Learning to live with Health Economics

Edited by H. Zöllner,
G. Stoddart and C. Selby Smith

Chapter V
Useful economic tools

WHO Regional Office for Europe

Copenhagen, 2003

Key Words

HEALTH ECONOMICS
DELIVERY OF HEALTH CARE – economics
HEALTH POLICY – economics
SOCIOECONOMIC FACTORS
HEALTH SERVICES ACCESSIBILITY
HEALTH CARE REFORM
COST–BENEFIT ANALYSIS – methods
OUTCOME ASSESSMENT (HEALTH CARE)
PROGRAM EVALUATION – methods
FORECASTING
TEACHING MATERIALS

EUR/03/5042783

Contents

5. Useful economic tools

5.1 Introduction

The final chapter of the learning materials provides modules on four useful tools in the economist's toolkit: health outcome assessment, costing, economic evaluation, and economic modelling and forecasting. These tools are relevant to each of the four main groups of potential users of the learning materials in terms of *appreciation* and *appraisal*. They can all benefit from knowing that such tools are available, their strengths and weaknesses, and when their use will be appropriate. However, the most senior users, and many of the concerned public groups, will tend to benefit from knowing they are available; where, when, and how they can best be used; their overall strengths and weaknesses; and the broad thinking which lies behind them; whereas the other two groups of potential users may have greater interest, in addition, in the more detailed aspects.

Module 5.2.1, by the late Dr Oliver Sangha and Dr Manfred Wildner from Munich, Germany, is concerned with outcome assessment in health care. This is an important topic, since it is essential to know what consequences follow, or are likely to follow, from specific actions in health care and other health-related systems. For example, scarce resources cannot be allocated efficiently if this information is not available for decision-makers (at all levels). Neither can appropriate decisions be made in relation to the pursuit of equity objectives. In recent years, growing interest has been shown in information about health outcomes by patients, providers, payers and policy-makers, and this seems likely to increase even further in the future. The authors stress that it is important to distinguish between efficacy, effectiveness and efficiency. There is also an important distinction to be made between process-related non-health dimensions and health outcomes. They recognize that health is a multidimensional construct, that individual and collective health care can be looked at separately, and that health care is only one among a number of determinants of health. They identify six major domains of health outcomes (the six Ds): disease (morbidity), death (mortality), discomfort, disability (limitations in functioning), dollars (costs) and dissatisfaction (preferences and satisfaction with care). Obviously, the appropriate measurement of health status requires the use of standardized instruments with proven psychometric properties, notably validity, reliability and sensitivity. Values are also important and the module argues that the objectives of health outcomes assessment are "based on equity and equality, quality of care, patient's autonomy and choices and responsiveness to patients". In their view, the explicit health rights of participants in the health care system, including patients and their carers, will play an increasing role in the future. This matter was discussed further in Module 4.4.2.

Module 5.2.2, on costing, was prepared by Professor Chris Selby Smith from Monash University in Australia. Since resources are scarce it is not possible to produce all the outputs which would be thought desirable. Thus, choices have to be made; and the cost of alternative courses of action is relevant to much decision-making in health care (consequences also need to be considered). For economists, costs refer to the opportunities foregone elsewhere because the resources are used for this particular purpose. Thus, costs are wider than financial expenditure alone. Other resources especially need to be considered if their opportunity costs are not adequately reflected in market prices. Examples include voluntary contributions; the time costs involved for patients and their carers; and services contributed by religious orders. Total costs are important; they can be viewed from various perspectives, such as the sources or types of costs or distribution of total costs. Changing the distribution of given total costs among the various parties to the complete resource allocation decision can alter the incentives they face and therefore the actions they take. Since the cost information sought by economists is often difficult to obtain, three stages can usefully be distinguished in costing studies: identification, measurement and valuation. Frequently the three stages become progressively more difficult. The module emphasizes that the purpose of collecting and analysing cost estimates is to contribute to improved decision-making. There are many purposes for which cost information is relevant; and the specific cost information required can only be determined by reference to the particular objectives of the decision-maker. While cost information is not the only input required, high quality decision-making is more likely to occur when adequate cost information is available. The module also notes that cost analyses can be undertaken, and cost information presented, in ways which are more or less helpful for decision-makers.

Module 5.3.1, on economic evaluation in health and health care (both in theory and practice), was prepared by Professor Michael Drummond of York University in the United Kingdom. Increasing pressures on health care budgets have led decision-makers throughout the European Region of WHO to search for methods of assessing the value for money from different health care treatments and programmes. In economic evaluation, programmes are compared in terms of their costs and their consequences, such as improvements in health outcomes and savings in health care resources. There are various forms of economic evaluation, such as cost-minimization analysis, cost–effectiveness analysis, cost–benefit analysis and cost–utility analysis. All of them include an analysis of costs, which underlines the significance of the factors considered in the previous module. It is important for all groups of potential users of these learning materials to understand the key methodological principles that are involved in economic evaluation approaches and to appreciate how they are applied (or misapplied) in the specific studies with which they are concerned. These principles include the consideration of an adequate range of alternatives, the use of good evidence of effectiveness, and allowance for uncertainty in the estimates of costs and benefits. The module includes a helpful checklist of matters to take into account when appraising an economic evaluation (or considering whether one should be undertaken). Economic evaluations have a range of uses in health care and related activities. For example, they can be used, generally in association with other policies, to encourage a rational diffusion and use of health technologies, including planning of specialist facilities, reforming payment schemes for institutions or health care professionals, and developing health care practice guidelines. Economic evaluations can also be used to assess health-producing measures in different sectors of the economy, including road safety, environmental protection and occupational health. For many of the potential users of these learning materials the way of thinking is likely to be at least as valuable as the

detailed arguments about costs and benefits, risk and uncertainty, and the distributional implications of the alternatives being compared. However, the way in which studies are conducted and reported can contribute to making them more or less useful for decision-makers elsewhere (e.g. reporting prices and quantities separately rather than only total expenditure).

Module 5.4.1, by Professor Reiner Leidl of the University of Ulm in Germany, considers economic modelling and forecasting. Economic models are a useful tool for the support of decision-making and policy development, since transparent models can structure problems, make explicit the assumptions used, and explore the consequences implied by particular decisions. Explanation, prediction and simulation are the main general purposes of economic models. The module contains a table which lists eight points to be checked by decision-makers who are considering modelling approaches. They are important since models are tools that have to be implemented properly and used adequately: "methodological expertise, expertise in the health problem investigated, and expertise in how to support decisions by model results is required." However, so long as these preconditions are met, models can significantly improve the available information and support decision-making in a transparent and rational way. The remainder of the module, recognizing that there are many different problems and many different types of model, considers three major approaches: a decision tree model, scenario analysis and disease modelling, and econometric models. This module is potentially valuable for each group of users envisaged for the learning materials, but detailed knowledge of the various approaches is likely to be particularly relevant to managers and health professionals. The most senior decision-makers and members of the various concerned groups are more likely to want to know that they are available, their strengths and weaknesses, where they are likely to be useful and how they can be appropriately incorporated into wider decision-making processes. In this sense the module is typical of many others included in these learning materials.

5.2 Outputs and inputs

5.2.1 Outcome assessment in health care

Oliver Sangha and Manfred Wildner[1]

Key messages

- In recent years, there has been an increasing interest in information about health outcomes from patients, providers, payers and policy-makers.
- The goal of health care is to protect, promote and preserve people's health. This requires standardized assessment of both organ morphology and function, as well as health status.
- To understand the concepts of health outcomes assessment, it is important to distinguish between efficacy, effectiveness and efficiency. The distinction between process-related non-health dimensions and health outcomes is also important.
- Moreover, it is important to recognize that health is a multidimensional construct, that individual and collective health can be looked at separately, and that health care is but one of many determinants of health.
- The objectives of health outcomes assessment are based on equity and equality, quality of care, patients' autonomy and choices, and responsiveness to patients. Observation of explicit health rights is likely to play an increasing role in the future.
- The main domains of health outcomes include the six Ds: disease (morbidity), death (mortality), discomfort, disability (limitations in functioning), dollars (costs), and dissatisfaction (preferences and satisfaction with care).
- Measurement of health status requires the use of standardized instruments with proven psychometric properties (validity, reliability, sensitivity).

Tutors' notes

This module introduces a valuable skill set which can be useful for all four groups of potential users of the learning materials. The economic way of thinking requires consideration of costs in relation to benefits for alternative courses of action by decision-makers at all levels of the health care system, and

[1] This module was prepared by Dr Manfred Wildner of the Bavarian Public Health Research Centre, LM University of Munich, Germany (e-mail : wil@ibe.med.uni-muenchen.de) and the late Dr Oliver Sangha, former Head of Research Unit at the Centre.

in related areas which affect health. This cannot be achieved without attention to the consequences, beneficial or adverse, from health interventions or their absence. Similarly, consideration of equity aspects of health care, health and wellbeing require that there be at least some broad attention to measures of the relevant outcomes.

The module provides a valuable introduction to this topic, including discussion of:

- the conceptual framework and theoretical concepts
- the objectives for which health outcome assessment is undertaken
- various methods and instruments of health outcome assessment
- criteria for choosing particular instruments.

The different groups of potential users of the learning materials could use the module in their separate groups. If so, the senior political and bureaucratic participants, and the members of the various concerned public groups, are likely to want a more general approach, while the other two groups may prefer a more detailed discussion. However, the module could also be used with participants who cross the boundaries of the different user groups, either in real life or by assuming roles in the context of the learning experience.

The first exercise is aimed at the level of *appreciation*. It can be used with separate (or mixed) groups of:

- policy-makers at the senior political and bureaucratic levels in health and health-related agencies;
- civil servants, other government staff in health or other agencies, and similar people in bodies such as nongovernmental organizations providing relevant services and voluntary, religious and charitable organizations;
- managers of health care facilities, such as hospitals, facilities for care of the elderly or community health centres, and managers in other health and health-related organizations;
- health care professionals, such as doctors, nurses, dentists, pharmacists or therapists;
- a wide range of other concerned public groups.

The second exercise is aimed at the level of *appraisal*. It can be used with a similarly broad range of potential users of the learning materials. It can be valuable for senior political and bureaucratic participants, and for the members of concerned public groups. However, managers, health care professionals and patients may be most interested in the details of where health outcome measurement can be usefully undertaken, how it can best be done and what implications it has. The former two groups of users may be especially concerned with how health outcome measurement relates to providing improved health for groups, whereas the latter two groups may be more focused on outcomes for particular individuals or small groups (e.g. their families or their patients).

Introduction

The purpose of this module is to provide an analytical exposé of different methodologies of assessing the health outcome of health care systems. In order to do so, the module offers background in two areas: firstly, it provides a brief analysis of theoretical concepts, which are important for the understanding of the relationship between health care and health outcomes. Secondly, it analyses the interrelationship between the broad objectives of health care provision and the patients' perspective.

In recent years, there has been growing interest in information about health outcomes for several reasons. Patients demand information to make informed decisions about their own care or the care of

their relatives. Health care providers are being made more and more accountable for what happens to patients. Doctors and hospitals are shifting their care towards evidence-based medicine with reliable data on efficacy and effectiveness of care. Finally, payers and policy-makers need to base their decisions about health care provision, insurance coverage and benefit planning on information about how policies might influence the health outcomes of individual patients and populations.

The goal of health care is to protect, promote and preserve people's health. This requires standardized assessment of both organ morphology and function as well as health status. Traditionally, measures of success or failures of health have addressed the four Ds – death, disease, disability and discomfort *(1)*. In particular, mortality and morbidity data have been widely used because they were most accessible from medical records, hospital databases or governmental sources. While mortality rates, life expectancy or the prevalence of disease provide significant inferences on population health, they say little about any other point on the continuum of dysfunction between perfect health and death or on individual health.

In recent years, a wide body of research has enabled sophisticated measurement of health status. Moreover, White's four Ds have been expanded by two additional Ds: dissatisfaction and dollars, to address patients' satisfaction and the efficiency of health care. Furthermore, the fulfilment of patients' rights, with respect to health care, has become more important in the global discussion of health outcomes.

Before the reader is introduced to selected methods and instruments of outcome assessment, we will provide a brief introduction to the theoretical concepts and objectives of health outcomes assessment.

Theoretical concepts

Definition of health

Good health is assigned the highest value in most societies. Definition of health as an operational and thus measurable concept has, however, been elusive. From antiquity health has been thought as a physical or mental state with assessments focusing on the presence or absence of diseases. Departing from prior definitions, Henry Siegerist stated in 1941, "… health is therefore not simply the absence of disease: it is something positive …" *(2)*. Building on Siegerist's definition, WHO's Constitution stated in 1947, "Health is a state of complete physical and social wellbeing, and not merely the absence of disease or infirmity" *(3)*. Since then, health status has incorporated measures of the physical, mental and social functioning of individuals.

In 1977, the World Health Assembly decided that the main social goal of governments and the World Health Organization should be the attainment by all citizens of the world by the year 2000 of a level of health that would permit them to lead a socially and economically productive life ("health for all").

Conceptual framework of health outcomes

Modelling is the basis for understanding health outcomes. All models of health outcomes are principally based on the WHO definition of health. WHO compares the aetiology and processes (pathology) of illnesses with the following three levels of disease consequences. **Impairment** is defined as any loss or abnormality of psychological, physiological or anatomical structure or function. It refers to the

level of an individual or organ system. Altered organ morphology or "damage" may cause organ dysfunction. Impairment is concerned with abnormalities of body structure and appearance and with organ or system function resulting from any cause. **Disability** is the physical and psychological functional limitation caused by an impairment which is described by an individual when there is a discrepancy between his or her capacity and an actual or perceived need for a specific function. Handicaps reflect the effects of disability, and of adaptations to it, on an individual's ability to perform social roles (e.g. work, parenthood) and thus the degree of social disadvantage conferred by the disability.

Although "outcome" is often used in a simple and global fashion, it is actually a complex construct composed of several independent dimensions. Hence, the use of "outcomes" in the plural form is intended to reflect the multidimensional nature of the term.

Monitoring of health care systems

Health outcomes can also be used to monitor health care systems. Health care systems have been described as all people and all activities whose primary purpose is to improve health. These systems may be formally integrated and centrally directed, or may consist of a multitude of particular services directed at promoting, restoring or maintaining health. Assessment of the quality of these services may be directed towards their structure, processes and outcomes. Evaluation of the performance of health care systems therefore takes account of measures of process and structure, such as responsiveness and fairness of financial contributions, in addition to health outcomes. Responsiveness relates to the non-health aspects and reflects the ability of the system to respond to all patients' needs swiftly. Fair financial contributions relate to the financial risks of households which should be distributed according to ability to pay and not according to the risk of illness or disability.

As diverse as the activities of health care systems are regarding structure and process, their common **outcome** is "producing health". Thus it may seem straightforward to monitor health care systems by assessing the overall health state, but it must be noted that:

- health is a multidimensional construct
- individual and collective health can be looked at separately, and
- health care is but one of many determinants of health.

The implications of the idea that health is a multidimensional construct have been detailed above. The necessity of a distinction between individual and collective health can be illustrated by the paradoxical fact that high quality health care may improve health at the individual level, while it increases at the same time the disease burden at the population level due to prolonged life in a less than perfect health state. While the individual may appreciate his or her health gain, the collective is experiencing an increase in prevalent disease – and vice versa. Moreover, the population's health may under certain circumstances improve at the expense of, or with disregard to, the needs of a minority within the population whose health status is deteriorating. These circumstances may be the consequence of, for example, rationing expensive services such as dialysis or transplantations.

Finally the fact that health is a "byproduct" of a number of activities and determinants must be considered. Examples of such diverse determinants are: the higher rate of cardiac death rates in a cold climate; the interaction between sanitation, climate and vector control; the influence of body height on fall-related deaths; cultural norms regarding food composition and the availability of healthy food; the influence of general education levels; population density and the geographical dispersion of health care infrastructure; the inverse association of road traffic death rates with increasing motor vehicle density; and the influence of legislation such as seat belt laws or speed limits and safety engineering standards.

In short, a distinction between process-related non-health dimensions and health outcomes is necessary. As regards health status, a multidimensional assessment should be performed, the absolute health status must be distinguished from the relative health gain related to health care system activities, and the health of individuals likewise distinguished from the health of populations.

Exercise 1

Discuss the ways in which resources are distributed to competing health care programmes in your country. Assuming that each programme should yield clearly defined benefits, elaborate on how you would define and measure health benefits (outcomes). Each participant may choose a specific perspective (e.g. patient, provider, payer, policy-maker).

Objectives of health outcomes assessment

Efficacy, effectiveness and efficiency

Health outcomes are primarily used in the contexts of efficacy, effectiveness and efficiency. **Efficacy** is the ability of an activity to achieve its goal under ideal or laboratory conditions. An example would be the successful transplantation of cartilage cells into defects on the surface of a joint in a specialized centre. This proven efficacy does not, however, guarantee the success of this procedure when it is performed by less specialized surgeons as part of their routine. **Effectiveness** hence describes the performance of a health care activity under conditions of "normal" health care. To go back to our example, the transplantation of cartilage cells may also be effective under these conditions when the surgeon is specially trained and increasingly skilful. However, this requires the allocation of specialized personnel, time in the operating theatre and the availability of tissue cultures. **Efficiency** therefore puts the health gain of an activity in perspective against the resources spent, which could also have been used for an alternative health activity: the patient could have had intensive physiotherapy instead and a knee replacement in the case of persistent problems.

It is important to note that the outcome measure for efficacy and effectiveness can be either traditional physician-centred clinical measures, for example a reduction in pain or improvement in functions such as the range of motion of a joint or multidimensional, or patient-centred health outcomes. Efficiency measures and values the costs for a certain health gain, e.g. the costs for each knee replacement avoided. If a cost–effectiveness analysis comparing alternative measures is planned, care must be taken that the same dimension of effectiveness is assessed for the concurrent measures. Efficacy, effectiveness and efficiency relate primarily to natural units of success as they are understood in clinical practice, e.g. reduction of pain or improvement in function. This is different from the approach of calculating so-called "utilities", a unifying concept for measuring health outcome for advanced economic analyses, which allows the comparison of the outcomes of very different diseases and treatments.

