

Nuffield Occasional Papers
Health Economics Series: Paper No. 7

Economic Evaluation and Health Care

John Cairns

Introduction by
John Wyn Owen



The Nuffield Trust
FOR RESEARCH AND POLICY
STUDIES IN HEALTH SERVICES

TITLES OF OCCASIONAL PAPERS: HEALTH ECONOMICS SERIES

1. *Mergers in the NHS: Made in Heaven or Marriages of Convenience?*
Maria Goddard and Brian Ferguson
2. *Devolved Purchasing in Health Care: A Review of the Issues.*
Peter Smith
3. *Going for Gold: The Redistributive Agenda
Behind Market-Based Health Care Reform.*
Robert Evans
4. *A Social Contract for 21st Century American Health Care:
Three Tier Health Care with Bounty Hunting.*
Uwe Reinhardt
5. *Who Pays for and Who Gets Health Care? Equity in the Finance
and Delivery of Health Care in the United Kingdom.*
Carol Propper
6. *Future Hospital Services in the NHS: One Size Fits All?*
Peter West
7. *Economic Evaluation and Health Care.*
John Cairns
8. *Managed Care: Panacea or Palliation?*
Alan Maynard and Karen Bloor

Nuffield Occasional Papers
Health Economics Series: Paper No. 7

Economic Evaluation and Health Care

John Cairns
Health Economics Research Unit
University of Aberdeen

Introduction by
John Wyn Owen

Series Editor
Professor Alan Maynard



The Nuffield Trust
FOR RESEARCH AND POLICY
STUDIES IN HEALTH SERVICES

Published by
The Nuffield Trust
59 New Cavendish Street
London W1M 7RD

ISBN 1-902089-08-1

© Nuffield Trust 1998

Publications Committee

Sir Derek Mitchell KCB, CVO
Professor John Ledingham DM, FRCP
Mr John Wyn Owen CB

CONTENTS

5	About the Author
6	Introduction
8	Foreword
10	Abstract
11	Preface
14	Evidence-Based Health Economics
22	Economic Evaluation – Timing and Form
28	The Monstrous World of Economic Evaluation
35	Issues Concerning Methods
43	Guidelines for Economic Evaluation
49	Conclusions
53	References

ABOUT THE AUTHOR

John Cairns is a graduate of the Universities of Aberdeen and York. He has been the Director of the Health Economics Research Unit, University of Aberdeen*, since 1993. His research interests include the economics of preventative health care and time preference for health.

* The Health Economics Research Unit is funded by the Chief Scientist Office of the Scottish Office Department of Health (SODOH). The views expressed in this paper are those of the author and not of SODOH.

Editor's Acknowledgements

This series of Occasional Papers was generously supported by the Nuffield Trust. In addition to the help and support from all the authors and the referees of their papers the editor would like to acknowledge the valuable and generous managerial and editorial assistance provided by Karen Bloor and Frances Sharp of the University of York.



INTRODUCTION

The government is once again undertaking a comprehensive health spending review. Sustainable financing of health care with appropriate mechanisms for individual community and national priority setting are important public policy objectives which have been under scrutiny over many years and must now be addressed with some urgency. The Trust has informed this debate in the past and will continue to do so.

These Occasional Papers offer the economists' contribution and should be of interest to policy-makers at the highest level as they strive to improve the effectiveness of the National Health Service, improve patient care and create the right incentives to reward efficient performance within inevitable financial constraints.

Paper 7 – *Economic Evaluation and Health Care* – by John Cairns, provides an overview of the state of economic evaluation of health care and considers the very pertinent question as to whether it is ready for the challenges of the next decade. The intended audience for the paper comprises those commissioning economic evaluations; those undertaking such evaluations; and those using the results. The author suggests that there is scope to learn from the rigorous approach of those seeking to improve the evidence base of medicine. Although economic evaluation with respect to health care has developed into a substantial body of work, it is difficult to mount a robust defence of many of the methods used. A priority for health economists must be to do much more to improve the evidence base of their methods.

The economic evaluation of health care has been labelled a half-way technology in that it has not yet reached an advanced stage where it can be applied routinely. It remains of benefit in specific cases; non-routine and therefore relatively expensive; dependent on specialist expertise; driven by intellectual curiosity; and having to justify itself continually

in an increasingly sceptical world. Cairns concludes that the way forward from this half-way position is not by promoting unwarranted images of consensus, but by continuing to develop and refine the methods.

John Wyn Owen
April 1998

FOREWORD

The application of economic analysis to health and health care has grown rapidly in recent decades. Alan Williams' conversion of Archie Cochrane to the virtues of the economic approach led the latter to conclude that:

“allocation of funds and facilities are nearly always based on the opinion of consultants but, more and more, requests for additional facilities will have to be based on detailed arguments with ‘hard evidence’ as to the gain to be expected from the patient’s angle and the cost. Few could possibly object to this.”*

During most of the subsequent twenty-five years many clinicians have ignored Cochrane’s arguments whilst economists busily colonised the minds of those receptive to their arguments. More recently clinicians and policy makers have come to equate, erroneously of course, health economics with economic evaluation. Thus the architects of the Department of Health’s R&D strategy have insisted that all clinical trials should have economic components and tended to ignore the broader framework of policy in which economic techniques can be used to inform policy choices by clinicians, managers and politicians. †

The purpose of this series of Occasional Papers on health economics is to demonstrate how this broad approach to the use of economic techniques in policy analysis can inform choices across a wide spectrum of issues which have challenged decision makers for decades. The authors do not offer ‘final solutions’ but demonstrate the complexity of their subjects and how economics can provide useful insights into the processes by which the performance of the NHS and other health care systems can be enhanced.

The papers in this series are stimulating and informative, offering readers unique insights into many aspects of health care policy which will continue to challenge decision makers in the next decade regardless of the form of government or the structure of health care finance and delivery.

Professor Alan Maynard
University of York

* Cochrane AL. *Effectiveness and Efficiency: random reflections on health services*. Nuffield Provincial Hospitals Trust, London, 1972.

† Maynard A and Chalmers I (eds). *Non-random Reflections on Health Services Research: on the 25th anniversary of Archie Cochrane's Effectiveness and Efficiency*. British Medical Journal Publishing, London, 1997.



ABSTRACT

Health economists appear to have been successful in persuading a sceptical and initially resistant world that economic evaluation has an important contribution to make to health care planning and research. There has been a huge increase in the number of published economic evaluations of health care, a sustained increase in the demand for health economists as research collaborators, and recent emphasis on using cost-effectiveness as a criterion in determining spending priorities. This paper considers whether or not economic evaluation in health care is ready for the challenges of the next decade.

Generally those producing economic evaluations have tried to base their conclusions on the best possible evidence, but the quality of the data upon which evaluations have been based and the standards of reporting have often been poor. In addition, there must be considerable concern about the evidence base in support of the methods applied. It is not clear that efforts to improve existing guidelines for the conduct of economic evaluation will do anything to encourage the development of the methods of evaluation and in fact the reverse appears likely. Guidelines for the reporting of economic evaluations are less likely to have such a discouraging effect on research and since they tend to encourage the provision of more detailed information have a positive contribution to make.

PREFACE

The primary purpose of this paper is to provide an overview of the state of economic evaluation of health care. Health economists have devoted considerable energies to educating or persuading a sceptical and initially resistant world that economic evaluation has an important contribution to make. Evidence of their substantial success is widespread, for example, the huge increase in the number of published economic evaluations of health care, the sustained increase in the demand for health economists as research collaborators, and the recent emphasis on using cost-effectiveness as a criterion in determining spending priorities. Success on this scale suggests that the economic evaluation of health care has 'arrived'. Substantial expectations have been raised (largely by health economists) and if the progress that has been made is to be sustained a new phase must be entered where economic evaluation meets its promise.

This forms the background to this paper. Is health economics, or more specifically economic evaluation, ready for the challenges of the next decade? The paper concentrates on the activity of health economic evaluation. The potential and actual contribution of economic evaluation must ultimately be judged with respect to its impact on decision making but to do so is beyond the scope of this paper. Thus the paper is about what health economists and others do rather than the value of these activities. Our starting point is the extent to which health economics is evidence-based. A range of alternative concepts of evidence-based as applied to health economics are explored and then an assessment is offered on the extent to which health economics measures up. The weak evidence base with respect to the methods used as part of economic evaluation sets the scene for a selective examination of some methods.

