Pergamon

European Journal of Cancer Vol. 31A, No. 9, pp. 1410–1424, 1995
Elsevier Science Ltd
Printed in Great Britain
0959-8049/95 $9.50+0.00

0959-8049(95)00286-3

Review

European School of Oncology Advisory Report to the Commission of the European Communities for the "Europe Against Cancer Programme" Cost-effectiveness in Cancer Care

C. Williams, D. Coyle, A. Gray, J. Hutton, T. Jefferson, G. Karlsson, K. Kesteloot, C. Uyl-de Groot and S. Wait

INTRODUCTION

Aim

This report is designed to examine the need for economic evaluation of cancer care, and to stress the importance that any such evaluation is carried out in a rigorous and effective manner.

Rationale

As pressures on healthcare budgets have grown worldwide, an increasing emphasis is being placed on identifying and improving value for money: that is, maximising the health gain from resources committed to healthcare by concentrating on interventions of proven effectiveness and cost-effectiveness.

To date, the cost-effectiveness literature across all areas of healthcare remains fairly small, despite strong growth in recent years: between 1960 and 1992, a total of 636 economic evaluations combining cost and outcome data were published, or less than two a month [1]. Another survey identified an annual average of 247 studies containing some form of evaluation between 1979 and 1990 [2]. The quality of these studies is known to be highly variable. Although more studies have been published with time this has only kept pace with the increasing number of clinical trials.

Contents

The reports will briefly discuss basic methods of economic evaluation. This is followed by a section discussing current methods of measuring and evaluating the effectiveness of care. Measuring and evaluating resources follows in the next section, and current practice of economic evaluation is covered in a section which reviews the general situation in medicine and then discusses two areas of cancer care—lung cancer and screening. The penultimate section discusses the use of results of economic evaluations. The final part consists of a series of recommendations applicable to a wide variety of individuals involved in all aspects of cancer care.

Correspondence to C. Williams at the Istituto Europeo di Oncologia, Via Ripamonti 435, 20141 Milan, Italy.
Revised 4 May 1995; accepted 24 May 1995.

METHODS OF ECONOMIC EVALUATION

This section briefly introduces concepts and techniques used in economic evaluation—for in depth discussion readers are referred to the short bibliography at the beginning of the reference section.

TECHNIQUES OF ANALYSIS

Economics offers a range of techniques for the evaluation of healthcare activities. Each has its place, but it is important to ensure that the appropriate method is being used in any particular context.

Cost minimisation analysis (CMA)

The least complicated approach to economic evaluation is to judge between alternatives on the basis of minimum cost. In doing this, it is essential to establish *a priori* that the treatment options being compared are identical in their effectiveness and risk of complications and side-effects i.e. that resource cost is the only significant difference between the options. Good examples of such studies can be found in the literature (e.g. [3] on long-term domiciliary oxygen therapy) but the approach is likely to have limited applicability in the cancer field because of the variable impact of treatments on patient survival and quality of life.

Nevertheless, this approach can be useful for comparing different treatment techniques with similar patient outcome. In radiotherapy treatment, for instance, one can choose between off-line or on-line *in vivo* dosimetry and portal imaging techniques. For each of these, a choice can be made on the basis of cost information [4, 5], since the patient outcomes are very similar with both techniques. For another quality assurance technique, the use of fixation masks, with no difference for patient outcome (i.e. reproducibility of the patient set-up) could be detected, implying that the choice between PCV (plastic) and thermoplastic (Orfit) masks can be based on cost data [6].

Cost-effectiveness analysis (CEA)

This is the commonest approach to economic evaluation in healthcare, enabling the comparison of alternatives with both differing cost and differing effectiveness: the latter being measured in natural units such as life years saved (for example, Ludbrooke [7] compared treatments for renal failure in this way). New medical treatments may be more effective and are often more costly, and their impact can be clearly expressed using the cost-effectiveness ratio, i.e. the extra cost of each extra life year gained by switching to the new treatment.

$$\Delta CE = \frac{C_1 - C_2}{E_1 - E_2},$$

where 1 is the new treatment and 2 is the current treatment.

Cost-effectiveness is a relative concept, and the ratio must always be calculated in the context of a comparison with existing or potentially competing treatments. Thus, new treatments for teratoma will concentrate on potential further improvement in survival over existing chemotherapy whereas in advanced non-small cell lung cancer improved palliation will be the main goal.

Cost–utility analysis (CUA)

The limitation of CEA is that it requires a single measure of effectiveness by which to compare procedures. Often treatments will have a range of outcomes, and these outcomes will be valued differently by different patient groups. Recent developments in economic evaluation have attempted to measure multi-dimensional outcomes, and to incorporate preferences in the assessment of the effectiveness of healthcare. Thus, an attempt is made to measure the effect of healthcare on the lives of patients (in economic terms, its utility to them) rather than just its impact on clinical variables. For example, to judge the success of chemotherapy in terms of tumour control, may overlook the aspects of the treatment which have most impact on patients. The most common utility measure used is the quality-adjusted life-year (QALY), which combines the benefits of survival and quality-of-life during the survival period. The results of cost–utility analysis (CUA) are also presented in ratio form as the cost per QALY gained from the adoption or expansion of a healthcare programme. As well as being a better measure of effectiveness, CUA also enables comparisons to be made between programmes for different patient groups, as the utility measures are not disease-specific.

Cost–benefit analysis (CBA)

The three approaches to economic evaluation described above are designed to answer specific questions regarding the choice between competing healthcare interventions, when a decision has already been made to use such treatments. Cost-effectiveness and cost–utility analysis can, therefore, be useful in allocating limited budgets between competing services. They do not address the more fundamental question of whether some healthcare activities are worth pursuing at all. The economic technique designed to deal with such questions is cost–benefit analysis (CBA). This differs from CEA and CUA in that it calculates the benefits in monetary terms as well as the costs. Thus, if some of the benefits are not directly purchased by people as consumers (as in the majority of healthcare), then proxy measures of monetary benefits must be calculated. The danger with using this approach is that items which are too difficult or controversial for the application of monetary values may be ignored, although they are important. The advantage of CBA is that it provides a measure of the rate of return to investment in healthcare which can be compared with the rates of return from other public and also private investments.

Early CBA studies in healthcare were often limited to those factors most easily quantified, and used poor proxies for the value of improved survival and length of life. More recently, renewed interest in CBA in health care has been generated by the successful use of the technique in other areas, such as environmental economics. The monetary value of health benefits can be estimated by measuring the patients' willingness-to-pay using questionnaire surveys [8]. This approach is the subject of much debate amongst economists, and is unlikely to become routinely applied in the near future [9].

STUDIES OF ECONOMIC ANALYSIS

Whatever the method of economic evaluation chosen, the analysis must follow the same framework. The basic stages are:

(i) problem definition
(ii) identification of options
(iii) identification of relevant costs and outcome
(iv) measurement of costs and effectiveness of each option
(v) allowance for differential timing of costs and benefits
(vi) assessment of risk and uncertainty
(vii) presentation of results.

Problem definition

This is a crucial stage as it dictates the subsequent conduct of the analysis. A key factor is to determine the perspective from which the analysis is being conducted. This governs which costs and outcomes are considered. For example, if the perspective of the health service is taken, then costs falling on other government agencies, such as social services, and patients may be ignored. An analysis carried out from the perspective of society as a whole would include these costs. The scope of the problem will determine the economic evaluation technique to be used. Thus, selecting the most appropriate imaging technique to make a given diagnosis may require CEA, but the choice between chemotherapy or surgery in the treatment of cancer may require CUA.

Identification of options

An economic study will only be valid if comparison is made between a new procedure and the best available alternative. In the case of new therapies, where no effective treatment currently exists, the assessment of costs and outcome must be made against current patterns of care, however ineffective.

Identification of costs and outcomes

The relevant costs and outcome measures will be influenced by the study perspective, the economic method and the healthcare procedure under analysis.

Measurement of costs and effectiveness

Good evidence of clinical effectiveness (or lack of it) is essential in an economic analysis. This should come from controlled clinical trials if possible. Prospectively collected cost data are also preferable, but caution should be taken in using cost data from clinical trials, as the care given is dictated by the trial protocol and may not match subsequent care delivered in routine situations.

Allowance for differential timing of costs and outcome

When comparing alternative healthcare programmes, the pattern of costs and benefits may differ significantly between options under review. For example, health promotion and preventative healthcare programmes incur immediate costs, but the main health benefits will only be realised many years into the future. Acute surgical treatment has a high immediate cost, but, if successful, gives immediate benefits to the patient. Long term maintenance drug therapy (e.g. for hypertension or peptic ulcer) may have a relatively low annual cost, but builds-up to a substantial amount over a lifetime. Simply adding total costs and benefits over time is not adequate because people have distinct preferences for delaying costs and bringing forward benefits—the flow through time matters as well as the absolute size of costs and benefits. To take into account this "time preference", in the economists' parlance, discounting of costs and benefits is necessary. In essence, this process reflects the declining importance in individuals' decision-making of outcomes occurring further and further into the future, by reducing their value (for analytical purposes) by a given percentage for each year ahead they are expected to occur. This produces estimates of costs and benefits valued at a single point in time, i.e. their present value. Thus, most decisions in lung cancer will not be discounted since the outcome (death) occurs rapidly. In breast cancer, discounting may be necessary for preventative or screening strategies, whilst it becomes less necessary the more advanced the disease is at the time of the intervention.