Quality of care

Assessment of the outcome of hospital treatment – mainly survival – has been the historical basis of current approaches to assure the quality of care. Quality management requires the feedback of meaningful outcome information into the assessment component of a "plan–do–check–assess" cycle. Such meaningful information may be simple survival information, for example following major surgery, or information on complication or adverse event rates such as for maternity services or immunization

programmes. It may also be comprehensive information on multidimensional health status assessment, e.g. in phase 3 and phase 4 trials of new drugs or in the management of chronic diseases. It has been demonstrated that a patient-centred approach to health outcomes assessment supplements the clinician-based assessment with important and at times very different outcomes information. Quality of care is ultimately determined by the patient's experience of improvement in the state of his or her illness and functional limitation. Modern multidimensional outcome measures are patient-centred and capture the patients' perspective well.

Responsiveness has been added more recently to the assessment of health systems. Responsiveness relates to non-health aspects of the provision of preventive services, care or non-personal services. It includes respect for the dignity of the person, confidentiality, autonomy to decide on treatment options, prompt attention, quality of the amenities (e.g. cleanliness, space, food), access to social support networks and free choice of providers. Hence rudeness in relation to patients, long waiting times, denial of access to care, unnecessary isolation or insensitivity to cultural values of connectedness, denial of choice and gag clauses in managed care systems all reflect poor responsiveness on the part of health care systems. Poor responsiveness may result in poor take-up of preventive services or treatment options. As responsiveness also reflects *expectations*, its assessment may vary with, for instance, the education level or socioeconomic status of the assessor. Moreover health systems often consist of compartments, for example defined by remuneration for services, so that their perceived responsiveness may be as heterogeneous as the system.

Some aspects of responsiveness, e.g. long waiting lists for surgery, are captured by health outcomes such as survival or quality of life. Other aspects, for instance respect for the dignity of the person need a special assessment, e.g. by satisfaction questionnaires or expert interviews.

Methods and instruments of health outcome assessment

Morbidity and mortality

Perhaps the most basic outcome measure of health care is death. The death rate of newborn babies (neonatal mortality rate) or of children under five years (infant mortality rate) and the maternal mortality rate are established indicators for the quality of a health system's performance. Other indicators are the standardized overall mortality rate within a population or the average life expectancy at birth. The influence of factors unrelated to health care on these measures has been pointed out earlier. The survival time is often an important outcome indicator following major therapeutic interventions, e.g. organ transplantations. Another approach to death as an outcome measure is to count the potential years of life lost and add them up for a population (setting an upper age limit). The total amount of potential years of life lost thus depends on the average loss of life years and on the frequency of these deaths in a population. Alternatively, the further life expectancy can be calculated from population-specific period life tables (period-expected years of life lost), from a cohort life table (cohort-expected years of life lost) or from an ideal standard (standard-expected years of life lost).

However, the consequences of many diseases are not adequately captured by focusing on reduced life expectancy. In particular, chronic diseases such as asthma, diabetes or rheumatism may have no significant influence on life expectancy while being a considerable burden to both the individual patient and the health care system. Traditionally, morbidity is captured by the physician as anatomical or physiological **impairment**: e.g. as a reduced forced expiratory volume, a reduced glucose tolerance or range of joint motion. If this impairment is experienced by the patient as functional limitation it is

called **disability,** e.g. the inability to run fast, stay without food for extended periods or do manual labour. If this disability interferes with the social role, e.g. the profession, a **handicap** is present. Due to their heterogeneity, organ-based clinical morbidity measures by and large have a narrow application in the context of the management of specific diseases.

An early attempt to capture **non-fatal health outcomes** more generally was the development of quality-adjusted life years (QALY). This method assigns weights between 1.0 (perfect health) and 0.0 (near death) to various health states, and sums up the QALYs by multiplying the time spent in each state by the respective weight. Frequently an additional age-weighting is used giving lower weights to years of old age. Murray et al. have promoted the concept of disability-adjusted life years (DALY) as a unifying concept to quantify the burden of disease within populations *(4)*. The focus on disability rather than the socially constructed handicap should ensure comparability of like conditions. This time-based health status measure aims to capture both loss of life years and loss of quality of life years, incorporating non-fatal health outcomes. Moreover, it aims to provide a unifying measure for cost–effectiveness analyses. Twenty-two indicator diagnoses were selected as a reference for the calculation of utilities, which are discounted for future life years in accordance with economic theory. An age-weighting gives highest weights to years of young adult age. A discussion of the necessary characteristics of a measure to qualify as a utility – utility independence, constant proportional trade-off and risk neutrality – goes beyond the scope of this introductory text.

Health status and health-related quality of life

Generally accepted dimensions of health status include physical, social, and emotional functioning. Two classes of measure can be used to assess health outcomes: generic measures and condition-specific measures. Among generic instruments, a further distinction is made between health profile and utility measures. **Health profile measures** are instruments that intend to measure all important aspects of health status or health-related quality of life. **Utility measures** are derived from economic and decision theory. They reflect patients' preferences for different health states. The chief element of utility measures are that they integrate utility measures and relate health states to death. The results from utility measures are frequently used as outcomes in cost–utility analysis.

In accordance with the WHO definition of health, **generic health status instruments** measure multiple aspects of health, including physical function, social function and pain. They are suitable for comparing health status across multiple diseases or the value of competing clinical programmes. Generic health status instruments are useful in the evaluation of subjects with multiple chronic conditions, since they can detect changes arising from different organ systems. This is of particular interest when interventions can have (adverse) effects on several organ systems.

Widely used representatives of generic health status measures include the Sickness Impact Profile (SIP), the Quality of Well-Being Index (QWB), the Nottingham Health Profile (NHP), the WHO Quality of Life Assessment (WHOQoL), the EuroQol (EQ)-5D, the Health Utilities Index (HUI) and the Short-Form 36 (SF-36).

The **Sickness Impact Profile** (SIP) *(5)* is a widely used general health status instrument containing 136 items answered true or false. Scores use predetermined weights based on rater panel estimates of relative severity of the dysfunction. The categories of ambulation, body care and mobility are aggregated into a physical dimension, and the four categories of emotional behaviour, social interaction, alertness behaviour and communication are aggregated into a psychosocial dimension. The remaining categories are work, sleep and rest, eating, home management, and recreation and pastimes.

The **Quality of Well-Being Index** (QWB), and an earlier version the **Index of Well-Being** *(6)*, assess mobility, physical activity and social activity. An interviewer asks what the patient did because of illness during the previous six days. Scoring for particular functions is based on preference weights derived from the normal population.

The **Nottingham Health Profile** (NHP) *(7)*, and its predecessor the **Nottingham Health Index** (NHI) *(8)*, assess perceived physical, social and emotional health with 38 items answered yes/no. It uses weighted scores from panels' judgments about the severity of individual items. The NHP covers physical mobility, pain, emotional reaction, energy level, sleep and social isolation, and can provide dimension-specific scores.

The **WHO Quality of Life Assessment** (WHOQOL)-100 *(9)* contains 100 questions over six broad domains of quality of life within which 24 facets are covered. The six domains include physical health, psychological health, level of independence, social relationships, environment, and spirituality/religion/personal beliefs. Four items are included for each facet, as well as four general items covering subjective overall quality of life and health, producing a total of 100 items in the assessment. There are many international translations including in Croatian, Dutch, English, French, German, Hebrew, Hindi, Italian, Japanese, Russian, Shona, Spanish, Tamil and Thai.

The **WHOQoL-BREF** contains 26 items, 2 from the Overall Quality of Life and General Health, and 1 item from each of the remaining 24 facets included in the WHOQOL-100. The WHOQoL-BREF is scored over four major domains: physical, psychological, social relationships and environment.

The **EuroQoL (EQ)**-5D *(10)* is a measure of health status for use in evaluating health and health care. It provides a simple descriptive profile and generates a single index value for health status on which full health is assigned a value of 1 and death a value of 0. EQ-5D has been especially designed to complement other health status measures such as the SF-36, NHP, SIP or condition-specific measures. The EQ-5D covers the following five domains: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. By combining different levels from each domain, the EQ-5D defines a total of 243 health states. These may be converted to a score using sets of values derived from general population samples.

The EQ-5D has been translated into several languages including Afrikaans, Catalan, Croatian, Czech, Danish, Dutch, English, Finnish, French, German, Hungarian, Italian, Japanese, Norwegian, Polish, Portuguese, Spanish, Swedish and Turkish.

The **Health Utilities Index (HUI)** *(11,12)* is a system for measuring health status and health-related quality of life, and producing utility scores. It is generic, preference-scored and comprehensive, based on an explicit conceptual framework of health status and health-related quality of life. Applications of the HUI require that data be collected to classify the health status of each subject at a point in time, relying on self-completed or interviewer-administered instruments in either self-assessment or proxy-assessment formats. The HUI documentation includes a health status classification system and formula for calculating (single and multi-attribute) utility scores, which in the Mark 3 version of HUI define 972 000 unique health states (by comparison, Mark 2 describes 24 000 states) that are based on eight attributes (vision, hearing, speech, ambulation, dexterity, emotion, cognition and pain) with five to six levels each.

By the end of 2000, over 300 investigators had used HUI in a wide variety of studies in over 20 countries worldwide, and more than 200 000 subjects had been assessed using HUI.

The **Medical Outcome Study Short Form 36 (SF-36)** *(13)* comes from a larger battery of questions administered in the Medical Outcomes Study. The SF-36 includes eight multi-item scales containing 2 to 10 items each and a single item to assess health transition. The scales cover the dimensions of physical health, mental health, social functioning, role functioning, general health and vitality. Forms cover a week or a month. The use of subscales is encouraged and it can be self-administered or interviewer-administered. The SF-36 is the most widely used general health status instrument and has been translated in many languages.

The SF-36 allows scoring of the eight subscales and the construction of two summary scales: the physical component summary and the mental component summary scales. Further evaluation of these two summary scales provided the foundation for the construction of an instrument that is much shorter than the SF-36 *(14)*. This short form, the SF-12, uses 12 items from the SF-36, and demonstrates satisfactory reproducibility of the physical and mental component summary scales. The SF-12 is likely to perform well enough for monitoring general populations; it does not, however, allow for the scoring of individual SF-36 subscales such as bodily pain or social functioning.

Disease-specific instruments are useful for measuring clinically important changes in response to treatments. Since these instruments include elements most relevant to a particular disease, they are usually more sensitive to subtle improvements in health status. Disease-specific instruments are available for many different diseases and afflictions. There are principally two types of condition-specific measure: (i) measures that focus on clinical signs, symptoms and tests, and (ii) measures that capture the impact of the disease or problem on the patient.

A comprehensive collection of condition-specific health status measures can be found in the literature *(15–18)* or on the internet (http://www.leeds.ac.uk/nuffield/infoservices/UKCH/home.html, accessed 6 November 2002).

Measurement of satisfaction with health care

Over the past decade, the legitimacy of patients' satisfaction as an outcome measure of health care has grown considerably. In the main, patients' satisfaction is considered an indicator of quality of care. Moreover, satisfaction is used to assess the performance of health care delivery at multiple levels, e.g. health system, organizational (hospital, health maintenance organization, outpatient clinic, etc.), service unit (laboratory, radiology, etc.) and individual (physician, nurse, therapist, etc.). Empirical research on patients' satisfaction has demonstrated various problems, including: (i) a lack of conceptual or theoretical models of the determinants of patients' satisfaction, (ii) methodological challenges and a lack of standardized approaches to assessment, (iii) few studies that compare care across settings, and (iv) a lack of consensus within the medical profession and policy-makers about what role patients' satisfaction should play.

Patients' satisfaction can address multiple aspects of care. The dimensions that are most frequently assessed include:

- interpersonal aspects of care (the way patients feel about those caring for them)
- accessibility, availability and convenience of care
- continuity of care
- physical setting
- technical quality of care
- efficacy
- financial considerations (costs).

The results of evaluations of patients' satisfaction depend on the measurement method selected (e.g. surveys, interviews, focus groups). Contrasting opinions have been expressed regarding appropriate measures of patients' satisfaction. Although few measures exist, comparative analyses of different methods/instruments to validate these measures are rare.

Economic outcomes

Economic outcomes cannot easily be reduced to expenditure on health interventions at different levels of care provision. On the other hand, such figures are most frequently used since they are readily available from provider, payer or governmental sources. In the past, utilization of health services has been used as a proxy of health status. However, it is difficult to interpret as a measure of health because of differences in access to care and other factors related to the population's utilization of health services. Cultural and economic factors in the population of interest may further distort the relationship between health and utilization information.

Economic outcomes, whether they reflect the monetary value of a health consequence or integrated measures of health status and patients' preferences (e.g. quality-adjusted life years), are the centrepiece of different methods of economic evaluations. All methods of economic evaluation have in common that they examine one or more possible interventions and compare the resources necessary to carry out such interventions (input) with their consequences or effects (output). The various methods of economic evaluation – cost-minimization analysis, cost–effectiveness analysis, cost–utility analysis and cost–benefit analysis – differ in the way they itemize and value inputs and consequences. Although they all value the inputs and consequences following the same approach – (i) identifying inputs and consequences, (ii) measuring them using appropriate physical units, and (iii) valuing them – difficulties can be encountered throughout the three phases. Some health care interventions have hidden or unknown costs. Not all inputs and consequences can be evaluated in appropriate physical units (e.g. some interventions have intangible consequences, such as pain reduction or improvement in physical function).

Valuing inputs and consequences is the most difficult aspect of conducting economic evaluations. This is because the only readily available measures of value, prices, exist only in true markets, and these cover only a minority of health inputs and consequences.

Choosing outcome instruments

In general, only instruments with demonstrated psychometric properties which have been published should be used. The attributes of any quantitative measure are validity, reliability, responsiveness and practical usefulness.

Validity refers to whether an instrument measures what is supposed to measure. Ideally, a measure would be compared with a standard, for example, comparing a suspicious nodule on a chest X-ray with a biopsy showing cancer (criterion validity).

For health status no reference or standard exists to judge the validity of a particular instrument. Instead, an assessment is made of the extent to which a measure is consistent with a theoretical concept (construct) concerning the phenomenon of interest (construct validity). Face validity (it "looks like" it measures what it intends to measure) and content validity (it represents the domain of interest) are other techniques to strengthen the validity of a construct.

Reliability is the extent to which a measurement yields the same result on repeated administration of the questionnaire under the same circumstances (reproducibility). If scores of a health status instrument have little random error, they are considered reliable.

Validity and reliability are the minimal criteria to differentiate individuals at one point in time. However, when used to evaluate changes over time, an instrument needs to be able to capture clinically meaningful changes. Sensitivity denotes the capacity of a measure to show any change whether it is meaningful or not.

Responsiveness, on the other hand, is the capacity to show a change that is clinically meaningful to the patient and/or the physician. The **responsiveness** of a measure is the criterion which ultimately determines the usefulness of any outcome measure in the evaluation of chronic conditions, but it is the measurement criterion least established for health status instruments.

Finally, the **practical utility** of a health status instrument needs to be assessed for a given setting.

In practice and research applications, the time needed to complete a questionnaire should be no more than 10–15 minutes to ensure compliance. In general, self-administered questionnaires are more practical than instruments requiring a trained interviewer. However, in multicultural populations, or where literacy levels are variable, a standardized interview might be the only way to obtain reliable information.

Strictly speaking, the validity and reliability of an instrument are characteristics of the instrument for a specific population and should be re-evaluated for a new population. This may not always be possible, but at the least, individual items should be inspected carefully to evaluate face validity and to make sure that all relevant outcomes and potential adverse consequences are included. The scale should cover the range of severity and the magnitude of the changes expected. A small pilot on individuals who are representative of those to be studied can be extremely informative.

Exercise 2

Imagine that you are asked to evaluate the effects of a defined health policy intervention (e.g. the implementation of a general health examination at the age of 40 years covered by health insurance). Define relevant outcome measures on the patient, provider and policy-maker level that would allow you to judge whether such an intervention is beneficial or not. Try to implement an evaluation strategy that will allow you to make inferences on short- and long-term effects.

The “costs” and “benefits” of health outcomes assessment

One of the principal tasks of health policy-makers is to decide how to translate health expenditure into more benefits. Increasingly decision-makers ask that every additional expenditure be justified according to expected outcomes. On the other hand, most countries have barely applied explicit criteria to establish a standardized set of health outcome measures and a methodology of assessment. Outcomes assessment on a broader level using morbidity and mortality – although frequently available from national data sources – are often insensitive to measure the effects of certain health interventions, particularly of those that target improvements in physical, mental and social functioning. In contrast, measures capturing multiple domains of health, such as generic or condition-specific measures, require substantial resources when applied on a larger scale. This has definitely held back a broad dissemination of such

measures. Moreover, until now, there has not been much empirical evidence on the true benefit of using population health status information in health policy decision-making. However, this will change when more data become available. Several countries have introduced measures of health status (such as the SF-36 or SF-12) into national or regional population-based health surveys, and an increasing number of health care providers are implementing patient-centred outcome measures to monitor the health of their patients.

Most multidimensional health status measures allow different modes of analysis, for example assessments on the level of health-related domains or an aggregation to global scores that reflect physical or mental health. Aggregated measures, however, may hide the underlying reality (e.g. defined aspects of health which have been targeted by a particular action). Hence, the selection of an outcome measure should be based on a clear sense of what it is anticipated will be measured and why. Outcome measures can be both generic and specific to a given problem. The generic measures are useful for looking at policy issues or reflecting the bottom effects of care on health status or aspects of quality of life.

Assessment of health status and patient-centred outcomes – on an individual or on a population level – is a continuous effort. Some outcomes (for example satisfaction with care) are very sensitive to policy actions, while others require time to achieve substantial changes (such as physical functioning). Since most health status and quality of life measures reflect a defined "time window" (e.g. the past month), it is necessary to implement several measurements to monitor the effect of a particular intervention. This is truly impractical on a population level (e.g. employing large-scale surveys), but manageable on the provider level. The challenge would be to aggregate this information into large databases to make it useful to policy and decision-makers.