PREFACE

A further theme addressed is the timing and form of economic evaluation. There are a number of stages at which economic evaluation can be undertaken and the stage at which it is undertaken will have important implications for the form it takes. The most appropriate form of evaluation to undertake or the appropriate timing of economic evaluation is not known. Similarly relatively little is known about how best to allocate scarce evaluative resources across the different stages of economic analysis. The discussion highlights the potential role for pre-trial economic evaluation and closes with an example of a 'quick and dirty' economic evaluation which is intended to concentrate thinking on where the boundaries lie with respect to what is acceptable in terms of the economic evaluation of health care.

Attention then shifts to the conflict between modelling and what might be loosely described as trial-based approaches which has generated particular controversy. The general view put forward is that the two approaches are more complementary than is often implied by the discussions which have taken place and that the central issue should be the appropriate use of both approaches. As with so many issues in the economic evaluation of health care, our ignorance is such that we cannot yet provide a convincing answer to what is the appropriate balance between trial-based and modelling approaches.

One of the main conclusions arising from the discussion of evidence-based health economics is that the evidence base with respect to many of the methods used in the economic evaluation of health care is weak. As a direct consequence there are many unresolved issues concerning the conduct of such economic evaluations. A number of areas are highlighted, namely: future health care costs; discounting; and the valuation of health and health care.

The emphasis on the unresolved nature of many of the issues of methods in turn underpins some of the criticism of too hasty a development of guidelines. A number of proposed guidelines are reviewed and the advantages and disadvantages of guidelines for the conduct of economic evaluation are discussed. However, the potential disadvantages of putting further effort into the development of guidelines are emphasised. Also the limited knowledge concerning the effects of guidelines and the costs and benefits associated with their implementation is highlighted. The general view espoused is that a premature move to guidelines may generate misleading conclusions and will make progress through research much harder. A more sanguine view is expressed about the scope for guidelines for the presentation of results.

The intended audience for this paper comprises those commissioning economic evaluations; those undertaking such evaluations; and those using the results. The intention is to bring a somewhat more critical tone to the appreciation of the contribution of economic evaluation. All participants in the growing evaluation industry will benefit from greater realism with respect to what can be achieved and a sharper focus on some of the limitations of economic evaluation and the areas requiring increased research efforts. While the focus of this paper is narrower than that of a recent critique by Maynard and Sheldon,¹ it stresses similar concerns with respect to the contribution of health economics.

EVIDENCE-BASED HEALTH ECONOMICS

A notable feature of debates over appropriate medical practice in recent years has been the rise of evidence-based medicine (EBM). Many clinicians in response to the recent emphasis on EBM would assert that good medicine has always been evidence-based. The recent prominence perhaps owes much to the need for apparently new themes to be enthusiastically embraced every so many years so as to provide opportunities for career advancement. The picture might be viewed as similar in health economics. Many health economists would aver that health economics has always been evidenced-based and some would argue that health economics has had a role in stimulating EBM. Certainly many economic evaluators would recognise a characterisation where the health economist takes a much more critical approach to the clinical evidence than do the clinical investigators. This may arise because the relative ignorance of the economist leads to the posing of fundamental questions, or it may be that the economist generally does not have a preferred answer when embarking on an evaluation.

The claim that health economics is and has to some extent always been evidence-based will be scrutinised here. But before deliberating on how evidence-based health economics is, it is necessary to define just what is or might be meant by the description evidence-based health economics.

What is evidence-based health economics?

There are two strong candidates: health economic aspects of evidence-based medicine; and health economics based on evidence.

EBM is not directly concerned with economic issues (in part because of its focus on individual patients). EBM, however, raises at least two broad economic issues. The first concerns the costs and consequences

of EBM interventions. To what extent are the pronouncements of EBM cost-effective? In short, we don't know. However, the emphasis on effectiveness and the lack of explicit consideration of resource use and, in particular, comparison of incremental benefits and incremental costs, must fuel the suspicion that much of EBM may not represent a cost-effective use of resources.

The second involves the questions of how and when to attempt to change practice (for example, by introducing guidelines). The conclusions of EBM may on occasion represent a more cost-effective use of resources but the various agents involved will not costlessly agree and implement such advice. Changing practice will be costly, on some occasions to such an extent that best evidence-based practice is an inappropriate use of resources. Moreover, there are generally alternative means of changing practice, each with potentially different costs and consequences, which should be considered when choosing how to implement change. This is largely unexplored territory for health economists and, as a result, there is a potentially important contribution to be made.

Health economists do not have a particularly strong record in terms of addressing implementation issues. The slight attention paid to implementation in economic evaluations may be partly explained by a wish to highlight the extent to which findings are generalised. Implementation usually raises a series of local issues.

While this first definition of evidence-based health economics is a defensible one, and there is clearly a legitimate role for economics with respect to EBM, a more natural definition is in terms of health economics based on evidence. At least three versions can be identified:

EVIDENCE-BASED HEALTH ECONOMICS

- ◆ evidence base of health economic evaluations;
- ◆ evidence base of the methods which health economists apply; and
- ◆ evidence base of more general pronouncements (e.g. regarding competition and the design of systems).

It is worth distinguishing these because the performance of health economics across them varies. It is perhaps novel to consider the evidence base for advice and for methods rather than the evidence base of decisions. However, health economists provide advice and supply methods of analysis rather than make decisions. Moreover, most non-economists assume that such advice and methods are firmly evidence-based.

How evidence-based are health economic evaluations?

Health economists would generally claim that they make 'conscientious, explicit and judicious use of the current best evidence'. Note that this claim refers to health economists and not to economic evaluations of health care. A distinction can be drawn between evaluations with an economic perspective undertaken with direct input from health economists and those not similarly blessed. The majority of evaluations containing an economic perspective, in most areas of health care, appear to have been undertaken without direct input from a health economist. There is a tendency for such studies to differ from those with a direct input.

It would be an interesting task to take the economic evaluations in a particular area and interrogate them from the perspective of whether they incorporated (when they were conducted) the current best evidence, and whether there are systematic differences between those

with and those without a direct health economics input. The NHS Economic Evaluation Database could provide much of the information required to complete such a task.²

It is appropriate to consider the nature of evidence for the purposes of economic evaluation. Economic evaluations are dependent on the quality of medical evidence but substantial progress can be made in the absence of good medical evidence. This is at any rate true where there is simply a lack of evidence. On the other hand, where the research evidence is biased it is unlikely that the economic evaluator can contribute in any substantial way. Whereas the randomised controlled trial (RCT) (and where possible a meta-analysis of comparable RCTs) is the gold standard for EBM, economic evaluations are based on a broader mix of evidence. A more flexible attitude towards evidence results primarily because health economics often addresses a more complex set of questions than how much more effective is A than B. The data demands of economic evaluation and the limitations of RCTs as a means of supplying the required data are discussed below. There is an issue here of whether more should routinely be demanded of economic evaluations with respect to the quality of the data which they use. The prevailing attitude is one of extreme pragmatism whereby most health economists are willing to make use of what is available (and do not generally ignore relevant data). The key point is that while the vast majority of economic evaluations would fail the 'RCT test' this does not mean that health economics is not evidence-based.

Evidence base of the methods which health economists apply

Health services research is generally multidisciplinary and there is a tendency for members of any one discipline to let those of others get

EVIDENCE-BASED HEALTH ECONOMICS

on with their allotted tasks. The non-statisticians do not generally question the power calculations of the statistician. Clinicians exercise clinical judgment not health economists. Similarly there is a general willingness to leave the 'health economics' to the health economist (if one is available). How much of this preparedness to trust the health economist is based on the belief that he or she will be using methods with a good evidence base and is it justified?

I shall leave the former question hanging and will consider the latter. The claim which will be made is that the evidence base of many of the methods routinely used by health economists is not satisfactory. There are numerous examples of methods in routine use with a questionable evidence base, including: quality-adjusted life-year (QALY); the modelling of programme costs; and discounting.

The QALY has been prominent in economic evaluations in the last ten or so years. At its heart is an adjustment for the quality of survival which enables quantity and quality of life to be combined in a single measure which can be expressed relative to resource use and compared across widely differing programmes. The quality adjustment has been based on various techniques: standard gamble; time-trade-off; magnitude estimation and rating scales.³ But relatively little is known about the properties of these various approaches. Choice of method appears to be driven more by analyst preference rather than by evidence. There has been research comparing two or more methods but the profession still has much to learn.

