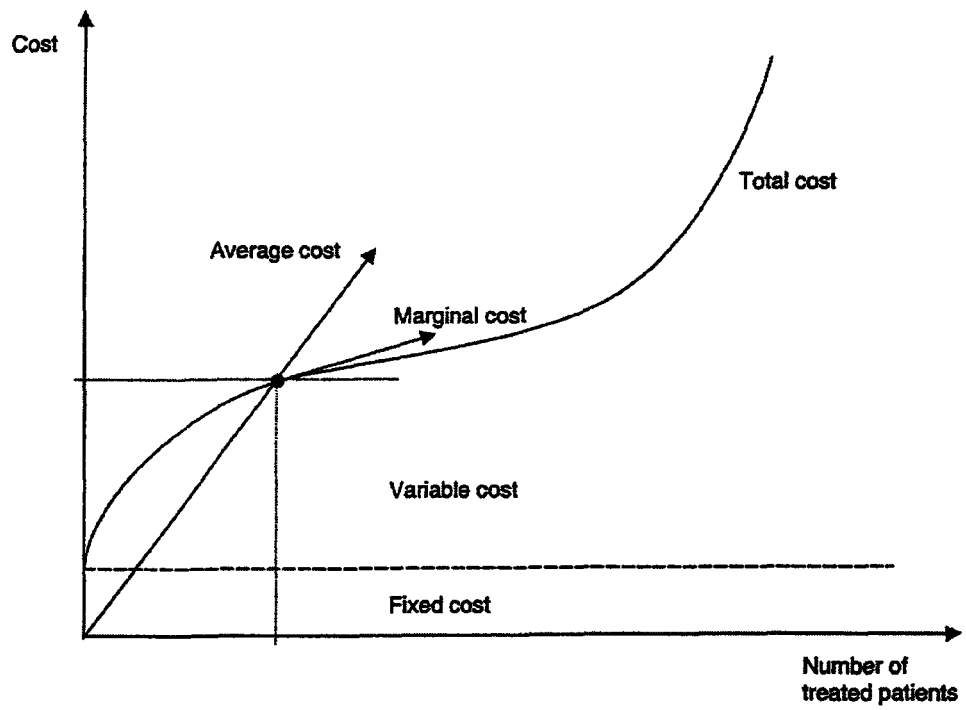
A checklist of different kinds of resources may be of some help. The different alternatives may be mapped in an economic matrix. Here are two different interventions listed, half and full dose of a new pharmaceutical, and one control, the current routine treatment. Also, the health effects in terms of function, length of life and quality of life has been listed in a similar manner. The HTA researcher is supposed to fill three copies (a, b, and c) of this matrix with a) the identification of what resources and effects are affected or changed by the program, for instance by just assigning tick marks in the appropriate cells, and b) assigning quantities (days, amounts, doses, numbers) to the ticked cells and then c) by multiplying the price to the quantities getting the cost for each cell. Then the matrix or matrices can be summarised and aggregated to the necessary level (patient group, hospital, geographical region, etc.)

Resource type	Half dose	Full dose	Control	Difference 1	Difference 2
R. Changes in resource use					
R1. Health care operation					
Inpatient care:					
Staff					
Equipment					
Materials					
Services from other depts.					
Outpatient care:					
Staff					
Equipment					
Materials					
Services from other depts.					
Pharmaceuticals					
Other (dental care, self care, etc.)					
Resources from relatives/friends:					
Time forgone					
Equipment					
Materials					
Payments for external services					
R2. Rehabilitation operation					
Therapy (e.g. Physiotherapy, Occupational, Psychotherapy)					
Training (On-the-job training, Education)					
Other rehabilitation					
Resources from relatives/friends (Time forgone, Equipments, Materials, Payments for external services)					
R3. Transportation					
Admission					
Acquisition					
Adaptation					
Operation					
Maintenance					
Resources from relatives/friend (Time forgone, Equipments, Materials, Payments for external services)					
Other Resources					
R4. Assistive devices					
Admission and prescription					
Acquisition					
Adaptation					
Operation					
Maintenance					
Disposal					
Other resources					
R5. Shelter/Accommodation					
Admission					
Acquisition					
Adaptation					
Operation					
Maintenance					
Miscellaneous					
Resources from relatives/friend (Time forgone, Equipments, Materials, Payments for external services)					
R6. Human capital time					
Physical					
Mental					
G. Changes in treatment effects					
G1. Life expectancy					
G2. Quality of life					
Dysfunction					
Disability					
Handicap (Participation)					
G3. Life expectancy of relatives/friends					
G4. Quality of life of relatives/friends					

6.1.3 *Marginal cost*

One of the consequences of the opportunity cost principle is the focus on marginal costs. The marginal cost is *the cost it takes to increase the treatment with one extra patient (or one day, hour or minute for one patient).* It can be proven that in an ideal situation the short-term marginal cost is the minimum amount of resources required to achieve a certain result. That means that if we attempt to estimate the differential marginal cost for the intervention and the control we would also know the minimum cost for achieving a change in the treatment results.

Figure 2. Fixed cost, variable cost, average cost and marginal cost.



In most cases the *total* cost looks like the curved line in the figure. Before any patients can be treated the capacity has to be built: buildings, administration, "hotel services", supporting laboratories etc. Those are the *fixed costs*. Often they are referred to as "overheads". This investment period is followed by a "breaking in" period when systems are tested, equipment calibrated etc. with a small number of patients. Every patient cost extra resources to treat and the sum of these treatment costs is the *variable* cost. The cost curve makes its way to the right in the figure, increasing the number of treated patients but also to an increasing variable cost. *The total cost is the sum of the fixed cost and the variable cost.* At normal throughput the cost is rather low per patient, but then, close to the capacity limits, cost will increase again due to congestion problems. Finally, when costs are running high (far to the right in the figure), a rebuilding has to take place and the cost curve starts from another beginning.

The slope of the cost curve is called the *marginal cost. It is the extra cost it takes to treat an extra patient. The average cost is the total cost divided by the number of treated patients.* In the figure both marginal and average costs are shown as angles to the curve and the origo respectively. Clearly the variable cost is larger than the marginal

cost all along the curve except for the point where they coincide, before the capacity limits.

This shows an important fact:

In some cases the desired marginal cost may be hard to assess or even define. *If the average cost is used instead of the marginal cost it implies an overestimate of the true cost. It can not be an underestimate.*

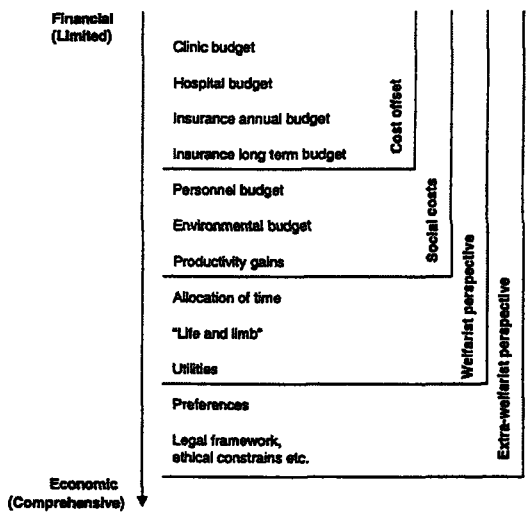
Finally, the time perspective plays a role for the marginal cost. In a short term perspective (days) only the time use of staff, certain materials and e.g. hospital bed use is variable (and then have an opportunity cost) and thus can be included in the marginal cost. In a longer-term perspective (years) almost all facilities are variable. Even buildings can be demolished or rebuilt, and the marginal cost will be larger since the variable cost will be a larger part of the total cost.

6.1.4 *Economic effects*

One of the most widespread misunderstandings of economic appraisal, and also an important cause to erroneous policy decisions is the perspective that economic evaluations are limited to the financial costs and effects, how to use health care money, and also only recognising effects that can show up in a budget.

This is not correct. Economics recognises economic consequences (in contrast to financial consequences), that is the usefulness of the resources for health care, and it aims at measuring our summarised satisfaction of the results in relation to these resources. There is a whole spectrum of consequences, which have different contents, from extreme financial to extreme economic perspectives. Consider the following.

Figure 3. Narrow and wide economic perspectives.



Cost offset: This is a perspective where the effects of the health care are not explicitly recognised. The only thing, which is explicitly at stake, is the *short-term costs*, how to economise and how to keep the budget. If the quality of care is affected no explicit measures are taken to control this quality.

Example 1: A new treatment is proposed for certain rheumatoid arthritis. The new drug is by far more expensive than the current treatment, but the results are also dramatically improved in terms of suppressed inflammation. However, the hospital management decides that the treatment is too expensive and should not be used. The improved treatment results are not discussed.

Even with possibilities to transfer costs over years, long term costs, the effects of the care are not in focus. The only concern is if investment can lead to lower future treatment costs.

Example 2: A national screening program for early detection of coeliac disease has been discussed. The cost of the screening program, for all children at the age of two, will be expensive, but a long-term reduction of secondary consequences, such as infertility, cancer and osteoporosis, can be expected. However, the program can not pay for itself even in the long run and is turned down by the government, in spite of the increased quality of life of the patients.

Social costs: The introduction of socio-economic evaluations and especially cost-benefit analysis came as a response to the need of prognoses of what was likely to happen when a large social project started, like a water irrigation project, underground railway, vaccination programs, etc. At the time the motivation was that many social projects had different costs and effects for different stakeholders in the project, but that the decision about how to perform the project should be made from its total merits, for the society in general.

Example: A new methodology is proposed at a neonatal intensive care unit. The costs per day for the new technique is marginally greater than the current technique, but the discharge can be done earlier. The hospital accepts the new technique. Unfortunately the new method leads to a slightly higher number of handicap. The cost for these handicap (special education, training and assistance) falls to a large extent on the local authority and not in the hospital budget. Socially the cost is greater with the new method compared to the current.

Welfarist perspective: Here is a clear dividing line. The earlier views only took into account how costs are distributed between individuals and over time. The welfarist perspective, however, recognise the health effects, as the *goal of the health care*, and the costs, i.e. the resources, as the means to reach this goal. Now, at first, it is meaningful to talk about a formal HTA study where costs are related to effects, the goal fulfilment of the treatment. Also, a cost-effectiveness analysis (CEA) or a cost-utility analysis (CUA) can not be done in the previous settings. The health effects, in terms of physical or mental function, are compared to the resource use in a ratio measure as in CEA and CUA.

Extra-welfarist perspective: If the goal of the health care can be defined also to care not only for your own health but also for others, which is the case in a tax based system or in a health insurance system with equal premiums, the health care follows the extra-welfare perspective. This has not a very different perspective from the welfarist view but there are some differences in the way quality-of-life issues are measured. Some researchers claim that the widespread QALY measure does not measure true preferences and therefore cannot be used in extra-welfare decisions.

For daily purposes and uncomplicated HTA studies it has little importance, but in cases where ethical issues are important it might be necessary to explicitly make a choice what scientific perspective to take. It will also have some influence on the choice of HTA technique, since those are not ethically neutral. See section 6.3.4 on evaluation techniques for comments on this issue.

6.1.5 Discounting

Costs and effects of health technology frequently do not appear at the same point in time. Typically the costs are in the beginning and the effects come many years later.

Example: Consider again the previously mentioned screening program for early detection of coeliac disease, thereby avoiding secondary consequences of undiagnosed and untreated gluten intolerance. The screening, all children at two years of age, is rather expensive, about 14 million PLN. The main benefits are less frequent cases of infertility, lymphomas, osteoporosis etc. These effects will appear 30 or more years ahead, causing intensive treatment of 29 million PLN not counting the loss of lives. It thus seems cost-effective to introduce the screening program. However, costs and effects need to be discounted to present value. This will reduce the value of the future positive effects considerably, to just over 2 million PLN. The screening turns out not to be possible to pay for itself.

The reason why costs and effects have to be valued differently, depending on the time they appear is that all individuals are more or less impatient. In economic terminology it has been described as if we all have an intrinsic "impatience rate". People would like to receive a benefit, money or otherwise, earlier than later, given a choice. People would also like to postpone sacrifices/costs if possible, *also in a hypothetical world with no inflation and no bank interests.* It is perceived as better to save a life today than to save one tomorrow in health care (also not the same life). This impatience, the individual *time preference*, seems to be different among people. Some people are very patient, they have a low time preference - some people are very impatient, they have a high time preference rate.

In the economising of taxpayers' or policyholders' money and other resources, as well as the legitimate claim of a speedy health care without unnecessary waiting lists this fact is recognised in economic appraisal by using a *social discount rate*. This simulates the effects of patient waiting times on costs and health effects.

The method of aggregation of individual discount rates to a social average has been discussed over time. There has been a certain agreement that the discount rate should be somewhere around five percent. Later some countries have made a policy statement that it should follow the long-term national bank interest (which can be interpreted as an extra-welfarist perspective). This idea, changing priorities of the health care according to the change in the national bank interest can be discussed, with the argument against that people have a long-term interest in an efficient health care system and to a great extent independent of current economic political fluctuations.