Once health status information is available, it is necessary to guide potential users in interpreting it. Stratified analysis should allow a closer look at particular groups of the population, and graphical displays may be a useful tool in monitoring longitudinal data and the achievement of predefined health status targets. Users should also be educated in the magnitude of changes in health status that might be achieved by certain actions. While a 5% improvement in physical functioning might sound small for the individual, it is a substantial change on the population level.

Implications

In order to assess outcomes in health care it is necessary to make a distinction between process-related non-health dimensions and health outcomes. To assess health status, a multidimensional assessment should be performed, the absolute health status must be distinguished from the relative health gain related to health care system activities, and the health of individuals likewise distinguished from the health of populations.

Generally, measurement instruments for both health and non-health outcomes should be standardized, reliable, valid and responsive.

Health status and quality of life outcomes should become commonplace in the measurement of benefits from health expenditure and in the assessment of the structure and process of health care delivery.

References

1. White, K.L. et al. The ecology of medical care. *New England journal of medicine*, **265**: 885–892 (1961).
2. Siegerist, H.E. *Medicine and human welfare*. New Haven, Yale University Press, 1941.
3. *Constitution of the World Health Organization*, Geneva, World Health Organization, 1985.
4. Murray, C.J.L. & Lopez, A.D. *The global burden of disease*. Boston, Harvard School of Public Health, 1996.
5. Bergner, M. et al. The Sickness Impact Profile: validation of a health status measure. *Medical care*, **14**: 57–67 (1976).
6. Kaplan, R.M. et al. Health status: types of validity for an index of well-being. *Health services research*, **1**: 478–507 (1976).
7. McDowell, I.M. et al. A method for self-assessment of disability before and after hip replacement operations. *British medical journal*, **2**: 857–859 (1978).
8. Hunt, S.M. et al. Measuring health status: a new tool for clinicians and epidemiologists. *Journal of the Royal College of General Practitioners*, **35**: 185–188 (1985).
9. WHO Quality of Life Group. The World Health Organization quality of life assessment (WHOQoL): development and general psychometric properties. *Social science & medicine*, **46**: 1569–1585 (1998).
10. Charro, F.T.D. & Rabin, R. EQ-5D from the EuroQol Group: an update. *Quality of life newsletter*, **22**: 3–4 (1999).
11. Feeny, D. et al. The Health Utilities Index: an update. *Quality of life newsletter*, **22**: 8–9 (1999) (reproduced in http://www.healthutilities.com/update.htm, accessed 2 December 2002).
12. Torrance, G.W. et al. Multi-attribute utility function for a comprehensive health status classification system: Health Utilities Index Mark 2. *Medical care,* **34**(7): 702–722 (1996).
13. Ware, J.E. Jr. & Sherbourne, C.D. The MOS 36-item short-form health survey (SF-36). A. conceptual framework and item selection. *Medical care*, **30**: 473–483 (1992).
14. Ware, J.E. et al. *SF-12: How to score the SF-12 Physical and Mental Health Summary Scales*. Boston, MA, The Health Institute, New England Medical Center, 1995.
15. Bowling, A. *Measuring health: a review of quality of life measurement scales.* Milton Keynes, Open University Press, 1991.
16. Bowling, A. *Measuring disease*. Milton Keynes, Open University Press, 1995.
17. McDowell, I.M. & Newell, C. *Measuring health: a guide to rating scales and questionnaires*. Oxford, Oxford University Press, 1996.
18. Spilker, B. *Quality of life and pharmacoeconomics in clinical trials*. Philadelphia, New York, Lippincott-Raven, 1996.

Further reading

Gold, M.R. et al. *Cost–effectiveness in health and medicine*. New York, Oxford University Press, 1996.

5.2.2 Costing[2]

Chris Selby Smith[3]

Key messages

- Costs refer to the benefits that are sacrificed elsewhere (foregone) when a given resource is used in the health care system. Resources have alternative uses: if they are not used in health care they could generate benefits elsewhere, for example in education, housing or the environment.
- Costs are wider than financial expenditure alone. Other resources especially need to be considered if their opportunity costs are not adequately reflected in market prices. Examples include voluntary contributions, the time costs involved for patients and their carers, and services contributed by religious orders.
- The total costs for a given health care activity should be compared with the benefits generated, to ensure that reallocation could not result in an increase in the total benefits achieved by society from the limited resources that are available. Thus, total costs must be estimated as accurately as possible.
- Total costs can be viewed from various perspectives. For example, different **types** of cost could be identified, such as the cost of staff compared to the cost of facilities, capital compared to recurrent costs, direct compared to indirect costs. Costs can be separated by reference to the **sources** from which they are met. They can also be distinguished by their **timing** and their level of **uncertainty**.
- Changing the distribution of the total costs for a given health care activity among the various parties to the complete resource allocation decision can alter the incentives they face and therefore the actions they take.
- The cost information that is, ideally, required by economists can be difficult to obtain. Three stages are often distinguished in costing studies: identification, measurement and valuation. Frequently the three stages are found to be progressively more difficult. As far as possible the approaches followed at each stage should be consistent in different studies, to facilitate comparisons and longer term learning.
- Cost information can be presented in ways which are more or less helpful for decision-makers. In general, cost information is not an end in itself but an aid to improved decision-making, better use of scarce resources in health care and improved outcomes.

Tutors' notes

This module provides material which is useful for:

- senior decision-makers in health care at national, regional or local level;
- managers of health care institutions;
- practitioners in the health care system;
- other decision-makers such as legislators; patients, their carers and families; and decision-makers in other sectors, which are competitive or complementary with health care in terms of resources used.

---

[2] Further references on costing matters are given at the end of the module. The book by Drummond et al. is particularly clear and useful (1). It has been drawn on extensively in preparing this module.

[3] This module was written by Professor Chris Selby Smith of Monash University, Melbourne, Australia (e-mail: Chris.SelbySmith@BusEco.monash.edu.au).

An improved *appreciation* of the material in this costing module is valuable for all participants in the health care system, especially those with important decision-making responsibilities, for whom *appraisal* is also particularly relevant. The ability to conduct costing *analysis* need not be so widespread, especially in terms of the technical detail. However, the broad way of thinking can have wide applicability.

The exercises focus on costing aspects at the level of: national or regional decision-makers (Exercise 1), managers of health care institutions (Exercise 2) and health care professionals (Exercise 3). Each exercise asks participants to consider how the resources that are required can be analysed, used effectively and, if possible, augmented, and what implications there are elsewhere, either inside or outside health care. Participants are asked to distinguish between situations where the available resources are expected to grow and those where they are likely to diminish.

It is suggested that tutors seek to focus the discussion on how the costing information can be used to inform and improve decision-making, e.g. to enhance effective choice, to ensure scarce resources are used as effectively as possible, and to throw light on the equity or inequity of the cost distribution; and the effect of the distribution and redistribution of costs on the incentives for stakeholders (such as governments, individual health care institutions, health professionals and patients) and their subsequent decisions (in health care and perhaps elsewhere).

Tutors are requested to collect relevant costing studies of which they become aware (at each level and for the variety of possible decision-making situations) and to provide them to WHO in Copenhagen. Over time, the additional case studies will be a valuable supplement to the existing material.

Introduction

Resources are scarce. As a result, it is not possible to produce all the outputs which would be thought desirable. Thus, choices have to be made in health care as in all areas of human activity. These choices are made on the basis of many criteria, some explicit and some implicit. Economic analysis seeks to identify and make explicit one set of criteria which are useful in deciding how to allocate the resources that are available among the various competing uses for them.

The cost of alternatives is relevant to much decision-making in health care. Policy-makers in the health sector at national, provincial/state or regional level frequently compare the costs (and the consequences) of alternative programmes. For example, they may compare preventive care with curative care, one way of treating a disease with another, or the extra cost of drugs compared with the reduced hospital costs attributable to them through shorter lengths of stay. Similarly, policy-makers at national, provincial/state or regional level will be comparing health programme costs (and consequences) with those elsewhere, say in education, transport or housing. The relative costs and consequences are also relevant for decision-making by the managers of health care institutions and (although probably to a lesser extent) by health practitioners.

Costs are relevant for each type of economic evaluation in health care, including cost-minimization, cost–effectiveness analysis, cost–benefit analysis and cost–utility analysis. In fact, costs tend to be relevant to most, if not all, health care choices. When costs vary widely they can be a powerful influence on the decisions that are taken. However, costs cannot be considered in isolation; they have to be weighed against the consequences arising from the different courses of action.

An important point to remember when embarking on a costing study is that, to an economist, cost refers to the sacrifice of benefits involved when given resources, say trained labour or financial resources, are consumed in activity A, say in a particular health care programme, rather than in activity B, say elsewhere in another health care programme or in an activity in another sector, such as education, housing or transport. The true cost or sacrifice being made is represented by what is foregone, what economists term the "opportunity cost". Therefore, in a costing study the analyst's attention should not be confined to financial expenditure alone. The activity may use other resources.

The consumption of some of these resources may not be adequately reflected in existing market prices. Examples include the time patients or their families have to take off work or spend travelling for a health care treatment, the time contributed by volunteers, the below market wage earned by some health care workers (such as those in religious orders) or clinic space which has been donated rather than purchased or leased.

Koopmanschap and Rutten *(2)* argue that indirect costs (e.g. production losses from ill health) can make up a substantial part of the costs of health care programmes, and propose a method of measuring them. They conclude that indirect costs tend to play an important role if health care programmes produce health care effects relatively quickly, if there is a considerable effect on (short-term) absence from work, and if a significant proportion of the target population is employed at the moment they benefit from the programme.

This module focuses on general matters. More detailed discussions are available in the items in the list of further reading at the end of the module. In many practical situations the options available to the analyst are limited by the availability of data. In the longer term it is worth considering the utility of improving the data, but in the short term the limitations of the data may seriously constrain the precision of the cost estimates.

It is also important to recognize that many costing issues can be context-specific. For example, the viewpoint adopted for the analysis can be significant. Thus, what patients spend on travel is a cost from the point of view of the patient (and society generally), but not from the perspective of the ministry for health (unless it offers reimbursement). Similarly, workers' compensation payments, which are a cost to the paying government and a gain to the patient (recipient), are a "transfer payment" and thus neither a cost nor a gain in society generally.

In some cases, certain costs are likely merely to confirm a result which would be obtained by consideration of a narrower range of costs. In such cases, it may not be worthwhile to complicate the analysis. However, such cost categories should be identified and some justification given for their exclusion (perhaps based on their small size, probable insignificant influence on the outcomes and the results of previous empirical work).

Total costs

To make a satisfactory estimate of total costs it is essential to know exactly what is to be costed. It is perhaps surprising that, in many costing studies, it is not possible to be clear about who is doing what, to whom, where and how often. Such clarity is a *sine qua non* for a satisfactory costing study. It also helps readers of a study to assess for themselves whether any relevant costs have been omitted. A similar point applies for consequences, but this is less relevant for the present module.

From the point of view of society as a whole, it is total costs for a given activity that are relevant. These are the total alternative opportunities that are being foregone to undertake this particular

programme. However, for individual participants in the health care activity only some costs may be directly relevant to them.

Total costs can be viewed from a variety of perspectives. This is considered further in the next section. However, one relatively simple and frequently used approach is to consider the total costs of a particular health care activity as consisting of three elements, as follows.

(i) **The costs of organizing and operating the programme, including dealing with any adverse events caused by the programme**. In some costing studies, identifying such costs involves little more than listing the resource items used in the health care activity. Variable costs should be included, such as the time of health professionals, supplies and the leasing of equipment. Overhead costs should also be included, including such items as light, heat, rent or capital costs. Valuing these elements tends to be more difficult for the capital costs than for the recurrent costs.

(ii) **Resources contributed by the patients or their families**. This would include the value of their time, for example while absent from work due to ill health or looking after family members when they are ill. These costs represent additional resources contributed to the treatment process. Patients may also make payments which cover some of the costs of organizing and operating the programme. To the extent that these contributions reduce the costs of (i) above, which were initially borne by other parties (such as the hospital or the government), patients are transferring costs between the various parties to the complete resource allocation decision rather than changing the total costs incurred by society generally.

(iii) **Resources consumed (or reduced use of resources) in other sectors as a result of the health care programme or activity**. Some health care programmes, such as those for the elderly, consume resources from other public agencies or the voluntary sector. Occasionally, it may happen that the operation of a health activity or programme changes the use of resources in the broader economy. If these factors are substantial they should be included in the economic analysis, although for many health care programmes they are likely to be insignificant.

Once the relevant range of costs has been identified for the particular health care activity or decision under consideration the individual items must be measured and valued. Here there are two elements: (i) measurement of the quantities of resources used, and (ii) the assignment of unit costs or prices.

The measurement of **quantities** of resources often depends on the context for the economic analysis. For example, if an economic study is being conducted alongside a clinical trial, data on the quantities of resources may be collected on an integrated base, for example through the case report forms. On the other hand, if the economic study is free-standing, the quantities of resources may be estimated by a review of case notes or from routine data systems. However, it may only be possible to estimate the quantities of some resources by developing special data collections, for example by asking patients directly or getting them to keep a diary.

Generally, market **prices** will be available for many of the items of resources. Where this is the case the costing process is simpler and more robust. Theoretically, the proper price for a resource is its opportunity cost, i.e. the value of the benefits foregone because the resource was not available for its best alternative use. However, the pragmatic approach to costing, which is normally adopted, is to use existing market prices unless there is some particular reason to do otherwise. For example, the prices of some resources may be subsidized by a third party, such as a charitable institution or foreign donor. Other resources may be provided by volunteers.

Although the costing of resource items is often relatively unambiguous, there are a number of issues which can arise in costing studies. Some of these are discussed in the section on identifying, measuring and valuing costs below.

Various perspectives

Total cost can be divided into its components in a number of different ways. For example, Creese & Parker *(3)* have made a useful analysis of costs for programme managers, including a classification of costs, in terms of capital and recurrent inputs, by reference to the activity or function for which the resources are used and by the source (or provider) of the resources. Creese & Parker define costs as the value of resources used to produce something, including a specific service or set of services as in a health programme. Complementary material is also available, such as the methodologies and worksheets developed by WHO to assist managers of HIV/AIDS programmes who want to use cost analyses to facilitate their own decision-making. More generally, WHO headquarters in Geneva is developing a template for costing across health programmes, to improve methodologies, facilitate comparisons and strengthen accumulative learning.

At a more specific level, costing can be considered in relation to particular aspects of the health care system. Abernethy *(4)* argues that an understanding of hospital costing systems and their strengths and limitations is essential if the information these systems can provide is to be used in economic analyses. For example, an understanding of cost behaviour is critical for the planning of new services or the expansion of existing services.

In relation to the development of clinical costing systems, say in hospitals, Abernethy identifies two fundamentally different approaches. One approach is based on "cost-modelling" principles and the other on "cost-finding" principles. The cost-modelling approach could, for example, be based on the Yale Cost Model, developed by a team of researchers from Yale University, where the cost object is a diagnosis-related group (DRG). This is a top-down approach to costing which uses general ledger data and predetermined allocation statistics to assign all hospital costs to a particular DRG. The final result of the cost allocation process is an average cost for a patient within the particular classification.

The cost-finding approach is based on the cost accounting principles that are used in the product costing systems developed and implemented in the manufacturing sector. This approach is often referred to as a patient-costing system, because the cost object is the patient. The approach is based on actual costing principles. Thus, the cost of a treated patient is determined by tracing the direct costs associated with the individual services received by the patient as well as an allocation of indirect patient and non-patient care expenditure (e.g. hospital overheads). This type of system adopts a bottom-up approach. The smallest cost object is the actual service, for example a laboratory test, received by a patient. The cost of this test can then be added to the costs of all other clinical services provided to a patient to arrive at a cost measure for the treated patient or other defined cost object.

In many costing systems in health care the absence of capital cost information is a significant limitation. Capital costs are particularly important in some types of service and in the treatment of certain types of patient.

The resources consumed by a particular programme in the health care system can also be considered in terms of sector, source of funds and type of cost. The total resources used in the programme are the same but they are viewed from different perspectives. In money terms, the quantities of each component are measured and the total cost is calculated by multiplying the quantities by the relevant prices for each component.

In reporting as well as calculating the costs it is important to show the units of each input and their prices separately (the "ingredients" approach). This facilitates comparisons, the generalization of results across settings, and the extrapolation of historical experience to new settings or new combinations of prices and quantities. Thus, work undertaken in one context can have maximum relevance for decision-makers in other contexts. The reservoir of knowledge *(5,6)* increases more rapidly, with benefits for other decision-makers and researchers.

Sectors

In terms of resource consumption by the sector which bears the costs, a threefold distinction can be made: the health care sector, participants in the programme and their families, and other sectors. This is the approach outlined above. Note that the net costs borne by a sector can be positive or negative, as when participants' true costs are more than compensated for, say by generous travel allowances or per diem payments. Resource consumption in the health care sector includes those costs borne by the sector, such as organizing and operating the programmes (ideally in terms of opportunities foregone). Note, however, that the costs to the health care sector include not only the initial programme but also any continuing costs associated with it. For many health activities there are substantial continuing costs.

Secondly, costs can be incurred by the participants in a health care programme (or their families). For example, participants may not be fully recompensed for the cost of participating (travel, attendance fees, accommodation, etc.). The leisure activities or work time of participants can be reduced, which in turn affects the valuation of the time contributed to the health care programme in terms of opportunities foregone. If participants were unemployed or underemployed the opportunities foregone tend to be less. Note that the opportunity cost of participation is not necessarily equal for each family member. For example, if in some countries or regions the opportunity cost of an absence from home of the mother of a family, say for an activity lasting a week or a fortnight, is substantially greater than for her husband or son, this is likely to affect programme participation by gender.

Thirdly, resources can be consumed in other sectors. For example, some health care programmes use resources in the education sector such as universities, technical institutes or research organizations. Many health care programmes also rely on resource inputs from the voluntary sector, with implications for such matters as costing, sustainability and continued availability for an expanding programme. To the extent that these resources are diverted from other worthwhile activities, opportunity costs will be incurred. These can be substantial.

Sources of funds

Four sources of funds are particularly relevant for health care activities in many less developed countries: national sources, whether public or private, including those from the health sector; WHO resources; other official sources of funds, such as the World Bank or similar regional bodies; and contributions from the voluntary and nongovernmental sector. Resources can be provided in cash or in kind. Since the sources of funds are additive, the total cost of a programme is the sum of the national contribution and the contributions from other sources.