In the context of cancer, the importance of the timing of benefit flows would be seen in an evaluation of strategies to reduce the loss of quantity and quality of life from, for example, lung cancer. Health promotion activities to reduce smoking by young people could have a major impact but the benefits would not be seen for many years. Alternatively, marginal improvements in the effectiveness of radiotherapy could give small but immediate benefits. The complexity or such analyses is further complicated, since the data can be examined from the viewpoint of society as a whole or from that of the patient.

Assessment of risk and uncertainty

In the absence of firm data, it is often necessary to make assumptions in estimating the costs of healthcare. Clinical effectiveness data may also be subject to uncertainty. This must be explicitly recognised and tested using sensitivity analysis. This will identify the key variables and assmptions, which, if changed, will alter the conclusions of the study.

Presentation of results

This should be done in a way which clearly shows whether the original problem defined has been solved. Where uncertainty has been shown to be present, ranges rather than point estimates should be presented. The results should be placed in their context, indicating the possibility of generalisation to other settings. For those studies designed to help health service decision-makers to allocate limited budgets, the marginal costs and benefits should in particular, be shown.

AREAS OF DEBATE

The above brief review of methods of economic evaluation, and the framework within which such evaluations should be conducted, is designed to give the reader an overall perspective on the process. Subsequent sections will discuss the development of methods in more detail. Needless-to-say, while the above broad framework is generally accepted by economists, there is intense debate on many specific points.

For example, there are many different approaches to the measurement of quality of life and the utilities to be attached to the different health states. This makes the interpretation of cost per QALY results very difficult [10, 11]. Although economists lay great emphasis on the need for strong evidence of clinical effectiveness (although effectiveness may differ in the community setting) from randomised controlled trials, they are ambivalent about the desirability of collecting economic data in the same way [12]. Some economists also question the appropriateness of discounting health benefits (as opposed to costs) as time preference may not apply to non-monetary costs and benefits. Amongst those who agree that discounting of benefits should take place, there is debate as to whether the current Treasury recommended rate of 6% is appropriate [13, 14].

With regard to cost, the general view is that the main focus should be on direct costs of healthcare, rather than the indirect costs of illness to society through the loss of working days. This remains a much debated area, and recent work has shown that, even if lost working time is to be included as a cost, the value placed on it should be much lower than that used in earlier studies [15].

Despite the healthy debate on these and other issues, methods of economic evaluation are sufficiently well established to provide an essential contribution to the debate on the development of healthcare services. This applies equally to cancer care as to other areas.

The checklist below serves as a guide to the key points for those assessing the quality of a completed evaluation or for those planning a new study.

Problem definition

Clear definition of problem, statement of perspective of analysis and analytical technique.

Options

There must be a comparison of at least two options (including the status quo).

Identification of costs and outcomes

All costs and outcomes relevant to the study perspective must be identified.

Measurement of costs and outcome

The best possible source of cost and outcome data should be used, e.g. randomised controlled trials may provide this, although it is important to note that these may vary from the costs in the community.

Differential timing of cost and benefit flows

Future costs and outcomes must be discounted at a rate appropriate to the study perspective.

Uncertainty

The impact of this should be demonstrated through the use of sensitivity analysis, which analyses the impact of different cost and/or outcome estimates on the final results.

Presentation of results

These should show the marginal impact on cost and effectiveness of each option.

Source: Adapted from Department of Clinical Epidemiology and Biostatistics [16] and Coyle and Davies [17].

MEASURING AND EVALUATING OUTCOME

Although this section, for brevity, concerns itself solely with randomised clinical trials (RCTs), evidence of treatment outcome is often derived from other data. Whilst RCTs remain the "gold standard", if data from such trials is not available, then useful information can be gleaned from non-RCT sources. The

same problems, noted below, of RCTs, of course, still apply, and are sometimes amplified. It should be emphasised that, whenever possible, expensive new therapies should be assessed in RCTs large enough to provide a clear result.

DATA FROM CURRENT RANDOMISED CLINICAL TRIALS (RCTs)

Randomised clinical trials have become the "gold standard" by which medical care is evaluated [18]. Despite many thousands of RCTs, many questions remain unanswered and apparently open. Even more damaging may have been the failure of the medical community to identify therapies that do work, and acceptance of treatments of little or no efficacy. Systematic reviews and consensus meetings based on such data will help to overcome the problem of RCTs, which are too small to give a reliable answer.

The basic reason for our failure to produce clear answers to questions has been the small size of most RCTs. For example, an ongoing systematic review of all chemotherapy trials, including a no chemotherapy control, (published and unpublished) in non-small cell lung cancer has found only 70 such RCTs. These include approximately 10000 patients in all. (Non-small cell lung cancer overview, Collaborators' conference, Cambridge, 24–26 September 1993.) Some trials had more than two arms so that the mean number of patients per arm is approximately 100. During the time period of these trials, several million people died of lung cancer—less than 1% of patients with non-small cell lung cancer are entered into chemotherapy trials.

It has taken cancer clinicians a long time to realise that large therapeutic gains are unlikely to be achieved by current therapies. Only now are large scale trials being contemplated. An alternative, currently available, is the tool of systematic review (overview or meta-analysis). Such a systematic review may be based on an exhaustive analysis of individual patient data from all published and unpublished RCTs. This approach is likely to be more accurate than such an exercise simply carried out on published data. Publication bias, patient exclusion, length of follow-up and rigour of the original analysis are all potential sources of error in a systematic review of published trials. Preliminary data from Stewart and Parmar [19] suggest that such reviews may overestimate treatment effects when compared to individual patient data overviews.

The utility of systematic review of individual patient data has been clearly demonstrated. Lau and colleagues [20] examined all of the RCTs for the therapy of myocardial infarction. These were looked at cumulatively so that the evidence (or lack of it) for each treatment could be examined over time. As an example, they found 33 RCTs (36974 patients) of intravenous streptokinase carried out between 1959 and 1988. When examined cumulatively, there was initial evidence of a favourable treatment effect after eight trials (1973, 2432 patients, $P = 0.007$) even though the majority of individual trials were negative. By 1977, the P-value was less than 0.001 (14 trials, 4314 patients). Despite this, clinical trials continued to be published up to 1988—large scale trials such as ISIS being demanded before clinicians would accept this approach. Similarly, the recent individual systematic review of early breast cancer management has shown unexpected survival benefit for a therapy discarded two decades ago—oophorectomy in premenopausal women [21].

As well as providing data to answer questions, systematic reviews often demonstrate the need for new large scale trials, and it is to be hoped that genuinely large trials that provide answers and data with tight confidence intervals will be carried out. Systematic reviews are likely to be seen more frequently, and the Cochrane Collaboration, through its Cancer Network, will act as a co-ordinator stimulating, supporting and making available data from systematic reviews of cancer management [22, 23]. The Cochrane Collaboration is an international network which will provide databases of all past and ongoing RCTs as well as past and present systematic reviews. These will provide, on line, updated systematic reviews of individual questions. The activities will encompass the whole of medicine—a truly monumental task.

When carrying out an economic evaluation, it is essential that there are reliable data on efficacy. In the absence of a large definitive trial (the usual case), selection of specific trials or a cursory examination of the literature is likely to lead to a biased result. The difficulty of obtaining reliable data and avoiding bias is demonstrated by the finding that computerised searches of databases, such as MEDLINE, only identify 50% of published RCTs found by hand-searching the literature [24].

Additional problems with the current literature include the use of inappropriate controls (for instance, trials of chemotherapy in non-small cell lung cancer comparing chemotherapy regimes when the benefit of chemotherapy was unproved) and the failure to measure appropriate endpoints. Although survival is highly important, concentration on this has meant that there is little or no data on the outcome of therapy on symptom control or quality of life—even when treatment is entirely palliative.

MEASUREMENT AND VALUATION OF QUALITY OF LIFE*

Cancer is an important cause of illness, and accounts for a high percentage of crude mortality in Western countries. In the last 20–30 years, cancer treatment has developed rapidly. It is now multidisciplinary, incorporating surgery, radiotherapy, chemotherapy and/or immunotherapy. Treatment, particularly radio- and chemotherapy, may have major side effects. Consequently, patients and those involved in their care have wondered whether improved probability of survival, if it exists, outweigh the burden of these severe side effects in all cases. Therefore, not only life years gained, but also the quality of those years is an important issue.

Objectives of Quality of Life research

The usual objective of carrying out Quality of Life (QoL) research is to gain insight into the reaction of patients to cancer and its treatment, and the inter-relations of these different reactions with the overall QoL. Furthermore, data concerning the impact of a certain therapy on the QoL may increase knowledge about the effectiveness of such therapy [25].

QoL studies generally give a description of the health status. Depending on the type of study, they deal with more or less objective features of health. Most studies incorporate the physical and psychological aspects of life, sometimes complementing these with aspects of social functioning. From the economic perspective, the objective of studying QoL is to determine the contribution of changes in QoL to a summary measure for the outcome of an intervention. When there is more than one effect parameter (e.g. survival and health status improvement) or when outcome is measured in more dimensions of QoL, a value judgement (e.g. utility weight) should be given to each of these parameters or QoL dimensions in order to arrive at such a

* We thank F.F.H. Rutten and G.J. Bonsel for permission to use some ideas from a jointly authored article.

summary measure. This makes it possible to make comparisons across different technologies and across different disease categories. The importance of different outcome parameters will depend on the disease type. Quality of life, for instance, assumes greater importance in advanced breast cancer, whereas survival is of paramount importance in malignant teratoma.