The discount rate is used as follows:

$$P = \frac{N}{(1+r)^t}$$

where P is the present value, N is the nominal value of the cost or effect, r is the discount rate and t is the number of years until the cost/effect appears. For costs/effects appearing over a number of years each year should be valued separately.

In the infancy of economic appraisal, debate about discounting of effects took place. Some readers claimed that there was no need for discounting health effects, mainly because they could not be sold or bought. Others maintained that the marketing of health effects was not the issue, but the representation of the intrinsic inconvenience to postpone effects was. It seems like there is now a common agreement that both costs and health effects should be discounted with the same rate. For economic studies the choice should therefore follow national guidelines if any, or international recommendations. If any deviations from such standards are made, they should be clearly stated, and in any case be supplied with a sensitivity analysis of the study results variation as a consequence of changes in the discount rate.

6.1.6 *Efficacy, Effectiveness, Cost-Effectiveness and Efficiency*

Efficacy has been discussed in Chapter Four. Companies applying to register pharmaceuticals in a country (or in the EU) are obliged to supply efficacy information. Most meta-analyses would supply compilations of efficacy information.

Long-term adverse reactions do generally not show up in efficacy studies since data gathering required to determine these would be too expensive. Therefore, special *effectiveness* studies need to be done to catch effects in routine use.

The alternative in an efficacy study may often be a placebo. In an effectiveness study there should be a relevant alternative, answering the question how much the effects are changing compared to the routine first treatment choice. *Cost-effectiveness* is then the routine effectiveness related to the cost difference of the alternatives.

What about the concept of *efficiency*? It is often used in contrast to the concept of productivity. The latter answers to *the question if something can be done faster than before or with fewer people*, which, in other words, is cost containment. But the concept of productivity does not say anything about the desirability of the treatment effects. For that the concept of efficiency has to be introduced. *Efficiency deals with the question if resources are used for the right purpose. This is known as goal fulfilment.*

Thus, productivity is a prerequisite for efficiency but not the other way around. Low productivity can never lead to good efficiency. Good efficiency always means good productivity.

6.1.7 *Need*

It is sometimes said that the health care resources should go to the people with the greatest need. This sounds like an agreeable position. However, this statement has to be qualified if it is not to be misunderstood.

The philosophical base for HTA and economic appraisal is utilitarianism. That means that economic appraisals give recommendations about how to put *resources to the best use*. "The best use" means in this case to give treatment to the patients that can utilise the treatment, but also not to give treatment to those who can not utilise it, because that would mean that some resources would be used without giving any or low effect. Consequently someone else could use them better.

This perspective is sometimes contrary to medical ethics, which states *the need of the patient* as the basic rule. The need is independent from the utility of the treatment. You can have a great medical need even if no treatment exists. Therefore, if the concept of need is used in economic studies, a statement should be added *about needs possible to address*. But the basic principle is still that the resources should go to those who could benefit the most, and not only to those who may benefit.

6.2 Effects of health care

As was stressed above, health economics stipulates that the costs for some health care activity or program should be related to the effects of this treatment and then compared to a realistic alternative. This is no different from other socio-economic evaluations like studying water irrigation or traffic safety. The need for this type of calculus emerged from the fact that there would not be a socially acceptable solution if all people acted in their own self-interest in such large-scale activities. Early cost-benefit analyses concentrated therefore on the economic aspects like difference between public and private values, pollution problems etc. There was little interest in the more subtle aspects like quality of life or utilities of broader values like a person's possibilities to participate in society or to execute his political rights. However, such broad effects are important in health care activities and there has been a continuous discussion how to handle them and to make them explicit for instance when handicap or rehabilitation activities are assessed.

Several early analysts had the opinion that CBA did not reflect the real valuations of health care programs if the utilities could not be given an even broader definition. The concept *extra-welfarism* came up in contrast to the common *welfarist* perspective, which started the socio-economic movement (Culyer, 1989). The key difference between the two perspectives came from the way the utility function was defined. In the welfarist perspective utility was defined in the traditional way. It was defined only by goods or services consumed. A massive criticism of this "consumerist" approach was raised in the public debate and in the late 1930s Bergson (1938) recognised also other *extra-welfarist* aspects, which also may be important to a person's utility, for instance being painless, free to choose, physically mobile etc. It might even contain aspects of relationships between individuals.

Later the extra-welfarist so-called "decision-maker" approach was taken (Sugden and Williams, 1978), which lies upon the HTA analyst to do the studies using the decision-maker's perspective. In this approach such immaterial values could be included under the plea that in a democratic society the decision-maker (e.g., the Minister of Health) is selected to make such non-consumerist choices and to make them deliberately.

Maybe the final outcome of health care is too subtle for the human mind to grasp but unfortunately decisions still have to made about how to use resources in health care to achieve as much health effects as possible. A number of different proxy or intermediate measures have been used.

- **More care activities** (e.g. operations) per year are a measure of higher productivity if the cost does not increase at the same percentage. Productivity does not necessarily imply increased cost-effectiveness because that requires also that the quality of care not be jeopardised.

- **Increased number of treated patients** is of the same kind as above.
- **Decreased number of treated patients** could be a measure if the treatment or program leads to fewer patients (example: preventive screening).
- **Lower treatment costs per patient** is less use of resources and not a health care effect.
- **Lower future treatment cost** is still less use of resources but in a longer time perspective.
- **Increased income of the government** as a consequence of paid taxes (from income and value added tax) because the patient can work more than before.
- **Increased income of the patient** because the patient can work more than before.
- **Increased productive ability at work** leads to better income.
- **Increased productive ability at home** does not discriminate those who prefer to work outside the labour market.
- **Increased leisure ability** includes also the time gains in normal activities for retired and other unemployed.
- **Successfully treated patients** according to established clinical standards.
- **Satisfied minister of health** who decides over investments in health care.
- **Satisfied patients** who are the main target group.
- **Increased patient length of life** values important lifesaving activities.
- **Increased patient quality of life** values patients' informed satisfaction.
- **Informed willingness to pay for the treatment** values what the patients actually think the care was worth.
- **Patient utility** measures the total value of the treatment in a full context of the patient's life situation but does not express the value in money terms.

In this list, there is a structure from the low validity measures in the top of the list to high in the low part. The scientific and financial possibility to find quantitative data to study the HTA problem will decide how far down the list is possible to go.

6.2.1 Health effects

As a community, we pay money and other resources to the health care and we require that the health care system use these resources to 1) prolong the length of life and 2) improve the quality of life - compared to a situation without the health care.

What is the value of life? The question is treacherous and misleading. In one perspective it speaks to the overall question of human values, purpose of life and human dignity. That is not what is asked for in economic appraisal. Instead, it is a question about how one can compare treatments where, in the one case for instance 100 people die per year and in the other 110. In some way or another the society has to value saving lives. If not, saving of lives will not be an issue in the decisions concerning how to use health care resources.

What is the value of prolonging life? Some attempts have been done to answer the question. In an English study (Jones-Lee et al., 1985) a number of people were asked to state *how much money they were willing to offer to reduce the risk of being killed in traffic*. The answers could be used to estimate the total social value of saving one statistical life, which was about £1m. In 2001 values it would be about 7.5m PLN.

The issue about the value of someone else's life is more difficult and also culturally dependent as of the value of healthy children, avoiding miscarriage, etc. This issue will not be discussed here. It is sufficient to recognise that the area (contingent valuation studies) is under development.

Consequently, if a society needs to assess health care or technologies in health care, it is necessary to study how many resources are at stake in a community, and what total effects in terms of length of life and quality of life is obtained from the use of these resources.

So far nothing is said about the desirability of a specific distribution of health effects among the population. For instance, is it more important to see to the health of children than to that of the elderly, or vice versa? Neither is there a notion of some public good other than individual health effects to be pursued. Should the gross national product increase as a result of health care? In that case, resources should be allocated to specific programs for the working population and care of the elderly reduced. If it is considered that care should be based on "merit", such "merit wants" have to be specified in separate evaluation procedures.

Two measures of a combination of length of life and quality of life have been proposed: the QALY and the HYE. QALY (Quality Adjusted Life Year) measures the length of life, but reduced according to the health related quality of that life. One year of full health is one QALY. One year with half health (however measured), or one half year with full health is 0.5 QALYs.

The HYE (Healthy Years Equivalent) measures the number of years of perfect health a person would find comparable to the probability of a certain illness for a specific time span.

For practical purposes there is not a great difference between the two measures. They will give roughly the same result. The HYE is theoretically more motivated than that of QALY but the QALY measure seems to be somewhat easier to use. In most cases the QALY or HYE has to be retrieved from questionnaires or interviews. Only occasionally such data is collected routinely in general registers. In this appendix we concentrate on the most common, the QALY.

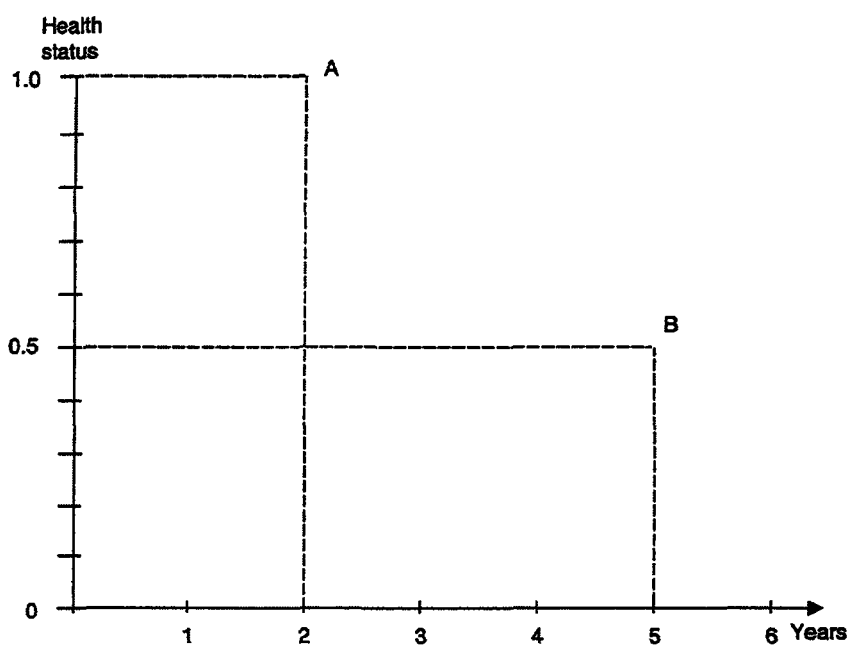
In many cases it is impossible to use the QALY or HYE type measure to reflect the primary health effect. One example is retrospective studies. The analysis in many case has great problems getting reliable answers when people are asked about health backward in time. Another case is when intermediate, clinical data is collected instead of collecting long-term health effects. For example, in introducing a new hypertension pharmaceutical, the effect in terms of mmHg reduction is measured instead of the long-term reduction in future cardiovascular disease. The evaluation then has to take a more limited CEA perspective instead of the CUA perspective, and the conclusions of the study will also be different and more limited.

6.2.2 *QALYs and TTO*

In the Time-Trade-Off (TTO) technique, QALYs are made operational. It can also be used in other frameworks but its application to QALYs is one of the most common. The idea of the QALY is simple and intuitive, though hypothetical. A person is asked to think about the following scenario: Consider one year of full health. It is reasonable to assume that if you were ill, one year of illness would be less valued then one year of full health. Thus if you hypothetically could choose, you would probably choose the full health year before the reduced health year. However, the question is not as simple when one full health year would be compared to *two* years of illness. The question then is: "Is it better to live *two* years with an *illness X* and then die, than *one* year of *full health* and then die?"

There are two dimensions, time and levels of health. Both are positive to the patient. But can an increased length of life compensate a reduced level of health? Maybe. In that case - the choice of a) the full health case or b) the increased time case depends on the perceived level of health (often compared to health related quality of life). For a very severe illness a length of one extra year would not suffice. For some minor health problem an extra year would make a large difference. In the figure two hypothetical perspectives are considered, **A**, which is full health for two years followed by instant death, and **B**, which is a reduced health valued to 50 % but five years length of life. To measure the QALYs the two dimensions are multiplied.

Figure 4. The QALY concept.



The total health content will be for **A**: 1.0 X 2 =2 QALYs, and for **B**: 0.5X5 =2.5 QALYs.

The difference between *health* and *health related quality of life* is somewhat blurred. There is an obvious difference between the two concepts.

Health is measured between a state of no value = 0, compared to being dead, and a full health value, which is valued 1 (100%). The quality-of-life scale has the same 0 value but basically no intrinsic upper limit. It thus has to be assumed that full health also leads to full health-related quality of life, an assumption that may be too presumptuous. On the other hand, the designer team of the TTO technique state that health states are conceivable being "worse than death ... and indeed states better than perfect health, if they exist" (Drummond, O'Brien et al., 1997, p. 171).

But there is also an ongoing discussion about the difference between quality of life values, utilities and preferences. This issue cannot be discussed here, but the interested reader can consult Drummond, O'Brien et al. (1997), page 146, for an introduction. It may be sufficient to know that the term *utility* should be reserved for states that incorporate uncertainty about the future.

6.3 Evaluation Techniques

This section will review the current assessment practice in the economic appraisal. It will concentrate on the dominant three techniques used in such appraisal and leave out experimental approaches.