There is a tendency for particular stakeholders to view costs in terms of what it costs them. Thus a national health authority which funds two thirds of one programme but only a third of another, whose total costs are equal, is likely to favour the latter even if the former programme has a higher ratio (up to twice as high) of overall benefits to overall costs (i.e. is a much more desirable programme on the basis of overall economic evaluation). To the national health authority one programme can look

less costly than the other, although in truth their total resource costs, including the costs borne by all parties, are equal. Other stakeholders are likely to act similarly.

From the donors' point of view, varying the levels of subsidy for particular programmes can provide differential incentives for national health authorities to act in particular ways. This would be a method of achieving maximum health care outcomes for the donors from their limited resource budgets, but it requires information on costs and consequences which is often not available. National health authorities can act in analogous ways to encourage particular activities by subordinate levels of government or by the private sector.

Another classification of health care costs, overall or by reference to particular programmes, focuses more on the internal sources of resources. For example, a distinction might be made in federal systems between resources contributed by the national authorities and the resources contributed by state, provincial and regional governments. Another distinction which is often drawn is between costs borne by the public sector and those borne by the private sector. In most advanced countries the proportion of total health care costs borne by the public authorities has been rising, but in some countries the private contribution is still as high as a quarter to a third. A further distinction is sometimes made within the private contribution to health care expenditure, between contributions through health insurance arrangements and contributions made directly to providers by patients or their families at the point of service.

Types of cost

Here, a major requirement is consistency, for example between the classification of costs in different programmes, in different countries (or regions) and for the various sources of funds. Comparisons tend to be unreliable if the data are not consistent. Consistency in the data is likely to be more readily achieved prospectively than retrospectively. It is important that double counting of costs does not occur. In the real world, it may be necessary to ascertain what classifications are most widely available and, so far as good practice permits, build on them. Note that some of these costs are represented by financial expenditure whereas others are opportunity costs. Ideally, the financial estimates would approximate closely to the true opportunity costs of using the resources for these activities. In practice, this is not always so.

Against this background a classification of cost headings could include such items as the professional staff in continuing employment who provide the health care activity. This would include the wages, salaries and on-costs[4] for the health care workforce – doctors, nurses, dentists and allied health professionals. The health care activity would represent the entire workload of some staff but only a part of it for others. In the latter case an estimate is required of the appropriate proportion of their salaries and on-costs which should be allocated to the particular health care activity which is being costed (including any follow-up activities). In many cases estimating the appropriate proportion is likely to involve judgement and approximations. In some cases the particular health care activity being costed will involve additional costs for the participation by members of the health workforce, such as for daily allowances, food or accommodation.

Secondly, the costs would also include the wages, salaries and on-costs of those professional staff who are temporarily employed to undertake the necessary work for the health care activity being costed. They may be employed solely for this health care activity, in which case costing is facilitated because there is no need to apportion their costs between more than one health care programme. The

[4] The costs of a person while engaged in a health care activity other than salaries/wages.

intention is to estimate the full cost attributable to the programme, again including any follow-up activities involved.

A third category of costs relates to facilities, equipment and materials. When there are dedicated items which are used solely for the particular health care activity being costed their total annual costs would be included. This may often be the case for materials. However, in the frequent cases where, for example, facilities and equipment are shared between health care activities or programmes the total costs would need to be apportioned between them.

A number of methods have been proposed to allocate such shared costs between health care programmes, such as direct allocation, step down allocation, step down with iterations or simultaneous allocation *(1, pp. 62–66 and the worked example at pp. 74–81)*. Judgement and approximations tend to be required. Often there is no unambiguously correct way to apportion such common costs between different users or health care programmes. Frequently, facilities and equipment will be used for a number of years, in which case the overall costs also need to be allocated over time. The cost of facilities, equipment and materials includes the recurrent costs of using them for the health care programme, say over an annual period, and a portion of the capital cost which equates to their use over that period for that health care programme.

A fourth element of cost relates to administration. This includes the wages, salaries and on-costs of administrative staff, including support staff, cleaners, cooks, janitors and those who work in the central administrative services such as finance and budgeting, human resources, planning, public relations and information technology support. There are also related costs, such as their accommodation, equipment and consumables. In theory, these overhead costs should be allocated, in an appropriate proportion, to the individual health care activity being costed. In practice, this often proves difficult and rough approximations are adopted for estimating these costs by health care programme.

Some implications

First, if the costs of a particular health care intervention (the point applies also to consequences) have been evaluated in a setting that is technically inefficient, while the costs of another intervention have been estimated in a setting that is technically efficient, conclusions on relative costs (and relative efficiency and their relations) can be biased. The confounding effect of variations in technical efficiency across study locations for the development of generalized comparisons of costs and consequences from alternative health care interventions, for example on cost–effectiveness league tables, needs to be minimized. At the same time, systematic variations in technical efficiency (due to such factors as health system characteristics or epidemiological patterns) should be incorporated.

Second, there can be stronger linkages between certain types of cost and certain sources of funds or sectors than others. It is worth looking out for such linkages when undertaking cost studies. For example, it may be that national or regional health authorities are more likely to finance local support services, participants are more likely to fund the opportunity costs of participating in health care activities, and external donors are more likely to finance visiting experts. If particular stakeholders focus on some costs rather than others, there can be a danger that no party is primarily concerned with the overall balance of costs (and benefits) for the health care programme or between programmes. Yet that is essential for achieving the optimal allocation of resources. It is a particularly serious problem if the national or regional health authority adopts a partial view, for example, if it focuses only on the costs it bears. Improved information on overall costs and their distribution among all the parties to the complete resource allocation decision tends to act as a counterweight to any partial viewpoints which may be fostered by prevailing arrangements.

Third, there is a challenging managerial and coordinating role to bring the disparate cost components together in a coherent whole so as to facilitate informed decision-making by national health authorities. Consideration of costs in terms of the different types of cost, the different sources from which the costs are defrayed, and the different sectors which bear the costs, raises implications for managerial decisions in the health care sector about alternative ways of providing health care in general or an individual programme in particular. Consider the following.

- At what scale should the programme be provided? Some costs, such as the initial development costs, may increase relatively little with growing participation in the programme. If the fixed costs are large, the average costs can fall sharply as participation increases. Other costs, such as the opportunity costs of participation, may fall relatively little. If the total costs are mainly fixed costs, average costs decrease markedly, whereas if total costs are mainly variable costs, average costs do not decrease much at all as the health care programme increases in size.
- What should be the breadth of geographical participation? Knowledge of the cost function can be most helpful in informing managerial decisions. For example, there may be a decreasing average cost of providing the health care activity as participation increases, but each participant may incur extra costs, for example for travel, accommodation and absence from employment. If the public authorities (or a donor organization) meet the cost of providing the programme, but participants have to bear the costs of attendance, equitable access (including geographical participation) is likely to be limited.
- What should be the balance between the capital and recurrent costs when providing particular health care programmes? Greater use of distance approaches or electronic participation increases some costs, including capital costs for infrastructure. However, such approaches could sharply reduce travel costs for (at least some) attendees or the recurrent costs of participants who previously had to be absent during working hours (assuming they had access to the necessary facilities). Capital costs may be paid from a different budget than recurrent costs. Knowledge of the cost function facilitates appropriate decisions by managers about the substitution possibilities between capital and recurrent costs.

Fourth, various equity aspects of health care programmes can be highlighted through consideration of resource use in terms of the different types of cost, the different sources from which costs can be defrayed, and the different sectors which bear them. Analysis of the costs of different programmes, or programmes provided in different ways or different places, can reveal interesting patterns. Are the most expensive or most heavily subsidized programmes accessed equally by the rich and poor, the powerful and the powerless, women and men, young and old? Do programmes involving an absence from home and workplace enable men and women to participate equally? Do programmes provided face-to-face in a particular location (say the capital city) enable non-metropolitan staff in that country or relevant health care workers in other countries to participate as much as programmes which are delivered by distance approaches or electronically? The total costs and their distribution, while not all the information required, often throw light on some of the reasons why there is unequal access to health care programmes and how cost aspects contribute to inequitable outcomes.

Identifying, measuring and valuing costs

The cost information required for the optimal allocation of health care resources is often difficult to obtain. Three stages can be usefully distinguished in costing studies: identification, measurement and valuation. While these occur simultaneously in some economic analyses, it is good practice to consider each as a separate phase of the analysis (and for users to evaluate the results in this way). The three stages often tend, in practice, to be progressively more difficult.

Identification

Even though it may not be possible to measure and value all of the costs and consequences of the health care programme under consideration, a full identification of the important and relevant ones should be undertaken. Identification consists of listing the likely resource implications of the health care intervention or activity as comprehensively as possible. Decisions can then be made about which effects should be included and which might reasonably be excluded. The perspective which is adopted for the study influences these decisions (e.g. society in general, particular groups or individuals who are affected, or other stakeholders) but, as mentioned previously, the perspective (at least) of overall society should generally be considered.

Measurement

Once the important and relevant costs have been identified, they must be measured in appropriate physical and natural units. For example, "measurement of the operating costs of a particular screening programme may yield a partial list of ingredients such as 500 physical examinations performed by physicians, 10 weeks of salaried nursing time, 10 weeks of a 1000 square foot clinic, 20 hours of medical research librarian time from an adjoining hospital, etc. Similarly, costs borne by patients may be measured, for instance, by the amount of medication purchased or the number of times travel was required for treatment, or the time lost from work while being treated." *(1)*.

Situations commonly arise where resources are used jointly by one or more programmes. These situations present a particular challenge. For example, in every hospital and in many other health care activities, numerous clinical services and programmes share centrally provided common overhead services, such as cleaning, administration, heating, light and power.

In general, there is no non-arbitrary solution to the measurement problem. "Costing in a multi-product firm is a difficult task, particularly when the final product is an amalgam of many intermediate products. There is no easy solution." *(4)*, especially where a large proportion of costs, as in many hospitals, is both indirect and fixed. However, users of results should satisfy themselves that reasonable criteria, such as number of square feet, number of employees or number of patients, have been used to distribute the common costs. Sensitivity analyses can also be helpful in showing whether (and where) different assumptions would make a significant difference to the cost estimates. Users should definitely ascertain that the shared costs have, in fact, been allocated to the participating health care activities or programmes.

The form in which the measured costs are reported also justifies attention. When making certain decisions it is helpful to know the distribution of costs, as well as their mean and median. There can be situations where the outliers are of particular interest. Aggregation of costs may result in more accurate costings for subgroups of interest being lost in the overall results.

Valuation

The sources and methods of valuation of the costs should be clearly stated. Costs are normally valued in units of local currency, based on prevailing prices of, for example, personnel, commodities and services. They can often be taken directly from programme budgets. All current and future programme costs are normally valued in constant dollars of some base year (usually the current year) in order to remove the effects of inflation.

The objective in valuing costs is to obtain an estimate of the opportunities foregone by using the resources in the particular health care activity rather than elsewhere. This may necessitate adjustments

to some apparent programme costs. For example, this would be the case for donated facilities, subsidized services or labour contributed by volunteers that are received by one programme but not another. It is also relevant where such services would not be available in larger quantities if the programme was to be expanded.

Valuing the cost of institutional care for a specific condition can present particular difficulties. The use of an average cost per day, calculated on the basis of the institution's entire annual case-load, is almost certainly an over- or underestimate of the actual cost for any specific condition. The difference can sometimes be quite large. For example, accommodating an extra patient in a hospital bed which would otherwise have been unused generally results in extra costs that are well below the average cost for the whole institution.

Top-down or bottom-up costing

In principle, and with great effort in practice, it is possible to identify, measure and value each depleted resource, such as drugs, nursing time, lighting and food, in treating a specific patient or group of patients. This yields a relatively accurate cost estimate but the detailed monitoring and data collection can be expensive and time-consuming. Another broad costing strategy is to start with the institution's total costs for a particular period and then to improve upon the method of simply dividing by the total number of patient-days to produce an average cost per day. Quite sophisticated methods of cost allocation to individual hospital departments or wards can be used. An intermediate method is to accept the components of the general average cost for hotel functions (since these are relatively invariant across patients) and to combine this with more precise calculations of the medical treatment costs which are associated with the specific patients in question. Of course, the effort devoted to accurate estimates depends upon their overall importance in the study. However, in general, unthinking use of average costs should be avoided.

Non-market items

There are particular problems when values are imputed for non-market items, and also when it is judged that existing market prices should be adjusted. In relation to the former, the major non-market resource inputs to health care programmes tend to be contributed services or facilities, volunteers' time, and patients/families' leisure time. One approach to their valuation is to use market rates. For example, unskilled wage rates might be used for valuing the time contributed by volunteers, on the basis of a value of lost leisure time of anything from zero through average earnings to average overtime earnings (on the grounds that this is the price that an employer must pay, at the margin, to buy some leisure time from the worker). A common practice is to value lost leisure time at zero in the primary analysis and then investigate the impact of other assumptions through sensitivity analysis. A slightly different approach is to identify and measure units of, say, voluntary input and document these when reporting results. The decision-maker can identify those programmes that rely heavily on volunteers and make appropriate adjustments if these are judged necessary.

In relation to the need to adjust market prices, it has long been recognized that, owing to imperfections, market prices may not reflect opportunity costs. For example, hospital charges will not equal costs if the cost of one activity is subsidized by another; and physicians' fees may not reflect the relative skill level for different procedures. Drummond et al. *(1)* suggest that before analysts adjust market prices, they should be convinced that two conditions are satisfied: first, that leaving prices unadjusted would introduce substantial biases into the study; and second, that there is a clear and

objective way of making the adjustments. These issues have been explored most extensively in the context of hospital charges in the USA *(7)*.

Note that if the economic study is being undertaken from the viewpoint of a third party payer, the actual charges may be more relevant than the costs. Yet even this approach is not necessarily clear cut, for instance when the third party does not pay the full amount billed. And from the point of view of society, the total opportunity costs continue to be relevant.

Two other matters are worth noting. First, the boundaries of a health care activity or programme, and thus its costs, are not always easy to define precisely. They may require careful consideration. For example, many health care programmes involve not only initial expenditure but also updating and refresher activities, which require further uses of scarce resources. Health care programmes may require changes in cooperating factors, such as operating procedures, work organization, technology and management (even changes in other sectors which are complementary to the health care activity), if the full benefit is to be obtained from the programme in terms of health care outcomes. In addition, the human capital created by many health care activities often yields benefits over a long period of time.

Second, how the resources to defray the costs incurred for a particular health care programme are raised can have implications for how they are spent. Both the aggregate level of health care expenditure and how it is distributed can be affected by how the required revenue is raised. For example, if health care consumers have to meet costs from their own pocket at the point of service they are likely to act differently (e.g. seek less services) than if the costs are defrayed from public revenues raised through taxation. Total health care costs will tend to be lower in the former case, and the distribution of total costs between the various parties to the complete resource allocation decision will also be different.

Handling capital costs, average and marginal costs, timing and uncertainty, and incentives inherent in the distribution of costs

Handling capital costs

Costs include capital as well as recurrent costs. Recurrent costs are included in the estimates of resource use for the time period in which they were used, e.g. the current year. But how should capital outlays be handled, such as those for land, buildings, equipment or other long-lasting assets? The flows of resources from these capital assets being used for the process of providing the health care activity should be added to the relevant recurrent costs, not the entire stock of the capital assets (which can be expected to continue to produce flows of useful resources for productive health care activities over a number of years). “Capital costs represent an investment in an asset which is used over time” *(1)*: in time, they wear out or depreciate (although land may not depreciate). Frequently, capital costs are not listed in the accounts or budgets of the organization (especially public organizations) because they have been funded in advance, sometimes by a one-off grant.

Capital costs can be thought of as comprising two components. The first component is the opportunity cost of the resources represented by the capital asset. For example, the land used for a hospital could have been used for something else, perhaps a school, public housing or a park. The cost is the foregone opportunity to use the resources in some other activity which would yield positive outcomes. This value is usually calculated by applying an appropriate interest rate to the amount of

resources invested. The second component of the capital cost represents the depreciation over time of the asset itself. Various accounting procedures, such as straight line or declining balance, can be used in the accounts of the organization. However, accounting practices may relate more to company tax laws governing the depreciation of assets than to the real change in the value of the asset. Note that if capital outlays relate to resources that are used by more than one programme, they require to be allocated between them.

Average costs and marginal costs

Economists draw an important distinction between **average costs** per unit of output in a given health care activity and **marginal costs**, which are the extra costs of producing one more (or less) unit of output. Where average costs are falling, for example where there are high fixed costs for the programme and participation is increasing, marginal costs are below average costs, sometimes well below. Conversely, if average costs are rising, for example in a poorly managed health care facility, marginal costs will be higher than average costs as participation in the health care activity increases. To remain viable a health care programme has to cover its average costs over a reasonable period of time, but in terms of small changes in the level of activity it is the marginal costs that should be related to the extra benefits likely to be received. It is desirable to record capacity utilization, so that the cost figures can be interpreted appropriately (and remedial action taken, if required).

When making a comparison between two or more health care activities (or a health care activity and an activity in another sector of the economy) it is worth asking what would be the extra costs (and consequences) of having a little more or a little less. In practice, it is important to recognize the difference between average and marginal costs (and to appreciate its potential significance for decision-making), although it is often the case that the issue can only really be explored in the context of specific locations or situations. For example, the extent to which costs can be saved when the average hospital stay for patients is shortened depends on the flexibility which is available locally and the time period over which the change is made. Freed resources will not always be redeployed efficiently. For example, if unemployment is high, reduced labour requirements may well not result in higher output elsewhere. Analysts have a responsibility to point this out explicitly where it is significant for the cost estimates and has implications for decision-making.

Timing and uncertainty

The cost estimates may need to be adjusted for timing and for uncertainty. This is not a problem if all the costs are incurred immediately, but often costs are spread over a period of time. In particular, alternative uses of resources may have different time profiles of costs (and consequences). The time profile of costs and benefits can also differ within a single health care programme. Future cost streams are reduced or "discounted" to reflect the fact that the resources spent or saved in the future should not be weighted as heavily in programme decisions as resources spent today. This is primarily due to the existence of what economists call time preference. Individuals and societies prefer to incur costs later (and receive benefits sooner) because they can benefit from them in the meantime.