Measurement instruments for description and valuation of QoL

Descriptive instruments can be divided into two main categories: the generic instruments and disease-specific questionnaires [26]. Generic instruments are developed to measure all health dimensions. They can be used over a wide range of disease categories, and are, therefore, suited to function as generally applicable descriptive instruments for measuring QoL improvement in economic evaluation. In this respect, they provide a reference for quantitative assessment of QoL changes in economic evaluation, and have been used in many economic appraisals. Generic instruments most frequently used are: the Nottingham Health Profile (NHP) [27], the Sickness Impact Profile (SIP) [28], the abbreviated RAND–Medical Outcome Study questionnaire (RAND–MOS 20) [29] and the newly developed Short Form-36 (SF-36) [30, 31]. Details and references for key instruments are included in Appendix 1.

Disease-specific instruments measure health problems specific for an illness or disease category. These instruments measure specific QoL aspects typical of a particular disease, and are, therefore, more sensitive to changes in the health states of patients with that disease. Using cancer as an example, these instruments can be divided into cancer-specific questionnaires (for all cancers) and specific cancer questionnaires (for one type of cancer). The following cancer-specific instruments are often used: the Functional Living Index for Cancer (FLIC) [32], Spitzer QL index [33], European Organization for Research and Treatment of Cancer Core Quality of Life Questionnaire (EORTC Core QoL Questionnaire) [34], Cancer Rehabilitation Evaluation System–Short Form (CARES-SF) [35, 36] and the Rotterdam Symptom Checklist (RSCL) [37]. Examples of specific cancer instruments are: the Breast Cancer Chemotherapy Questionnaire [38], the Performance Parameters for head and neck cancer [39] and the Linear Analogue Self-Assessment (LASA) of Voice Quality for laryngeal cancer [31, 40].

For the description of QoL in economic evaluation, generic instruments suffice for overall comparison and testing, but the inclusion of specific instruments may be advisable, particularly for reasons of explanation, if changes in dimensions are small or conflicting. It is not always possible to gain insight into how changes in health have taken place if generic instruments have been used.

When strategies score differently on different dimensions and when one needs a summary measure of outcome, comparable across a range of interventions, it is necessary to get valuation of health states in one single dimension representing an overall judgement of QoL of a health state. This then allows calculation of "quality adjusted life years", where a life year gained is "adjusted" using this value judgement for QoL. To acquire valuations, respondents are requested to value the particular health state. The values are derived from descriptors and usually range from 0 to 1 (0 = "worst imaginable health state", and 1 = "best imaginable health state"). By means of modelling techniques, a multi-attribute utility function may be derived to predict values of any compostion of health states. Methods of acquiring valuations are: Direct Rating (e.g. determining a point on a line with clearly defined endpoints), Standard Gamble (choosing between options under uncertainty) and Time Trade-Off (choosing between options with different durations of a particular health status) [41, 42].

Furthermore, there are different basic approaches to measuring the utility of health status. One approach is to find people with the particular health state, and ask them to value such health states in terms of a utility. Using such subjective utility assessments is most in line with welfare theory. Another approach is to describe the condition, usually by an abbreviated written scenario, to people who do not have the condition and ask them to make the appropriate valuation [41]. These people are thought to represent public opinion. Other approaches include, for instance, using the opinion of the patient's family or that of health professionals.

Which approach is preferable depends on the viewpoint of the study. When a study is conducted from the societal viewpoint and is pertinent to public policy decision-making, the appropriate utilities are those of the general public. If a fixed set of utility values is available and if a generic instrument is used to measure health outcomes in an economic evaluation, this adds to the comparability of results from economic evaluations and fits with the public sector context of decision-making, supported by results from an economic evaluation. The problem with soliciting value judgements from persons without relevant disease experience is that the person often does not properly understand what the described health state is like.

Choice of measurement instruments

In clinical trials, the choice of the measurement instrument depends on criteria such as psychometric features (internal consistency, reliability, validity, etc.), specific patient features, feasibility and the possibility of gaining insight into the QoL aspects of the disease itself and into the side effects of treatment. QoL measurement in economic evaluation should allow comparison across different technologies and diseases. This requires the inclusion of a generic, a cancer-specific measurement instrument and a valuation instrument. The precise combination of instruments depends on the required psychometric and specific patient features and on feasibility (see Appendix 1 for an overview).

Interview schedule and follow-up time

Ideally, the data collected should describe QoL of patients before, during and after treatment. The choice of the time scale depends on the natural history of the disease and on the timing and pattern of treatment. During intensive treatment periods, it may be important to have frequent measurement of QoL, and very occasionally intervals, as short as 1 or 2 days, may be desirable [48]. If there is a survival benefit, it is important to take into account QoL during life years gained. In this respect, it is necessary to make a distinction between curative and palliative treatment modalities. In both modalities, QoL measurement is very important, but the outcome indicators differ. In palliative treatment modalities, QoL is the single most important outcome indicator. For curative treatment modalities, the most important indicator is the assessment of QoL of life years gained, which is estimated by a separate observation of survival and QoL.

As economic evaluation generally requires lifelong follow-up, a modelling approach may be necessary to predict lifelong consequences. By specifying a simulation model, or a MARKOV-model defining various states characterised by health status and/or specific treatment (associated with a cost-profile),

the follow-up period of the actual clinical trial may be reduced. It will be assumed that each state can be characterised by a unique quality of life. The following disease states in cancer may be distinguished: disease-free, partial remission, relapse (symptom-free), relapse (symptoms) and terminal state. The measurement and valuation of QoL of patients in the terminal phase of cancer may be difficult, because these patients may be "off protocol" and difficult to approach. Acceptable compliance rates require motivation of the patient, sophistication of the procedure and a permissive attitude towards the ideal interview schedule.

It should be obvious from the above discussion that QoL measurements are not easy to perform nor to interpret. Drummond [49] for example, includes a thorough discussion of the (dis) advantages of such measurements; e.g. the QoL estimates can only be as reliable as the underlying mortality and morbidity data, different measurement methods may yield different results, QALYs value quantity (life years) and quality of life in an identical way. Despite the methodological and practical difficulties of outcome measurement, such measures can clearly contribute to improved resource allocations decisions, if applied and interpreted correctly.

MEASURING AND VALUING RESOURCES

Introduction

The basic idea of economic evaluations is that resources are limited and have an alternative use. Resources used for a specific cancer treatment, e.g. radiation for breast cancer patients, could be used in other cancer treatments, in non-cancer treatments within the health care sector or even outside the health care sector. The true economic cost for a resource is the value this resource would have in its best alternative use, i.e. the opportunity cost. The monetary cost is made up of two components, volume (quantity) as well as valuation (pricing). In most cases, the market price of a resource is the relevant value in monetary terms, but in some cases the price does not correctly reflect the value of the alternative use; sometimes market prices do not even exist. However, a non-existing market price does not mean that the cost is zero; a "shadow price" which reflects the value of the resource in its best alternative use has to be used.

When valuing and reporting the costs, it is useful to identify the quantity of the resources used, and the price separately, before reporting the cost (quantity × unit price).

Types of costs

In identifying the relevant costs, it is useful to make a distinction between direct and indirect costs. Traditionally, direct costs are defined as the resources involved in the treatment of the disease. Typical direct medical costs include costs of staff time, drugs, equipment and buildings. However, usually more resources are associated with the treatment than pure medical costs controlled by the healthcare sector. Direct non-medical costs, such as the patient's travel costs, housekeeping costs, extra diet costs and extra costs for the patient's family, are relevant and ideally should be incorporated. The importance of non-medical costs varies between treatments and patient groups. For example, in cost comparisons between home care and hospital care for terminally ill cancer patients, it is important to include non-medical costs. As home care requires participation of the patient's family in the treatment process, there are good reasons to assume that non-medical costs, to a larger extent, are associated with home care. Indirect costs represent resources lost due to the treatment and illness. Traditionally, indirect costs are valued as the changes in the productivity of the patient in the labour market—the human capital approach. This means that the indirect costs are estimated as the value of the goods and services not produced due to cancer. However, there are good reasons to include changes in the productivity of leisure activities as well; otherwise time costs for patients who do not participate in the labour market, e.g. retired people, children and housewives, are estimated to be nil. By using the "willingness to pay" approach, these time costs can be estimated, either by explicitly asking individuals how much money they are willing to spend for achieving a better productivity on leisure activities, or, if possible, by deriving the value by observing their behaviour (revealed preferences).

The indirect costs primarily stem from changes in the patient's health status. The patient's quality of life changes due to changes in labour productivity and leisure activities might, at least to some degree, reflect the effectiveness of the intervention measure. Hence, there is a risk that productivity changes will be included both in cost and outcome. In order to avoid double-counting, indirect costs have to be handled with care.