In many cases an assessment has to take the perspective that resources are limited and that no extra money is available for new technologies (or for re-assessments of existing technologies). This will force decision-makers; either they are doctors or administrators, to set *priorities* - to choose, to decide what is going first and what has to wait. It will also lead to the problem of choosing what technologies should be used for the treatment of one diagnosis and what treatment needs to be abandoned. In the basic benefits package as well as in hospital formulary committees this leads to the common problem that if something comes in, something else has to go out.

If only the cost is considered no expensive treatment would be undertaken, not even if it could save lives. It would even be discussed if such treatments should be undertaken which would lead to further illness, for example saving people from dying in cardiovascular disease. The result would be that the prolonged life would lead to increased risks of getting cancer later in life, and probably more expensive health care. Only considering the costs of care will not lead to optimal health care.

On the other hand, one could go by the extreme Hippocratic oath and only regard the health effects to be important. The patient in front of the doctor would have total access to his/her expertise, and the patients in the waiting room would be basically disregarded. No restrictions would be put on the expensive treatment the doctor can prescribe. Only the good results would be counted. It is likewise believed that this situation would not lead to optimal care.

In economic appraisal, both the resource use and the health effects are appreciated in a dual estimate, most often a ratio between costs and effects (for instance "Cost per

successful operation", "Cost per quality adjusted life year"). Thus, it takes into account the resources used for a proposed technology and relate them to the medical results that can be achieved. Then the same procedure is done with other, already existing technologies and the priorities can be set.

The costs, the resources needed for applying a health care technology are measured in money terms. But the medical results, the health effects, are most often not directly measurable in money terms. They have to be measured in physical terms or immaterial terms like number of saved lives or years of life, number of healthy days or even quality of life. There is a certain risk if the evaluations are not done properly of comparing "apples and pears". The three techniques shown below handle the problem of comparison in three different ways, and the need for different treatments of the comparison problem is also why the three techniques are needed.

Two related concepts should be mentioned. Firstly, the Cost-of-illness or Burden-of-illness technique is not an evaluation at all. It quantifies and values the consequences to society of a disease (but also other phenomenon like smoking or traffic accidents); how much it costs in total, and in most cases draws the conclusion how much could be saved by reducing the cause of the illness. This analytic technique could be defended to draw attention to the funding of research programs or screening activities, but it does not compare realistic alternatives.

The second example is the cost-minimisation analysis (CMA). It only compares the costs for two or more ways of getting to the same result. In this appendix we will consider this case being a special case of the CEA. It has all the characteristics of a cost-effectiveness analysis below, but it regards the effects being equal. The problem with this approach is that it is not wise in some cases to state, already before the study design, that the effects are going to be equal. It might turn up that so is not the case, and then the research project has been to narrow, and a new study has to be done. It is better to start from a CEA point of view and then, maybe, conclude that the effects happened to be similar.

6.3.1 Cost-effectiveness analysis (CEA)

This technique presents results as a ratio between costs in money terms and effects in physical terms. The effects could be of any type as long as they are quantitative.

Example: In a study of a new treatment for reducing blood pressure the costs for treatment A can be 820 PLN per patient and another treatment B 780 PLN per patient. In treatment A the average reduction of diastolic blood pressure is 6 mm Hg, and in treatment B the reduction is 12 mm. Using a CEA the A treatment shows a ratio of 820/6 =136.7 PLN per mm. Treatment B shows a ratio of 65 PLN per mm. Treatment B is then more cost-effective than A.

This technique is the most common. The popularity stems from the fact that many clinical outcomes can be presented and related to the cost for achieving these effects, and compare the result with other treatment methods. The attractiveness is also reinforced by the increasing number of CEA studies published, which makes new studies even more comparable, at least on the surface.

However, there are some shortcomings with this evaluation technique. In the example above, the CEA technique regards every millimetre equally important, regardless if it is

from 120 to 119 or if it is from 85 to 84, even though each millimetre contributes differing proportions. If it is believed that it is more important to reduce high blood pressures then moderate, another evaluation technique might be a better choice where the values are better expressed. Actually CEA does not value the effects at all. It is just a counting of the physical numbers.

Furthermore, there are some mathematical properties of ratios that will cause problems. If a treatment leads to no change in the result for one patient but still draws considerable treatment cost the result will be a non-defined ratio. It may help if the ratio is turned upside down. Then the result will be defined and equal to zero, but still there will be problems of evaluating this individual's result in comparison with other patients. In such cases the solution is to sum all costs and all effects respectively for all patients in the group and then take the global ratio between costs and effects. This reduces however the possibilities of statistical reliability testing.

Another nominator/denominator problem, which is common in the CEA studies, is the choice of what should be included as an effect and what is not. For instance, a rehabilitation program shows better result than traditional care in terms of health effects, but also in terms of reduced convalescence time (CT). Depending on how the convalescence time is defined - either as a reduced treatment resource (R) (it is a part of the treatment), or as a gained effect (E), we will get different CEA results. Say that CT is five in some units, R is 50 in the same units and E is 10. Then clearly if CT is seen as a result:

$$CEratio_1 = \frac{E+CT}{R} = \frac{10+5}{50} = 0.3$$

but if SL is seen as a reduced resource/cost:

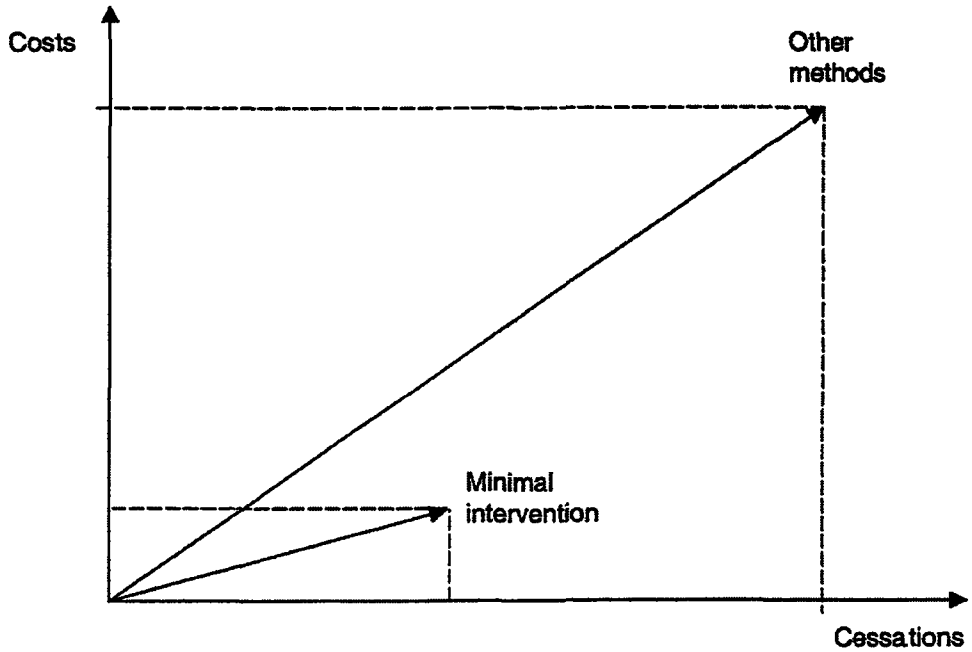
$$CEratio_2 = \frac{E}{R-CT} = \frac{10}{50-5} \approx 0.22$$

It is therefore important to be clear about the definitions of the costs and effects. In this methodological appendix the view is taken that costs are items that can be used to produce something valuable and that have an alternative value.

Less convalescence time could certainly be used to produce either salaried work or homework or even leisure time. *Effects on the other hand are seen as primary health effects. They can not be used for production.* (We have to disregard the sublime case that good health can be used to produce a "good life", a concept used by Becker (1965) and Lancaster (1966)).

A fourth problem could be elucidated by another example. A meta-study was done of different treatment methods for smoking cessation (SBU, 1998). The result showed that a "minimal intervention" strategy was the most cost-effective. At a doctor's visit the patient is recommended to stop smoking in a simple identification and counselling session, less than 5 minutes. With this method nine percent of the patients stopped smoking compared to other methods, leading to 12 - 17 percent cessation. However, although a less favourable result using this simple technology is so cheap that the minimal intervention therapy shows to be cost-effective. The reason can be shown in the following figure.

Figure 5. The cost-effectiveness of smoking cessation.



The example shows that the minimal intervention strategy is inferior to other methods in the effects, but the cost is also almost negligible. As a result the ratio between costs and effects is more favourable (the arrow has a lower inclination). This does not necessarily mean that all other methods for smoking cessation should be banned. An analysis of subgroups, for instance pregnant women, may show other results.

The CEA technique leaves a simplified picture of the problem of choice but that may often is sufficient. It often leaves a large discretion to administrative and political bodies to set limits and express values. For example it is a cumbersome task to investigate from a scientific point of view how much money a saved life is worth. Likewise, it is no scientific ways the limit can be set were to stop the acceptance of cost-effective treatment. In several such cases it have to be a political decision. From a certain point of view this can be regarded as leaving important issues to less scientific judgements, but on the other hand it may also reduce the claims to the formal evaluations of pretending to be more scientific then they are apt to.

6.3.2 *Cost-utility analysis (CUA)*

In many cases there is a need for a more elaborated analysis. In CEA political decision-makers have to state where to draw the line, i.e. what to treat or not to treat. Also, CEA has the disadvantage of not being very comprehensive and global. It is hard to find common measures to compare for instance treatment of decubitus with heart transplants. Instead we would like to quantify the effects, not in just any quantities, but in a value that can be used globally for several illnesses, and would hold its value between hospitals and also over time. In the beginning of the HTA movement the Cost-benefit analysis (see below) was the economists' answer to such needs, but their enthusiasm was not shared by the medical profession, who regarded CBA as a dubious and even unethical technique.

Instead, some efforts were made to find alternatives that did not use money values and willingness to pay for subtle treatment effects, saved lives and quality of life. One of the

most successful attempts was created by Torrance (1972), developing a technique (TTO, Time-Trade-Off technique) to ask people about their values of time instead of money (see section 6.2.2). At about the same time the synthetic measure of a QALY was invented by H. Klarman (1968). *QALY stands for Quality Adjusted Life Year and attempts to compare the hypothetical time a person is spending in an illness or a treatment to the value of the corresponding (also hypothetical) time having full health.* Thus one QALY is equivalent to one year of full health. One half QALY is 6 months in full health or one year of reduced quality of life to 50 %. In this way, it is thought, the concept of quality of life can be given a more quantitative meaning.

If these or other similar values are used for evaluation purposes the term Cost-utility analysis is used. The idea is that the QALY values represent the individual's utility or quality of life, and that priorities, i.e. the policy choice among treatments should be made according to such (well-informed) individual preferences. The CUA technique has gained great importance and is now the preferred technique in some countries for example to introduce new pharmaceuticals into subsidised insurance schemes.

CUA combines the simplicity of a single measure for health care effects with the global applicability of many health conditions. However, it is not simple to use and to interpret and it also has to be used with some discretion not to overestimate its importance as a single effect measure.

CUA is especially useful when the effects of a treatment are multiple, for instance several clinical outputs and/or effects in social life and participation, and also if they include negative effects from adverse events and reactions. CUA has the ability to measure all effects in one single measure, generally the QALY, and also to weight the different effects against each other in a way that reflects the relative importance of them.

Example: A study was made of zanamivir, an antiviral drug for treatment of influenza. The aim was to investigate if a high-risk population would benefit from the drug and to what cost. Both the treatment group and the current therapy control group included also other traditional symptom-alleviating treatments and optional hospital treatment. A population model was constructed and the results show that the cost for the treatment with zanamivir is on average $A148 and for the control group $A112. Using a quality of well-being scale a utility measure of 0.5579 was assigned to "a day with influenza symptoms", assuming a utility of 1 with no influenza symptom. The number of influenza days in the treatment group and control group were 5.55 and 8.08 respectively, which results in an incremental cost per QALY gained by treatment of $A11,715. However, a sensitivity analysis shows that the results are sensitive to assumptions about the true number of influenza cases, the hospitalisation rate, the onset of symptom relief and the utility "normal" value assumption about perfect health.

A number of instruments exist for estimation of quality of life. Figure X shows the different categories (Brodin & Persson, 1995).

Figure 6. Categories of quality-of-life instruments.

Generic instruments	Health Profiles		Nottingham Health Profile (NHP)
			Sickness Impact Profile (SIP)
			Short Form 36 (SF-36)
	Utilities	Classification in predetermined utility measures (decomposed approach)	Rosser-Kind index
			IHQL
			15D
			Euro-Qol (EQ-5D)
			QWB
			MMHCS
		Patient's own estimate (holistic approach)	Rating Scale
			Standard gamble
			Time trade-off (TTO)
			Willingness-to-pay (WTP)
Disease specific instruments			A great number

Primarily Quality-of-Life instruments are classified by the application. Global or generic instruments are intended to use in several situations where quality of life needs to be assessed. Disease specific instruments, on the other hand, are designed to pick up the essential qualities in a limited set of diagnoses for instance the St. George's Respiratory Questionnaire (SGRQ), which concentrates on lung-related illnesses like asthma and COL.