Time preference arises for various reasons: individuals (and to a lesser extent societies) may have a short-term view of life: the future is uncertain – with positive economic growth, individuals and societies expect to be more wealthy in the future, and since most individuals appear to have a positive rate of time preference, a positive return can usually be obtained when making a riskless investment. Note that the notion of wanting to postpone costs (or preferring benefits today) extends beyond money transactions to goods and services that cannot easily be traded.

Also, life contains uncertainty and imprecision, and so do cost estimates. Careful analysis can help to identify critical methodological assumptions or areas of uncertainty. Analysts often attempt to rework their analyses, employing different assumptions or estimates to test the sensitivity of their results and conclusions. If large variations in the assumptions do not produce significant alterations in the cost results, they tend to have more confidence in the original estimates. If the converse occurs, more effort is required to reduce the uncertainty and improve the accuracy of the variables that are particularly significant for the cost estimates. Such sensitivity analyses are an important element of a sound costing exercise. Where the cost data are stochastic (i.e. have a mean and variance), rather than the point estimates which have often been used in the past, tests of statistical significance can be performed or confidence intervals presented.

Incentives inherent in the distribution of costs

The distribution of total cost among the various parties to the complete resource allocation decision, that is the costs they bear, can have important incentive effects and affect the actions of decision-makers. To take a simple example, a particular health care activity may clearly be a worthwhile use of scarce resources when viewed in terms of its total costs in comparison with its total benefits. However, if the costs are all borne by one party and the benefits are all received by another party, the former will be unenthusiastic. If they are powerful, either economically in a market situation or politically in a public-provision situation, they may well block the effective implementation of the programme. Of course, many real life situations are not as blatant as this, and reallocation of the costs (and benefits) can promote a less confrontational situation. Nevertheless, the distribution of total costs generates a particular pattern of incentives for participants which is likely to influence their decisions (such as whether to participate in the programme). Participants may react to incentives which exist but were not consciously intended, or to incentives resulting from cost distribution patterns which have been proactively designed to encourage particular actions, such as participation in screening programmes. Knowledge of the distribution of costs, as well as of their total size and composition, is necessary for effective health care decision-making, whether this is aiming at achieving an efficient use of scarce resources or equitable outcomes.

Presentation of cost information

The **purpose** of collecting, refining and analysing the cost estimates for different health care programmes, or for the health sector as a whole, is to improve decision-making. Cost information is not the only input required for decision-making in health care, but high quality decision-making without adequate cost information is generally impossible.

There are many purposes for which cost information is relevant, and the particular cost information required can only be determined in the light of the specific objectives of the decision-maker. Abernethy notes that "costing information may be required to determine the cost of a particular patient for cost reimbursement purposes, or for comparing costs of different diagnosis-related groups, or for determining the cost of a laboratory test to establish a price. Alternatively, the management of a hospital may wish to use costing information for developing a clinical budget and subsequently to monitor performance. These decisions require different types of cost" *(4)*. No decision-makers are likely to achieve their goals unless they are clear about what they want to achieve and carefully marshal the means (including cost information) by which they can reasonably expect to get there.

Different decision-makers will, perfectly legitimately, have rather different objectives, or a rather different balance between different objectives. For example, politicians or bureaucrats in central agencies at national level are likely to be concerned with the appropriate balance between the overall public and private sectors, and between health care and other competing uses for resources, while politicians and bureaucrats in the health sector at state, provincial and regional level are likely to give a greater priority to health care, their geographical area, and perhaps public sector activities. Those decision-makers responsible for particular health care institutions, such as hospitals, facilities for the care of the aged or community health centres, are likely to focus more on obtaining what they regard as their fair share of resources and using them effectively to produce efficient and equitable health care outcomes. Individual practitioners and the consumers of health care services are likely to focus more on their specific concerns within the overall structures and incentive patterns established by higher level decision-makers. Cost information is relevant to the decisions of all these participants in the health care system. However, their focus differs as does the cost information they want, when they get it and how they use it.

The **way** in which cost information is collected, analysed and presented can make it easier or more difficult for decision-makers to use it effectively. "The importance of identifying the limitations of the costing information and the potential effects on the results of economic analysis cannot be overstated" *(4)*. For example, it has been increasingly recognized that disaggregated information is required by decision-makers, as well as the final summary result. They need to know how the information was obtained, what assumptions had to be made, what approximations were employed, how sensitive the cost results were to different (plausible) assumptions or approaches and so on, in order to know how much reliance to put on the final cost estimates and their components. Writers of economic studies are now encouraged to report prices and quantities separately, rather than expenditure totals alone. Similarly, if costing studies are not publicly available they are not likely to be subject to the same level of critical scrutiny. In the long run this militates against improvements in technique, the growth of shared knowledge and accumulative learning, all of which contribute to growth in the reservoir of knowledge available for future users, whether policy-makers, practitioners or researchers *(6)*.

Thirdly, it can be helpful to decision-makers – depending on precisely which questions they are concerned with – to have cost information not only about total costs, but also about their composition and distribution. It is necessary to know the magnitude of **total costs**, for example, when decision-makers in central agencies, the government or parliament are considering whether to use scarce resources in the health care sector or elsewhere, or comparing alternative uses for resources within the health care sector. When other things, especially the expected benefits, are similar, activities that are less demanding in their use of resources will be preferred over more expensive activities. The **composition** of costs, for example between different types of expenditure or between different sources, can also influence decision-making. For example, if an existing activity is heavily dependent on contributed facilities or on voluntary labour, it may not be possible to expand its scale without a significant increase in costs. Or if foreign donors provide many more resources for some health care activities than for others, the former will tend to be preferred over the latter by domestic decision-makers (assuming other things are equal). The **distribution** of costs can influence the actions of individual parties to the complete resource allocation decision by changing the incentives they face to participate or withdraw. If the existing pattern of incentives resulting from the current distribution of costs (and consequences) is judged to be undesirable from the point of view of efficiency or equity, it may be possible to reallocate the costs (even if the total costs to society are unaltered) to achieve more satisfactory outcomes.

Finally, the linkages between costing studies and health care decision-making can often be improved. Decision-making in health care is a complex process, involving many actors, and costing information is only one input into their decisions (often it is not the most important input). Research and investigation, including for costing purposes, is a domain which has many differences from decision-making, whether at policy-making or practitioner level. Sometimes decision-makers are concerned about the quality of costing studies, including the large number of assumptions that are made. Health care practitioners, who usually have some knowledge of biomedical research, tend to be more comfortable with evidence based on randomized controlled trials rather than on modelling studies. Thus, the type of costing study can affect where it has an audience and how it might be used in decision-making. Sometimes studies, particularly those conducted at a distance, do not adequately reflect the concerns of local decision-makers. For example, many economic studies in health care do not examine the costs of implementing the preferred course of action. Other studies appear to assume that savings, such as in reduced hospitalization costs, can be realized easily. However, from a local decision-making perspective these can both be significant issues. It is important that decision-makers have adequate access to costing information and that the results are communicated in a way that busy decision-makers can readily understand.

Research, including improved knowledge about costing matters, is cumulative. Much research does not stand on its own as isolated work but adds to that which existed before. This accumulating body of knowledge (it can also be diminished or unavailable) can contribute to the improvement of particular techniques or methodologies as well as to the creation of a climate of opinion and the development of a set of ideas, so that at any given time certain ideas, approaches or ways of thinking are "in" while others are not. The outputs of research and development, including costing studies, also include human capital, such as research skills and attitudes and trained personnel.

The potential for different emphases by researchers and decision-makers underlines the importance of linkages between them, the desirability of two-way flows of information throughout the costing study (not solely when it is completed) and the benefits from collaboration, based on diverse contributions including costing expertise, to improve health care processes and outcomes. An emphasis on linkages increases the mutual responsibilities of the parties. Enduring linkages, based on sustained mutual esteem and an understanding of the potential contribution of each party, are critical to enable costing studies and approaches to contribute effectively to high quality decision-making by policy-makers and practitioners.

Exercise 1

You are a senior decision-maker at national level in the health ministry. You are seeking:

a. to obtain as many resources for the health sector as you can in competition with potentially valuable uses of the resources in other sectors, such as education, transport, defence and public order;
b. to ensure that the resources provided for the health sector are used in the most efficient, effective, economical and equitable fashion.

What cost information would you use for each purpose, if it was available? If it was not available, to what extent would you seek to develop it?

Exercise 2

You are a manager of health services, say a hospital, facility for care of the aged or community health centre. How would you use the information available to you on costs to combine resources and produce health care outputs in the most efficient and equitable way?

What other information on costs would be useful? To what extent would the extra cost of generating it be justified?

To what extent would your use of costing information encourage continuing improvements in the future?

Do actions taken by you either increase or decrease the costs borne elsewhere in the health care system (including by patients or their carers)?

Finally, if the resources available to you are shrinking, would you use cost information in the same way as when they are growing? If not, what would be the differences and why would they occur?

Exercise 3

You are a practising health care professional, say a doctor, nurse, pharmacist or dentist. How can you use the resources which are available to you to produce the best outcomes for your patients?

To what extent can you reallocate the resources to which you have access? In what ways would reallocation improve the outcomes for your patients (and any other stakeholders)?

Are there any ways in which the available resources could be augmented by you? Are there constraints on your practice which increase costs, especially if they do so without improving outcomes?

Do actions taken by you either increase or decrease the costs borne elsewhere in the health care system (including by patients or their carers)? Do you take these factors into account when making your decisions?

Finally, if the resources available to you are shrinking rather than increasing, how would you decide which activities to reduce and to what extent? How would your approach differ from the approach you would follow when resources are increasing (and why)?

References

1. Drummond, M.F. et al. *Methods for the economic evaluation of health care programmes,* 2nd ed. Oxford, Oxford University Press, 1997.
2. Koopmanschap, M.A. & Rutten, F.F.H. Indirect costs in economic studies: confronting the confusion. *In*: Selby Smith, C. & Drummond, M.F., ed. *Economic evaluation in Australian health care.* Canberra, Australian Government Publishing Service, 1995, pp. 58–67.
3. Creese, A. & Parker, D. *Cost analysis in primary health care.* Geneva, World Health Organization, 1994.
4. Abernethy, M.A. Costing systems in hospitals: relevance for economic evaluation studies, *In*: Selby Smith, C. & Drummond, M.F., ed. *Economic evaluation in Australian health care.* Canberra, Australian Government Publishing Service, 1995, pp. 44–57.
5. Buxton, M.J. & Hanney, S. *Assessing payback from Department of Health research and development.* Brunel University, Uxbridge, Health Economics Research Group, 1994, Vol. 1.
6. Buxton, M.J. & Hanney, S. *Assessing payback from Department of Health research and development. Second Report*. Uxbridge, Brunel University, Health Economics Research Group, 1997 (2 volumes).
7. Cohen, D.J. et al. Economics of elective coronary revascularization: comparison of costs and charges for conventional angioplasty, directional atherectomy, stenting and bypass surgery. *Journal of the American College of Cardiologists*, **22**(4): 1052–1059 (1993).

Further reading

Boyle, M.H. et al. Economic evaluation of neo-natal intensive care of very-low-birth-weight infants. *New England journal of medicine*, **308**: 1330–1337 (1983).

Gold, M.R. et al, ed. *Cost–effectiveness in health and medicine.* Oxford, Oxford University Press, 1996.

Horgren, C.T. *Cost accounting: a managerial emphasis,* 5th ed. Englewood Cliffs, NJ, Prentice Hall, 1994.

Krahn, M. & Gafni, A. Discounting in the economic evaluation of health care interventions. *Medical care,* **3**: 403–418 (1993).

Luce, B.R. & Elixhauser, A. Estimating costs in economic evaluation of medical technologies. *International journal of technology assessment in health care*, **6**: 57–75 (1990).

Ramsey, R.H. Activity-based costing for hospitals. *Hospital and health services administration*, **39**: 385–396 (1994).

Selby Smith, C. & Drummond, M.F., ed. *Economic evaluation in Australian health care*. Canberra: Australian Government Publishing Service, 1995.

Selby Smith, C. & Drummond, M.F. R&D's impact on decision-making by practitioners and managers: intersectoral and international perspectives. *In*: Baldry, J. ed. *Economics and health: 1999,* Sydney, University of New South Wales, School of Health Services Management, 2000.

5.3 Evaluation

5.3.1 Economic evaluation

Michael Drummond[5]

Key messages

- There are a number of criteria that might be used in evaluating health care treatments and programmes.
- Economic evaluation is concerned with assessing efficiency or value for money.
- There are many forms of economic evaluation (e.g. cost–effectiveness analysis, cost-benefit analysis) but all compare the costs and consequences of treatments and programmes.
- It is important to understand the key methodological principles of economic evaluation, including the consideration of an adequate range of alternatives, the use of good evidence of effectiveness and allowance for uncertainty in estimates.
- Economic evaluation can be used in association with a range of policies to encourage the rational diffusion and use of health technologies, including planning specialist facilities, reforming payment schemes for institutions or health care professionals, and developing health care practice guidelines.
- Economic evaluation can be used to assess health-producing measures in different sectors of the economy, including road safety, environmental protection and occupational health.

Tutors' notes

A wide range of groups within the health care system would benefit from understanding more about economic evaluation.

The core exercise in this module (Exercise 1) is aimed at the level of (critical) *appraisal* of economic evaluation methodology. It can be used with the following groups:

- civil servants and other governmental technical staff
- health service managers
- health care professionals (e.g. doctors, nurses).

---

[5] This module was prepared by Professor Michael Drummond, Centre for Health Economics, University of York, York, United Kingdom (e-mail: chedir@york.ac.uk).

Particular benefit can be achieved by running the exercise with a multi-professional group.

The exercise requires the group to have access to a published study relevant to their own setting. The choice of the study is left to the tutor, but as a fallback the paper by Mark et al. *(1)* can be used, as a worked answer is provided in the book by Drummond et al. *(2)*.

The second part of the module (including Exercise 2) is aimed at the level of *appreciation* and can be used with the following groups:

- policy-makers (e.g. elected officials)
- civil servants and other governmental technical staff
- health service managers.

It may also be used with health care professionals, but would probably need to be structured around a specific issue, such as the purchase of a piece of equipment, the listing of a drug on the formulary, or the development of a health care practice guideline.

The third part of the module will be of most interest to individuals with broader responsibilities for health and health care. It is aimed at the level of *appreciation* and could be used with policy-makers and civil servants. It could also be of interest to such officials in finance and ministries other than health.

Introduction

There are a number of criteria by which health care treatments and programmes can be evaluated. These include effectiveness, equity, access and efficiency. Different actors outside and inside the health care system (e.g. politicians, managers, health care professionals and patients) will place different emphases on the various criteria. For example, politicians may be particularly interested in equity or fairness in the distribution of health care resources, professionals will normally be most interested in effectiveness and managers will be most interested in the budgetary consequences of health care interventions.

Increasing pressures on health care budgets have led decision-makers to search for methods of assessing the value for money from health care treatments and programmes. In economic evaluation, programmes are compared in terms of their costs and consequences. The consequences typically include improvements in health outcomes and savings in health care resources.

There are a number of forms of economic evaluation, each following the same general methodological approach but differing in the extent to which the health outcomes are measured and valued. For example, in **cost–benefit analysis** attempts are made to value all the costs and consequences in money terms. On the other hand, in **cost–effectiveness analysis** the health outcomes are measured in the most appropriate natural units, such as life-years gained or disability-days avoided.

Economic evaluation is a multidisciplinary activity, to which many health care professionals can contribute. For example, clinicians and epidemiologists can advise on the quality of evidence on the effectiveness of the interventions being evaluated. Conversely, administrators and finance personnel can provide data on the costs of interventions. Usually, economic evaluation requires a synthesis of information from a number of sources, including clinical trials, observational studies and routinely available data sets.

It should be remembered that economic evaluation is an aid to health care decision-making, not a substitute for decision-making itself. Indeed, it incorporates a number of important value judgements and its contribution is to make these more explicit for the person ultimately taking the decisions.

The first part of this module deals with aspects of the methodology of economic evaluation, since it is important that health professionals and policy-makers understand how to tell a well conducted study from a poor one. The second part deals with the application of economic evaluation in health care decision-making, by exploring the potential decision-making mechanisms in which economic evidence could be used. Finally, the third part of the module extends the discussion to a broader range of interventions to improve health, including not only health care but also investments in other sectors of the economy such as road safety and environmental protection.

Methodological features of economic evaluation

The basic components of economic evaluation are set out in Fig. 1. The methodological features of economic evaluation have been well documented elsewhere *(2,3)*.

Fig. 1. Components of economic evaluation in health care

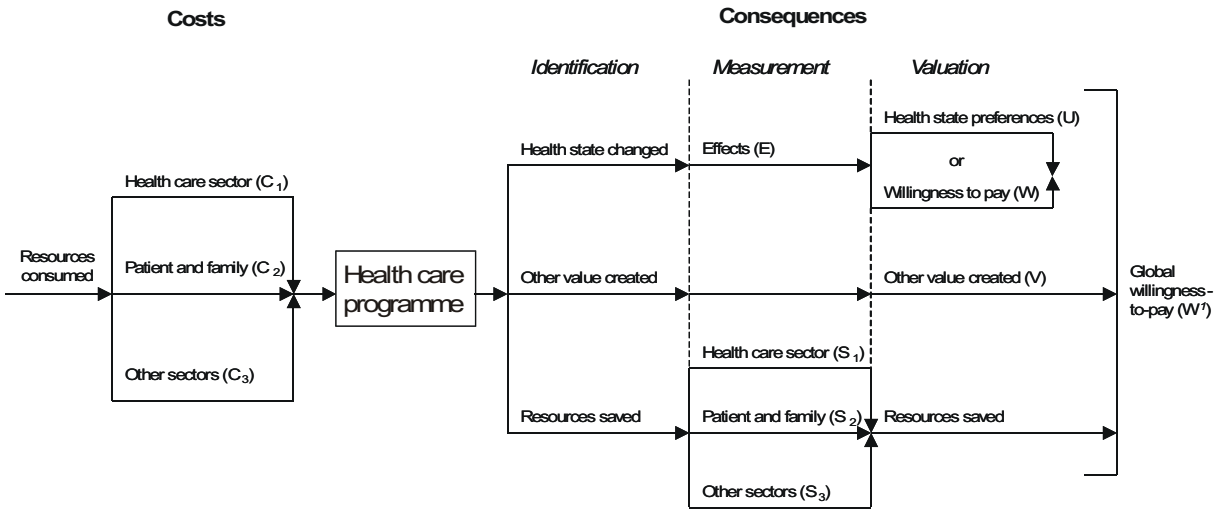


Table 1 contains a checklist of questions to ask of a published study. The main features of this are as follows.