When presenting cost data, it is useful to make a distinction between fixed and variable costs. Fixed costs do not change with the volume of the activity (e.g. the number of procedures and patients), while variable costs do. In identifying the costs per unit of activity (e.g. per patient), it is further necessary to distinguish average costs from marginal costs. The marginal cost refers to the extra cost of 'producing' (e.g. treating) one additional unit (e.g. patient). In situations where the fixed costs are all relatively high (e.g. in radiotherapy), the marginal cost for an extra patient will be much heavier than the average cost per patient.

Perspective and scope of costing

The costs so far have been classified from the type of resources. However, the classification can be performed from the perspective of who pays. This perspective is important because the costs might vary. In order to correctly calculate the relevant costs of a healthcare intervention, the viewpoint of the analysis should first be identified.

In a traditional cost–benefit analysis, which takes "society as a whole" as the viewpoint, all costs, direct as well as indirect, should be included regardless of who pays and regardless of who benefits. From this viewpoint, a treatment is "profitable" if the total benefits exceed the total costs. However, other perspectives may be relevant. From the perspective of national healthcare system, the objective is to allocate resources to achieve the maximum health benefit within the limits of resources available for the healthcare sector. Costs of resources other than those controlled by the healthcare system are not included. From the hospital viewpoint, only costs for resources paid by the hospital are relevant; the insurer takes account only of costs charged to him, and from the patient's viewpoint only the costs for the patient are of interest. Which viewpoint is relevant, depends upon the aims of the study. Obviously, if the perspective is not clearly identified, cost data may be used in a misleading way. The fact that a cost is not paid by, for example, the hospital does not mean that the cost is zero. Too narrow a perspective might lead to suboptimal health care resource allocation decisions.

However, the societal perspective is not without controversy either. In most countries, the provision of healthcare is also guided by ideas of justice and equity which imply, for example, that priority setting must not be based on the patient's socio-economic status. However, the amount of indirect costs clearly

depend on the status of the patient. This is hence an argument for excluding the indirect costs and taking into account only the costs paid by the national healthcare system, i.e. the perspective is limited only to one sector. These considerations emphasise the need for a systematic presentation of the data so that others can analyse the data in different ways.

Time scale

When calculating the costs of a healthcare intervention, it is important to take a sufficiently long timespan into account. Many cancer patients sooner or later relapse and will need further treatment. Ideally, the timespan should be long enough to include the costs of all these therapies—preferably the patients should be followed until death. Too short a timespan can lead to doubtful or incorrect conclusions, since part of the relevant costs, on which conclusions should be based, would be omitted. Obviously, it is not always possible, nor feasible, to take such a long timespan, but one should be aware of the possible erroneous conclusions, based on an analysis of short duration.

Collating data

Cost-effectiveness studies most frequently assemble information on the resource implications of interventions by: (1) retrospectively using data from individual RCTs, literature surveys, or epidemiological and clinical advice, to derive probabilities attached to treatment pathways and events along them; (2) collecting descriptive resource data from previous costing studies or published unit cost averages which enable these pathways and events to be costed; and (3) evaluating the identified resource implications of each treatment path and associated probabilities.

It is possible to collect economic data prospectively in an RCT, and to perform economic analysis alongside the clinical and epidemiological aspects of the trial. The potential benefits of doing so have been set out by Drummond and Stoddart [50] and by Mugford and Drummond [51]. They argue that prospective collection of cost information may be desirable if the resource consequences of the treatment, gaining widespread acceptance, are likely to be substantial, or if diffusion is thought likely, or if resource constraints are likely to force decision-makers into choices between alternatives. Using these three criteria, Mugford and Drummond found that almost half of 100 perinatal trials sampled should have, but did not, incorporate economic analysis.

The main difference between retrospectively and prospectively collected data is that prospectively collected data offer better opportunities to examine variations in cost between patients, and, therefore, to report the distribution of cost data, thus allowing opportunities to examine the statistical significance of cost differences between patient groups.

To ensure consistency and comparability of cost data between study settings, a number of steps should be taken:

(i) where possible, resource data should be presented in the form of volumes, such as in-patient days, number and length of consultations, or drugs consumed and prices separately. Summary measures such as drug-days (number of drugs × number of days on drugs) are more useful than no information. Where space does not allow presentation of such data, authors should still make clear that it is available. Such volume data facilitate comparison between practice settings.

(ii) Where these resources are valued, the basis of this should be made clear. In particular, it should be stated whether valuation is based on prices charged or on actual resource costs; if the perspective of the analysis is that of the healthcare financing institutions, charges are relevant, but if the viewpoint of the hospital is taken, it is wrong to "approximate" the cost of an intervention with charge data. In this situation, the actual costs of the resources used should be calculated; whether any consumption taxes, such as value added tax, are included or excluded; which year and currency the values are expressed in; and the basis for any adjustment in the reported year (for example, whether a consumer price index or some other form of index was used to inflate or deflate values).

(iii) If resources are valued in a currency other than the currency of the territory in which the study was performed, or if the study covers more than one territory, the method used to translate valuations into other currencies (in particular, the date and basis such as exchange rate or GDP purchasing power parity) should be explicit.

(iv) Where cost-effectiveness ratios are being reported, the numerators and denominators as well as the ratios should be attached to the results.

REVIEW OF CURRENT PRACTICE: ECONOMIC EVALUATION—A GENERAL REVIEW

Current concerns over exploding healthcare costs and scarce resources have led to a steady increase in studies wholly or partly containing economic evaluations. Elixhauser counted a total of 88 CBAs and CEAs carried out in 1979. By 1990, the annual average was 247, with the overall 1979–1990 total being 1802 studies [2]. This increase, on the one hand, has led to attempts at cataloguing economic studies in databases, and, on the other, at attempts to develop methods of systematically reviewing a group of studies on the same subjects and of generalising their results. Backhouse and associates [1] have published the content of their database, containing 1987 economic studies published between 1964 and mid-1992. A similar database has also been recently published by Elixhauser for the period 1979–1990 [2]. Both publications provide a very useful starting point for any search of the economic literature as they incorporate subject and author indices.

Such a plethora of available studies would, on the surface, indicate an adequate and reliable coverage of the main topics of medical practice. To assess such qualities, however, a systematic review of the topic is needed, but at present there are very few such reviews of economic evaluations. Those that exist all point to considerable gaps in the overall reliability of methodology used in available economic evaluations.

Gerard reviewed 51 cost–utility analyses (CUAs) carried out in 14 different countries and found source validity in only 41% of studies [11]. In discussion of her study, Gerard expressed doubts as to the methodological validity of aggregating results of CUAs in so-called QALY league tables, a process which has the aim of indicating "best buy" procedures for clinicians and managers. Gerard's doubts are shared by Mason and associates [52] and Drummond [53].

This far from satisfactory picture is further confirmed by another study which carefully reviewed methods used in 77 economic analyses with different subjects [54]. The authors recommended that readers should be careful in interpreting results and should scrutinise methods used carefully. Such a suggestion, however, is unlikely to be ideal for busy clinicians with no training in health economics.

Poor technical execution was also found by Adams and

colleagues [55] in the crucial area of economic analysis nested within RCTs. Such a finding raises the real problem of the waste of scarce resources in ill-conceived and poorly executed economic evaluations, and of a clear requirement to standardise international economic methodology [12]. Additionally, only 0.2% of RCTs contained economic analyses.

Jefferson and Demicheli reviewed 90 published and unpublished studies on the economics of introducing vaccination against Hepatitis B (HB) [56]. They found that 92% recommended vaccination on economic grounds, but a sizeable minority (38%) showed major methodological inconsistencies. The authors also found very wide variations in basic variables used in the reviewed studies, for instance, the within-country incidence of HB used to assess the economic burden of the disease in the same population and in the application of basic rules of economic evaluation. An example of the confusion on the subject is given in Table 1, where the spread of costs due to HB is shown, converted into U.S. dollars. Such a large variation is unlikely to have been caused wholly by different patterns of resources used.

The table also illustrates a possible way of attempting to pool results from different economic studies in order to reach a verdict on whether an intervention is economical or not. At present, efforts are being made to develop such a methodology further [57].

There is probably only one overview on the economics of cancer literature where an attempt has been made at generalising results of different studies [58]. In this overview, Brown and Fintor examined the cost-effectiveness of screening for breast cancer and used a computerised model to "reconcile" differences in cost-effectiveness ratios from two studies. Encouragingly, they concluded that cost-effectiveness estimates for breast cancer screening reported by different studies are not as arbitrary as they seem [58].

In summary, although the science of systematic reviews of economic evaluations is in its infancy, the current methodological validity of international literature is open to question, and readers should beware of uncritically accepting results from published works without prior scrutiny by a health economist.

REVIEW OF STUDIES SPECIFIC TO CANCER

As discussed above, there has been an increase in the number of studies on the cost-effectiveness of cancer interventions—be they screening, treatment or prevention. This section provides a critical review of the methodological standards of a selected number of recent studies pertaining to a specific common disease area—cancer of the bronchus and a specific intervention type, screening. This part of the report focuses on those studies which include both elements of costs and of effectiveness. The purpose of this review is to highlight methodological inadequacies in these studies, using the checklist of guidelines discussed earlier as a reference standard. Thus, they are selected as examples of problems and it is important to note that, because of this, the strengths of the studies have not been highlighted.