Secondly the generic instruments are grouped into profiles and utilities. The profiles give a number of aspects important to your quality of life, for instance mobility, sleeping problems, energy etc. But often the profile can not weight the different dimensions into one single value, and this property is needed to do the CUA-ratio. This is on the other hand the big advantage of the utility instruments. A problem with these instruments is allegedly that they often are insensitive to small changes and should not be taken all too seriously.

Thirdly, the utility instruments are grouped into decomposed and holistic approach. The former uses a reference population in a calibration procedure to assess general scores to the different questions. This is done using any of the instruments in the holistic group. The instrument is then used with these predetermined scores as a proxy to assess the quality of life of the HTA population sample. This two-step procedure is rather inexpensive and does not try the patients' patience. The holistic approach, on the other hand, uses the instruments like the TTO above to get a true and real value of the quality of life of the patient directly. This is a tedious procedure and the logic and rationale is often difficult to communicate to the patient or others involved in the questionnaire.

Very often a CUA is accompanied by a CEA. Although more expensive the advantage is that the probably less specific quality-of-life instrument is combined with a more clinically oriented instrument to pick up effects closer to the concept of production effectiveness. In this way both the productivity aspects of a smooth operating health care activity and the consumer appreciation of a desired effect can be mapped better than only using a simple CUA.

6.3.3 Cost-benefit analysis (CBA)

The cost-benefit analysis measures both costs and health effects in monetary terms. Thus, in CBA all conceivable costs and consequences could be encompassed in the same units. It gives CBA the strength of not only presenting ratio measures of different treatment alternatives, but also actually telling how much one alternative is preferred

before others. In other words, CBA like CUA can encompass a larger span of activities than the CEA technique. CBA helps to set even broader priorities than just the choice between two methods of treatment for one diagnosis. In an ideal world the CBA technique would give the necessary global information to make decisions also how to allocate resources to health care in general. The political judgement would be reduced to executing the will of the people expressed in WTP terms and subject to the budget available.

CBA often concentrates on financial effects, like for instance increased income due to increased working capacity, in turn due to treatment of a disease. But also more subtle effects can be included, like the values of avoiding lethal traffic accidents - values far beyond the treatment of the accident itself (see section 6.2.1). How much should be included in a cost-benefit analysis depends on the perspective of the study, which the client is, and who is the financer.

Often it is a straightforward process to identify and quantify the costs and effects, but sometimes it takes quite some efforts, especially when the consequences have no market prices. Examples of such non-market values are loss of lives, years of life or loss of quality of lives. It can also be the costs of recruiting and replacing burnt-out staff in a hospital, cost that you will not see in the annual accounts but in decreased performance.

In many cases such values can be revealed by asking people in a structured way what they would be willing to pay for avoiding such a (statistical and hypothetical) event. The common term for such interviews or questionnaires is Willingness-to-pay (WTP) studies. A rarely used variant of this is a Willingness-to-accept (WTA) study - to ask what a person would accept in compensation to get deprived of a certain health care service.

But the theory does not correspond to practice. Apart from often having statistically poor reliability (as also the CUA) the validity can be questioned. The questionnaire generally asks for the sum of money the respondent hypothetically would accept to pay for a certain treatment, but frequently neglects other restrictions to produce the health effect, like the patient's time, physical and mental suffering of the treatment etc. To only ask for the money price would give an answer but would in some cases still not result in a realisation of the treatment if the patient could choose, because lack of time or feeling uneasy about the pain.

Some promising research is devoted to CBA. Interesting concepts are for instance the contingency valuation techniques (for a short review, see Drummond et al. section 7.2.4).

6.3.4 Conclusion

The three techniques for compiling HTA results are all different in their applications. Four aspects will be mentioned:

Effect focus: CEA provide a rather narrow focus. Only closely related effect can be thoroughly assessed. CUA works best in situations of pronounced wellbeing or utility aspects. CBA has still a tendency to focus on the budgetary consequences and financially measured effect.

Scientific power: All three techniques are inflicted with uncertainty. However, CEA can be assessed better than CUA and CBA because of better validity. Also the reliability is defined in a better way since the "normal" statistical techniques can be used without problems. When it comes to CUA and CBA it is more uncertain how the reliability should be checked.

Degree of instrumentalism: In theory the decision-making process can be made entirely mechanic in the CBA case. The decision needs no democratic process or supervision from political levels to make the decisions about what health care activities or even how much of each, should be performed. CUA still needs the political level to state the value of a QALY, but the technique is a full value technique. CEA on the other hand needs political decisions to be decisive about the volumes of different health care activities that should be done.

The assessment techniques are not ethically neutral: Depending on which technique is chosen, the four ethical principles of autonomy, beneficence, non-maleficence and justice will be met differently. Probably CUA is the one technique that satisfies autonomy the most due to the way utility is derived directly from the patient. CEA may be more sensitive to the beneficence and non-maleficence, and finally CBA may address the justice clause differently (not necessarily better) than the others, especially if justice is defined in a utilitarian sense.

Thus, the use of one particular technique is guided by the way data can be achieved and what assumptions have to be made about the cost and effect causal relationships. Furthermore, the choice will then have implication not only for the interpretation of the results and the policy following these results, but also which ethical aspects will be affected, harmed or promoted. There lies a large responsibility on the users of HTA (policy-makers) how to use these instruments and how to instruct producers of HTA what analysis to perform.

6.4 How to collect data

6.4.1 Outcome data

The collection of outcome data has been discussed in chapter 4 and in other parts of this appendix. Relevant and appropriate methodology is the first requirement for doing a trustworthy assessment of health care procedures, and relevant and appropriate data is the second. Data about the effects suffer from both validity and reliability problems. In general the preferred method of getting effect data is to use a randomised controlled trial (RCT). This design attempts to measure the outcomes in an *efficacy* type situation, where all possible factors affecting the outcome needs to be controlled and the effect of the change in technology can be certified.

In many cases it is not possible to do formal RCTs. In that case several second-best choices can be made:

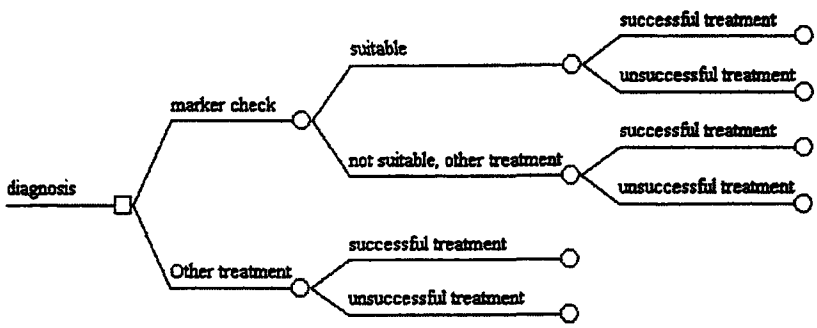
1 Randomised controlled trial;
2 Controlled trial with pseudo-randomisation;
3 Controlled trial without randomisation;
4 Cohort prospective study with parallel control;
5 Cohort prospective study with historical control;
6 Cohort retrospective study with parallel control;

7 Epidemiological retrospective case/controlled study;
8 "Before/After" study.

For an HTA study, however, also other needs should be met. The effectiveness as it is defined in this appendix focus on the routine use of the technology. In some cases, therefore, separate HTA studies have to be done to check effectiveness.

In many cases the RCTs are used to draw conclusions about the probable future outcomes of disease or disease treatment. Using probabilities and statistical techniques likely prognoses can be done of future outcomes, although no such outcome has been measured. Every prognosis uses a statistical model and these techniques have consequently been grouped under the common term "modelling". On example is the decision tree. It details different possible outcomes of a treatment, attaches probabilities, costs and effects to each branch and then calculates likely end outcomes. In Figure 7 such a model is done for the possible outcomes of a treatment with or without a certain DNA marker, which identifies the patient's alleged fitness for the treatment.

Figure 7. Decision tree for use of a DNA marker.



6.4.2 *Cost data*

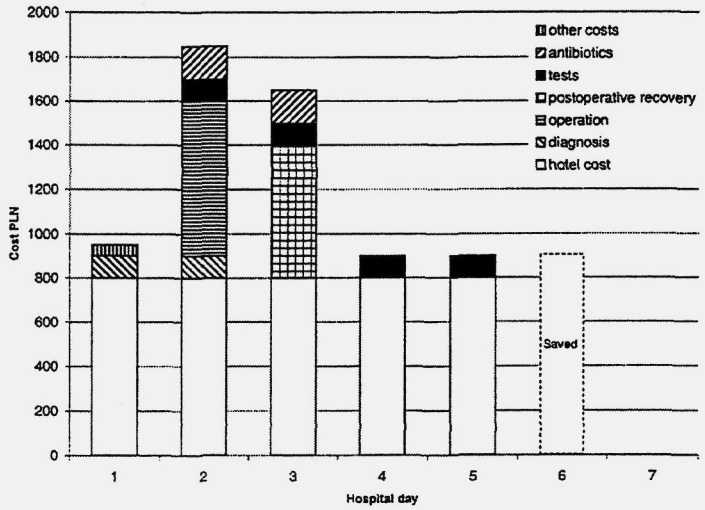
Experience shows that cost data is not a great problem from a methodological point of view, but that knowledge about the right way of collecting and use cost data could be improved. In section 5.6 and 6.1.1 the need for an opportunity cost framework was stressed, and also for marginal costs for estimations of the true resource consumption. What has not been discussed so far is how to get close to the marginal cost if important checks are not fulfilled.

6.4.3 *Fixed cost, variable cost, marginal cost (again).*

The following actual example is slightly arranged but shows a general pattern of hospital costs.

A certain surgical operation method has been changed, so that it can be shortened from six to five hospital days. The decrease has been possible through a more precise check-up in the postoperative days. See Figure 8.

Figure 8. Hospital stay for operation.



The costs are composed of the following parts:

Hotel cost: This is the cost that can be attributed to one patient being in the hospital without treatment. It includes the ward administration, e.g. registration of the patient, and also meals, laundry etc. It does not include general hospital administration, which can be judged not to be affected by the stay of one extra patient.

Diagnosis: This is the cost for primary diagnostic procedures like the introductory blood tests, the specific tests for the operation, x-rays etc.

Operation: This is the marginal cost for the operation theater, the staff involved, the wear and tear of instruments and equipment and supplies. The capital cost of the operation theatre is not included unless it reflects the actual opportunity cost of the space. This may be the case if the operating theatre is fully booked, but it is probably not equal to the rent of the space.

Postoperative recovery: This is the cost for recovery in intensive care or special units not at the normal ward. The same principles hold as for the operating theatre.
Tests: In this case secondary diagnostic procedures for monitoring the health status of the patient during the stay and also the postoperative tests.

Antibiotics: In this case antibiotics was a major cost and it was reported separately. It includes the cost of the pharmaceutical, the cost of intravenous equipment and staff for operating this equipment. However, the cost for the test of the right dosage is included in the test costs.

Other costs: Some consultancies from specialists from other departments are included here, as well as a preoperative intravenous catheter.

The *total cost* for six hospital days is 7,150 PLN. The *average* day care *cost* is 1,192 PLN before the change. The cost reduction from six days to five is 900 PLN. The new average day care cost is 1,250.

The operating surgeon fails to recognise the *marginal cost*. He claims that by saving one day the clinic can save the average cost of 1,192 PLN (a considerable annual save, which would be used to buy a new diagnostic device). However, the marginal cost of

the last day is only 900 PLN, and the actual savings thus 24 % less. Another thing that also surprised the surgeons when they saw the annual economic report after the change was that the average day cost had increased with almost 5 % despite the change. They failed to recognise that the first days now had a greater impact on the total cost (and consequently on the average cost) than before. The higher productivity in treating this group of patients led to a higher cost per day, a false representation of the true cost per case.

Definitions:
Total cost is the cost for the entire treatment period. Total cost per day is consequently all costs for one day (and not the average total cost).

Average cost is the total cost divided by the number of days, patients or other units of interest.

The marginal cost is the change in total cost by increasing (or decreasing) the treatment with one day, one patient or other unit of interest.

6.4.4 *Direct and indirect costs, taxes.*

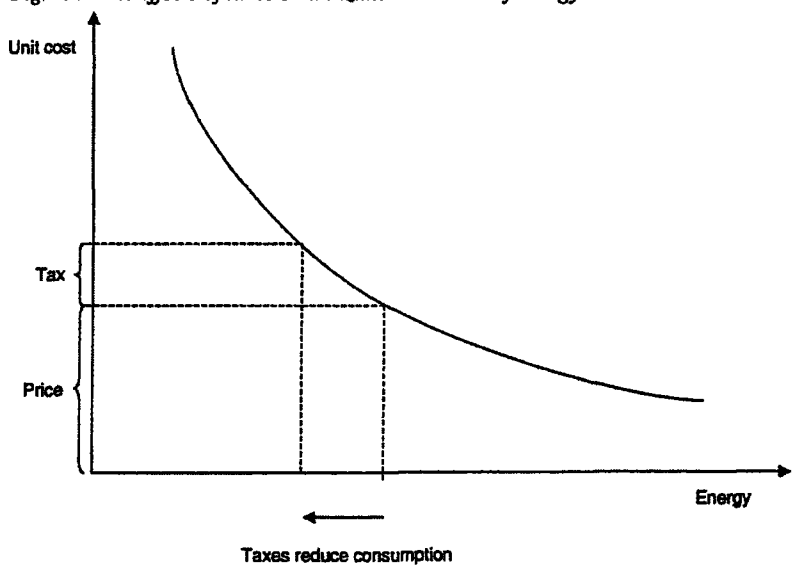
Ever since the end of the sixties the HTA community has used to categorise costs as being direct and indirect. This distinction is nothing but a convention and stems from an accountancy habit, *describing direct costs as those directly involved in the production process*. The cost of an aspirin pill would be a direct cost for treating headache. *Indirect costs are not involved in the treatment but may be a secondary consequence.* The most common example is a patient who takes time off from work for health care treatment. The lost income is then an indirect cost of the treatment.