The measurement of benefits is notoriously difficult. Surely a more satisfactory picture emerges when we look at what many would regard as the staple diet of health economists – costing. But yet again despite admitted progress the record is not good. A nice example was provided

recently by Fenn *et al*⁴ in the context of modelling programme costs. They highlighted how economists were typically unsophisticated in their estimation of the costs of treatment and emphasised the scope for methods developed for analysing survival data to be applied to total costs. A further example with respect to costs is the widespread reliance on mean or median values and a failure to consider the distribution of costs. Coyle⁵ notes how few pharmacoeconomic studies have adopted appropriate statistical analysis of costs (and effects).

The standard practice in general is to discount all costs and benefits at a constant and common rate and that the evidence on individual preferences, such as it is, does not support this traditional approach. This example is possibly rather different from the previous ones in that we possibly adopt this policy, knowing that it is unsupported by evidence, on normative grounds (and no amount of evidence will shift this preference).

That there are unresolved issues with respect to the conduct of economic evaluation and with respect to the methods used is no bad thing. It can stimulate and sustain interest in the economic evaluation of health care. Where economists provide a disservice is when they fail to research rigorously their methods, and fail to be entirely honest regarding the shortcomings of those methods. Lifting the lid on the black box will not reassure clinical collaborators but it is essential if economists are to be appropriately self-critical and forward-looking.

Evidence base of health economic recommendations

The record with respect to policy recommendation is not auspicious as will be seen from a number of examples given below. However, this seems less of a problem than with respect to methods. This is for two reasons: first the record is not particularly poor compared with other

EVIDENCE-BASED HEALTH ECONOMICS

professional groups offering advice, and second, it seems to be a more fundamental shortcoming to be employing unsatisfactory methods.

There are many instances where health economists have exhibited a willingness to make policy recommendations in the absence of a particularly strong evidence base, for example with respect to: competition among health care providers; competition among third party payers; and giving GPs budgets. However, a discussion of this interesting topic would lead too far away from our primary concern – the economic evaluation of health care interventions.

Summary

Health economists would probably claim that their work is evidence-based. Such a claim might be supported with respect to economic evaluation which is by nature evidence-based. Although it would not be difficult to find instances where the economic evaluation was not based on the best available information, such instances will be becoming rarer now that clinical colleagues are becoming more familiar with and committed to systematic review of the evidence. Some evidence on this sanguine summary would be welcome. There appears to have been little formal study of how health economists use or fail to use clinical evidence.

An evidence-based claim is much harder to sustain with respect to the evidence base for many of the methods routinely employed which is quite clearly unsatisfactory. Similarly, but of less relevance in the current context, the evidence base for many of the policy pronouncements that some economists are apt to make is also unsatisfactory.

There is scope to learn from the rigorous approach of those seeking to

improve the evidence base of medicine.⁶ Although economic evaluation with respect to health care has developed into a substantial body of work, it is difficult currently to mount a robust defence of many of the methods used. A priority for health economists must be to do much more to improve the evidence base of their methods. Given the strong demand for economic evaluations (and the money available) it is not surprising that the emphasis has been on meeting this demand rather than on undertaking the more fundamental research into the nature of the methods used. The profession will benefit in the future from investing rather more resources currently in developing new, and refining existing, tools.

ECONOMIC EVALUATION – TIMING AND FORM

Economic evaluation of health care can take a number of forms. The traditional classification of course comprises: cost-minimisation; cost-effectiveness analysis; cost-utility analysis; and cost-benefit analysis.⁷ Economic analysis can potentially inform decision making without recourse to any of these formal techniques. Such analyses are generally referred to as partial evaluations and tend not to feature heavily in the published literature, although in practice these less formal approaches may be having a greater impact on decision making than the more visible peer-reviewed tail of the distribution. Attention in this paper is focused on the published end of the spectrum but it is important to remember the potentially significant contribution made by less formal, partial analyses, particularly when considering the different stages at which economic evaluation (of some kind) may have a role to play.

This section considers the timing and form of economic evaluation. It emphasises that economic evaluation can take many forms and also how little is known about how best to allocate scarce evaluative resources across the different stages of economic analysis. It highlights the potential role for pre-trial economic evaluation and closes with a nice example of a ‘quick and dirty’ economic evaluation which may help in setting the boundaries of what is acceptable in terms of the economic evaluation of health care.

Stages of economic analysis

Sculpher *et al*⁸, when considering the role of economic evaluation in health technology assessment, identify four stages at which economic analysis can take place. These are: early developmental (stage I); maturing innovation (stage II); close to widespread diffusion (stage III); and moving into practice (stage IV). Stage I comprises the systematic review of evidence relating to the cost and effectiveness of

existing practice, and the use of informal clinical opinion to assess the potential value of the new technology. Stage II includes modelling studies using data from existing clinical studies, and pilot studies of economic data collection alongside controlled trials. Stage III involves economic data collection alongside RCTs, and refined modelling studies using systematic overviews of clinical data. Finally, Stage IV involves economic data collection alongside pragmatic trials, and modelling studies to generalise results to other settings or to extrapolate to the long term.

Since the questions to be answered and the methods used at each stage differ, these analyses are to be seen as complementary and economic evaluation as iterative. Lest the cost of such seemingly enhanced economic evaluation appears prohibitive, the authors are careful to point out that evaluative resources may be saved at later stages as a result of the earlier analyses, and all four stages will not necessarily be required in every case. This is an excellent and persuasive start. However, opportunity cost is as relevant a concept with respect to economic evaluation as it is in any sphere of economic life. The next stage must be to learn more about the optimal allocation of scarce research resources in order to get the maximum benefit from these resources. The authors' key message that economic evaluation should take place early and often fails to address this issue squarely.

Davies *et al*⁹ present a re-iterative structured decision making process for the commissioning and design of economic evaluation. They emphasise that economic evaluation can indicate whether it is worthwhile pursuing the intervention and associated research; whether there are circumstances in which it could be efficient; whether it will be efficient in routine practice; and efficient research designs. Their structured process goes some way towards recognising

ECONOMIC EVALUATION – TIMING AND FORM

opportunity costs (for example, there are questions concerning whether additional research offers value for money).

The economic analysis to be undertaken at stages II and III have been compared by a number of authors. The debate which has sometimes become polarised into one between economic modelling and economics alongside RCTs is considered below. Partly as a result, the relative value of economic analysis undertaken at different stages has received insufficient consideration. Moreover, relatively little attention has been given to the scope for stage I and IV analyses, and relatively few studies (compared with stage II and III analyses) have been reported. The novelty of any economic analysis (pre-, alongside or post-trial) has now in most areas of medical care worn off. However, until relatively recently, as Adams *et al*¹⁰ are frequently quoted to observe, ‘of about 50,000 randomised trials undertaken over a 22 year period, only 121 included economic analyses’.

Pre-trial economic evaluation

The scope for and potential value of pre-trial economic evaluation is increasingly being recognised. A number of claims can be made for pre-trial evaluation. There may be scope for such an analysis to assist decision making with respect to whether or not a trial should be undertaken. Resources for research are of course scarce and the opportunity cost of undertaking one particular evaluation is in terms of the foregone benefit from other research. Pre-trial evaluation may also be able to assist with the design of any trial, for example, with respect to the options to evaluate, the sample size, what information to collect, and the choice of end-points. Finally there may be a role in terms of informing clinical decision making and resource allocation while waiting for trials to be completed.

A recent study of screening for *H pylori* to prevent gastric cancer (Parsonnet *et al*¹¹ 1996) illustrates some of the scope for pre-trial evaluation. The authors primary purpose was to evaluate whether a clinical benefit could translate into a cost-effective clinical strategy. Incidence and prevalence rates, and sensitivity and specificity of the screening test were taken from the literature and combined with costs based on Medicare charges. Their base case finding was that the screening programme evaluated would cost \$25,000 per year of life saved. This estimate was then used as support for undertaking a trial. As part of an extensive sensitivity analysis the authors then identified which variables had a large, moderate or little impact on the estimated cost-effectiveness. The efficacy of the intervention in preventing cancer and the age at which screening was performed were shown to have a large impact, and the costs of screening and *H pylori* treatment had a moderate impact. Whereas the eradication rate, the specificity and the sensitivity of the assay, the relative risk of cancer associated with *H pylori*, the prevalence, the reinfection rate, the rate of adverse reactions, life expectancy with cancer, and the costs of cancer treatment had little impact. Clearly such information has the potential to inform the design of a trial and any accompanying economic evaluation by showing for which variables it is more or less important to collect good data.