Table 1. Cost per case due to HB infection, summary of descriptive statistics

	Type of cost		
	Direct	Indirect	Indirect intangible
Number of studies	36	20	11
25th percentile	650	1.834	14.164
Median	1.793	3.584	54.272
75th percentile	3.768	12.473	127.592

All costs are in U.S. dollars adjusted from year of publication to 1993 values using the U.S. Retail Price Index. Source: Jefferson and Demicheli. Is vaccination against Hepatitis B efficient: a review of the world literature. *Health Econ* 1994, 5, 25–38 [56].

Economic evaluations of treatment for lung cancer

After a reasonably detailed literature search, only a few studies concerned with the treatment of lung cancers were identified. The majority of these studies pertained to treatment with chemotherapy, although one concerned the cost-effectiveness of radiotherapy and two concerned the staging of lung cancer. The methods and results of these studies are outlined in Appendix 2a. Whilst they have made a major contribution to the literature, there is a need for improved data. Since this paper is critically examining the use of economic evaluation in cancer, the rest of this section concentrates on the weaknesses of these studies. The major methodological failings of these studies are as follows.

–The evidence of the effectiveness does not always appear appropriate. The preferred source of evidence of the effectiveness of interventions is results obtained from a large, randomised, clinical trial. Most of the selected studies obtained evidence of the effectiveness of interventions from RCTs [see 61, 63]. However, in other studies, it is unclear from where evidence of the effectiveness of the alternatives evaluated was obtained [65]. Even when data were obtained from RCTs, it is not clear whether these are representative of the literature.

–Not all relevant costs of treatment are included in some studies, thus questioning of the reliability of study results. In an economic evaluation, all the costs relating to treatment should be considered. What the relevant costs are depends on the perspective of the study. However, in the study by Dillman and associates [59] both the costs of treating side effects and of follow up care were excluded. This may be of importance, as, for instance, the treatment which was found to be cost-effective did involve greater hospitalisation due to side effects.

–Certain studies did not include all relevant consequences, especially the effect of treatment on the quality of life of individuals. Quality of life is often inappropriately omitted from economic evaluations of cancer therapies, for example, in the study by Rees [65] of the cost-effectiveness of various cancer therapies, including palliative care, for carcinoma of the bronchus, the benefit from treatment was measured solely by the prolonging of life and the effect of treatment on the quality of a patient's life was disregarded. Similar concerns exist over other studies [59, 63] as the omission of quality of life considerations questions the validity of study results. However, the study by Goodwin and associates [61] does illustrate how the credible valuation of the quality of life of patients can be incorporated within a study. QoL is all the more important since the great majority of treatment in this disease aims at palliation rather than cure.

–The methods used by most of the published studies to measure and value resource use are not ideal. Most of the studies reviewed adopted a methodology similar to Pashko and Johnson [64], whereby the average use of resources attributable to a procedure are measured, and an average cost per unit for each form of resource employed is calculated. An average resource cost per procedure can then be estimated. In the study by Rees [65] the

methods for the measurement of resource use were not made clear. An alternative, more accurate, method is to monitor a sample of patients throughout their course of treatment (either prospectively by the use of case record forms, or retrospectively by examining hospital charts), measuring resource use as it is employed, and from this analysis derive a marginal cost for the procedure. The studies by Houston and associates [62] and Eddy [60] adopted this approach.

–In almost all of the studies selected, there is a failure by authors to include any analysis of the sensitivity of their results. However, the studies by Goodwin and associates [61] and Jaakkiminen and associates [63] include several such analyses.

Economic evaluations of cancer screening programmes

There are relatively few comprehensive economic evaluations of cancer screening programmes or policies. Most pertain to the area of breast cancer screening, due to the important body of literature on randomised and case controlled clinical trials in this field. Economic analyses of cervical cancer or colorectal cancer screening initiatives are more scarce. The majority of economic analyses of breast cancer screening are based on theoretical models, which combine efficiency results from former clinical trials with actual costs of screening programmes or clinical practice. Analyses of cervical and colorectal cancer screening programmes either base estimates on theoretical models or are limited to very small numbers of cases.

Among the studies presented in Appendix 2b, the following comments can be made regarding methodology.

–The evidence for effectiveness is not always based on RCTs and, if so, the results of RCTs are not readily applicable to the situation. Joseph and associates [66] for example, based their measurement of the HemoQuant test's effectiveness in screening for bowel cancer on a sensitivity value obtained in a study of only 30 patients. Obuko and associates [75] used a combination of sources for estimates of efficiency, yet failed to address the possible discrepancies in the calculation of these estimates. Van der Maas and colleagues [72] and de Koning and associates [73] suitably addressed the problem of the applicability of clinical trial results, by combining results from two respected clinical trials for the measurement of outcome in their analysis. The reliability and applicability of results from clinical trials performed several years prior to the economic evaluation of a screening programme must be assessed with respect to possible changes in technology and in population demographics, which could limit their relevance to the current situation.

–A comprehensive description of alternatives is not always provided, rendering interpretation of results and applicability of conclusions difficult. In their evaluation of cervical screening policies, Parkin and Moss [69] analysed seven alternative policies in terms of cost-effectiveness, yet the distinction between the alternative policies is, at times, unclear throughout the text. In the study by Johnson and colleagues [67], no alternative situation was discussed. In the study by Mandelblatt and associates [68], the number of screened patients included in the study was very small (7), and the logical comparison group, non-screened patients of similar age, was excluded from all analyses.

–Economic studies of screening policies often fail to include all relevant costs of screening (especially capital costs) and all relevant consequences (especially morbidity). Little consideration is given to the effect on patients' quality of life. Indeed, the only study which incorporated quality of life measurements was the study by de Koning and associates [73] on breast cancer screening, all other studies being on cost-effectiveness. The study by Joseph and associates [66] considered mortality as the only possible negative consequence of Hemoccult and HemoQuant screening, and limited costs to the charge for the screening procedure, excluding all capital costs. Johnson and colleagues [67] took the total cost of the programme and divided it by the number of cancers detected in the study group to obtain a cost-effectiveness ratio. Inadequacies in the estimation of costs and outcomes question the credibility of economic evaluations.

–Costs and consequences are not always adjusted for differential timing. Most of the studies in Appendix 2b include appropriate discounting of costs and consequences of screening, although no discounting was done in the case of Johnson and colleagues [67], and only outcomes were discounted in the study by Parkin and Moss [69].

–The costs and outcomes of different screening alternatives are rarely studied in the context of an incremental analysis, and only total sums and average costs are presented and discussed. The articles by Van der Maas and de Koning are two examples of studies which included detailed incremental analysis in their estimates of cost-effectiveness.

–The discussion in these papers presents the authors' recommendation for screening, yet fails to discuss its implementation or to place the study results within the context of healthcare priorities. Since the goal of economic evaluation is to aid policy makers in choosing among competing policies, the conclusion of most studies either advocates for or against the implementation of a given screening policy. However, very few studies mention which difficulties might arise in the implementation of policies, or which competing policies should also be analysed in terms of cost-effectiveness. An example can be found in Mandelblatt's recommendation to implement cervical cancer screening programmes for elderly women. The author never addresses the issue of cervical screening in younger women (under 70 years) and neither does she compare the cost-effectiveness ratio in women over 70 years to that obtained in studies of younger women.

Other, also methodological, problems with breast cancer screening studies are discussed further by Skrabanek [77].

USE OF RESULTS OF ECONOMIC EVALUATIONS

Use by policy-makers

The policy-maker's aim is that of choosing and implementing a "best buy" strategy to deal with a particular aspect of cancer, be it prevention, treatment or care for cancer patients and their families. This is particularly important in the field of neoplastic diseases which requires a considerable amount of resources. Such a requirement is especially important when the policy-maker is in a purchasing role, and must make best use of available resources in commissioning a new service, or must look at making best alternative use of available resources.

If a decision-making process is to be of good quality, it must be based on reliable assumptions. If the reliability of international economic literature is open to question, where should the decision-maker go for advice?

Probably the quickest way of getting such advice is to ask an economist, in other words, somebody who not only has the technical knowledge required to provide sound advice but also whose everyday work is economic evaluation. If they are not

involved in the specific field of enquiry, a local economist should be able to direct the enquirer to economists active in the field. Alternatively, some national or international organisations are beginning to develop a strategy to review and disseminate results of economic evaluations. These include the Cochrane Collaboration, through an increasing number of centres around the world, the University of York, U.K. and the offices of Health Technology Assessment, being established in many countries (e.g. Sweden, Canada and Australia). The EORTC has, in addition, recently opened a department for economic evaluations, based at the Data Centre in Brussels.

Use by clinicians

The results of economic evaluation of therapies should not be regarded as being simply in the domain of planners and "purchasers". In our cost-constrained world, decisions on healthcare spending may involve everyone. As expensive new therapies develop, healthcare planners are not always going to make hard decisions on whether to pay for them. Often they will allow clinicians to use such treatments, provided that they are paid for out of the current budget. Clinicians will, in this situation, need to understand any available economic evaluation and be able to compare it with economic evaluations of other therapies they provide. Only then will they be in a position to decide which treatments to use. In cost-constrained health systems, management does not have sufficient expertise or manpower to make decisions based on economic evaluations on the myriads of interventions in many areas of medicine. Inevitably, this role will often be devolved to the individual clinicians or departments.

Use of the results of economic evaluations

Comprehensive cost-effectiveness studies can be expensive to perform, and may be hard to repeat if attached prospectively to a large randomised controlled trial. Given the comparatively small number of studies performed each year, and the likelihood that their results will have a limited lifetime before being overtaken by changes in efficiency or by technical progress, it is important to find ways of making maximum use of their results.