The convention of dividing costs into direct and indirect has caused discussion and confusion. Some economists have gone into details about the actual value of lost production, supplying techniques for valuing complex corporate organisations were individual productivity can be hard to assess (e.g. in teamwork). Also, the salary as a measure of productivity has been criticised. In most cases, however, it has become the standard to value lost productivity to the patient's salary, sometimes gross of income tax, sometimes net of tax, depending on whose values are to count, the society's, the company's or the patient's.

The simplistic division of direct and indirect costs sometimes has led to analytical errors and even flaws in decision making. Despite a resistance to change habits, a movement is going on to get out of this logical straitjacket (see Drummond et al. (p. 23-24)).

The tax value problem has been a problem in many HTA studies. For instance income taxes leads to a problem of *transfer costs*. The employer pays a salary tax in Poland and many goods and services are also imposed with considerable VAT taxes. In the health care sector many goods and services get reduced VAT taxes, from the stipulated 22% to only 7%. Some goods and services can even get a tax of 3 or even be totally relieved from VAT. But if the original 22 % holds the patient or the insurance company will experience a considerable expense rise due to taxes. It is generally assumed that taxes will be redistributed to the taxpayers, maybe even to the patients, and if so the problem is reduced. However, the original price effect of the tax is still there, see Figure 9.

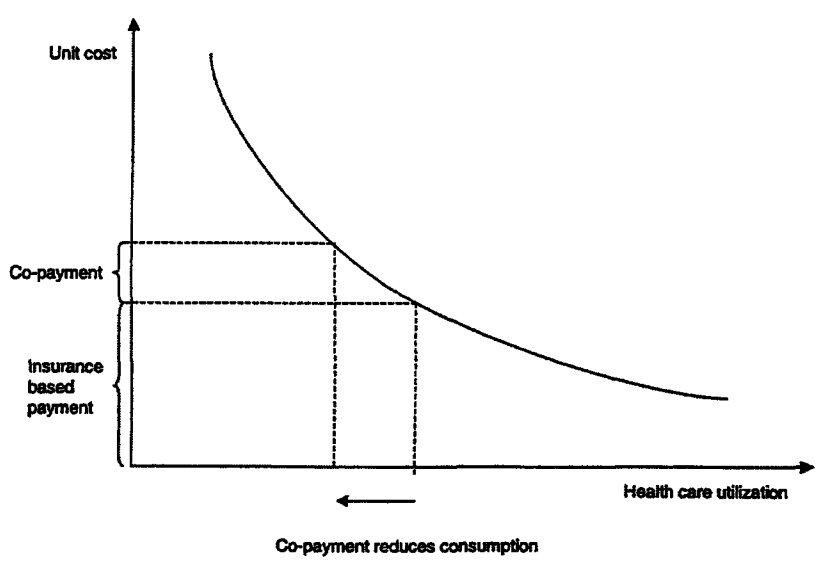
Figure 9. The effects of taxes on utilization- the case of energy



The case above shows a general economic "law" that says: *If the price goes up the consumption will go down.* That is, when the unit cost rises, regardless if it is a rise in production cost or a governmental tax, the utilisation of this good (or service) will decrease. The same law applies not only for energy, other goods, but also services. A rise in the salary of employed people will have the same effect, but also a rise in the employer's labour tax and pension premiums. All these extra costs will decrease consumption.

What is even more important from the HTA point of view is that also co-payments in the BBP will have the same effect. *This is a reduction of the part that would have been accepted by HTA based insurance but would not be realised.* It will thus have effect in the sense that it will not supply efficient health care to the appropriate number of people. Once again:

Figure 10. The effects of co-payment on health care consumption



Furthermore, co-payment will be especially critical to people with low income.

Another effect of a tax is also that the patient and the provider will meet different prices. The HTA assessment then comes into the problem of choosing which price to use; the provider's or the patient's. According to the economic theory there is an argument for choosing a point in between but in most cases the analyst tend to neglect the problem and use gross-of-tax values.

According to the economic theory there is an argument for choosing a point in between. This point should be proportional to the demand and supply elasticity of the good in question, according to the formula

$$a(GOT)+(1-a)(NOT)$$

where *GOT* is the gross-of-tax price and *NOT* is the net-of-tax price, and

$$\frac{a}{1-a}=\frac{-e_d}{e_s}$$

where e_d ed and e_s are demand and supply elasticities respectively.

For special purposes, when tax issues can be regarded as a main analytical problem the technique above could be useful but for small health care programs tax effects probably play a less important role for incentives and the success of a new health care technology.

6.5 Conclusions

It seems obvious that use of resources in health care is the key policy question at this time, both in Poland and in the world. Therefore, economic appraisal, as a part of HTA, has an important part to play in policy decisions.

Assuming that a technology is efficacious and (relatively) safe, questions concerning its costs and resource use are inevitable. An HTA that does not bring costs into study may be of little use to the policy-maker.

Economic methods are fairly well-defined and validated, and are becoming more so. Therefore, expertise in health economics is definitely needed in Poland to help guide decision-making throughout the system.

7 Assessing Social Implications of Health Technology

The social implications of a new or existing health technology can be the most challenging and difficult aspects of evaluation. Any decision to develop of use a health technology inevitably rests on value judgements. Social and cultural factors are completely intertwined in questions concerning the place of technology in health. At the same time, the methods for assessing social implications of health care technology are relatively undeveloped and few mechanisms exist to take action based on the results of such evaluations (Heitman, 1998).

Society does not seem to have enough confidence in social scientists and other dealing with social issues to ensure that results of this type of assessment are used in policy making.

It can be forgotten, in this rational and scientific age, that culture and society underlie all actions in health. Life is obviously an important value in all societies, yet different societies value different lives differently, as indicated that some countries have had age cut-offs in dialysis programs, assuring that some young and some old people with serious kidney disease would die. The attitude toward life is also expressed in the controversy surrounding abortion in some countries. Such social values explain much that occur in health care that may not appear to be rational.

Technology assessment itself cannot be totally objective or value-free. As an activity carried out by human beings, it too is influenced by social and cultural values.

This Chapter will give a brief introduction to what should be a much more active and dynamic field of assessment.

7.1 Some Interactions between Social Values and Health Technology

The interaction between technology and the broader environment is very dynamic. Cultural factors such as intellectual curiosity, tolerance for new ideas, and social values favouring efficiency and productivity have obviously greatly influenced the course of technological change (Cassell, 1986). Culture has affected the "ideas, facts, modes of thought and observation, instruments, images, language, values, symbols, and rituals of biomedical science and technology" (Fox, 1986). Technology has also influenced the broader environment. For example, it has led to changes in work and leisure, sexual behaviour, and nutrition.

As the rate of technological change has increased dramatically during the last decades, individuals and social systems show signs of not being able to adapt. Some have accepted any technological change uncritically, and have later seen that technology adoption was associated with unanticipated, and often indirect, implications (Illich, 1976). On the other hand, some have become anti-technology, and have ignored the many benefits that modern technology has brought to humans, including the benefits of health technology (Pirsig, 1975).

Some scientists see modern scientific knowledge as totally transforming our ways of being and of seeing ourselves. The discoverer of the structure of DNA, Francis Crick, for example, says, "we are here because we have evolved from simple chemical

compounds by a process of natural selection . . . ". He goes on to conclude that this destroys the "Christian" and "literary" "nonsense" that has guided our lives (quoted in Kaye, 1986). Genetics and brain research seem to be leading to the power to "reconstruct ourselves." What kind of human beings would then be good? The discoveries of modern biology " . . . seem to force upon man a transformation . . . of his self-understanding and his view of his place in the whole" (Kass, 1971).

Health technologies and their patterns of use reflect subtle interactions among various social values, such as the value a society places on health, on innovation, on financial security, on technological progress itself. Why do societies provide health care or health insurance, for example? One important factor is clearly the need for a sense of solidarity, in the face of decline in the importance of the family and religion. We need a sense that we are willing to take care of each other.

The definition of health and illness shape a society's conceptions of necessary or desirable technologies (Cassell, 1986). The changing view of mental illness is an example. Debates over how to care for those with AIDS is another. Cultural forces have led to a widening of the definition of health, so that the World Health Organisation now says that it is a "state of complete physical, mental and social well-being." The danger is a redefinition of human problems to make them problems for professionals, the process of "medicalisation" (Fox, 1977; Woolf et al, 1985).

Another important concept is the "right to health" or the "right to have care". In European societies, including Poland, such concepts as equity and solidarity have meant that the entire population has relatively good access to health. The United Nations Commission on Human Rights has sought in recent years to promote the idea that basic human rights include social, cultural and economic rights, including the right to health (UN Commission, 1999). Recently the Commission has analysed the issue of health in some details.

One of the most visible aspects of the interaction between values and health technology concerns resource allocation and the possibility of rationing (Cookson and Dolan, 1999; Daniels, 1993; Hall, 1997). Some have characterised the present health system as overly oriented to "rescue", meaning dealing with serious and life-threatening illness as a priority, which inevitably means less emphasis on primary care and preventive strategies (Heitman, 1998; Hadorn, 1996). A related argument is that priority must be given to the young, since the old have already lived long lives and have relatively little to contribute to future society (Callahan, 1990). Some have argued that costs and cost-effectiveness inevitable will play a larger role in allocating resources (Drummond, 1990). Others have advocated using practice guidelines as a way to allocate resources (Giacominni et al 2000). All of these alternatives are obviously related to the values and ethics of health.

7.2 Social Values and the Development of Health Technology

Social values affect which technologies will be developed, and, through the resources available for that development, the speed with which they will be developed. Social values are one factor determining the questions to be addressed. The special concept of the heart in Western society, for example, delayed research on approaches to heart disease for many years (Swazey and Fox, 1979).

A common effect of social forces on technology development occurs during the early stages of clinical investigation (Fox, 1986; Fox, 1974). Who is to be the first recipient of a possibly dangerous innovation? If potential risks are high, it may not be possible to continue the investigation. This helps explain why investigators have often been the subjects of their earliest experiments. It also helps explain why prisoners have often been subjects of experiments. With increasing concern about such issues, it is often no longer acceptable to use captive populations such as prisoners.

The question of human experimentation, brought to the fore so poignantly by the actions of the Nazis in World War II, has led to increasing protection of subjects of research. For example, the Declaration of Helsinki, a formal code of ethics for the guidance of doctors in clinical research was adopted by the World Medical Association in 1964 and extended in 1975 and 2000. Many national organisations have developed their own codes and guidelines, including the National Institutes of Health, the British Medical Research Council, and the Association of American Medical Colleges (Silverman, 1985, p. 155). For example, the Helsinki Declaration calls for a specially appointed independent committee to review the protocol in advance of any experimental procedure involving human subjects).

With adequate safeguards, the public in Western countries (at least) supports human experimentation. In the United States, studies have shown that a large majority of those with serious illnesses such as cancer and heart disease, as well as healthy members of the population, believe that patients should serve as research subjects (Silverman, 1985, p. 164). One serious issue is how to recruit patients for a human experiment. The burden should ideally be distributed over the entire community. In practice, however, human subjects of research are often the disadvantaged, such as prisoners or the economically deprived.

7.3 Social Values and the Assessment of Health Technology

Technology assessment itself is a social and cultural activity and is a part of social and cultural institutions. This has many implications. Perhaps the most important is that the tools of technology assessment - which can give the illusion of objectivity and rationality - must be applied with great caution.

Assessment inevitably selects aspects of a problem to consider. In this process, problems can be reduced to terms that mistake their underlying structure and their fundamental nature (Office of Technology Assessment, 1982). The collection of data implies selection of what data to collect; the analysis requires value judgements of which aspects to emphasise. And the limited attention to social values in assessment itself indicates cultural priorities.

The limited place of assessment of health technology in the past is a reflection of social values: the value that lay people placed on the role of the physician and the conviction of physicians that their success was in large part a personally-based art. As the technology of much of health practice has increased, evaluation itself naturally has a larger place. Publicised failures in health technology have promoted assessment. Successes have helped ease pressures for assessment.