(i) The **study question** should be clearly stated. In particular, it should be clear whose point of view is being considered when costs and consequences are assessed. (Possible viewpoints are those of the hospital, health care system, government or third party payer, patient and family, or society as a whole.) Normally the viewpoint of society as a whole is preferred.

(ii) The **alternatives** for evaluation should be clearly described. Normally a new health care treatment or programme should be compared with current practice or a widely used existing treatment. Comparisons can be made in terms of health, money or utility.

(iii) The **effectiveness** of the alternatives being compared should be reliably assessed. In the case of health care treatments, the most reliable evidence comes from randomized clinical trials, although some modelling may be required. Module 5.4.1 on economic modelling considers how to adapt clinical trial results to reflect regular practice or to extend results beyond the end of the trial (e.g. to lifetime).

(iv) The **costing** should reflect the viewpoint adopted. The relevant costs should first be estimated in physical units of resources consumed (e.g. hospital days, visits to a physician) before being valued using a set of prices or unit costs relevant to the setting concerned. In some settings (e.g. hospitals) certain resources are used to produce a number of joint outputs. For example, the heating plant of the hospital services a number of clinical departments. Therefore, when costing a particular clinical intervention or treatment, only some of the resources will be unambiguously attributable to that intervention. Others will be shared resources, sometimes called "overheads". Therefore, in costing a given treatment, either a method for allocating shared resources (or overheads) is required, or attention should be focused on the additional resources, at the margin, that are required to provide the intervention concerned. (Benefits can also be wider than health narrowly defined: for example, externalities or health and development.)

(v) The study should allow for **differential timing** in costs and consequences, through discounting to present values, and for **uncertainty** in estimates, either through sensitivity analysis or statistical test.

(vi) The **presentation of results** should include an incremental analysis of costs and consequences, comparing one alternative with another, and comments on the major weaknesses in the study. If comparisons of cost–effectiveness are made with other studies, care must be taken to ensure that they employ similar methodologies.

Table 1. A check-list for assessing economic evaluations

1.	Was a well defined question posed in answerable form?
1.1	Did the study examine both the costs and effects of the service(s) or programme(s)?
1.2	Did the study involve a comparison of alternatives?
1.3	Was a viewpoint for the analysis stated and was the study placed in any particular decision-making context?
2.	Was a comprehensive description of the competing alternatives given (i.e. can you tell who did what to whom, where, and how often)?
2.1	Were any important alternatives omitted?
2.2	Was (should) a "do-nothing" alternative (be) considered?
3.	Was the effectiveness of the programmes or services established?
3.1	Was this done through a randomized, controlled clinical trial? If so, did the trial protocol reflect what would happen in regular practice?
3.2	Was effectiveness established through an overview of clinical studies?
3.3	Were observational data or assumptions used to establish effectiveness? If so, what are the potential biases in results?
4.	Were all the important and relevant costs and consequences for each alternative identified?
4.1	Was the range wide enough for the research question at hand?
4.2	Did it cover all relevant viewpoints? (Possible viewpoints include the community or social viewpoint, and those of patients and third-party payers. Other viewpoints may also be relevant depending upon the particular analysis.)
4.3	Were capital costs, as well as operating costs, included?
5.	Were the costs and consequences measured accurately in appropriate physical units (e.g. hours of nursing time, number of visits to physicians, lost work-days, gained life-years)?
5.1	Were any of the identified items omitted from measurement? If so, does this mean that they carried no weight in the subsequent analysis?
5.2	Were there any special circumstances (e.g. joint use of resources) that made measurement difficult? Were these circumstances handled appropriately?

6.	Were the costs and consequences valued credibly?
6.1	Were the sources of all values clearly identified? (Possible sources include market values, patients' or clients' preferences and views, policy-makers' views and health professionals' judgements.)
6.2	Were market values employed for changes involving resources gained or depleted?
6.3	Where market values were absent (e.g. voluntary labour), or market values did not reflect actual values (such as clinic space donated at a reduced rate), were adjustments made to approximate market values?
6.4	Was the valuation of consequences appropriate for the question posed (i.e. has the appropriate type or types of analysis – cost–effectiveness, cost–benefit, cost–utility – been selected)?
7.	Were the costs and consequences adjusted for differential timing?
7.1	Were the costs and consequences which will occur in the future "discounted" to their present values?
7.2	Was any justification given for the discount rate used?
8.	Was an incremental analysis of costs and consequences of alternatives performed?
8.1	Were the additional (incremental) costs generated by one alternative over another compared to the additional effects, benefits or utilities generated?
9.	Was allowance made for uncertainty in the estimates of costs and consequences?
9.1	If data on costs or consequences were stochastic, were appropriate statistical analyses performed?
9.2	If a sensitivity analysis was employed, was justification provided for the ranges of values (for key study parameters)?
9.3	Were study results sensitive to changes in the values (within the assumed range for sensitivity analysis, or within the confidence interval around the ratio of costs to consequences)?
10.	Did the presentation and discussion of the results of the study include all issues of concern to users?
10.1	Were the conclusions of the analysis based on some overall index or ratio of costs to consequences (e.g. cost–effectiveness ratio)? If so, was the index interpreted intelligently or in a mechanistic fashion?
10.2	Were the results compared with those of others who have investigated the same question? If so, were allowances made for potential differences in study methodology?
10.3	Did the study discuss the extent to which the results could be generalized to other settings and patient/client groups?
10.4	Did the study allude to, or take account of, other important factors in the choice or decision under consideration (e.g. distribution of costs and consequences, or relevant ethical issues)?
10.5	Did the study discuss issues of implementation, such as the feasibility of adopting the preferred programme given existing financial or other constraints, and whether any freed resources could be redeployed to other worthwhile programmes?

Exercise 1

Select a published economic evaluation of health care treatments or programmes relevant to your own setting and assess its quality using the checklist given in Table 1. In answer to each of the main ten questions give your response (yes/no/can't tell) and provide a brief commentary on the major strengths and weaknesses of the study.

Using economic evaluation in health care decision-making

The existence of studies with good methodology is a necessary but not sufficient condition for economic evaluation to be useful in health care decision-making. In addition, it is necessary to identify decision-making mechanisms in which economic evidence could, in principle, be used. These are likely to differ from setting to setting, but a number are discussed below, with examples of how the results of economic evaluation have been used.

Planning specialist facilities or specific technologies

This mechanism is obviously most relevant to the "big ticket" technologies (technologies with a single, large financial commitment, such as major items of equipment) and to those health care systems where central or local government has the power to influence decisions about the location of (say) open heart surgery units, neonatal intensive care or specialist diagnostic facilities. Although such power exists primarily in predominantly public health care systems, or those with a national health insurance plan, there may also be opportunities to influence decisions in "liberal" health care systems if the development of specialist facilities either requires significant medical research funding or a large number of patients whose bills are paid by the government.

There are a number of ways in which economic analysis could contribute to decisions about the number and location of specialist facilities. First, there is the question of the optimum size of such facilities, where information about the shape of the long-term average cost curve would be useful, although presumably the costs (borne by the health care system or patients) of travel to specialist facilities should not be neglected. This suggests examination, by economic analysis, of another choice: that of transporting patients to specialist facilities as an alternative to providing more facilities closer to a greater number of centres of population.

Excluding technologies from public reimbursement

This mechanism can be applied to both big ticket and small ticket technologies. A number of countries have organizations which decide on the suitability of new technologies for public funding. In addition, health care insurers in some countries are guided by a central organization (e.g. the Sickness Fund Council in the Netherlands). In principle, such agencies could consider evidence on costs together with effectiveness when taking decisions about the size of the health insurance "envelope."

There is evidence that this is beginning to happen. An early example is the Netherlands where the Health Insurance Executive Board has commissioned a number of economic evaluations *(4)*. However, the problems should not be understated. It is important that such bodies have clear remits with respect to considerations of cost–effectiveness. Also, whether or not a particular technology is the most cost-effective approach to the treatment of a patient may often depend on the specific circumstances, such as the severity of the patient's condition or the diagnostic and therapeutic procedures that have already been applied.

In Australia and Canada (Ontario), guidelines have been proposed for the pharmaceutical industry on the preparation of economic analyses to be included in submissions to the government committee deciding on the reimbursement of pharmaceuticals *(2)*. It has to be shown that a new drug gives good value for money before it is listed in the national or provincial formularies. These policy initiatives demonstrate that governments are beginning to take value for money evidence seriously and that guidelines for undertaking studies can be specified. The company wanting to market a product is

increasingly being made to bear the burden of proof that the product is safe and clinically effective and provides value for money.

Reforming payment schemes for health care institutions (especially hospitals)

One of the most significant reforms over the past few years has been the movement towards prospective reimbursement for hospitals, the most well known scheme being that based on diagnostic related groups (DRGs) operated by Medicare in the United States. Therefore, the calculation of reimbursement rates could take note of evidence on the relative cost–effectiveness of alternative treatment methods for clinical conditions, and this evidence should be more actively disseminated. At present there is perhaps too much of a tendency to set the rates and leave the hospitals to cope with the consequences. This is potentially inefficient, especially if hospitals take decisions based on their own costs and benefits rather than on those of the community at large. This reaffirms the importance of carrying out economic evaluations from a number of viewpoints, including that of society as a whole, so that appropriate incentive structures can be devised for the key actors in the health care system, as mentioned above.

Changing payment systems for health care professionals

In countries where physicians are paid by a fee-for-service system, or where special additional payments are made for some services, there have been concerns that the payment system leads to inappropriate use of technology. Some analysts suggest that this system leads to supplier-induced demand. Others are concerned that the rewards to the physician may be relatively higher for time spent using expensive technology than for time spent talking to the patient or counselling. Given these concerns, it is surprising that there has been relatively little study of fee schedules and few attempts to change them. For example, it would be interesting to study whether there are consistent incentives (implicit in the schedules) to encourage physicians to spend their time using expensive technology, whether physicians are consciously aware of these incentives, and whether they influence their behaviour. This would be an important precursor to studies of how the fee schedule could be used more aggressively to change clinical practice in the direction of greater cost–effectiveness, by withdrawing payment for procedures known to be inefficacious and by offering attractive fees for procedures for which benefits are known to exceed costs. The cost–effectiveness approach can also be useful in health care systems where physicians are mainly paid by salary.

Developing health care practice guidelines

A few years ago the WHO Regional Office for Europe reviewed the schemes operating in a number of countries, with a view to the potential for incorporating economic criteria *(5)*. Two schemes were of particular interest: the Scandinavian model health care programmes, where guidelines are developed for the management of key diseases such as hypertension; and the medical audit schemes developed by the National Association for Quality Assurance in Hospitals in the Netherlands (the CBO), where groups of physicians are provided technical support to review local clinical practices. In both cases there was evidence that economic criteria could be incorporated into the development of guidelines and that attempts were being made to assess the impact of guidelines in terms of cost–effectiveness.

There are other examples of economic appraisal being used to help develop guidelines recommended by medical bodies, such as the work by Eddy *(6)* on cancer screening and that by the Royal College of Radiologists *(7)* in the United Kingdom on routine skull X-rays for patients admitted to casualty departments with head injuries. Against the background of increasing pressure on health care budgets, there is no reason why more studies could not be encouraged. The influence of professional

bodies and medical opinion leaders has probably been under-exploited by those undertaking economic evaluation and those funding health services research.

The relevance of economic evaluation to decisions in clinical practice is further explored in a series of articles published in the *Journal of the American Medical Association (8,9)*. These use the example of developing guidelines at the hospital level for the use of tissue-type plasminogen activator (t-PA) in the treatment of acute myocardial infarction.

Exercise 2

Consider the decision-making mechanisms outlined above and any others relevant to your own setting.

(a) Are there any examples, in your setting, of economic evaluation evidence being used in decision-making? If so, how was the evidence used and what was the outcome?

(b) Do you see any greater potential for using economic evaluation evidence in your setting? What are the major barriers to the more widespread use of economic evaluation and how could these be overcome?

Economic evaluation of health-producing measures in different sectors of the economy

Modules in Chapter 2 discussed the linkages between health, health care and the wider economy. It is clear from this discussion that activities in other sectors can either produce or reduce health. Therefore, if the objective is to improve health, it would be unwise to concentrate solely on activities in the **health care sector**. It might be more efficient, at the margin, to invest in health-promoting measures outside the health care sector.

In the United States, Tengs et al. *(10)* have begun to address this issue. They produced a ranking of more than 500 lifesaving interventions, ordered in terms of their incremental cost per life-year saved. These included not only interventions in the health care sector but others in road safety, environmental protection and occupational health. An abridged version of their data is given in Table 2. It can be seen that the implied price to save a life-year varies greatly between different interventions.

Exercise 3

Consider the data in Table 2.

(a) What are the major advantages and drawbacks of this type of analysis? In the event that the drawbacks outweigh the advantages, what type of analysis would you propose instead?

(b) Assuming, for a moment, that the data in the Tengs et al. analysis apply to your setting, what mechanisms would have to be put in place to ensure a more rational use of resources (in improved health) across sectors?

Table 2. Cost per life-year of life-saving interventions in different sectors

Intervention	Cost/life-year (US $)
Fatal injury reduction	
Mandatory seat belt use and child restraint law	98
Smoke detectors in aeroplane lavatories	30 000
Widen shoulders on rural two-lane roads to five feet versus two feet	120 000
Seat belts for passengers in school buses	2 800 000
Flammability standard for children's clothing size 7-14	15 000 000
Toxin control	
Reduced lead content of gasoline from 1.1 to 0.1 grams per leaded gallon	50
Ban asbestos in brake blocks	29 000
Benzene emission control at pharmaceutical manufacturing plants	460 000
Ban asbestos in thread, yarn, etc.	34 000 000
Radionuclide emission control at coal-fired industrial boilers	260 000 000
Health care	
Measles, mumps and rubella immunization for children	50
Beta-blockers for myocardial infarction survivors	850
Postsurgical chemotherapy for premenopausal women with breast cancer	18 000
Annual mammography for women aged 55-64 years	110 000
Sickle cell screening for newborns	65 000 000

Economic evaluation: linking theory and practice

This module has discussed both the methodological features of economic evaluation and its potential for use in decision-making. The following conclusions can be drawn.

(i) Health care decision-making inevitably involves a number of social, economic and political considerations. Assessments of the cost–effectiveness of alternative treatments and interventions will, therefore, only ever form part of the overall decision.

(ii) Despite this modest aspiration for economic evaluation, it is nevertheless important that the results of studies are reliable. The methodological principles outlined in Part 1 of this module should, therefore, be adhered to.

(iii) In addition to adhering to sound methodological principles, economic evaluations need to be made relevant to the health care decision-maker's own setting. This issue is discussed further in Module 4.3.1 on the development and diffusion of health technology.

(iv) Several mechanisms for using economic evaluation in health care decision-making have been discussed. Although several instances of the use of economic evaluation can be cited (e.g. in decisions about the reimbursement of pharmaceuticals in Australia), in most cases the potential for use is greater than actual use.

Some of this unfulfilled potential may be due to the problems with economic evaluation methodology discussed above. However, the main reason for lack of use is that decision-making procedures in health care do not easily incorporate evidence of any description. Therefore, those in decision-making positions consulting this module should reflect on the decision-making procedures operating in their own countries in order to assess the potential for incorporating economic and other forms of evidence.

References

1. MARK, D.B. ET AL. Cost-effectiveness of thrombolytic therapy with tissue plasminogen activator as compared with streptokinase for acute myocardial infarction. *New England journal of medicine,* **332**(21): 1418–1424 (1995).
2. DRUMMOND, M.F. ET AL. *Methods for the economic evaluation of health care programmes,* 2nd ed. Oxford, Oxford University Press, 1997.
3. GOLD, M.R. ET AL., ed. *Cost-effectiveness in health and medicine*. New York, Oxford University Press, 1996.
4. HAAN, G. & RUTTEN, F. No cure, no pay: an acceptable way of financing fertility treatment? *Health policy,* **13**: 239–249 (1989).
5. *Guidelines for health care practice in relation to cost–effectiveness.* Copenhagen, WHO Regional Office for Europe, 1981 (Euro Reports and Studies No. 53).
6. EDDY, D. *Screening for cancer: theory, practice and design*. Englewood Cliffs, NJ, Prentice-Hall, 1980.
7. ROYAL COLLEGE OF RADIOLOGISTS. Costs and benefits of skull radiography for head injury. *Lancet,* **ii**: 791–795 (1980).
8. DRUMMOND, M.F. ET AL. Users' guides to the medical literature. XIII: How to use an article on economic analysis of clinical practice. A: Are the results of the study valid? *Journal of the American Medical Association*, **277**(19): 1552–1557 (1997).
9. O'BRIEN, B.J. ET AL. Users' guides to the medical literature. XIII: How to use an article on economic analysis of clinical practice. B: What are the results and will they help me in caring for my patients? *Journal of the American Medical Association,* **277**(22): 1802–1806 (1997).
10. TENGS, T.O. ET AL. Five-hundred life-saving interventions and their cost-effectiveness. *Risk analysis,* **15**(3): 369–389 (1995).