Quick and dirty

It is important to recognise that economic analyses come in all shapes and sizes and can vary substantially in terms of level of sophistication. The quick and the dirty do not generally feature in peer-reviewed journals. In a challenging paper published recently by Barr *et al*¹², the authors present an economic evaluation of allogeneic bone marrow transplantation (BMT) aimed at ‘providing an evidence-based

ECONOMIC EVALUATION – TIMING AND FORM

approach to the formulation of clinical policy'. The paper is challenging because although the authors make a case for the need for such work to inform clinical policy and they are careful to identify the limitations of their study, it is based on 10 BMT patients and eight 'controls', and a minimum follow-up period of only 18 months. Furthermore, in deriving QALY estimates, the quality adjustment was undertaken by getting a convenience sample of ten healthy volunteers to score directly descriptions of each patient's experience. Also the authors eschew available RCT and controlled clinical trial evidence on effectiveness and rely on outcome data for a single institution.

Where should we stand regarding such quick and dirty evaluations? One response would be to say that we need to consider the alternatives to such a study. To what extent can studies of this type provide useful information not available in the literature (at a reasonable cost)? However, we should be wary of overusing the argument that it is better than nothing or better than current practice. It is a powerful argument but it might also be a recipe for making much slower improvements in the quality of economic evaluations than might be possible or desirable. The saving grace of the study in question is possibly the scope it offers to inform local decision making using local data, in an environment where there is a paucity of good economic evaluations of large trials.

Summary

The arguments for undertaking economic analysis early and often are likely to be persuasive for most health economists. Recognition of the opportunity cost of more of any given form of economic evaluation is also widespread. We know relatively little about the relative costs and benefits of undertaking economic analysis at the four different stages.

Suppose £1 million has been earmarked for the economic analysis of the detection and treatment of colorectal cancer, how should this be allocated over the four stages of economic evaluation in order to provide the greatest benefit? Without getting into the obvious question of how is benefit to be measured in this context, it appears clear that the current state of our knowledge leaves us ill-equipped to provide a convincing answer to this question. This is despite the existence of two major RCTs and a number of economic evaluations. We are still learning the best ways to incorporate economics into the design and conduct of trials and probably know even less about pre-trial economic evaluations. Stage four evaluations have barely been considered, possibly because of the more limited scope they appear to have to affect decision making.

THE MONSTROUS WORLD OF ECONOMIC EVALUATION

The previous section highlighted the scope for economic evaluation to take a number of forms and to contribute at various stages with respect to the development and diffusion of a new treatment or service. Inevitably because of the lack of appropriate data much of the evaluation involves modelling. This section focuses on the conflict between modelling and what might be loosely described as trial-based approaches which have generated particular controversy. The general view put forward is that the two approaches are more complementary than is often implied by the discussions which have taken place and the central issue should be the appropriate use of both approaches. As with so many issues in the economic evaluation of health care, our ignorance is such that we cannot yet provide a convincing answer to what is the appropriate balance between trial-based and modelling approaches.

Count Dracula and Frankenstein's monster

Bernie O'Brien has recently distinguished two broad approaches to economic evaluation.¹³ These he has graphically described by analogy with Frankenstein's monster and Count Dracula. His analysis is in the context of the internal and external validity of pharmacoeconomic studies but most of it is relevant to the broader field of the economic evaluation of health care. In the first approach (Frankenstein's monster) the economic evaluation combines information from a number of sources, and in the second (Count Dracula) the economic evaluation feeds off a prospective randomised controlled trial.

His main theme is that although some of the relevant economic issues can be addressed in an RCT context it is likely to be necessary to undertake some modelling in order to answer policy relevant economic questions. The argument is advanced by way of illustration

by seven potential problems with RCT data: not comparing the most appropriate options; use of 'gold standard' measures of outcome which would not be available in practice; emphasis on intermediate rather than final outcomes; inadequate patient follow-up or sample size for purposes of economic evaluation; protocol-driven costs and outcomes; difficulties with geographical transferability of trial evidence; and the use by trials of selected patient and provider populations.

It is not only health economists who, while recognising the strengths of RCTs, are critical of their shortcomings. Black¹⁴ makes a strong case for observational studies by arguing that RCTs are (in certain circumstances) unnecessary, inappropriate, impossible and inadequate. Although these remarks were not made in the context of economic evaluation and were not made with economic modelling in mind, they have resonances for the discussion of economic evaluation and RCTs.

Modelling in economic evaluation

A recent briefing paper from the Office of Health Economics¹⁵ summarises many of the salient points regarding modelling in economic evaluation as opposed to undertaking economic analysis in association with RCTs.

Attitudes towards economic modelling vary depending in part on the proposed purpose of the modelling. Resistance to modelling is probably greatest where the analysis is regarded as predicting the cost-effectiveness of an intervention and thus potentially having a fairly direct effect on policy. Opposition to modelling is less strong where it is viewed as a means of identifying and highlighting issues which require the exercise of judgement or further research.

THE MONSTROUS WORLD OF ECONOMIC EVALUATION

The opposition to modelling is probably explained in part by a clash of two approaches – the largely non-experimental tradition of economics and the more experimental tradition of clinical science. Those from an experimental tradition are likely to be much more concerned by whether or not it is possible to assess accurately reliability and bias. They will be particularly concerned by the absence of standardisation in the use of modelling approaches to economic evaluation.

At the outset, it is important to recognise that it is not just health economists who model.¹⁵ For example, it is quite common to model clinical outcomes when trying to move from efficacy to effectiveness, and when trying to work out the implications of intermediate outcomes for long-term outcomes. Also where economic evaluation takes place alongside an RCT the need for some modelling is not generally removed. For example, the use of measures such as QALYs is likely to be dependent on modelling which has taken place outside of the trial setting. It will also generally be the case that the results of an economic evaluation undertaken alongside an RCT will need to be reworked for a number of different practice settings and countries. This can often be feasibly achieved via modelling rather than through repeating the trial elsewhere.

The ultimate defence of modelling is generally that it is necessary to consider realistically what would be the alternative. Where it is a case of saying that we simply do not know whether or not a particular intervention is a good idea, making use of the best data available in order to provide some guidance seems relatively attractive. It is of course important that the limitations of the data used and of the modelling approach are made clear so that inappropriate confidence is not placed on the results of any modelling exercise. For many purposes an approximate answer available now or in the near future is likely to

be of considerable assistance to decision makers.

Limitations of modelling

Sheldon¹⁶ emphasises that decision analytic models are particularly prone to bias because they combine information from a number of sources. He highlights a number of common errors in model construction. The problem is not simply that there will be biases but rather the difficulty of identifying the extent and direction of bias. Often the only option is to rely on the face validity on the model (and in more complex models this may be very difficult to assess).

There is not much scope to validate models by their predictions because generally modelling is being used to inform decision-making prior to certain information being known. There is little routine re-examination of conclusions of models once new information is available. There are a number of reasons for this: to some extent attention has moved on to the next set of decisions; academic incentives are for new work rather than possibly confirmatory replication; it may be quite some time later before there are appropriate data to enable validation; and it is probably easier to influence new decisions rather than change minds about practice once accepted.

Because of these problems Sheldon¹⁶ suggests restricting the use of modelling in economic evaluation to:

- ◆ 'to permit statistically valid analysis of reliable data...
- ◆ to carry out sensible combination of reliable information on effectiveness, cost and other parameters ... to compare whole treatment strategies...

THE MONSTROUS WORLD OF ECONOMIC EVALUATION

- ◆ to identify better where there are gaps in knowledge, to assess how important these are ... and how best to get the necessary information ... and whether further evaluation is likely to be worthwhile.' (p.9)

Note the main omission from this list – to provide cost-effectiveness information with which to inform policy making (although the very last part of the third purpose might be interpreted as straying into the area of providing cost-effectiveness information).

Trial-based economic evaluation

It is notable that the case against relying on trials for data tends to receive far wider discussion than the case for. The prime argument for trial based economic evaluation is that the design and conduct of RCTs offer the best means of determining efficacy. There are also some practical advantages in terms of there being processes in place for data collection (this is more usually turned on its head and argued to raise problems of overburdening patients and investigators). There are reasons to believe that the more rigorous procedures for the conduct of the trial will result in the collection of better quality economic data.