The transferability of the results of a cost-effectiveness study may be considered in a number of ways:

(1) The perspective of the study. A study may have been performed from one of a number of perspectives, such as that of a third party payer, the health sector, the public or private subcomponent of the health sector, or society as a whole. This may restrict the transferability of the results to another setting, with different divisions of responsibility or with different information requirements. Transferability would therefore be facilitated by reporting, where possible, the results from a standard perspective, which may be the baseline perspective adopted by the study or a supplementary part of the analysis. The health sector as a whole would probably raise the fewest problems of transferability, although this perspective may also pose problems at least for studies in, for example, the field of terminal care or home care, since whether such types of care fall under the 'healthcare' or 'social services' may vary widely between countries. Further work is required in this area before prescriptive guidelines can be agreed.

(2) The care level of the study. A cost-effectiveness study performed in a teaching hospital or trial setting may be based on substantially higher unit costs than would be the case in a non-teaching hospital. Therefore, some indication of the extent to which the costs reported can be regarded as typical of the health sector as a whole would be useful.

(3) Similarly, there may be regional differences in costs within a country that result in the reported results being typically low or high for the country as a whole. Some indication of this would also be useful.

(4) The time the study was performed. A cost-effectiveness study may make use of information on costs, treatment paths and outcomes, all from different sources and potentially different times. These data, and in particular the technology and organisation of care in the treatment paths, must be made explicit in order that obsolescence of the study can be judged.

(5) The country of the study. To facilitate wider use of the study, it may be tempting to consider ways of transferring results from one country to another. To facilitate this, it is essential that the guidelines set out above in "Measuring and valuing resources", as well as items 1–4 on transferability are closely adhered to. Efforts to achieve this are in their infancy, and more effort is needed on the methodologies and empirical procedures relevant to international transferability.

RECOMMENDATIONS

(1) When any clinical trial is being planned, the potential economic outcome of the intervention should be considered. The requirement for economic evaluation should also be examined. Even when this is considered inappropriate, thought should be given to collecting basic data which might be used for a later economic evaluation.

(2) If an economic evaluation is being planned, the advice of a health economist should be sought at the outset.

(3) Funding for RCTs is limited. Ask yourself is the design efficient; is it likely to answer the question or make a useful contribution; would the money be better spent elsewhere?

(4) Do not try to replicate locally an economic evaluation study you have found in the literature without involving a trained economist. Listings of health economists will be available from national Offices of Health Technology Assessment and in Europe from the EORTC offices in Brussels.

(5) Where possible, resource data should be presented in the form of volumes, such as in-patient days, number and length of consultations or drugs consumed. After the identification of the cost of each of the resources, the costs can be calculated. Authors submitting cost-effectiveness studies should ensure that, where available, results report ranges and confidence intervals around the cost data used, and that sensitivity analyses on the main dimensions of the study are performed and reported. Results in the form of a single point estimate of cost-effectiveness for a given intervention should be discouraged. Numerators and denominators of cost-effectiveness ratios should be reported.

(6) There is a need for dialogue between clinicians, policy-makers and health economists. The present gap is liable to lead to inappropriate health evaluations and ill-informed discussion. The end result may be poorer healthcare for patients.

(7) There is a need to ensure methods used in studies are clear. In all published studies, the methods used in measurement and evaluation of all costs and consequences must be made clear. This will allow policy-makers to determine whether the study's results are applicable to their policy arena, and

to help other researchers avoid duplication of research and to facilitate the extrapolation of studies to other settings.

However, this has to be balanced with editors' desires to constrain the space available to report such studies. Often when researchers are criticised for not accurately displaying the methods adopted, this is the result of editors being parsimonious with the space available to present the work and not with any desire on behalf of the authors to conceal their methods.

Editors and authors must consider how to make a detailed description of methods employed more available. Study methods could be made available if they are published as an appendix to the main article. However, given space constraints, this is an unlikely option. As the trend to make journals available on-line gathers pace, one alternative is to include within the on-line version of the journal details of the methods employed. Another alternative is for editors to insist that authors make available a full description of the methods to all those requesting them. Even if the full methods are not contained in the main body of text within their journal, they must still be made subject to peer review. There is a need for a European Workshop to discuss standardisation of study methods and reporting, with a special emphasis on transferability.

(8) Editors should ensure that a health economist is included among referees for all papers containing an economic evaluation component.

(9) Co-ordination of teaching and education. In order to achieve optimal understanding and co-operation between health economists, clinicians and policy-makers, each party must be able to appreciate the objectives, methodology and priorities of the others' vision of the healthcare sector. This mutual understanding and respect can be achieved by integrating elements of health economics into medical school curricula, and similarly by teaching health economics students basic medical concepts. On a European Community level, a forum for dialogue between all parties can be achieved in the form of symposia, guidelines recommending the inclusion of economic aspects in evaluations of oncological interventions, or encouragement of multidisciplinary approaches in cancer care research.

(10) Economic evaluation of cancer care is complicated by the poor quality of the evidence currently available on treatment efficacy. The lack of data on quality of life and symptom control further complicates assessment of efficacy in what is often palliative therapy.

(11) Analyses of the cost-effectiveness in cancer care must be based on economic evaluation studies. As there are, so far, few good quality studies, the information and knowledge from existing studies is of great importance. Where possible, health economics should provide information on ongoing studies to a central database. The EC has an important role to encourage this, particularly overcoming the problems of recognising and translating non-English language papers.

(12) Health policy-makers should be considering the need for an economic analysis of any new expensive therapy. They may choose only to recommend such therapies for licensing and widespread use if the economic evaluation is favourable.

BIBLIOGRAPHY ON METHODS OF ECONOMIC EVALUATION

Department of Health. *Assessing the Effects of Health Technology: Principles, Practice, Proposals.* Department of Health, London, Department of Health, 1992.

Drummond MF, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes.* Oxford, Oxford Medical Publications, 1986.

Drummond MF, Teeling-Smith G, Well N. *Economic Evaluation in the Development of Medicines.* London, Office of Health Economics, 1987.

Luce B, Elixhauser A. *Standards for the Socioeconomic Evaluation of Medical Products.* Springer, 1990.

* Boyle MH, *et al.* Economic evaluation of neonatal intensive care of the very low birthweight infants. *N Engl J Med* 1983, **308**, 1330–1337.

* Drummond MF, ed. *Economic Appraisal of Health Technology in the European Community.* Oxford, Oxford University Press, 1987.

* Drummond MF, Davies L. Economic analysis alongside clinical trials: revisiting the methodological issues. *Int J Technol Assessment in Health Care* 1991, 7, 561–573.

* Williams AH. The economics of coronary artery by-pass grafting. *Br Med J* 1985, **291**, 326–329.

1. Backhouse ME, Backhouse RJ, Edey SA. Economic evaluation bibliography. *Health Econ* 1992, **1**, (suppl), 1–235.
2. Elixhauser A, ed. Health care cost–benefit and cost-effectiveness analysis (CBA/CEA). From 1979 to 1990: a bibliography. *Med Care* 1993, **31** (suppl), JS1–149.
3. Lowson K, *et al.* Costing new services: long-term domiciliary oxygen therapy. *Lancet* 1981, **1**, 1146–1149.
4. Kesteloot K, Dutreix A, Scheuren E van der. Quality assurance in radiotherapy—economic criteria to support decision making. *Int J Technol Asses Health Care* 1993, **9**, 274–285.
5. Kesteloot K, Dutreix A, Scheuren E van der. A model for calculating the costs of *in vivo* dosimetry and portal imaging in radiotherapy departments. *Radiother Oncol* 1993, **28**, 108–117.
6. Weltens C, Kesteloot G, Vandevelde R, Van den Bogaert W. Comparison of plastic and Orfit masks for patient head fixation during radiotherapy; precision and costs. *Int J Radiat Oncol Biol Phys* 1995, in press.
7. Ludbrook A. A cost-effectiveness analyses of the treatment of chronic renal failure. *Appl Econ* 1981, **13**, 337–350.
8. Johanessen M, Jonsson B. Economic evaluation in health care: is there a rate for cost-benefit analysis. *Health Policy* 1991, **17**, 1–23.
9. Hutton J. Cost–benefit analysis in health care expenditure decision-making. *Health Econ* 1992, **1**, 213–216.
10. Gerard K. Cost–utility in practice: a policy-maker's guide to the state of the art. *Health Policy* 1992, **21**, 1–31.
11. Gerard K. Cost–utility in practice: a policy maker's guide to the state of the art. *Health Policy* 1992, **21**, 249–279.
12. Drummond MF, Brandt A, Luce B, Rovira J. Standardising methodologies for economic evaluation in health care. *Int J Technol Assessment Health Care* 1993, **9**, 26–36.
13. Parsonage M, Neuberger H. Discounting health benefits. *Health Econ* 1992, **1**, 71–75.
14. Cairns J. Discounting health benefits: another perspective. *Health Econ* 1992, **1**, 76–79.
15. Koopmanschap MA, VanInevald BM. Towards a new approach for estimating indirect costs of disease. *Soc Sci Med* 1992, **34**, 1005–1010.
16. Department of Clinical Epidemiology and Biostatistics, Mc Master Health Sciences Center, How to read clinical journals: VII: to understand an economic evaluation (Part A). *Can Med Assoc J* 1984, **130**, 1428–1434.
17. Coyle D, Davies L. How to asses cost-effectiveness: elements of a sound economic evaluation. In Drummond MF, Maynard A, eds. *Purchasing and Providing Cost-effective Health Care.* Edinburgh, Churchill Livingstone, 1993, 66–79.
18. Sacks H, Chalmers RC, Smith H. Randomised versus historical controls for clinical trials. *Am J Med* 1982, **72**, 33–240.
19. Stewart LA, Parmar MKB. Meta-analysis of the literature or of individual patient data: is there a difference? *Lancet* 1993, **341**, 418–422.
20. Lau J, Antman EM, Jimeneq-Silva J, Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med* 1992, **327**, 248–254.
21. Early Breast Cancer Trialists' Collaborative Group. Systemic treatment of early breast cancer by hormonal, cytotoxic or immune