5.4 Modelling

5.4.1 Economic modelling and forecasting

Reiner Leidl[6]

Key messages

- Economic models can support decision-making and policy development.
- Transparent models structure problems, make explicit the assumptions used, and explain the consequences implied.
- Decision models help a rational decision-maker to choose the best strategy among clearly defined alternatives.
- Scenario models provide an approach to reasoning in situations characterized by uncertainty, complexity and even lack of data. Based on "what-if" reasoning, they describe possible developments. Disease models describe epidemiological and economic processes that are linked to disease and health intervention in populations. Both types of model can provide insight for policy-makers.
- Econometric models describe a set of statistical techniques for quantitative empirical analysis. They serve analytical purposes, for example by measuring the influence of determinants on a target variable or by forecasting the variable's values.
- In order to ensure high-quality support for decision-makers, both the methods and the results of a model must be used properly and adequately. To achieve this requires methodological expertise, expertise in the health problem investigated, and expertise in supporting decisions.

Tutors' notes

Key learning objectives of this module are:

- to recognize the relevance of modelling in the support of decision-making in health care, the quality criteria for good modelling approaches, and the limitations of modelling providing information to decision-makers (General remarks);

---

[6] This module was prepared by Professor Reiner Leidl, Department of Health Economics, University of Ulm, Germany (e-mail: Reiner.Leidl@mathematikuni-ulm.de). The author is grateful to the members of the WHO Working Group on the Development of Health Economics Training Modules and to the staff members of the Department of Health Economics for comments on earlier versions of this module.

- to understand how a simple, well defined and static economic decision problem can be quantitatively formulated, how its result can be calculated and how it can be interpreted for use in decision-making (Decision analysis);
- to develop an idea of how complex decision problems in health care and in health policy can be simplified and quantitatively tackled by modelling, and how models can then be further extended in a module-based approach (Scenario analysis and disease modelling);
- to develop an idea of the role of traditional statistical modelling techniques (i.e. of econometrics in the case of economic issues) in analysing data and testing theoretical hypotheses on the basis of well established quantitative methods (Econometric models).

For a better understanding of the various modelling approaches and their possible limitations when used to support decision-making, it is recommended that participants discuss the practical examples included in the text. Emphasis should be put both on the technical aspects of modelling and calculation, and on the implications of using this information for decision-making. It is important to stress all critical aspects of the modelling approaches, but it should also be considered on what information decision-making is based when modelling is not used. Together with the examples, it is also suggested that participants should try to solve the exercises, the main results of which are given at the end of this module. Tutors are further recommended to consider the texts suggested in the references and further reading (if they are not covered already).

General remarks on economic modelling

This module is intended to give readers an *appreciation* of a broad range of quantitative methods in economic analysis. Modelling issues are investigated from an applied perspective only, with an emphasis on support for decision-making in health care. Accordingly, readers who are not familiar with the formal side of quantitative analysis may develop an understanding of the expertise required when decisions are to be supported by sound scientific work. A formal treatment of the quantitative methods introduced is beyond the scope of this paper. Readers are, however, invited to consider the simple examples (and the exercises) presented in this module in order to develop a more concrete understanding of the methods underlying economic modelling results. A first understanding of modelling and its potential and limitations is best achieved when working with its tools.

To some, modelling economic issues may seem to be one of the folklore dances celebrated in academic ivory towers. But consider the following questions. Will the general use of a new treatment, for example the eradication of *Helicobacter pylori* in every dyspeptic patient, lower or increase health care costs? What will be the impact of an ageing population on health care financing in the coming decades? Which are the most important determinants of health care expenditure? The answers to such questions may be of considerable relevance to decision-makers in the health care system. The issues at stake may, however, be complex, involve long-term developments, and offer some data seemingly suitable for analysis. Short consideration quickly reveals that the immediate responses to the above questions can only rely on “off the cuff” estimates, the basis of which remains unclear. Experiments and empirical analysis of their results can hardly be conducted with these types of problem. Thus, cases exist where modelling may be the only realistic alternative left for scientifically-based reasoning.

Rational decision-making should be based on the best available evidence. Modelling is a transparent and rational way of reasoning in complex situations in which other appropriate evidence is not available. There are also limitations to modelling approaches and, in consequence, to the use of their results. Yet before completely disregarding such approaches because of their limitations, it should

be considered on what type and on what quality of information decisions can be based when modelling is not used. Models may produce extra information for decision-makers which can best be used both by taking into account its limitations and the evidence available otherwise.

There is a broad range of modelling methods that can be used on health care issues, and there are many different types of economic decision at different levels of the health system which can be supported by modelling. Whether or not a modelling approach is appropriate to inform a decision cannot be generally stated but will depend on the type of result available and on the decision at stake. The results of modelling aim to inform decision-makers on the economic aspects of the problems needing their decisions, for instance the costs and effects involved in alternative treatment options, the capacity of hospitals beds needed, or the major determinants of a policy parameter such as health care expenditure. Several major types of approach to modelling will be introduced in this module and examples of their application shown.

Explanation, prediction and simulation are the main general purposes of economic models. Even before explaining, predicting and simulating, however, modelling produces some important advances. Models require the investigator to structure a problem clearly, to make the precise question investigated transparent and to explain the methods used. Application of a model must include a section on the strengths and on the weaknesses of the approach chosen. Models thus clarify the way and the methods by which answers to complex problems are found. High-quality models do so in a way that is fully understood by the model's audience.

As with other approaches used to inform decision-making, models may have some disadvantages. They may distract the analyst or the reader to technical issues, and they may move the audience away from the issue at stake. Bad models may miss the main relevant points to be investigated. Even worse, some models may make assumptions that are too simple to reflect reality, and may thus suggest policy conclusions drawn from an inadequate analytical framework. In addition to all these problems, models may still give the impression of sound science due to the formal elegance of the approach they present. Accordingly, decision-makers should be well aware of the quality of the models available – which, in the end, they can use or disregard in their decisions.

Besides the general pros and cons of modelling, there are a number of technical aspects which must be considered in approaches to modelling. A highly salient starting point is whether or not the model's theoretical structure appropriately reflects the underlying causal structure of the issue investigated. This theoretical structure must be set up before a quantitative analysis is conducted – there are statistical search techniques that help to identify the best theoretical structure for a model. Accordingly, it needs to be decided whether the purpose of the model is to explore the best theoretical structure, or whether it is empirically to test a well founded theory. The latter tends to be of more interest to decision-makers than the former (which may feature an earlier phase in scientific work).

Models should clearly state their purpose and justify their theoretical basis. Complex models may consist of a number of modules which should be consistent with respect to the definitions and variables that they employ. In any case of empirical modelling, the quality of data and indicators used must be carefully assured. This includes an investigation of criteria such as validity, reliability, completeness and representativeness. Decision-makers should regard these quality criteria when evaluating the information which a model contributes.

In general, the model should be implemented as transparently as possible. It should thus clearly state its structure, the methods used, data input, analytical procedures, results and interpretation. Assumptions must be discussed and well founded. Results need to be tested. In stochastic models

(which contain probability elements), this can be achieved by calculating and reporting confidence intervals. In deterministic models, the assumptions used should be varied within plausible ranges. Such an exercise is called sensitivity analysis. This analysis can be conducted with individual variables, or with combinations of more than one variable, of a model. In any case, the variables selected and the ranges calculated in sensitivity analysis should be well justified. The effects of these variations on results should be carefully investigated and reported. Decision-makers should be aware of the range of uncertainty surrounding the information basis of their decisions.

Last but not least, it should be clearly stated for which decision context the model's results are considered useful, and which restrictions can be foreseen. Some models may directly support the choice between alternative decisions, and some may be more suited to guide policy-makers through situations which may have to be tackled. In general, decision-makers must check whether or not the approach of a model is appropriate and useful in the context where the decision is to be taken. As an example of study focus, the cost–effectiveness resulting from a decision model that takes into account the age structure of an insured population could be more useful to a health insurer than results derived from an experimental study that would render harder evidence but feature patients of a different age structure. Decision-makers must assess the appropriateness of models' results for their intended uses.

There may be situations where models are the only transparent and rational way to deal with complex issues. In order to assure the methodological quality of models and their suitability to the issue being investigated, modelling approaches should be confronted with the same scientific rigour as the traditional empirical analysis of an experiment. Both in the analysis of experiments and in modelling, careful consideration must be given to the design issues, appropriate structure of analysis, and quality of data and indicators. Decision-makers can use the checklist in Table 1 for an instant control of the information provided by a modelling approach. They should be aware that a competent scientific assessment of a specific model would require detailed technical expertise which cannot be gained from this overview. What can be gained, however, is a recognition of the important issues that need to be assessed and managed when supporting decisions on the basis of economic models.

Table 1. Eight checkpoints for decision-makers looking at modelling approaches

1.	Is the model adequately designed to contribute to the issue of interest?
2.	Is the model backed by relevant theory and by a confirmed structure?
3.	Does the model use high-quality data and indicators?
4.	Are individual modules linked with each other consistently?
5.	Are model design, methods, inputs and analyses well documented?
6.	Is uncertainty dealt with appropriately?
7.	Is the model analysis appropriate for the decision context?
8.	How does the quality of model results compare with other sources of evidence?

In sum, models must be implemented properly and used adequately. In order to assess the methods and use of models in health care, a decision-maker needs methodological expertise, expertise in the health problem investigated, and expertise in how to support decisions by model results. Given these preconditions, models can significantly improve the state of information available, and support decision-making in a transparent and rational way.

The further structure of this module

There are, of course, many different problems and many different types of model. Health care models can be directed at individual patients or at populations. Static models focus on one point in time, dynamic models integrate time as a variable and consider the course of the model over time. Complex models may describe and integrate many different areas relevant to a problem, whereas simple models may concentrate on the core issue only and disregard other aspects. From the wide variety of models, this module introduces three major modelling approaches and briefly describes the basic technical aspects of each. This is followed by a simple exercise.

The section on decision analysis introduces the quantitative analysis of a static decision problem. The example used is a simple decision-tree model describing the choice of a medical intervention from an economic point of view. The section on scenario analysis and disease modelling discussed more general forms of models. A simple example deals with demographic impact on costs and financing in a health system. The final section deals with traditional econometric analysis. This method is widely accepted for analysis and prediction. The example used refers to an often investigated problem, the explanation of the level of health care expenditure.

Decision analysis

Decision analysis is directed at the best choice between clearly defined decision alternatives. It is concerned with the detailed and transparent description of all parts of a decision. This includes the events that can occur, the impact of these events, the type and number of relevant decision-makers, their preferences, and the rules that are being applied in the decision-making process. When such an approach is applied in situations of uncertainty, with complicated sequences of events, multivariable preferences and the participation of several decision-makers, analysis can become very complex. Presented here is a simple model which incorporates the basic idea of decision analysis: a decision tree. A simple decision-tree model is characterized by a sequence of events which occurs only once and does not include any feedback loops. More complicated types of model where, for example, cohorts of patients pass through different stages of disease but may recur to earlier stages, are not dealt with here.

In a decision-tree model, the problem to be addressed must be clearly defined. This includes specification of the choices that can be made, the possible events and sequences of events that can occur, the probabilities of events occurring, and the impacts of interest that are linked to these events. In economic evaluation, these impacts would be the effects on costs and health. By computing the probability-weighted outcomes of the various possible choices ("strategies"), the best strategy (aiming, for example, at minimum cost per health effect gained) can be chosen. Many assumptions may be incorporated into the decision tree. In order to assess the impact of assumptions on results, it is strongly recommended that a variety of assumptions be investigated up to possible extreme values. However, the ranges of assumptions investigated in this so-called sensitivity analysis should be well justified.

Decision-theoretical models have the potential to combine the best available data in order to enable rational comparisons of choices in complex decision situations. They have been criticized because they are often not based on data measured in experiments, but on (sometimes crude) assumptions. It has been recommended that these models should be used in cases where experiments with direct measurements are not feasible. As far as possible, data from experimental studies should be used as inputs to decision models.

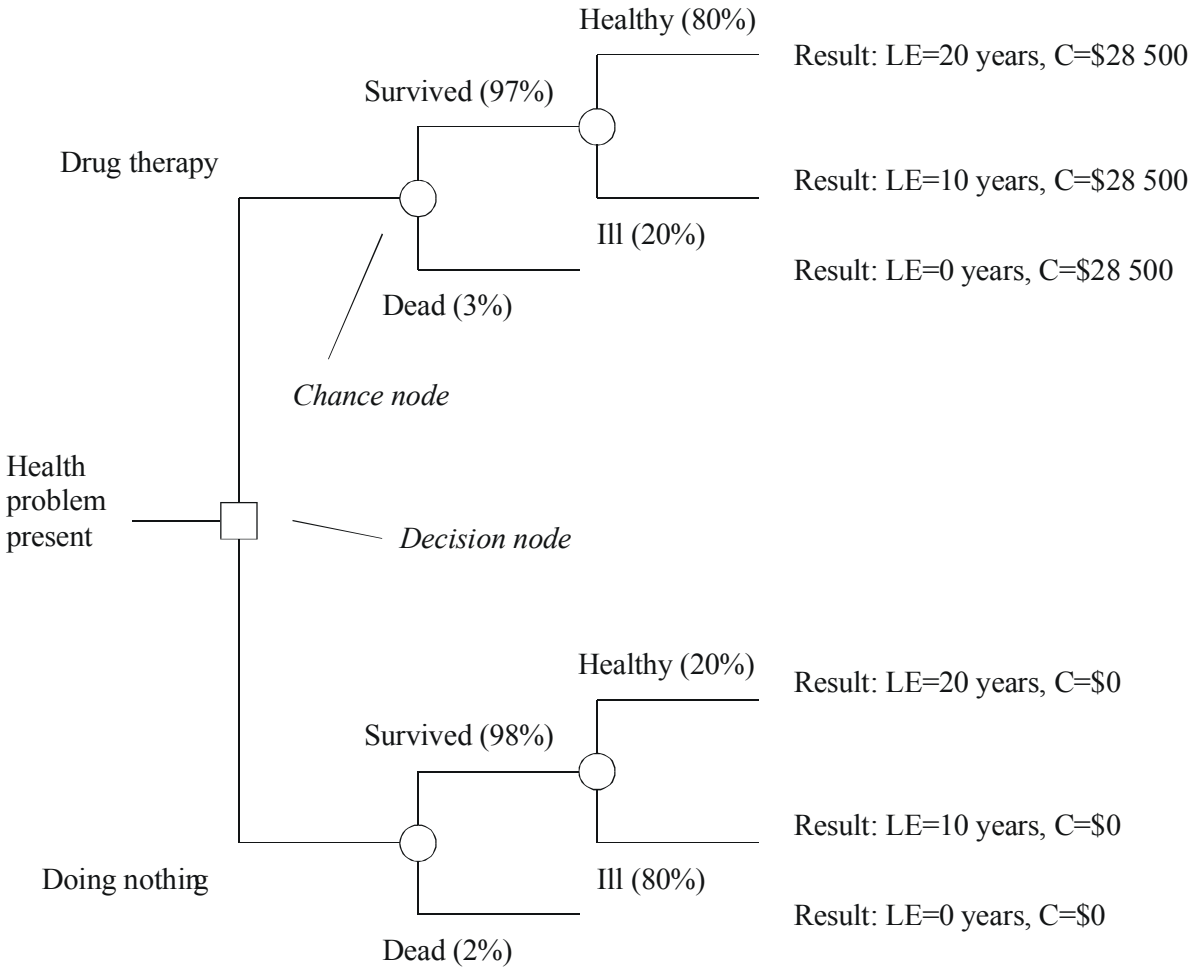
Example 1: A decision tree

Assume that, in a particular population, a disease can be dealt with either by a new drug therapy with some severe side effects, or by doing nothing. Doing nothing incurs a 2% mortality from the disease, while just one fifth of the survivors regain full health. With drug therapy, mortality increases to 3% due to side effects, but four fifths of the cases will be completely cured. For those who achieve full health, another 20 years of life can be expected on average. Those surviving but not being fully cured remain chronically ill and have a life expectancy of 10 years. Drug therapy costs US $28 500.

These data can be used to support a decision on the provision of the drug in a health care system. It is estimated that for quite a number of the currently provided interventions in this system, the costs of a life-year gained by medical care exceed US $5000. Given this, can a clear suggestion be made as to whether or not this new drug therapy should be provided?

Starting with the above data, the decision problem can be exactly structured according to the decision tree in Fig. 1. The events following the choice for drug intervention or for doing nothing are reported in terms of mortality and survival, the latter differentiated by survival in full health and survival in the state of chronic illness. The probabilities of each event are reported. The total probability of an event at the end of a branch can be calculated by multiplying the probabilities of the events passed through retrospectively from the end of the branch back to the root of the tree, i.e. the starting choice on whether or not to use the drug.

Fig. 1. A decision tree for the economic evaluation of a drug treatment



Multiplying total probabilities with the outcome (costs and effects) linked to each branch, the total outcomes for each choice can be summed up. Thus, the expected outcome for each strategy can be calculated.

The additional life expectancy (D LE) in years which is gained by drug therapy and the additional costs (D C) incurred can thus be computed. Together, the two indicators render an incremental cost-per-life-year-gained ratio for the drug intervention, as opposed to the option of doing nothing:

Δ LE = [0.97*(0.8*20+0.2*10)+0.03*0] years – [0.98*(0.2*20+0.8*10)+ 0.02*0] years = 5.7 years
Δ C = US $28 500
Δ C/Δ LE = US $5000 per life-year gained.

As the cost-per-life-year-gained does not exceed the cost accepted in a number of existing interventions, the model's result would not go against a provision of this treatment from an economic point of view. However, the assumptions in the model would have to be discussed in detail and would have to be varied in order to find out how stable this result is. Another aspect might be to investigate the impact of therapy on health-related quality of life. For further detail on economically assessing medical intervention, the reader is referred to Module 5.3.1 on economic evaluation.

Exercise 1. Economic modelling to minimize treatment cost

During the 1990s, there was intense discussion on how to use the newly discovered drug treatments to eradicate *Helicobacter pylori* (hp). Hp is thought to be an important determinant of dyspepsia, peptic ulcer and even gastric cancer. Among the issues discussed is the question of whether eradication treatment should be generally offered to all dyspeptic patients found to be hp-positive, irrespective of whether there is evidence of further disease. A previous study, which describes a simple decision-tree model and a restricted economic evaluation *(1)*, has reported the following data (only the branch of hp-screening and follow-up treatment is modelled and some assumptions have been adapted to ease computations):

A test that perfectly identifies all hp-infected individuals costs US $50. Of the dyspeptic patients tested, 30% are expected to be hp-positive. Using a drug therapy costing US $218 on those testing positive, hp will be eradicated in 60% of cases. Eradication prevents chronic dyspepsia and peptic ulcer in 10% of the patients and gastric ulcer in 0.1% of the patients (and has no preventive effect in the rest of the patients). Treatment costs are estimated to be US $5000 for chronic dyspepsia, US $7000 for peptic ulcer and US $30 000 for gastric cancer.