One way of characterising the choice between these seemingly competing approaches is in terms of a trade-off between internal and external validity.¹² The RCT is strong on internal validity (freedom from bias) but weak on external validity (generalisability), whereas modelling approaches are weak in terms of internal validity but potentially stronger in terms of external validity. The naturalistic pragmatic trial might be offered as a compromise between the two approaches.¹⁵ The criticisms of RCTs as a basis for economic evaluation might be viewed as arguments for more naturalistic pragmatic trials rather than necessarily as arguments for economic

modelling. Naturalistic pragmatic trials would appear capable of addressing all but one of the seven problems identified by O'Brien.¹³ The exception concerns the geographical transferability of trial evidence. Certainly they could offer normal settings with an appropriate choice of comparators. In principle, they could have follow-up of all patients and avoid difficulties with protocol-driven costs and outcomes. Sample sizes could be increased and duration of follow-up lengthened to those required by the economic evaluation. Thus there is some prospect of providing internal and external validity. However, several of the changes, such as leaving doctors with more discretion and using less heavily selected patient populations, will require larger sample sizes and will increase the cost and complexity of the trial.

However, there are arguments (other than cost arguments) in favour of some modelling.¹⁶ Modelling can extend RCTs and naturalistic trials to simulate final health outcomes, opportunity costs and artificial study arms. A modelling approach will be useful when RCTs are impossible or impractical. Also of course prior to trials taking place any economic evaluation must have recourse to modelling. As already noted, there is considerable scope for modelling to inform the design of subsequent trials (for example, to identify what information is required and assist in the selection of sample sizes). By modelling in advance of a trial, gaps in our knowledge can be identified, their importance assessed and further evaluation can be better targeted.

However, the inevitability of some modelling leaves open the question of how much modelling is desirable. The answer to this will be strongly influenced by views regarding the validity of modelling, a subject on which not a great deal is known.

THE MONSTROUS WORLD OF ECONOMIC EVALUATION

Summary

It is clearly not a case of whether or not to undertake economic modelling. It appears inevitable that in most situations it will be necessary to undertake some modelling. The key issues are the extent of modelling and the emphasis to be placed on modelling. In determining the appropriate use of modelling it is necessary to ask what value is added by the activity. Also care must be taken to assess the validity of models. As is the case elsewhere with respect to the economic evaluation of health care there is much to be learned. Indeed, it is possibly particularly the case with respect to economic modelling. Thus we return to our recurring theme – the (relative) ignorance of those undertaking economic evaluations of health care.

ISSUES CONCERNING METHODS

It was noted above, in the discussion of evidence-based health economics, that the evidence base with respect to many of the methods used in the economic evaluation of health care is weak. As a direct consequence there are many unresolved issues concerning the conduct of such economic evaluations, the existence of which should lead to some hesitancy with respect to the adoption of guidelines for the conduct of economic evaluations. In this section a number of areas are highlighted where further research is required before guidelines can be embraced with any enthusiasm, namely: future health care costs; discounting; and the valuation of health and health care. This is an illustrative rather than inclusive listing of a number of key topics. The aim is simply to provide support for the claim that there are many unresolved issues.

Future health costs

The treatment of future health care costs is an interesting and challenging one. The key issues are which future health care costs ought to be included and what are the effects of different practices with respect to the inclusion of future costs. Few would argue against the inclusion of future treatment savings as part of the benefit arising from, say, a screening programme. Most would wish to take account of them on the cost side producing a net cost of the intervention which would then be compared with some measure of effectiveness. True there is an issue regarding the extent to which future savings are appropriately compared with the commitment of additional resources currently. This is more than an issue of discounting or a marginal versus average cost issue. It concerns the extent to which future savings will be matched by actual resources released. Notwithstanding this issue there is a willingness to include these future effects in the analysis.

ISSUES CONCERNING METHODS

The situation appears to change when considering increased health care utilisation and costs as a result of a successful intervention. The patient does not die of A but dies of B several years later. In practice most studies have come down in favour of not including such increased future health care costs in the analysis. This widespread practice is not easy to defend, particularly if the estimate of the benefit of the intervention has included life-years gained as a result of the intervention. These life-years are generally only gained if further health care costs are incurred associated with treating some other condition. No-one seems to have difficulty including future health care costs associated with the condition originally treated. A defence of these discriminatory practices might be the entirely pragmatic argument that the information will not be available and thus we simply do what we can. It is true that to do otherwise would require an estimate to be made of the probability of the range of treatments that might be appropriate for any given individual. The task would be complex. It involves not only making a probabilistic assessment regarding the pattern and timing of future disease but also requires assumptions concerning what future disease would or would not be treated, how would it be treated and with what success.

For example, consider giving diabetic patients an antioxidant in order to reduce the risk of coronary heart disease and stroke. If such an intervention were successful, fewer patients would have fatal ischaemic events. Some might simply experience these events at a later stage than they otherwise would have. But what of the proportion who would no longer ever experience these events but rather live on to die of something else? One alternative might appear to be to make an estimate of the contribution to quality and or quantity of life arising from the original intervention in the absence of further treatments. But such an estimate

would be difficult to make and artificial. A refinement that might work in some instances would be to make cost-effectiveness estimates contingent on the treatment of various future health problems. Although it is clearly a formidable challenge to model the health consequences and resource costs of treating these other conditions.

In a recent paper Garber and Phelps¹⁸ draw a distinction between 'related' and 'unrelated' future costs and demonstrate that the ranking of programmes is unaffected by the consistent inclusion or exclusion of 'unrelated' future costs. This appears to offer, if not a resolution, some comfort. However, as noted by Weinstein and Manning¹⁹ it is virtually impossible in practice to disentangle the two categories of cost. Moreover, Meltzer²⁰ in the same issue of the *Journal of Health Economics* demonstrates that it does matter whether or not future health care costs are included, and (more ominously for the analyst) that future *non-health care* costs must also be considered.

Given these theoretical and practical concerns it is not surprising that the Washington Panel recommended that the treatment of future costs be left to the discretion of the analyst and where appropriate be subject to sensitivity analysis.²¹ A further unresolved (and related) issue is the appropriate treatment of productivity losses. The Washington Panel argue for taking such losses into account via their impact on the quality of life,²¹ whereas Brouwer *et al*²² argue strongly for taking account of the costs of lost productivity in the numerator rather than the denominator of the cost-effectiveness ratio.

Discounting

Discounting is the means by which future costs and benefits are weighted so as to take into account their timing. The underlying notion is that a lower weight should be attached the further in the

ISSUES CONCERNING METHODS

future an event occurs. Discounting practices often play a central role in determining the relative cost-effectiveness of different interventions. If evaluations are undertaken on an incorrect basis the quality of decision making will suffer and health service efficiency will be reduced. Moreover, confusion or lack of agreement over standard discounting practice potentially undermines the credibility and value of economic evaluation. Furthermore, an improved understanding of how individuals view future costs and benefits could be valuable with respect to the design of policies for the promotion of health.

There are a number of issues with respect to the treatment of future costs and benefits which are to some extent unresolved:

- ◆ what discount rate(s) should be used?
- ◆ should all costs and benefits be treated similarly (discounted at the same rate)?
- ◆ the appropriate discounting model.

The first is important for a subset of projects where the ranking is particularly sensitive to the choice of discount rate. While there is more or less agreement that the social discount rate should take into account the social opportunity cost of capital and the social time preference rate, the rates agreed vary between countries. The only drawback of such variety arises with respect to the making of international comparisons but this can be fairly easily overcome by reporting results for a range of the leading rates.

To economists working outside of health economics it might appear odd that there is any question whatsoever regarding the differential discounting of health care costs and health benefits. Any stragglers should have been neatly corralled by the admonition of Fuchs and

Zeckhauser²³ 'Self-respecting economists should not adjust discount rates for externalities stretching to the future or use different rates because it is health that is being valued.' However, the issue of whether or not future health effects should be discounted at the same rate as future costs has been raised. Moreover, current advice in England and Wales²⁴ (to use 6% for financial values and 1.5% – 2% when health effects are quantified in physical units) is currently at odds with that elsewhere, for example, Canada²⁵ and the USA²⁶, (5% and 3% respectively applied to *all* costs and benefits).