The ethics of the processes of evaluation have developed rapidly during the past decades. The ethics of controlled clinical trials have been hotly debated (Barber, 1976;

Silverman, 1985; Heitman, 1998). One frequent statement by some clinicians is that it is not ethical to test an "established" technology, even if there is no clear evidence of its benefit. Those who believe in assessment answer that clinicians cannot just do as they feel best; their practice must be based on sound scientific evidence (Lambert, 1978). Is it ethical to subject a patient to a treatment whose effect is unknown and which may do harm?

Bradford Hill, one of the developers of the method of the randomised controlled trial, wrote (Quoted in Burckhardt and Kienle, 1978):

> It must be possible ethically to give every patient admitted to a trial any of the treatments involved. The doctor accepts, in other words, that he really has no knowledge at all that one treatment will be better or worse, safer or more dangerous, than another ... If the doctor does not believe that, if he thinks even in the absence of any evidence that for the patient's benefit he ought to give one treatment rather than another, then the patient should not be admitted to the trial...

The problem for clinical trials, as in all human research, is balancing the welfare of the individuals in the trial against the potential benefit to future patients. The Helsinki accords raise a particular problem for clinical trials. The accords state that each subject should be adequately informed about the aims, methods, hazards, and benefits of the study and should grant freely-given consent in writing. However, in a randomised trial, for example one involving a placebo, the benefits and hazards are not predictable for the individual patient. Still, the very justification for a randomised trial is insufficient information to permit a rational informed choice (Silverman, 1985, p. 166). Other problems arise when the community or the parents must give approval. Current opinion concerning children is that parents cannot consent, but can give permission; still, older children should make their own decisions in this case (Levine, 1981).

7.4 Social Values and the Use of Health Technology

Use of a technology depends on the complex interaction of many forces. A sick person visits a physician seeking care. The physician, based on the complaints and an examination, uses knowledge from experience and medical science, to prescribe a diagnostic and therapeutic plan. The care is provided in an institution, and will usually be paid for by a public program that includes certain constraints. Educational programs attempt to change physician behaviour to certain interventions. Patient demands are stimulated by press coverage of certain treatments, and physician responses are conditioned by such public demand. But the entire decision-making process occurs in a social and cultural context that may be very difficult to define. This is indicated by variations in use of technology within a country and from country to country (Paul-Shaheen et al, 1987; Wennberg, 1984).

The major test for the medical profession is whether the ends of the profession are so humanitarian that experts may be given autonomy (Freidson, 1972). Ideally, actions are based on reliable evidence and patients and physicians define "good" in the same way. Physicians often choose for patients, however, raising the possible problem that the patient would define good in a different way. McNeil (1985), for example, has shown that some patients prefer radiotherapy for larynx cancer even though the outcome in terms of life expectancy is worse than with surgery because radiotherapy preserves the

ability to speak normally. In the conflict between value judgements of a health provider and a patient, surely the patient's values should ordinarily determine the treatment.

A particular example concerning technology is focusing on the physical symptoms and biological problems a patient has and ignoring the human being. Disease occurs in a human being, and information and knowledge bearing on it also can be psychological and social (de Vries, 1981).

One obvious reflection of social values is laws and regulations dealing with technology. In Sweden, interventions that can be made available to large sub-groups of the population are favoured over those of benefit to few people, for example. Sweden was rather late in accepting brain death as the standard for declaring that a person was dead. For these reasons, Sweden offered heart transplant rather late, compared to other developed countries.

The values of the individual may conflict with the values of society in the case of any technology. For example, vaccines produce "herd" immunity that protects the entire population from epidemics. From the standpoint of the individual in a society with adequate immunisation rates, however, the most rational course is to refuse the vaccine. For this reason, most societies require immunisation of children, for example, before entering school.

Social values seem to have their greatest effect when the technology deals with processes with deep human meaning. Thus, technologies dealing with birth, death, and reproduction have been most controversial in recent years. Technologies in these areas interfere, or seem to interfere, with natural processes, and may be very disturbing to moral feelings.

There has been considerable progress in the field of bio-ethics during the past few years (Engelhart, 1996; Gillon, 1994). One useful formulation has been put forward by Beauchamp, in what he called "the four principles approach" (1994):

1 Beneficence (the obligation to provide benefits and balance benefits against risks);
2 Non-maleficence (the obligation to avoid the causation of harm);
3 Respect for autonomy (the obligation to respect the decision-making capacities of autonomous persons;
4 Justice (obligations of fairness in the distribution of benefits and risks).

Beauchamp states that these principles are considered prima facie: they are always binding unless they conflict with obligations expressed in another moral principle, in which case of balancing of demands of the two principles is necessary. The principles often do not give clear guidance. Rules must also be formulated to operationalise the rules. Filling in of details to overcome apparent moral conflicts.

7.5 Nature of Social Implications

Social effects of a health technology can take place at any time in its life cycle: when a new technology is introduced, when a technology is in general use is (especially if it is applied to a different purpose, or when its use for the same purpose is significantly increased, or late in its life cycle).

A social impact or implication can be characterised best by specifying certain aspects of its effects. The most obvious is the nature of the effect itself. For example, a change in societal sex ratios, an increased amount of adolescent sexual activity, or an unresolved legal and moral confusion over the definition of life or death are examples of effects that could have (and have) resulted from widespread use of certain health care technologies. Beyond the effect itself, one can ask, what individuals or groups are affected? Which are benefited? Which harmed? Another issue is the conditions giving rise to the effect. Such conditions could be technical, or they could be social and cultural. For example, brain death has been defined for some years through use of the electroencephalogram (EEG), but some countries, including Sweden, have come to accept this definition rather later than others.

Social implications often are indirect effects, they may occur in the long run, and they are generally unintended. They may also occur as the result of other effects, thus are especially indirect. These have been called "higher order" effects. For example, the first order effect of a contraceptive pill might be considered to be the ability to prevent conception reliably. A second order effect might be loosening the tie between women and the childbearing function. A higher order effect might be increased participation of women in the labour market. Another higher order effect might be loosening of societal attitudes toward sexual behaviour, the "sexual revolution."

A particular type of social effect in today's world of limited resources concerns access to and distribution of health services. In many countries, new technologies diffuse first into private institutions, meaning that the poor are left with less access. To a significant extent, any technological advance fed into a delivery system that discriminates against some people is likely to heighten such discriminatory practices. Thus, technology can worsen problems of equity and help to break down solidarity patterns.

7.6 Techniques for Assessing Social Implications

The techniques for assessing social implications are varied, but generally fall within the area of social sciences such as sociology. Given the nature and complexity of these effects, such assessments will almost always be multidisciplinary efforts. The social scientist must clearly understand the technical nature of a technology to assure that an analysis is sound. This means that engineers and physicists, and as well as physicians and nurses, have part to play in such assessments.

The assessment of social implications has lagged in all countries, and there is no validated method for assessing values or ethics in health care (Heitman, 1998). The Office of Technology Assessment, working with several social scientists, suggested a number of questions that could be asked regarding a specific technology to reveal its effects (Office of Technology Assessment, 1976). The questions were intended to be a guide to thinking about such effects. Social assessments involve more than applying a certain method or tool; they also involve creative thinking and "brain-storming."
The list presented here is illustrative and not exhaustive.

- What are the medical aims, technical characteristics, and developmental state of the technology in question?
- What medical problems is the technology designed to solve, and how severe are these problems? What contribution will be technology make? Diagnose an early form of disease? Treat a life-threatening condition? Correct a functional problem?

- Is the technology a major or minor innovation? Will it radically alter medical practice or will it modify and improve established procedures?
- How soon can development and adoption of the new technology be expected? Can development and adoption be speeded or slowed by policy mechanisms?
- How effective is the procedure? Has its medical efficacy been assessed? How will efficacy be assessed? Are there reasons to think (animal evidence, e.g.) that the technology will be efficacious? Are controlled clinical trials possible? Underway? If controlled trials are not possible, what methods will be used to assure effectiveness? What are the potential or proved dangers of the technology to individuals using it?
- What are the implications of the technology for the patient's life?
- What will be the quality of life of the patient? Normally active? Moderately restricted? Physically crippled?
- What psychological effects can be anticipated? Guilt? (because of burdens on the family) Anxiety? Feelings of dehumanisation? Dependency?
- What are the implications for the patient's family?
- What will be the costs to the family, both financial and non-financial? How will the technology affect family structure? Will there be any physical dangers to the immediate family? Will the device or procedure be psychologically acceptable to the family? Will active co-operative or assistance of family members be necessary on a continuing basis?
- What are the implications for society in general?
- Will the technology affect demographic characteristics of the society? For example, can changes in sex ratios or age distribution in the population be anticipated? Will the technology affect reproductive capability of patients and thus change the genetic pool and the prevalence of genetic disease?
- Will use of the technology by an individual create threats to the environment?
- Will introduction of the technology challenge important beliefs and values of the society about birth, gender, bodily integrity, personal identify, marriage and procreation, respect for life, right to live, right to die, responsibility for each other? Will introduction of the technology result in changes in any of these values?
- Will the technology alter any basic institutions of society (e.g. schools, recreational facilities, prisons)?
- What are the implications for legal and political systems?
- Will problems of justice, access, or fairness arise? Will they lead to legal action?
- Will the manufacturer be liable for damages resulting from failure of the technology (in the case of devices or drugs)? Will liability extend only to damage to the individual or will it cover environmental effects as well?
- Will use of the technology require changes in legal definitions of such concepts as death or suicide?
- What are the implications for the economic system?
- What is the projected or present overall monetary cost of adopting the technology?
- What are the implications for programs of disability or life insurance? Pension funds?
- Who will pay? Will government support be required for development and/or use of the technology?
- How will the technology affect the national economy? Will development and use produce jobs? How will this affect overall productivity? Will productivity in health be affected by the technology? Will the tax structure and rates be affected? Will imports or exports be affected?

If all of these questions were answered for a new technology, the result might be called a "comprehensive technology assessment," since technology assessment refers to any policy-oriented study of the implications of technology (Office of Technology Assessment, 1976).

7.7 Conclusions

Usable information on social implications of health technology is not easy to find. Information on social implications is generally not quantitative, and much of it is not in professionally published sources. Much of the available information has been developed by the lay press, both newspapers and journals. This information seems surely to heightened consciousness of social impacts in the public mind, yet there is no definitive evidence that it is so. Still, except for economic implications, there is little research-derived, organised knowledge on the social implications of health technology.

Certainly, consideration of social implications has played a part in policy decisions in the past (Reiser, 1988). Debates on the treatment of end stage renal disease and the frequent exclusion in early days of elderly people from such programs is an example of serious ethical issues coming into the public debate. The issue of treatment of very low birth weight infants is another example (Gould et al, 2000; Hack et al, 1996). More recently, genetics has been the subject of a great amount of political debate. Still, there is no body of literature that could be characterised as a serious attempt to study the social aspects of health technology.

Increasingly in the future, as technology increasingly interferes with such processes as life, death, and reproduction, and as resource pressures raise problems of solidarity and equity, systematic considerations of social aspects of technology will be necessary. Such studies seem to deserve much more emphasis in the future than they have received to the present (Office of Technology Assessment, 1982; Liberati et al, 1997).

8 Guidelines for HTA studies (Liberati et al, 1997; Polish Pharmacoeconomic Guidelines)

Definition of the policy question. The assessor should engage in a process of communication with the commissioner of the study to better articulate the policy question. Communication with user groups is necessary at an early stage. Generally speaking, the Ministry of Health will be an important user group. The timing of the study is a key issue, since an HTA that comes after the policy decision has been made is of little use. The relations between study performer and the commissioner of the study should be transparent.
Definition of the research question. The policy question must be specified in terms of research questions. These research questions then need to be related to specific sets of scientific evidence, specific tools, and specific measures of effectiveness. The extent or quality of the knowledge base relevant to the problem, including epidemiology of the health problem and natural history of the relevant condition, just is examined. For an HTA to be relevant to all users, the communication with them mentioned in point 1 is critical.

Current stage of development and use of the health technology. An HTA needs to describe the technology clearly. It should include a description of the current pattern of use, trends in use, and current indications for use. Where possible, the assessment should attempt to identify what factors might be determining uses, such as clinical perceptions, lack of alternatives, and market pressure.

Technical characteristics of a drug or device may be needed in an HTA. The content, structure, and importance of this information depend on the type of technology and the goals of the assessment.

Efficacy, safety, and effectiveness of the health technology. The evidence of efficacy and safety will usually be based on a systematic review of the published results of research. Ideally, HTA studies should focus on effectiveness and not of efficacy. Data should be collected as close as possible to routine use. Clinical research often uses **surrogate endpoints**, instead of true health outcomes; such endpoints need to be viewed with scepticism unless they have been empirical validated. **Psychological and social effects** are seen as increasing important as outcomes of health technology, although these are not widely used in HTA at present. The patients to whom the new treatment is directed must be clearly defined. Although **prospective research** is increasingly being done as part of HTA, the results need to be viewed in the context of existing knowledge. Until recently, reviews of health care research have been carried out in an ad-hoc fashion, and the results of such reviews have been of questionable validity. A review must be systematic in order to make it more informative and less susceptible to bias. Guidelines for carrying out **systematic reviews** are available. Systematic reviews carried out by HTA agencies and others should be very helpful in examining these issues.