Given these data, and cost minimization as the aim, should the hp-screening and treatment strategy be introduced or not? What limitations do you see in the underlying calculation?

Scenario analysis and disease modelling

Decisions regarding health policy sometimes have to be taken in situations of significant uncertainty, complexity, and scarcity of data and lack of evidence about interventions. The early phase of the global HIV epidemic was one example. Scenarios may provide a helpful basis to deal with such situations. Scenario analysis is characterized by "what-if" reasoning. Using a variety of assumptions, possible developments are transparently described. Developments mostly concern the future but

sometimes also refer to the empirically not known, or not well understood, past or present. When applied to health care systems, scenario analysis often features modelling at the population level, including a description of the population and its health problem, and usually involves the investigation of future trends or alternative interventions.

Disease models are usually applied to difficult health problems. Disease models require:

(i) a sub-model on the epidemiological spread of the disease in the population, for example a risk factor model in a stochastically emerging disease, or a model of spread for an infectious disease; and

(ii) a sub-model on progress of the disease dealing with the future course of the disease in an individual after onset; this typically includes both the natural history of the disease (i.e. its course without any intervention) as well as its development following health care interventions.

Scenario analysis and disease modelling may become quite complex. They usually require detailed knowledge of the problem investigated, such as the determinants of future need for physicians or the impacts on health of early interventions in HIV infection. In the following example, a simple scenario analysis is described. Important aspects, or modules, for more comprehensive analysis are then looked at.

Example 2: Scenario analysis of demographic impact on health care financing

The populations of a number of countries are ageing. To a significant extent, this results from declining fertility rates. Changes in size and composition of a population will influence health care utilization as well as costs. Furthermore, a common way of financing health care is the pay-as-you-go system: the working population has to earn the income from which care is financed for the total population, including care of children and the elderly. Scenario analyses of future health care utilization, cost and financing may provide highly salient information for health policy-makers.

Consider the following fictitious population (Table 2) of 106 million people, which experiences a 10% reduction in births over the period 1998–2018. In the simplified example, health care costs are indicated for children (0–15 years), the working population (16–64 years) and pensioners (65+ years). According to the pay-as-you-go system, each member of the working population has to contribute US $1925 (= total costs/working population) to health care in order to cover the costs incurred in this population. Substituting population figures for 1998 by population projections for 2018 provides the reference scenario for future costs and contributions. Everything else is kept constant in this exercise – epidemiology, health care technology, capacity for care, and the preferences and behaviour of all participants. These assumptions are very restrictive, but the factors listed cannot be predicted simultaneously for the next two decades. This leaves the simple methodology introduced as a starting point that indicates the "pure" demographic effect *(2)*. Extension of the analysis could involve changing one or more of the above listed factors in alternative future scenarios.

Exercise 2

Start from the scenario described in Table 2. What is the 20-year growth rate of contributions? What impact would the inclusion of technical progress have on this rate? In order to quantify this impact, assume the following alternative scenarios resulting from increasing costs of health care technology over the two decades (other things being equal):

Table 2. Demography, health care cost and financing in a population, 1998 and 2018

Age groups (years)	Population in millions	Cost/capita in US $	Total cost in billion US $	Contribution in US $
		year 1998		
0–15	10	1 000	10	–
16–64	80	1 000	80	1 925
65+	16	4 000	64	–
		year 2018		
0–15	9	1 000	9	–
16–64	66	1 000	66	2 470
65+	22	4 000	88	–

(i) all health care costs in the population grow by 35%, which corresponds to a yearly growth rate of a little more than 1.5%; and

(ii) health care costs for children and the working population grow by 35% but those for the elderly grow by 70%, the latter corresponding to a yearly growth rate of almost 2.7%.

In the light of the problems that emerge from scenarios such as those described above, it is often suggested that a capital funding system would be a better alternative. In this approach, total health care costs for each individual over his or her whole lifetime are calculated and then covered by a constant premium rate over the whole lifetime. From a modelling perspective, is it easy and realistic to choose this option?

In example 2, projections of a "demographic module" have been combined with figures from a current "cost module". Two assumptions regarding changes in technology and, subsequently, costs provided the basis for further scenarios. Modelling other than demographic and cost changes may require the use of additional modules, such as an "epidemiology module" when investigating the effect of changes in disease incidence and prevalence; a "health care utilization module" when investigating the impact on health care capacities; or an "effects on health module" when analysing cost–effect relationships at the population level. These five modules are listed in Table 3. This table also includes examples of the relevant data and indicators that must be identified when empirical information is collected for use in the model. One important requirement in scenario analysis and disease modelling is that the different models to be combined match with respect to their theoretical scope, technical definition of indicators, and representativeness of the data used.

The five modules may be used in quite different combinations, and may be part of a variety of modelling approaches. Table 4 gives an overview of the main types of modelling approach, although it does not claim to be exhaustive. The first three types show that economic modelling of the impact of demographic change can be directed at health care needs, at the development of costs, or at the description of health care costs and effects. For the next three types of model, only the cost modelling type is shown:

- epidemiological models emphasize the relevance of changes in disease patterns by including an epidemiological module;
- trend models focus on technological change and include projections of what current costs per case will be in the future; and
- disease models concentrate fully upon the developments in a cohort of patients once a disease has already started, while leaving out the demographic module.

Table 3. A conceptual approach to scenarios and disease models in 5 modules

	Module 1 Demography	Module 2 Epidemiology	Module 3 Health care	Module 4 Costs	Module 5 Effects
Theory	population models	risk factor models, development of disease models	concepts of care appropriate to need and to the health care system	costs according to measurement concepts from different perspectives	clinical endpoints, survival modelling, quality of life and utility concepts
Data	population statistics, socio-economic data	register data, epidemiological and clinical studies	data on diagnoses, services and prescriptions from insurers and providers	expenditure data from insurers' data bases, cost data from providers' accounting systems	data from clinical or observational studies
Indicators	size of population by sex and age groups	incidence, prevalence, progression of disease, mortality	diagnosis-specific capacities of care, frequency and intensity of services, length of stay	total cost per case or per period	cases prevented, life years gained, QALYs produced

Table 4. Modules used in different types of scenario or model

Type of modelling approach	(1) Demography	(2) Epidemiology	(3) Health care[a]	(4) Costs	(5) Effects
a) demographic projection of health care needs projected	projected[b]		current[c]		
b) demographic projection of costs	projected			current	
c) demographic projection of costs and effects	projected			current	current
d) epidemiological projection of costs	projected or current	projected		current	
e) trend cost scenario	projected			projected	
f) disease model of costs		current[d]		current	

Note: "Projected" means that a description of the future course of the module is made, "current" means that the state as of today is used.

[a] utilization
[b] general and clinical epidemiology
[c] "Current" means that the state as of today is used
[d] General and clinical epidemiology

Within all the types of modelling presented, comparison of a basic scenario with a further alternative is possible, e.g. comparing intervention versus no intervention, or new intervention versus old intervention. The models introduced can also be considered starting points for further analysis.

Exercise 2 referred to models type (b) and (e) in Table 4, but combined these with the financing issue in the context of pay-as-you-go systems. Other examples can easily be identified and discussed using the framework of Table 4.

Scenario analyses do not claim to generate predictions in the sense in which statistical extrapolations of time series data do. Rather, scenarios link possible sets of assumptions with the consequences these assumptions imply within the modelling framework. Scenario models have, however, been criticized because they are often hard to validate. Another problem is that their construction or the data they include may not feature the scientific rigour of other methods. Disease models intend to describe, analyse and predict the epidemiological and economic processes linked to disease and health care in populations *(3,4)*. First-quality standards that may apply for this type of modelling have recently been developed (see Sonnenberg *(5)* for a focus on decision-type models). In general, the scientific rigour of scenario analyses and disease models can be assured by applying traditional concepts for the testing of scientific methodology: consistency of the theoretical approach; validity and reliability of data and indicators; matching of the different modules that are combined; and representativeness of the data used for the decision context that is to be supported by the scenario analysis.

Econometric models

Econometrics has been defined as the science and art of building and using models that quantitatively reflect the real world. As with any other modelling, econometric models should be based on good theory. Econometric methods comprise a large set of models for empirical analysis and forecasting, such as simple and multiple regression analysis, simultaneous equation systems, time series analysis and several other types. The data useful for sound econometric analysis must fulfil a number of requirements such as sufficient number of observations, certain distributional properties, and completeness of data in multivariable models.

An important group of econometric models is explanatory models. These models relate a dependent variable, the variations of which are to be explained, to one or more independent variables, whose variations are to explain that of the dependent variable. After an econometric model has been run, a number of assumptions must be checked, for example the properties of the variable gathering up the unexplained variability, the so-called “error term”. This term should not be correlated with the independent variables. Another example is that two (or more) independent variables should not be correlated with each other, a case called (multi-) colinearity. Furthermore, specification tests can be used to assess the appropriateness of the model structure. Scientifically, econometrics is probably the most developed field among the methods discussed in this module.

In practice, however, this means that skill and thought are necessary when using econometrics. In a number of situations, the conditions required for this type of modelling may not be met. The appropriateness of what is being modelled thus has to be assessed in practice. In some situations, decision-makers may ask for faster information support than an econometric model may be able to render on the basis of the data available. For example, how much burden will the financing of a new drug treatment put on our budget, and will it be worthwhile? For such questions, models of the type discussed earlier could be considered.

In the following discussion a very simple example is considered based on regression analysis. The emphasis of the example is on understanding the approach to the analytical problem and the impact of results on decision-making at the macro-level of a health care system. Technical issues of the analysis are not dealt with here.

Example 3: Explanation of health care expenditure

A number of econometric studies have tried to explain health care expenditure at the level of the health care system. All the studies include income as a determinant of health care expenditure. In theory, it seems intuitively plausible that income will explain – at least in part – the level of health care expenditure. In addition, many other determinants have been considered, for example, variables indicating the organization of the health care system (e.g. whether or not it features a national health service, or remuneration by a budgeting or a fee-for-service system), or factors such as consumption of alcohol and tobacco. In many studies, the approach to explain health care expenditure is to run a regression of national per capita health care expenditure (HCE, as the independent variable) on income, expressed as gross domestic product per capita (GDP, as the dependent variable). The typical units of observation in such analyses are countries, thus featuring an international cross-sectional analysis.

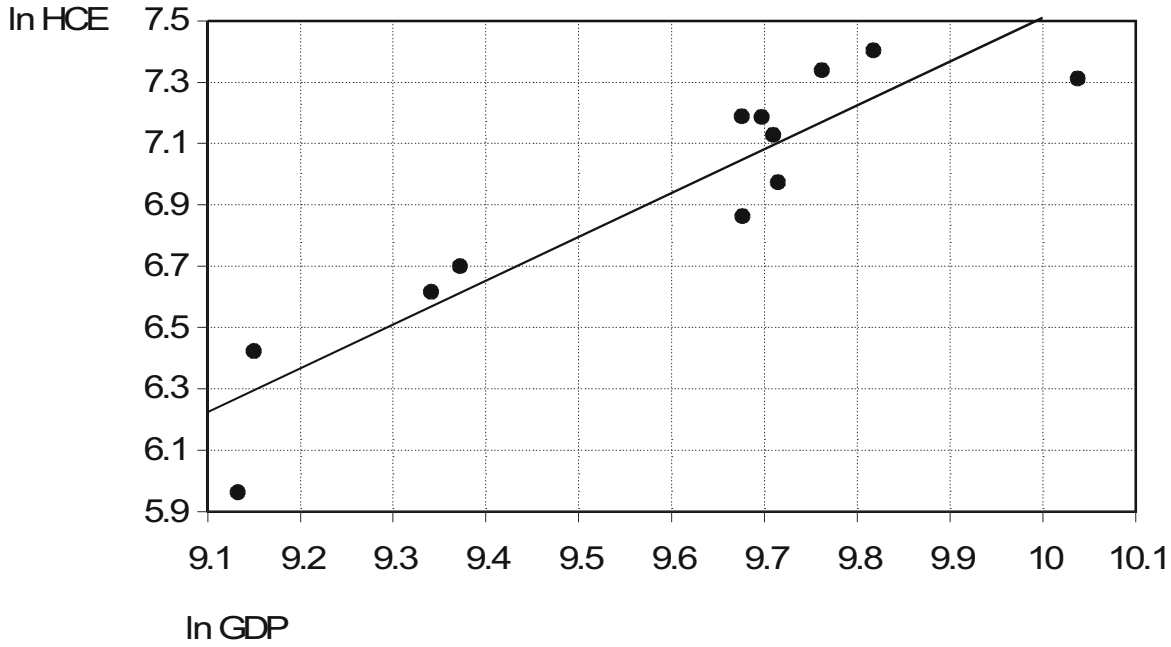
Fig. 2 shows an example of a basic result found in many studies. This shows the observation points for HCE and GDP for the twelve member states of the European Community in the year 1990 (data from the OECD health database as of July 1997). The econometric model uses a frequently found specification, employing the logarithms (ln) of HCE and GDP instead of baseline values *(6)*. The regression line minimizing distance to the points of observations has a slope of 1.43. For the model specification used and the set of countries observed, this means that if the level of GDP grows by 10%, the level of HCE will increase by 14.3%. The slope is thus called the income elasticity of health care expenditure. In this model, 84.4% of the total variance of HCE is explained by GDP. Looking at the development of baseline per-capita income, historical observation over the last forty years shows that it took poorer member states between less than one up to more than two decades to reach the mean income in the European Community which existed at the beginning of an observation period *(6)*.

Exercise 3

Discuss the appropriateness of the above econometric analysis. What conclusions would you draw from the results reported concerning the following three issues:

- the determinants of health care expenditure
- the possibility of cost-containment
- the impact of economic growth on the health care system?

Fig. 2. Gross domestic product per capita (GDP) and national health care expenditure per capita (HCE) in the European Community in 1990 (1990 purchasing power parity US$)



Note: From left to right, the observations display the following EU member states: Greece, Portugal, Ireland, Spain, Netherlands (upper point), United Kingdom, Italy, Belgium (upper point), Denmark, France, Germany and Luxembourg.

References

1. Consensus conference on economic modelling. *Pharmacoeconomics,* **17**(5): 443–513 (2000).
2. Leidl, R. Health economic issues relevant to countries with aging populations. *World health statistics quarterly*, **4**5(1): 95–108 (1992).
3. Leidl, R. European integration, economic growth, and health care expenditure. *In:* Leidl, R., ed. *Health care and its financing in the single European market*. Amsterdam, IOS Press, 1998 pp. 38–58 (Biomedical and Health Research Series, Vol. 18).
4. Postma, M.J. et al. Hospital care for persons with AIDS in the European Union: assessment of current and future impact controlled for severity stages. *Health policy,* **41**(2): 157–176 (1997).
5. Sonnenberg, A. Cost-benefit analysis of testing for *helicobacter pylori* in dyspeptic subjects. *American journal of gastroenterology*, **91**(9): 1773–1777 (1996).
6. Weinstein, M.C. et al. Forecasting coronary heart disease incidence, mortality, and cost: the coronary heart disease policy model. *American journal of public health*, **77**(11): 1417–1426 (1987).

Further reading

To the author's knowledge, there are no textbooks on the methodology of economic modelling in health care which cover all three approaches presented here. Suggestions for some further reading for each approach are given below.

1. On **decision theory**, there are a number of general textbooks. Special reference to health economics is found, for example, in:

   KEELER, E. Decision trees and Markov models in cost–effectiveness research. *In:* Sloan, F.A., ed. *Valuing health care*. Cambridge University Press, 1995, pp. 185–205;

   RITTENHOUSE, B. *Uses of models in economic evaluations of medicines and other health technologies*. London, Office of Health Economics, 1996.

2. No comprehensive methodological textbook is, unfortunately, available on **scenario analysis and disease modelling**. An instructive methodological overview of scenarios to support health policy, using the example of AIDS, is found in:

   JAGER, J.C. & VAN DEN BOOM, F.M.L.G. Scenario analysis, health policy, and decision-making. *In:* Kaplan, E.H. & Brandeau, M.L. *Modelling the AIDS epidemic: planning, policy, and prediction*, 1994, pp. 237–252 (Chapter 13).

Important elements of constructing disease models come from decision analysis, dynamical mathematical modelling, and basic epidemiological methods, each a field for which general textbooks exist.

3. On **econometrics**, there are numerous textbooks available. A very brief, nice overview on statistical tools for health economics, including the econometric ones, is found in:

   FOLLAND, S. et al. *The economics of health and health care*, 2nd ed. Englewood Cliffs, Prentice Hall, 1997 (Chapter 3).

   A comprehensive introductory text not directed at health care is:

   PINDYCK, R.S. & RUBINFELD, D.L. *Econometric models & econometric forecasts*, 3rd international edition. New York, McGraw-Hill, 1991 (4th edition forthcoming).

Quantitative solutions to the exercises

***Exercise 1*:** The screening and treatment strategy for helicobacter pylori saves US $106. This is derived by multiplying monetary outcomes and probabilities in the decision tree, progressing backwards from the ends of the branches to the root of the tree: – US $50 + 0.7 * $0 + 0.3 * {- US $218 + [0.4 * US $0 + 0.6 * (0.1 * US $5000 + 0.1 * US $7000 + 0.001 * US $30 000 + 0,799 * US $0)]} = $106.

***Exercise 2*:** The baseline 20-year growth rate of contributions in the table (1998: US $1925.0; 2018: US $2469.7) is 28.3%. Adding technical progress raises the growth rate of contributions to 73.20% in case (i), referring to a 2018 contribution of US $3340.9), and to 97.44% in case (ii), referring to a 2018 contribution of US $3800.8. The average yearly ("compound") growth rates are 1.25% for baseline, 2.78% for case (i) and 3.46% for case (ii), respectively.