The question of whether or not future health effects should be discounted at the same rate as future costs has generated considerable interest in recent years, and much of the empirical work that has been undertaken has aimed to inform this debate. There are difficulties in estimating such rates and debate concerning whether or not it is possible to identify individuals' rates of time preference with respect to future health effects.²⁷

The literature on discounting in health care has been almost universally based on the Discounted Utility (DU) model despite a growing body of evidence at odds with the predictions of the DU model. Loewenstein and Prelec²⁷ identify four intertemporal choice anomalies that run counter to the predictions of the DU model. These anomalies are explained by Loewenstein and Prelec in general terms with reference to future consumption, and supported by evidence from monetary choices. However, there is no reason to suppose that they are any less in evidence when the outcomes are in terms of health. As a result the model underlying standard discounting practice does not appear to be the basis for individual decision making.

The authors of the various studies of time preference rates with respect

ISSUES CONCERNING METHODS

to health have acknowledged that their results, particularly with respect to the impact of delay on implied discount rates, have been at odds with the DU model. However, there has been little systematic examination of functional forms and investigation of alternative models. One explanation for this lack of interest is that the DU model has very strong normative appeal. For many it would appear a desirable model for evaluating health care choices. Individuals may not behave according to it but this is no reason why governments and other public bodies should not.

Thus discounting is a nice example of the tension apparent elsewhere in the application of cost-effectiveness analysis to health care. As Weinstein and Manning¹⁹ note 'For costs, CEA is willing to use opportunity costs (which arise out of an economy with a particular distribution of income), but for health outcomes it steps back and treats outcomes independently of the income distribution'. Some may feel slightly uncomfortable incorporating some individual health state preferences while systematically ignoring individual time preferences.

Valuation of health and health care

Health economics has made a major contribution to the valuation of health and health care. However, despite the progress made in the last 25 years the evidence base for the methods used in this area is perhaps surprisingly low. This generalisation could be supported with a range of examples but three will suffice:

- ◆ methods by which health state preferences are elicited
- ◆ the quality-adjusted life-years (QALY) versus healthy-years equivalents (HYE) debate
- ◆ contingent valuation of health outcomes and health services.

Ever since the initial conceptual breakthrough²⁸ that yielded the quality-adjusted life-year, a variety of methods for eliciting health state preferences have been explored. These include: standard gambles; time trade-offs; category rating and visual analogue scales; and person trade-offs.³ Different researchers prefer different methods. Relatively little work has compared two or more methods and much of this has been characterised by fairly small sample sizes. The usual finding is that the different methods do not produce scales which differ in a systematic and reproducible fashion. One fundamental problem is with respect to establishing the validity of any particular measure. A further difficulty is with respect to whether or not the various methods of eliciting preferences generate scales with interval properties.

Although there is no agreement as to how to make the adjustment for quality, there is widespread support (among health economists) for the notion of QALYs. There is discussion as to how the information should be used, illustrated by the debate over cost per QALY league tables. While the measurement of QALYs is clearly the dominant method, alternatives to QALYs have been proposed and vigorously defended. The main debate has been between QALYs and HYE_s.^{29,30} That there has been a debate might be taken to be indicative of a concern over the evidence base of the methods used. However, the very limited recourse to empirical data that has been one characteristic of this debate might suggest otherwise.

Although the research efforts have been less widespread than in the case of QALYs and largely restricted to the last 10 to 15 years, a similar story might be recounted with respect to contingent valuation and specifically with respect to willingness-to-pay. O'Brien and Gafni³¹ present an excellent overview of progress to date. They stress that there are a large number of issues concerning how to collect contingent

ISSUES CONCERNING METHODS

valuation information. These include: what type of measure to use; what questions to ask and of whom; and how to ask questions in order to minimise bias and increase precision. They conclude that there are a growing number of studies ‘where the practice of measurement is not consistent with the principles of CBA’.

There will be increasing pressure in this area to develop guidelines, in part, because of the prominence of those developed by the National Oceanic and Atmospheric Administration (NOAA) panel³² on contingent valuation for environmental damage assessment. However, a sufficient body of good quality studies do not yet exist to enable satisfactory judgements to be made regarding the appropriate conduct of contingent valuation studies in health care.

Summary

The important message is not that progress has been slight but rather that it continues to be made and many research issues remain. This is of course good news for those who are primarily interested in the development of methods but less so for those who want to evaluate programme X or product Y. As ever a balance must be struck. Too vigorous self-criticism may undermine the campaign to get economic evaluation more routinely used in health care decision making. Too little will slow progress with respect to developing more useful tools with which to play a greater role in informing such decision-making. The implications for guidelines for economic evaluation of the necessarily continuing debate over methods are considered next.

GUIDELINES FOR ECONOMIC EVALUATION

As interest in the economic evaluation of health care increased, it was inevitable that guidelines for these economic evaluations would be developed. In fact, the early guidelines predate the massive expansion of economic evaluation activity seen in the past 15 years.³³ Guidelines are valuable in a number of ways. First, they encourage the analyst to be more explicit. Second, and partly as a result of increased explicitness, they increase the comparability of studies. Third, more speculatively they may bring benefits in terms of quality control. Thus at first sight they appear to be quite a reasonable notion. However, as will be emphasised below there are also potential disadvantages of putting further effort into the development of guidelines. The limited knowledge concerning the effects of guidelines and the costs and benefits associated with their implementation is also highlighted.

Guideline family

The guideline family has a growing number of members. They can usefully be classified into three broad groups:

- ◆ guidelines covering virtually every stage of an evaluation (for example, DoH and Washington Panel)
- ◆ guidelines with respect to the conduct of specific parts (for example, willingness-to-pay)
- ◆ guidelines concerning the review and presentation of studies.

The first group could be further subdivided into guidelines for the conduct of economic evaluations which will then play a role with respect to the reimbursement of pharmaceuticals, and the rest. The former include Australian guidelines and those for Canada. A recent and notable example of the latter comes from the Panel on Cost-

GUIDELINES FOR ECONOMIC EVALUATION

effectiveness in Health and Medicine (the 'Washington Panel') convened by the US Public Health Service.²¹ It is notable for the depth of the discussion of the literature underlying their recommendations.

A key element in the recommendations of the Washington Panel is the Reference Case which is a 'standard set of methodologic practices that an analyst would seek to follow'. The analyst would retain flexibility by being able to augment the Reference Case but comparability would be secured across studies, at least via the Reference Case. At first sight the Reference Case concept, that all evaluations should be in part conducted on a common basis so that different studies can be readily compared, is appealing. Whether in practice it performs well will depend on a number of factors. To what extent will varied approaches be discouraged by the introduction of a standard, and is more lost or gained by such standardisation? Will trust in, or support for, the Reference Case be such that decision makers may be encouraged to make comparisons which are not strictly warranted? Will the Reference Case come to be regarded as the most appropriate indication of cost-effectiveness rather than simply facilitating comparison between studies? These important questions are empirical issues upon which we are not in a position to draw firm conclusions.

Formal guidelines of the second type, those pertaining to specific elements in the conduct of economic evaluations, are rare but less formal guidelines are present at the end of most reviews of the state of the art, for example, of willingness to pay³⁴ or of sensitivity analysis.³⁵

The best developed guidelines for the review and presentation of studies are those recently produced for the *British Medical Journal*.³⁶ These are aimed at assisting reviewers of submitted papers and at informing potential authors as to what is expected. It is anticipated

that the combined effect for the two groups will be to raise the standard of the papers appearing and expedite the review process. To the extent that such guidelines are essentially reporting guidelines, they can be distinguished from guidelines on the conduct of economic evaluation. They should be capable of promoting explicitness and enhancing comparability without necessarily having too restrictive an effect on the conduct of the evaluation. The most important issue with respect to the use of reporting guidelines is how flexibly they are applied. There may be relatively few papers which could receive an entirely clean bill of health from referees fully applying the *BMJ* guidelines and still be of a suitably short length and written in a not too technical style to satisfy other editorial considerations.