Qualitative synthesis. A qualitative review or synthesis considers all results of research and experience, taking into account the reliability of these studies, but also helping to highlight and explore differences. A qualitative analysis of evidence is an essential step in the assessment of the effectiveness of an HTA. It is within the framework of the qualitative overview that any quantitative synthesis should take place. **Expert opinion**

can be an important part of this synthesis process. In many cases, existing knowledge will not be adequate for settling important policy questions. In other cases, the ultimate question will rest on individual and societal values regarding the costs and benefits of a technology. In both instances, systematic elicitation of informed opinion can be a valuable part of HTA.

Economic evaluation. The economic evaluation component of an HTA should be conducted within a common methodological framework. Except for prospective clinical trials with simultaneous collection of economic data, there may be little motivation to undertake the economic analysis until some evidence of efficacy is found. Clear guidelines for carrying out economic analysis are available. It is desirable to use a **social perspective** in HTA studies. The patients to whom the new treatment is directed must be clearly defined. Subgroup analysis, e.g. by patient groups, subdiagnoses, severity or co-morbidity, should be performed. The economic analysis should be done over all defined populations. **Alternative treatment** should be analysed. The current standard treatment, that is the treatment seen as the routine everyday first choice, should be the control alternative. Costs and effects should be reported as incremental, as differences between two or more alternatives. Total values should also be presented if possible. **Cost identification:** From a social perspective the treatment costs inside and outside the health care sector is the basis of the analysis. Also other costs, secondary to the treatment, inside and outside the health care sector should be reported separately. **Cost estimations:** The use of human and other resources may first be described in natural units, for instance hours, activities, nursing days, doses of medicines. All international reports used for estimation should be validated for Polish use. **Cost valuation:** Economic definitions should be used, and also commonly agreed unit price data, to facilitate comparisons. **Discounting future costs and effects:** Future costs and effects should be discounted with the same percentage. If no other guideline can be motivated, use 5 percent. A **sensitivity analysis** is always an important part of the economic evaluation.

Analytical period: The period of analysis should be long enough to give a valid and reliable judgement. If **modelling** needs to be done structure, assumptions and technique should be presented. Modelling is used especially in two cases: a) to get effectiveness results from efficacy data and b) transfer of results from studies in other countries. Assumptions in the modelling should be carefully explained and motivated. Modelling must be used with caution. It is not a substitute for reliable data obtained from clinical trials but is complementary.

Reliability and validity: In reporting the methods of analysis all assumptions should be described, ordered and motivated. Weak points in the assumptions should be discussed. A sensitivity analysis should be used to show the influence of assumptions.

Effects on the organisation of health services: Health technology both affects and is affected by the organisational structure and other aspects of health services. When new health technology is introduced, organisational structure may change or need to be changed, for example, to accommodate new routines. Efforts to educate and motivate the staff members are frequently necessary. No validated methods are available in this area; analysis of experience and "brain storming" are the main approaches to such issues.

Ethical and social aspects: In some cases, despite evidence that technology is efficacious and cost-effective, ethical and moral perspectives make it difficult to reach agreement on the preferred course of action. Therefore, an HTA often must frame claims and concerns in such a way that ethical dimensions are explicitly articulated. Although methods are lacking in this area, it is becoming increasingly important.
Reporting results: All results should be presented in a detailed manner, and thereafter aggregated in a standard form issued by the government. The use of clear, **simple language** is important. **Present conclusions, options, and (sometimes) recommendations. The scientific basis for conclusions and recommendations must be clear.**

Fully reference all sources of information, including published literature and tools such as evidence tables.

A clear-cut strategy for dissemination and implementation is essential. Development of the strategy should begin early in the assessment process.

9 References

Advisory Group on Health Technology. Assessing the effects of health technologies, principles, practice, proposals. Assessment for the Director of Research and Development. London: Department of Health, 1992.

Apfel RJ, Fisher SM. To do no harm: DES and the dilemmas of modern medicine. London: Yale University Press, 1984.

Arnstein S. Technology assessment: opportunities and obstacles. IEEE Transactions on Systems, Man, and Cybernetics 1977; SMC-7: 571-585.

Bailar JC, Mosteller F eds. Medical uses of statistics. Waltham, Massachusetts: NEJM Books, 1986.

Banta HD, van Beekum WT eds. Anticipating and assessing health care technology. Volume 8. Potentials for home care technology. Dordrecht: Kluwer Academic Publishers, 1988.

Banta HD, Buch Andreasen P. The political dimension in health care technology assessment programs. International Journal of Technology Assessment in Health Care 1990; 6:115-124.

Banta, HD, Gelijns, A, eds. Anticipating and assessing health care technology. Volume 5. Developments in human genetic testing. Dordrecht: Kluwer Academic Publishers, 1988.

Banta H, Luce B. Assessing the cost-effectiveness of prevention. Journal of Community Health 1983; 9:145-152.

Banta HD, Luce B. Health care technology and its assessment, an international perspective. Oxford: Oxford University Press, 1993.

Banta HD, Thacker SB. Assessing the costs and benefits of electronic fetal monitoring. Obstetrical and Gynecological Survey 1979; 34:627-642.

Banta HD, Thacker SB. The case for reassessment of health care technology, once is not enough. Journal of the American Medical Association 1990; 264:235-240.

Barber B. The ethics of experimentation with human subjects. Scientific American 1976; 234:25-34.

Battista R, Banta HD, Jonsson E et al. Lessons from the eight countries. Health Policy 1994; 30: 397-421.

Battista R, Hodge M. The evolving paradigm of health technology assessment: reflections for the millennium. Canadian Medical Association Journal 1999; 160: 1464-1467.

Beauchamp T. The 'four-principles' approach, In R Gillon ed. Principles of health care ethics. New York; John Wiley & Sons, 1994: pp. 3-12.

Becker G. A theory of the allocation of time. Economic Journal 1965; 75: 493-517.

Bergein Y, Tugwell P. Introduction: needs-based technology assessment. Who can afford not to use it? International Journal of Technology Assessment in Health Care 1995; 11:647-649.

Bergson A. A reformulation of certain aspects of welfare economics. Quarterly Journal of Economics 1938; 52: 310-334.

Blume S. Cochlear implantation: establishing clinical feasibility 1957-1982. In Rosenberg N, Gelijns A, Dawkins M eds. Sources of medical technology: medical innovation at the crossroads. Washington, DC: National Academy Press.

Bonsel GJ. Methods of medical technology assessment with an application to liver transplantation. Doctoral thesis. Erasmus University Rotterdam, 1991.

Briggs AH. O'Brien BJ. The death of cost-minimisation analysis? Health Economic Letters, Volume 4, Number 4, 2000.

Brodin H, Persson J. The use of cost-utility analysis in the European Commission's TIDE program. International Journal of Technology Assessment in Health Care 1995; 11: 276-283.

Brouwer WBF, Koopmanschap, MA. On the economic foundations of CEA. Ladies and gentlemen, take your positions! Journal of Health Economics 2000: 19: 439-459. 2000.

Buch Andreasen P. Consensus conferences in different countries. International Journal of Technology Assessment in Health Care 1988; 4:305-308.

Burkhardt R, Kienle G. Controlled clinical trials and medical ethics. Lancet 1978; 2:1356-1359.

Bush JW, Fanshel S, Chen MM. Analysis of tuberculine testing program using a health status index. Socioecon Plann SCI 1972;6: 49-68.

Byar D, Simon RM, Friedewald WT et al. Randomized clinical trials. New England Journal of Medicine 1976; 295:74-80.

Callahan D. What kind of life: the limits of medical progress. Washington DC: Georgetown University Press, 1990.

Calltorp T. Consensus development conferences in Sweden: effects on health policy and administration. International Journal of Technology Assessment in Health Care 1988; 4:75-88.

Carlsson P, Jonsson E, Werko L et al. Health technology assessment in Sweden. International Journal of Technology Assessment in Health Care 2000: 16: 560-575.

Carlsson P, Jorgensen T. Scanning the horizon for emerging health technologies. Conclusions from a European workshop. International Journal of Technology Assessment in Health Care 1998; 14: 695-704.

Coates J. Technology assessment. In Teich A ed. Technology and man's future. New York: St. Martin's Press, 1977:251-270.

Committee on the Life Sciences and Social Policy, Assembly of Behavioral and Social Sciences, National Resarch Council. Assessing biomedical technologies: an inquiry into the nature of the process. Washington, DC: National Academy of Sciences, 1975.

Cook D, Sackett D, Spitzer W. Methodologic guidelines for systematic reviews of randomised controlled trials in health care from the Potsdam consultation on meta-analysis. Journal of Clinical Epidemiology 1995; 48: 167-171.

Cookson R, Dolan P. Public views on health care rationing: a group discussion study. Health Policy 1999; 49: 63-74.

Coyle D, Davies L, Drummond L. Trials and tribulations: emerging issues in designing economic evaluations alongside clinical trials. International Journal of Technology Assessment in Health Care 1998; 14: 135-144.

Cranovsky R, Matillon Y, Banta D, et al. EUR-ASSESS project subgroup report on coverage. International Journal of Technology Assessment in Health Care 1997; 13: 287-332.

Culyer AJ. The normative economics of health care finance and provision. Oxford Review of Economic Policy 1989; 5(1).

Culyer A, Horisberger B. Medical and economic evaluation: a postscript. In: Culyer A, Horisberger B eds. Economic and medical evaluation of health care technologies. Berlin: Springer-Verlag, 1983: pp. 347-358.

Daddario EQ. Statement of the chairman. In: Technology assessment. Washington, DC: U.S. Congress, House of Representatives, Committee on Science and Astronautics, Subcommittee on Science, Research, and Development. 90th Congress, 1st Session, 1967.

Daniels N. Rationing fairly: programmatic considerations. Bioethics 1993; 7: 224-233.

Donaldson M, Sox H eds. Setting priorities for health technology assessment, a model process. Washington, DC: National Academy Press, 1992.

Drummond M. Allocating resources. International Journal of Technology Assessment in Health Care 1990; 6: 77-92.

Drummond M, Stoddart B. Economic analysis and clinical trials. Controlled Clinical Trials 1984; 5:115-124.

Drummond M, O'Brien B, Stoddart G, Torrance G. Methods for the economic evaluation of health care programmes. New York: Oxford University Press, 1997.

Drummond MF, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. Oxford: Oxford Medical Publications, 1987.

Eddy D. Clinical decision-making: from theory to practice. The individual vs society. Is there a conflict? Journal of the American Medical Association 1991; 265: 1446, 1449-1450.

Eddy D. Designing a practice policy, standards, guidelines and options. Journal of the American Medical Association 1991; 266:3077-3084.

Eddy D. Oregon's plan, should it be approved? Journal of the American Medical Association 1991; 266:2439-2445.

Eddy D. Practice policies, what are they? Journal of the American Medical Association 1990; 263:877-880.

Eisenberg J. Substituting diagnostic services, new tests only partly replace older ones. Journal of the American Medical Association 1989; 262:1196-1200.

Elixhauser A, Luce B, Taylor W, Reblando J. Health care cost-benefit and cost-effectiveness analysis from 1979 to 1990: a bibliography. Submitted for publication, 1992.

Engelhard H. The foundations of bioethics. Oxford: Oxford University Press, 1996.
van Everdingen JJE. Consensus ontwikkeling in de geneeskunde (Consensus development in medicine). Thesis, University of Amsterdam. Antwerp: Bohn, Scheltema en Holkema, 1988.

Feeny D. Neglected issues in the diffusion of health care technologies: the role of skills and learning. International Journal of Technology Assessment in Health Care 1985; 1:681-692.

Field M, Lohr K eds. Guidelines for clinical practice. Washington, DC: National Academy Press, 1992.

Fineberg HV. Gastric freezing: a study of diffusion of a medical innovation. In: Committee on Technology and Health Care. Medical technology and the health care system. Washington, DC: National Academy of Sciences, 1979: pp. 173-200.

Fineberg HV, Bauman R, Sosman M. Computerized cranial tomography: effect on diagnostic and therapeutic plans. Journal of the American Medical Association 1977; 238:224-227.

Fink A, Brook R, Kosecoff J et al. Sufficiency of clinical literature on the appropriate use of six medical and surgical procedures. Santa Monica, California: RAND, 1991.

Fox RC. The medicalization and demedicalization of American society. Daedalus 1977; 106:9-22.

Fox RC. A preface. International Journal of Technology Assessment in Health Care 1986; 2:189-194.

Fox RC, Swazey JP. The courage to fail. Chicago: The University of Chicago Press, 1974.

Geertsma RH, Parker RC, Whitbourne SK. How physicians view the process of change in their practice behavior. Journal of Medical Education 1982; 57:752-761.

Giacomini M, Cook D, Streiner D et al. Using practice guidelines to allocate medical technologies. An ethics framework. International Journal of Technology Assessment in Health Care 2000; 16: 987-1002

Gillon R ed. Principles of health care ethics. New York; John Wiley & Sons, 1994.

Gold M, Siegel J, Russell L, Weinstein M. Cost-effectiveness in health and medicine. Oxford, Oxford University Press, 1996.