therapy. 133 randomised trials involving 31,000 recurrences and 24,000 deaths among 75,000 women. *Lancet* 1992, **339**, 1–15 and 71–85.
22. Herxheimer A. The Cochrane Collaboration: making the results of controlled trials properly accessible. *Postgrad Med J* 1993, **69**, 867–868.
23. Williams CJ. Knowing just how effective is treatment: implications for care of cancer patients. *Clin Oncol* 1995, **7**, 42–45.
24. Dickenson K, Hewitt P, Mutch L, Chalmers I, Chalmers TC. Perusing the literature: comparison of Medline searching with a perinatal trials database. *Controlled Clin Trials* 1985, **6**, 306–317.
25. Haes JCJM de, Knippenberg FCE van. The quality of life of cancer patients: a review of the literature. *Soc Sci Med* 1985, **20**, 809–817.
26. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989, **27**, S217–S232.
27. Hunt SM, McEwen J, McKenna SP. Measuring health status. London, Croom Helm, 1986.
28. Bergner M, Bobbitt RA, Carter WB, Gilson BS. The sickness impact profile: development and final revision of a health status measure. *Med Care* 1981, **19**, 787–805.
29. Stewart AL, Hays RD, Ware JE. The MOS short-form general health study—reliability and validity in a patient population. *Med Care* 1988, **26**, 724–735.
30. McHorney CA, Ware JE, Rogers W, Raczek AE, Rachel JF. The validity and relative precision of MOS short and long form health status scales and Dartmouth COOP charts. *Med Care* 1992, **30**, MS253–MS265.
31. Uyl-de Groot CA, Bonsel GJ, Rutten FFH. Measurement and valuation of quality of life in economic appraisal of cancer treatment. *Eur J Cancer* 1995, **30A**, 111–117.
32. Schipper H, Clinch J, McMurray A, Levitt M. Measuring the quality of life of cancer patients. The functional living index—cancer: development and validation. *J Clin Oncol* 1984, **2**, 472.
33. McDowell I, Newell C. Measuring health–a guide to rating scales and questionnaires. Oxford, Oxford University Press, 1987.
34. Aaronson NK, Bullinger M, Ahmedzai S. *A Modular Approach to Quality of Life Assessment in Cancer Clinical Trials. Recent Results in Cancer Research*. Berlin, Springer, 1988.
35. Ganz PA, Schag CAC, Lee JJ, Sims MS. The CARES: a generic measure of health-related quality of life for patients with cancer. *Qual Life Res* 1992, **1**, 19–29.
36. Schag CAC, Ganz PA, Heinrich RLH. Cancer rehabilitation evaluation system—short form (CARES-SF). *Cancer* 1991, **68**, 1406–1414.
37. Haes JCJM de, Knippenberg FCE van, Neijt JP. Measuring psychological and physical distress in cancer patients: structure and application of the Rotterdam Symptom Checklist. *Br J Cancer* 1990, **62**, 1034–1038.
38. Levine MN, Guyatt GH, Gent M, *et al*. Quality of life in stage II breast cancer: an instrument for clinical trials. *J Clin Oncol* 1988, **6**, 1798–1810.
39. Lansky SB, List MA, Ritter-Sterr C, *et al*. Performance parameters in head and neck patients. *Proc Am Soc Clin Oncol* 1988, **7**, 156 (abstr 603).
40. Llewellyn-Thomas HA, Sutherland HJ, Hogg SA, *et al*. Linear analogue self-assessment of voice quality in laryngeal cancer. *J Chron Dis* 1984, **37**, 917–924.
41. Drummond MF, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. Oxford, Oxford University Press, 1989.
42. Torrance GW. Utility approach to measuring health-related quality of life. *J Chron Dis* 1987, **40**, 593–600.
43. EUROQOL study group. EuroQoL, a new facility for measurement of health-related quality of life. *Health Policy* 1990, **16**, 147–161.
44. Essink-Bot ML, Bonsel GJ, Maas PJ van der. Valuation of health states by the general public: feasibility of a standardized measurement procedure. *Soc Sci Med* 1990, **31**, 1201–1206.
45. Bennett K, Torrance GW. McMaster Rheumatoid Arthritis Utility Measurement Questionnaires. McMaster University, Dept. of Clinical Epidemiology and Biostatistics, Hamilton, Canada 1990.
46. Torrance GW, Boyle MH, Horwood SP. Application of multi-attribute utility theory to measure social preferences for health states. *Oper Res* 1982, 1043–1069.
47. Rosser R, Kind P. A scale of valuations of states of illness: is there a social consensus? *Int J Epidemiol* 1978, **7**, 347–358.
48. Fayers PM, Jones DR. Measuring and analysing quality of life in cancer clinical trials: a review. *Stat Med* 1983, **2**, 429–446.
49. Drummond M. Output measurement for resource allocation decisions in health care. In McGuire A, Fenn P, Mayhew K, eds. *Providing Health Care: The Economics of Alternative Systems of Finance and Delivery*. Oxford, Oxford University Press, 1991, 99–119.
50. Drummond MF, Stoddart GL. Economic analysis and clinical trials. *Controlled Clin Trials* 1984, **5**, 115–128.
51. Mugford M, Drummond MF. The role of economics in the evaluation of care in pregnancy and childbirth. In Chalmers I, Enkin M, Kierse MJNC, eds. *Effective Care in Pregnancy and Childbirth*. Oxford, Oxford University Press, 1989.
52. Mason J, Drummond MF, Torrance G. Some guidelines on the use of cost-effectiveness league-tables. *Br Med J* 1993, **306**, 570–572.
53. Drummond MF, Torrance G, Mason J. Cost-effectiveness league tables: more harm than good? *Soc Sci Med* 1993, **37**, 33–40.
54. Udvarhelyi S, Colditz GA, Rai A, Epstein AM. Cost-effectiveness and cost–benefit analyses in the medical literature. Are methods being used correctly? *A Int Med* 1992, **116**, 238–244.
55. Adams ME, McCall NT, Gray DT, Orza MJ, Chalmers TC. Economic analysis in randomised controlled trials. *Med Care* 1992, **30**, 231–238.
56. Jefferson TO, Demicheli V. Is vaccination against Hepatitis B efficient? A review of world literature. *Health Econ* 1994, **5**, 25–38.
57. Mugford M, Jefferson TO, McGuire A. How can we generalize from the results of economic evaluations? Paper presented at the United Kingdom Health Economics Study Group 1994 (unpublished).
58. Brown ML, Fintor L. Cost-effectiveness of breast cancer screening; preliminary results of a systematic review of the literature. *Breast Cancer Res Treat* 1993, **25**, 113–118.
59. Dillman RO, Seagren SL, Propert KJ, *et al*. A randomized clinical trial of induction chemotherapy plus high dose radiation versus radiation alone in stage III non-small cell lung cancer. *N Engl J Med* 1990, **323**, 940–945.
60. Eddy RJ. Cost effectiveness of CT scanning compared with mediastinoscopy in the preoperative staging of lung cancer. *J Can Assoc Radiol* 1989, **40**, 189–193.
61. Goodwin PJ, Feld R, Evans WK, Pater J. Cost-effectiveness of cancer chemotherapy: an economic evaluation of a randomized trial in small-cell lung cancer. *J Clin Oncol* 1988, **6**, 1537–1547.
62. Houston GA, Griffith JE, Sanders JA, *et al*. Staging of lung cancer. A cost-effectiveness analysis. *Am J Clin Oncol (CCT)* 1985, **8**, 224–260.
63. Jaakkimainen L, Goodwin PJ, Pater J, *et al*. Counting the costs of chemotherapy in a National Cancer Institute of Canada randomized trial in non-small cell lung cancer. *J Clin Oncol* 1990, **8**, 1301–1309.
64. Pastiko, Johnson, 1992.
65. Rees GJG. Cost-effectiveness in oncology. *Lancet* 1985, **ii**, 1405–1408.
66. Joseph AM, Crowson TW, Rich EC. Cost effectiveness of HemoQuant versus hemoccult for colorectal cancer screening. *J Gen Int Med* 1988, **3**, 132–138.
67. Johnson MG, Jolly PC. Analysis of a mass colorectal cancer screening program for cost-effectiveness. *Am J Surg* 1987, **154**, 261–263.
68. Mandelblatt JS, Fahs MC. The cost-effectiveness of cervical cancer screening in low-income elderly women. *JAMA* 1988, **259** 2409–2413.
69. Parkin DM, Moss SM. An evaluation of screening policies for cervical cancer in England and Wales using a computer simulation model. *J Epidemiol Commun Health* 1986, **40**, 143–153.
70. Mooney G. Breast cancer screening: a study in cost-effectiveness analysis. *Soc Sci Med* 1982, **16**, 1277–1283.
71. Forrest P. Breast cancer screening, report to the Health Ministers of England, Wales, Scotland and Northern Ireland. Department of Health and Social Security, 1987.
72. Van der Maas PJ, *et al*. The cost-effectiveness of breast cancer screening. *Int J Cancer* 1989, **43**, 1055–1060.
73. De Koning HJ, Van Ineveld BM, *et al*. Breast cancer screening and cost-effectiveness: policy alternatives, quality of life considerations and the possible impact of uncertain factors. *Int J Cancer* 1991, **49**, 631–637.
74. Eddy DM, Hasselblad V, *et al*. The value of mammography screening in women under age 50 years. *JAMA* 1988, **259**, 1512–1519.