Drummond³⁷ identified three areas of continuing methodological debate. The key point for Drummond was that ‘several areas of methodological debate still remain. Therefore it would currently be inadvisable to apply cost-effectiveness guidelines in a rigid or mechanistic way’. (This may appear somewhat ironic given his leading role in the development of the *BMJ* guidelines). He was on balance optimistic about the requirement for the use of economic evaluation prior to reimbursement of pharmaceuticals but noted that it was likely to raise searching questions about economic evaluation methods. These searching questions about economic evaluation methods have barely been asked by economic evaluators or those using evaluative information. As a result the areas of continuing methodological debate remain as relevant now as five years ago.

Concerns over guidelines

Guidelines at first sight appear to be a quite reasonable notion. It is the case that the recent emphasis on guidelines, not just in health

GUIDELINES FOR ECONOMIC EVALUATION

economics but also throughout medicine, has possibly engendered a certain amount of guideline fatigue. Few would argue against the desirability of explicitness, greater comparability and improvements in the quality of evaluations.

However, there are a number of reasons why guidelines (or more precisely continuing efforts to develop guidelines) should not necessarily be regarded in a positive light. First, they presuppose a fair measure of consensus and imply particular levels of knowledge. Second, they may have unsought effects on behaviour, for example, by discouraging research. Third, they take resources to develop and these resources obviously have an opportunity cost. Fourth, there is a sparsity of evidence regarding their effectiveness (and cost-effectiveness). Fifth, they may entrench sub-optimal practice.

With regard to implied consensus, it would be misleading to suggest that there is not consensus on many issues. The point is rather that there are also numerous important areas where there is not agreement. Of course some of this disagreement will reflect ideological differences on which individuals might not ever be expected to wholly agree, but much of the disagreement is simply because we lack knowledge and more research is required: or possibly we have failed to assimilate information which is available to us. Whichever is the case, undertaking further research or attaining greater familiarity with what is known requires time and energy and a recognition that to do so is important. Further refinement of guidelines competes for these scarce resources and emphasises consensus rather than the unresolved issues.

The theme of research being discouraged requires further discussion. There is a danger that as a consequence of the development and promulgation of guidelines that certain practices become 'set in stone'

and alternative approaches become harder to develop and implement. It might become harder to obtain funding if the protocol diverges significantly from guidelines. It may become harder to publish research based on methods not approved of by whichever guidelines currently hold sway. The discouragement of alternative approaches could thus become institutionalised. This would probably, on balance, be desirable if there were not so many unresolved issues and if the economic evaluation of health care was not at such a relatively early stage of development. However, given that it is, there appears a danger that in attempting to encourage good practice we discourage good research.

The opportunity cost of the development of guidelines for economic evaluation is a theme developed by Maynard³⁸ in a paper subtitled *reinventing the wheel?*. His main contention is that ‘investment in the improvement of guidelines and development of consensus about best practices yields a small marginal product’. One problem he highlights is the way in which the development of guidelines diverts attention away from unresolved issues. Similarly, instead of seeking to refine guidelines, efforts might give a greater return monitoring the practice of economic evaluation. A range of studies which review the conduct of economic evaluation have reached similar conclusions regarding the small proportion which attain the highest standards.³⁹ Indeed dissatisfaction over the standards of economic evaluation is one of the engines which drives interest in guidelines. However, it is clearly not the lack of guidelines which is responsible for low quality evaluations, in that guidelines have been available for years. Thus, if the quality of many evaluations is a major concern, the preferred policy for raising standards is unlikely to involve the further development of guidelines, but rather needs to consider why, when guidelines have existed for

GUIDELINES FOR ECONOMIC EVALUATION

such a long time, they have had so little impact.

What evidence have we on the effectiveness (and cost-effectiveness) of guidelines? With respect to guidelines on medical practice, a considerable volume of research is under way (on effectiveness if not cost-effectiveness). Rather less is known about the effectiveness of guidelines for economic evaluation. The casual observation of continued dissatisfaction with the quality of economic evaluations despite long-established guidelines suggests that guidelines in this area are not particularly effective in changing practice.⁴⁰ If this is the case fears about a deleterious impact on research are possibly unfounded. However, concern that the development and refinement of guidelines is not a good use of scarce resources grows.

At least one prominent health economist has voiced concern that some of the interest in the development and use of guidelines is driven by commercial rather than scientific motives.⁴¹ In short, guidelines may be of potential benefit to the pharmaceutical industry but are they of potential benefit to a wider community?

Summary

The issue is not whether or not guidelines for economic evaluation are a good or bad thing. Rather the issues are: what efforts should be made to augment or revise existing guidelines, or develop new guidelines?; what should the guidelines include and what should they exclude?; and how should guidelines be implemented? In addressing these issues it is important to recognise that relatively little is known about the impact of such guidelines.

CONCLUSIONS

The application, in the last three decades, of the techniques of economic evaluation to health care developments can be portrayed as a success story, particularly if the focus is on the growth of activity and on influencing the way people speak, and possibly think, about health care. If a different criterion is adopted, namely the impact of all of this activity on decision-making, a quite different picture is possibly formed. Even focusing on the activity of economic evaluation rather than the value of the activity, as has this review, there are grounds for questioning the nature of the health economists' success.

While there are no doubt exceptions, generally those producing economic evaluations have tried to base their conclusions on the best possible evidence, although the quality of the data upon which evaluations have been based and the standards of reporting have often been poor. However, there must be considerable concern about the evidence base in support of the methods applied. We must be careful not to claim too much for economic evaluation. When properly applied it can be a valuable addition to informed decision making. However, in winning acceptance for this approach or in stimulating demand for it, it is all too easy to play down its limitations. There is clearly an important balance to be struck between singing its praises and undermining its position by setting overly exacting standards and being negative and dismissive regarding failures to attain such standards.

The positive message is that sustained progress is being made. More and more is known about how to undertake economic evaluations. The expert base is broadening. One important question concerns the best way to further this process and in particular the role of guidelines, checklists, etc. in raising standards and encouraging continuing development of the economic approach.



CONCLUSIONS

The view expressed in this paper has been characterised by a limited enthusiasm for the development of guidelines for economic evaluation. They do have a role, particularly with respect to the reporting of evaluations, but the over-riding concern must be that too widespread and mechanistic use of guidelines will undermine much needed progress in research. We must not attempt to set guidelines with respect to the conduct of economic evaluations in areas where there are many unanswered research questions, several of which have been highlighted. There is the danger of falling into error by making pronouncements on best or most appropriate practice when we simply do not have adequate information.

Furthermore, there is a dynamic aspect to be considered. Increased emphasis on research rather than the application of standardised approaches offers benefits in terms of improved decision making, not just with respect to the current urgencies but stretching into the future. A questioning and active research culture will stimulate future advances on a scale which is unlikely to be matched in a culture where the emphasis is on the prescription of rules for the conduct of economic evaluation. Moreover, we must of course acknowledge scarcity with respect to research and evaluative capacity and practice what elsewhere we preach. The return to further efforts at standardisation must be questioned in the light of the considerable opportunities to advance our knowledge elsewhere.

Such negativity or caution concerning guidelines must be tempered by an awareness that it might certainly be viewed as self-serving for those furthering research careers to argue against steps to make routine and more widely understood their stock in trade.

If health economists were to evaluate clinical guidelines they would

consider whether or not the proposed guidelines represented cost-effective health care and would also be interested in the relative costs and benefits of different ways of changing clinical practice. Is there any good reason not to ask the same questions in the case of guidelines for economic evaluation? Evidence on the cost-effectiveness of guidelines would be particularly welcome.

Economic evaluation of health care has developed quite significantly in the past thirty years. The growth in the demand for economic evaluations has been stimulated by some jurisdictions requiring economic evaluation in the case of new pharmaceutical products. Another stimulus has been the continuing interest in cost containment manifest most recently in a commitment to obtain value for money from the resources devoted to health care. While there is no reason to suppose that the future prospects for the application of economic evaluation are anything but good, further progress depends on not only improving the standards of economic evaluation but also continued research into the methods of evaluation. If further requirements for economic evaluation are introduced, the rising demand should be viewed as an opportunity to improve the quality of evaluations and not simply increase their number.

Efforts to improve guidelines for the conduct of economic evaluation might have some positive effect on raising standards but are a fairly indirect approach. It is not clear that they will do anything to encourage the development of the methods of evaluation – the reverse appears likely. Guidelines for the reporting of economic evaluations are less likely to have such a discouraging effect on research and since they tend to encourage the provision of more detailed information have a positive contribution to make. However, reporting guidelines could also have potentially negative effects if they are interpreted and

CONCLUSIONS

applied without some flexibility.