Goodman C, Baratz S eds. Improving consensus development for health technology assessment: an international perspective. Washington, DC: National Academy Press, 1990.

Gould J, Benitz W, Liu H. Mortality and time to death in very low birth weight infants. Pediatrics 20o00: 105: E37.

Granados A, Jonsson E, Banta HD et al. EUR-ASSESS project subgroup report on dissemination and impact. International Journal of Technology Assessment in Health Care 1997; 13: 220-286.

Greer AL. The state of the art versus the state of the science: the diffusion of new medical technologies into practice. International Journal of Technology Assessment in Health Care 1988; 4:5-26.

Groot LMF. Diffusion of medical technology, a Dutch case study. Presented at the EEC Workshop on Regulatory Mechanisms Concerning Expensive Health Technology, London, 22-25 April 1986.

Hack M, Friedman H, Fanaroff A. Outcomes of extremely low birth weight infants. Pediatrics 1996; 98: 931-937.

Hadorn D, The Oregon priority-setting exercise: cost-effectiveness and the rule of rescue revisited. Medical Decision Making 1996; 16: 117-119.

Hailey D, Menon D. A short history of INAHTA. International Network of Agencies for Health Technology Assessment. International Journal of Technology Assessment in Health Care 1999; 15: 236-242.

Haley LRW. Managing hospital infection control for cost-effectiveness: a strategy for reducing infectious complications. Chicago: American Hospital Publishing Inc., 1986.

Hall M. Making medical spending decisions: the law, ethics, and economics of rationing mechanisms. Oxford: Oxford University Press, 1997.

Heitman E. Ethical issues in technology assessment: conceptual categories and procedural considerations. International Journal of Technology assessment in Health Care 1998; 14: 544-566.

Henshall C, Oortwijn W, Stevens A, et al. Priority setting for health technology assessment. Theoretical considerations and practical approaches. A paper produced by the EUR-ASSESS Project. International Journal of Technology Assessment in Health Care 1997/ 13: 144-185.

Hilborne L, Leape L, Kahan J et al. Percutaneous transluminal coronary angioplasty. Santa Monica, California: RAND, 1991.

Illich I. Medical nemesis: the expropriation of health. New York: Pantheon Books (Random House, Inc.), 1976.

Institute of Medicine. Consensus development at the NIH: improving the program. Washington, DC: National Academy Press, 1990.

Jacoby I. The consensus development program of the National Institutes of Health. International Journal of Technology Assessment in Health Care 1985; 1:420-432.

Johnsson M. Evaluation of the consensus development program in Sweden: its impact on physicians. International Journal of Technology Assessment in Health Care 1988; 4:89-94.

Jones-Lee, MW. Hammerton, M. Phillips, PR. The value of safety: Results of a national sample survey. Economic Journal, 95, 49-72. 1985.

Kanouse D, Winkler J. Kosecoff J et al. Changing medical practice through technology assessment. Santa Monica, California: RAND, 1989.

Kass LR. The new biology: what price relieving man's estate? Science 1971; 174:779-788

Kaye HL. The biological revolution and its cultural context. International Journal of Technology Assessment in Health Care 1986; 2:275-283.

Klarman H, Francis, J, Rosenthal G. Cost-effectiveness analysis applied to the treatment of chronic renal disease. Medical Care 1968; 6: 49-54.

Kosekoff J, Kanouse D, Rogers W, McCloskey L, Winslow C, Brook R. Effects of the National Institutes of Health consensus development program on physician practice. Journal of the America Medical Association 1987; 258:2708-2713.

Lambert EC. Modern medical mistakes. Bloomington, Indiana: Indiana University Press, 1978.

Lancaster K. A new approach to consumer theory. Journal of Political Economy 1966; 74: 132-157.

Lasch K, Maltz A, Mosteller F, Tosteson T. A protocol approach to assessing medical technologies. International Journal of Technology Assessment in Health Care 1987; 3:103-122.

Levine R. Ethics and regulation of clinical research. Urban & Schwarzenberg, 1981.
Liberati A, Sheldon T, Banta HD. EUR-ASSESS project subgroup report on methodology. Methodological guidance for the conduct of health technology assessment. International Journal of Technology Assessment in Health Care 1997; 13: 186-219.

McNeil B. Values and preferences in the delivery of health care. In: Committee for Evaluating Medical Technologies in Clinical Use. Assessing Medical Technologies. Washington, DC: National Academy Press, 1985: pp. 535-541.

Marcus SH, Grover PL, Revicki DA. The method of information synthesis and its use in the assessment of health care technology. International Journal of Technology Assessment in Health Care 1987; 3:491-493.

Meinert C. Toward more definitive clinical trials. Controlled Clinical Trials 1980; 1:249-261.

Moses L. Framework for considering the role of data bases in technology assessment. International Journal of Technology Assessment in Health Care 1990; 6:183-193.

Mulrow C, Cook D eds. Systematic reviews, synthesis of best evidence for health care decisions. Philadelphia: American College of Physicians, 1998.

Murphy M, Black N, Lamping D et al. Consensus development methods and their use in clinical guidelines development. Health Technology Assessment 19998/ 2: 1-83.
National Heart and Lung Institute. The totally implantable artificial heart. Report of the Artificial Heart Assessment Panel. Bethesda, Maryland: National Institutes of Health, 1973.

NHS Centre for Reviews & Dissemination. Review of the research on the effectiveness of health service interventions to reduce variations in health. York, England, CRD Report 3, 1995.

NHS Centre for Reviews and Dissemination. Undertaking systematic reviews of research on effectiveness: CRD guidelines for those carrying out or commissioning reviews. CRD Report 4. York: University of York, 1996.

Nord E. Cost-value analysis in health care: making sense out of QALYs. Cambridge: Cambridge University Press, 1999.

Office of Technology Assessment. Assessing the efficacy and safety of medical technologies. Washington, DC: US Government Printing Office, 1978.

Office of Technology Assessment. Cost effectiveness of influenza vaccination. Washington, DC: U.S. Government Printing Office, 1981.

Office of Technology Assessment. Development of medical technology: opportunities for assessment. Washington, DC: US Government Printing Office, 1976.

Office of Technology Assessment. Evaluation of the Oregon Medicaid proposal. Washington, DC: US Government Printing Office, 1992.

Office of Technology Assessment. Identifying health technologies that work: searching for evidence. Washington: US Government Printing Office, 1994.

Office of Technology Assessment. The implications of cost-effectiveness analysis of medical technology. Washington, DC: US Government Printing Office, 1980.

Office of Technology Assessment. The implications of cost-effectiveness analysis of medical technology. Background paper # 1. Methodological issues and literature review. Washington, DC: US Government Printing Office, 1980.

Office of Technology Assessment. The quality of medical care, information for consumers. Washington, DC: US Government Printing Office, 1988.

Office of Technology Assessment. A review of selected Federal vaccine and immunization policies. Washington, DC: US Government Printing Office, 1979.

Office of Technology Assessment. Strategies for medical technology assessment. Washington, DC: US Government Printing Office, 1982.

Oortwijn W. First things first. Priority setting for health technology assessment. Leiden, 2000.

Oxman A ed. The Cochrane Collaboration handbook, VI: Preparing and maintaining systematic reviews. Oxford: Cochrane Collaboration, 1994.

Paul-Shaheen P, Clark J, Williams D. Small area analysis: a review and analysis of the North American literature. Journal of Health Politics, Policy and Law 1987; 12:741-792.

Petitti D. Meta-analysis, decision analysis, and cost-effectiveness analysis, methods or quantitative synthesis in medicine. Oxford: Oxford University Press, 1994.

Pirsig RM. Zen and the art of motorcycle maintenance. New York: Bantam Books, 1975.

Preston T. Coronary artery surgery: a critical review. New York: Raven Press, 1977.

Reiser SJ. A perspective on ethical issues in technology assessment. Health Policy 1988; 9: 297-300.

Robert G, Stevens A, Gabbay J. 'Early warning systems' for identifying new healthcare technologies. Health Technology Assessment 1999; 3: 1-97.

Rossum W van. Besluitvorming en Medisch Technologisch Aspectenonderzoek, Een analyse van de Gebruikswaarde van TA-Studies naar Hart- en Levertransplantaties en In Vitro Fertilisatie (Decision-making and Medical Technology Assessment, An Analysis of the Utility of TA Studies of Heart and Liver Transplant and In Vitro Fertilization). Appeldoorn: Ziekenfondsraad, 1990.

Rossum W van. Medical technology assessment, an analysis of the perceived value of the Dutch technology assessments of heart and liver trasnplantation and in vito fertilization (in Dutch). Medisch Contact 1990; 45:509-511.

Sachs et al. Meta-analysis of randomized clinical trials. New England Journal of Medicine 1987; 316:450-5.

Safran C. Using routinely collected data for clinical research. Statistics in Medicine 1991; 10:559-564.

SBU. Metoder for rokavvanjning. SBU-rapport nr 138. Stockholm, 1998.

Schwartz D, Flamant R, Lellouch J. Clinical trials. New York: Academic Press, 1980.

Scitovsky AA. Changes in the costs of treatment of selected illnesses, 1971-1981. Medical Care 1985; 23:1345-1357.

Showstack JA, Hughes Stone M, Schroeder SA. The role of changing clinical practices on the rising costs of hospital care. New England Journal of Medicine 1985; 313:1201-1207.

Showstack JA, Schroeder SA, Matsumoto MF. Changes in the use of medical technologies, 1972-1977. A study of 20 inpatient diagnoses. New England Journal of Medicine 1982; 306:706-712.

Silverman WA. Human experimentation: a guided step into the unknown. Oxford: Oxford University Press, 1985.

Sisk J. Introduction to measuring health care effectiveness. International Journal of Technology Assessment in Health Care 1990; 6:181-182.

Spitzer WO. Report of the task force on the periodic health examination. Canadian Medical Association Journal 1979; 121:1193-1254.

Stevens A, Packer C, Robert G. Early warning of new health care technologies in the United Kingdom. International Journal of Technology Assessment in Health Care 1998; 14: 680-686.

Sugden R, Williams A. The principles of practical cost-benefit analysis. Oxford: Oxford University Press, 1978.

Swazey J, Fox R. The clinical moratorium. In Fox F. Essays in medical sociology: journeys into the field. New York: John Wiley & Sons, 1979.

Thacker SB. The efficacy of intrapartum electronic fetal monitoring. American Journal of Obstetrics and Gynecology 1987; 156:24-30.

Thacker S, Stroup D. 2000. Continuous electronic heart rate monitoring for fetal assessment during labor. Cochrane Database Syst. Rev. CD000063.

Tierney W, McDonald C. Practice databases and their use in clinical research. Statistics in Medicine 1992; 10:541-557.

TNO. Report 3. Methodology of the cost-effectiveness analysis of particular groups of services. Health technology assessment for policy-makers. Leiden, 2001.

Torrance G, Thomas W, Sackett D. A utility maximization model for evaluation of health care programs. Health Services Research 1972; 7: 118-133.

Tugwell P, Sitthi-Amorn C, O'Connor A et al. Technology assessment. Old, new, and needs-based. International Journal of Technology Assessment in Health Care 1995; 11: 650-662.

United Nations High Commission for Human Rights. Substantive issues arising in the implementation of the International Covenant on Economic, Social and Cultural Rights: General Comment No. 14 (2000). The right to the highest attainable standard of health. Geneva, 1999.

US Preventive Services Task Force. Guide to clinical preventive services: an assessment of the effectiveness of 169 interventions. Baltimore: Williams & Wilkins, 1989.

US Preventive Services Task Force. Guide to clinical preventive services. Baltimore: Williams and Wilkins, 1996 (second edition)
de Vries MJ. The redemption of the intangible in medicine. London: Institute of Psychosynthesis, 1981.

Warner K, Luce BR. Cost-benefit and cost-effectiveness analysis in health care, principles, practice, and potential. Ann Arbor, Michigan: Health Administration Press, 1982.

Weinstein MC, Fineberg HV. Clinical decision analysis. Philadelphia: WB Saunders Co., 1980.

Welch HG, Fisher E. Cost-containment efforts in the public sector: Oregon's priority list. In: Gelijns A ed. Technology and health care in an era of limits. Washington, DC: National Academy Press, 1992: pp. 63-75.

Wennberg JE. Dealing with medical practice variations: a proposal for action. Health Affairs 1984; 3:6-32.

Woolf S, Battista R, Anderson O et al. Assessing the clinical effectiveness of preventive maneuvers: analytic principles and systematic methods in reviewing evidence and developing clinical practice recommendations. The Canadian Task Force on the Periodic Health Examination. Journal of Clinical Epidemiology 1990; 43:891-905.

Wortman PM, Saxe L. Assessment of medical technology: methodological considerations. Appendix C in Office of Technology Assessment. Strategies for medical technology assessment. Washington, DC: US Government Printing Office, 1982: pp. 127-149.

Wortman PM, Vinokur A. Evaluation of NIH consensus development process. Phase I: final report. Ann Arbor, Michigan: Center for Research on Utilization of Scientific Knowledge, University of Michigan, 1982.

Wortman PM, Yeaton WH. Using research synthesis in medical technology assessment. International Journal of Technology Assessment in Health Care 1987; 3:509-522.