75. Obuko I, Glick H, *et al.* Cost-effectiveness of mass screening for breast cancer in Japan. *Cancer* 1991, **67**, 2021–2029.
76. Gravelle HSE, *et al.* Breast cancer screening and health service costs. *J Health Econ* 1982, **1**, 185–207.
77. Skrabanek P. The cost-effectiveness of breast cancer screening. *Int J Technol Assessment Health Care* 1991, **7**, 633–635.

APPENDIX 1

Description and valuation of quality of life

	Reliability (test–retest)	Internal consistency	Validity	Duration of interview
Generic measurement instruments:				
Nottingham Health Profile (NHP) [27]				
Physical mobility	0.77–0.88	0.81	Content	10 min
Pain		0.77	Construct	
Sleep		0.69		
Social isolation		0.42		
Emotional reaction		0.75		
Energy		0.68		
Sickness Impact Profile (SIP) [28]				
Physical	0.90–0.91	Overall:	Content	20–30 min
Psychological	0.79–0.84	0.81–0.94	Construct	
Overall	0.75–0.92	Per dimension: 0.23–0.80 (NL)		
Rand Medical Outcome Study (RAND–MOS 20) [29]				
Physical functioning	0.78	0.85	Content	3–4 min
Role functioning	0.69	0.80	Construct	
Bodily pain	0.58	–		
Social functioning	0.75	–		
Mental health	0.83	0.89		
General health perception	0.85	0.91		
Cancer specific measurement instruments:				
Rotterdam Symptom Checklist [37]	0.82–0.88	0.81–0.93	Content Construct	8 min
Cancer Rehabilitation Evaluation System-Short Form (CARES-SF) [35, 36]	0.86	0.61–0.85	Concurrent	Unknown
Valuation instruments				
Quality of Well-being Scale (QWB) [33]	0.90	Unknown	Content Construct	18 min (descr) Unknown (val)
EuroQoL [43, 44]	0.69–0.94	Not relevant	Content	1 min (descr) 20 min (val)
McMaster Utility Measurement Questionnaire (MUMQ) [45]	0.66–0.94	Not relevant	Content	45 min (descr + val)
Torrance's Health State Classification System [46]	0.86–0.94	Unknown	Unknown	Unknown
Rosser & Kind-index [47]	0.79–1.00	Unknown	Content	10 min (descr) > 1.5 h (val)

APPENDIX 2A

Economic evaluations relating to lung cancer management

Reference	Country	Form of analysis	Study area	Alternatives evaluated	Main results
Dillman *et al.* [59]	U.S.A.	CEA	Induction chemotherapy plus radiotherapy for patients with non-small cell lung cancer	Radiation plus chemotherapy versus radiation alone	The inclusion of induction chemotherapy leads to a 33% increase in the costs of therapy but a 100% increase in mean survival after 2 and 3 years.
Eddy [60]	Canada	CMA	Staging of lung cancer—evaluating the mediastinum before surgery	CT scan versus no CT scan	The approach of using a CT scan to determine the need for mediastinoscopy with nodal biopsy rather than giving these tests routinely reduced the costs of treating these patients.
Goodwin *et al.* [61]	Canada	CUA	Treatment of extensive small cell lung cancer	Chemotherapy regimes	The most expensive regimen (cyclophosphamide, vincristine and doxorubicin) was more cost-effective, as it was more efficacious and savings were made in other resources.
Houston *et al.* [62]	U.S.A.	CMA	Staging of lung cancer—evaluating the mediastinum and assessing metastases	CT scan versus no CT scan	CT scans appeared to be no more accurate than current routine tests and the use of CT scans was more expensive.
Jaakkimainen *et al.* [63]	Canada	CEA	Treatment of advanced non-small cell lung cancer	Chemotherapy regimes versus best care	Both drug regimes were cost effective when compared to best care. The less expensive regimen (cyclophosphamide, doxorubicin and cisplatin) was more cost effective and cost saving.
Pashko *et al.* [64]	U.S.A.	CMA	Chemotherapy treatment in small cell lung cancer	Chemotherapy regimes	The regime involving the oral administration of a larger dose of etoposide on study days 2 and 3 rather than a single intravenously administered dose was less costly. Both regimes had equal efficacy.
Rees [65]	U.K.	CEA	Palliative treatment for carcinoma of the bronchus	Radiotherapy versus no intervention	Palliative care had a higher cost per "national patient benefit year" than curative therapies for other cancer therapies.

CEA, cost-effectiveness analysis; CMA, cost-minimisation analysis; CUA, cost–utility analysis; CT, computed tomography.

APPENDIX 2B

Economic evaluations relating to cancer screening programmes

Reference	Country	Form of analysis	Study area	Alternatives evaluated	Main results
Joseph *et al.* [66]	U.S.A.	CEA	Colorectal cancer screening	HemoQuant versus Hemoccult screening	Despite higher sensitivity of the HemoQuant test, its high cost is only justified if it can provide a specificity similar to that of the Hemoccult test.
Johnson and Jolly [67]	U.S.A.	CEA	Mass colorectal cancer screening	Self-administered stool blood test (no alternatives)	The studied programme costs per cancer detected are relatively high ($9670). The authors advocate mass screening programmes for colorectal cancer to be targeted towards at-risk populations.
Mandelblatt and Fahs [68]	U.S.A.	CEA	Cervical cancer screening in low-income, elderly women	Single Papanicolaou test for cervical cancer versus no test (hypothetically) in the same population	The proposed programme would both be efficient and cost-saving, thus supporting cervical cancer screening of elderly women.
Parkin and Moss [69]	U.K.	CEA	Cervical cancer screening	Seven screening policies recommended in the U.K. since 1966 (using different screening schedules and targeting different age groups)	The best cost-effectiveness ratio obtained corresponds to screening at 5-year intervals of women over 35. Extension of screening to women under 35 leads to loss of efficiency.
Mooney [70]	U.K.	CEA	Breast cancer screening, women aged 40–59 years	Mammography, thermography and clinical examination versus mammography and thermography (6 alternatives in total)	Mammography (single reporting) combined with a single clinical examination showed the most attractive cost-effectiveness ratio of the alternatives studied.
Forrest [71]	U.K.	CUA	Mass breast cancer screening by mammography, women aged 50–64 years	Mammography at 3-year intervals in women aged 50–64 years versus no screening	Mass screening by mammography at 3-year intervals would result in 695 QALYs per 100 000 women invited for screening.
van der Maass *et al.* [72]	The Netherlands	CEA	Breast cancer screening by mammography, women aged 50–70 years	Four alternatives based on different screening schedules: 4-year, 2-year, 1,3-year and 1-year frequencies, are compared to a no-screening situation in women aged 50–70 years	Mammographic screening of women aged 50–70 years at 2-year intervals allowed for savings of $4850 per life-year saved, a ratio which compares favourably to that of screening at either shorter or longer intevals and to that of other health interventions.
de Koning and Van Ineveld [73]	The Netherlands	CEA/CUA	Mass breast cancer screening by mammography, women aged 50–70 years	The invitation schedules used in the van der Maas study are compared for their respective effect on quality of life and QALYs	Mass screening of women aged 50–70 years would appear to affect quality of life very little, as the cost per QALY differed little from the cost per life-year gained unadjusted for quality. The authors recommended the adoption of a national screening policy based on a 2-year schedule.
Eddy *et al.* [74]	U.S.A.	CEA	Mass breast cancer screening by mammography, women aged under 50 years	Annual mammography for women aged 40–49 already having had a physical examination versus physical examination alone	Reduction in the costs of treatment in screened women under 50 years would be outweighed by extra costs of screening and necessary biopsies.
Obuko *et al.* [75]	Japan	CEA	Breast cancer screening, women aged 30–80 years	Physical examination alone (current Japanese policy) versus combinations of mammography and physical examination versus no screening	According to the computer simulation model, screening by mammography alone provided the best (lowest) cost-effectiveness ratio of the studied alternatives.
Gravelle *et al.* [76]	England and Wales	CA/CEA	Breast cancer screening, women aged over 40 years	Physical examination and mammography versus conventional management (no screening)	The authors concluded that the introduction of screening would increase costs for the NHS, yet the cost per life-year saved was very low.

Abbreviations as for Appendix 2a.