The economic evaluation of health care has been labelled a half-way technology in that it has not yet reached an advanced stage where it can be applied routinely.⁴² Thus it remains 'beneficial in specific cases; non-routine and therefore relatively expensive; dependent on specialist expertise; driven by intellectual curiosity; and, quite rightly, having to justify itself continually in an increasingly sceptical world'. We do not move forward from this half-way position by promoting unwarranted images of consensus but rather by continuing to develop and refine our methods. However, where agreement is possible, say with respect to reporting standards, it should be exploited. Also, any such agreement might be further exploited, as recently suggested by an influential commentator,⁴³ by the development of an audit infrastructure.

REFERENCES

1. Maynard A, Sheldon TA. Health economics: has it fulfilled its potential? In: Maynard A and Chalmers I, eds. *Non Random Reflections on Health Services Research on the 25th Anniversary of Archie Cochrane's Effectiveness and Efficiency*. London: BMJ Publishing, 1997.
2. NHS Centre for Reviews and Dissemination. *Making cost-effectiveness information accessible: the NHS economic evaluation database project*. York: University of York, 1996.
3. Torrance GW. The measurement of health state utilities for economic appraisal. *Journal of Health Economics* 1986; 5: pp1-30.
4. Fenn P, McGuire A, Backhouse M, Jones D. Modelling programme costs in economic evaluation. *Journal of Health Economics* 1996;15: pp115-26.
5. Coyle D. Statistical analysis in pharmaco-economic studies: a review of current issues and standards. *PharmacoEconomics* 1996; 9: pp506-16.
6. Chalmers I, Altman DG, eds. *Systematic Reviews*. London: BMJ Publishing, 1995.
7. Drummond MF, O'Brien B, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: OUP, 1997.
8. Sculpher M, Drummond M, Buxton M. The iterative use of economic evaluation as part of the process of health technology assessment. *Journal of Health Services Research and Policy* 1997; 2: pp26-30.
9. Davies L, Palmer S, Richardson G, Torgerson D. Economic evaluation: what? when? how? Paper presented to the Health Economists' Study Group, Liverpool, 1997.
10. Adams ME, McCall NT, Gray DT, Orza MJ, Chalmers TC. Economic analysis in randomized controlled trials. *Medical Care* 1992; 30: pp231-43.
11. Parsonnet J, Harris RA, Hack HM and Owens DK. Modelling cost-effectiveness of *Helicobacter pylori* screening to prevent gastric cancer: a mandate for clinical trials. *The Lancet* 1996; 348: pp150-4.
12. Barr R, Furlong W, Henwood J, Feeny D, Wegener J, Walker I, Brain M. Economic evaluation of allogeneic bone marrow transplantation: a rudimentary model to generate estimates for the timely formulation of clinical policy. *Journal of Clinical Oncology* 1996;14: pp1413-20.
13. O'Brien B. Economic Evaluation of Pharmaceuticals. Frankenstein's Monster or Vampire of Trials? *Medical Care* 1996; 34: DS99-DS108.

REFERENCES

14. Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996; 312: pp1215-8.
15. Office of Health Economics. *The pros and cons of modelling in economic evaluation*. OHE Briefing No. 33, London, 1997.
16. Sheldon TA. Problems of using modelling in the economic evaluation of health care. *Health Economics* 1996; 5: pp1-11.
17. Buxton MJ, Drummond MF, van Hout BA, Prince RL, Sheldon TA, Szucs T, Vray M. Modelling in economic evaluation: an unavoidable fact of life. *Health Economics* 1997; 6: pp217-27.
18. Garber AM, Phelps CE. Economic foundations of cost-effectiveness analysis. *Journal of Health Economics* 1997; 16: pp1-31.
19. Weinstein MC, Manning WG. Theoretical issues in cost-effectiveness analysis. *Journal of Health Economics* 1997; 16: pp121-8.
20. Meltzer D. Accounting for future costs in medical cost-effectiveness analysis, *Journal of Health Economics* 1997; 16: pp33-64.
21. Gold MR, Siegel JE, Russell LB, Weinstein MC. *Cost-effectiveness in health and medicine*. Oxford University Press, Oxford, 1996.
22. Brouwer WBF, Koopmanschap MA, Rutten FFH. Productivity costs measurement through quality of life? A response to the recommendation of the Washington Panel. *Health Economics* 1997; 6: pp253-60.
23. Fuchs V, Zeckhauser R. Valuing health – a “priceless” commodity. *American Economic Review* 1987; 77: pp263-8.
24. Department of Health. *Policy Appraisal and Health*. HMSO, London, 1995.
25. Torrance GW, Blaker D, Detsky A, Kennedy W, Schubert F, Menon D, Tugwell P, Konchak R, Hubbard E, Firestone T. Canadian guidelines for economic evaluation of pharmaceuticals. *PharmacoEconomics* 1996; 9: pp535-59.
26. Gafni A. Time in health: can we measure individuals’ pure time preferences? *Medical Decision Making* 1995; 15: pp31-37.
27. Loewenstein GF, Prelec D. Anomalies in intertemporal choice: evidence and an interpretation. *Quarterly Journal of Economics* 1992; 107: pp573-97.
28. Torrance GW. *A Generalized Cost-Effectiveness Model for the Evaluation of Health Programmes*. Hamilton, Ontario: Faculty of Business, McMaster University, 1970.

29. Gafni A, Birch S. Economics, health and health economics: HYE versus QALYs. *Journal of Health Economics* 1993; 11: pp325-39.
30. Symposium on QALYs and HYE. *Journal of Health Economics* 1995; 14: pp1-45.
31. O'Brien B, Gafni A. When do the "Dollars" make sense? Toward a conceptual framework for contingent valuation studies in health care. *Medical Decision Making* 1996; 16: pp288-99.
32. National Oceanic and Atmospheric Administration. Report of the NOAA panel on contingent evaluation. *Fed Reg.* 1993; 58: pp4607-14.
33. Williams A. The cost-benefit approach. *British Medical Bulletin* 1974; 30: pp252-6.
34. Morrison GC, Gyldmark M. Appraising the use of contingent valuation. *Health Economics* 1992; 1: pp233-43.
35. Briggs A, Sculpher M. Sensitivity analysis in economic evaluation: a review of published studies. *Health Economics* 1995; 4: pp355-71.
36. Drummond M, Jefferson T. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996; 313: pp275-83.
37. Drummond M. Cost-effectiveness guidelines for reimbursement of pharmaceuticals: is economic evaluation ready for its enhanced status? *Health Economics* 1992; 1: pp85-92.
38. Maynard A. Economic evaluation techniques in healthcare: reinventing the wheel? *PharmacoEconomics* 1997; 11: pp115-8
39. Udvarhelyi, IS, Colditz GA, Rai A, Epstein AM. Cost-effectiveness and cost-benefit analysis in the medical literature: are methods being used correctly? *Annals of Internal Medicine* 1992; 116: pp238-44.
40. Mason J, Drummond M. Reporting guidelines for economic studies. *Health Economics* 1995; 4: pp85-94.
41. Evans RG. Manufacturing consensus, marketing truth: guidelines for economic evaluation. *Annals of Internal Medicine* 1995; 123: pp59-60.
42. Hutton J. Economic evaluation of healthcare: a half-way technology. *Health Economics* 1994; 3: pp1-4.
43. Reinhardt UE. Making economic evaluation respectable. *Social Science and Medicine* 1997; 45: pp555-62.

During the last 25 years economists have been successful in persuading researchers and policy makers that economic evaluation can make an important contribution to informing clinical decision making and policy choices. The research output of those involved in economic evaluation is increasing rapidly: economic imperialism is on the march!

However, despite the vigorous endeavours of those involved in evaluation to base their conclusions on the best possible evidence, the evidence base remains deficient and at times inadequate to support the demanding methods used.

The practice of economic evaluation continues to improve but serious methodological problems remain unresolved. John Cairns argues that these problems may be ignored if the authors of practice guidelines focus on the generation of consensus. The role of economic evaluation in informing health care choices is fundamental. There is a risk that 'guidelines mania' will create a false aura of consensus about methods. There is a need not only to improve methods but also to audit practice systematically. Economic evaluation in health care remains an evolving technology which needs to be applied with care and flexibility. If practised intelligently, it can be an important tool for informing choices. If done badly, it can corrupt the knowledge base and waste resources.