

DEVELOPMENTS IN ECONOMIC EVALUATION IN HEALTH CARE: A REVIEW OF HEED

Clive Pritchard, OHE*

Contents

1	Introduction	1
2	Content of the database	3
3	Topics of evaluation	5
4	Methods of evaluation	6
5	Sponsorship of studies	11
6	Review of a sample of studies	11
7	Conclusions	16
	References	17
	Appendix 1: HEED sample references	18

a comparison between the two databases) provide complementary sources of evidence to clinical databases such as Medline and Embase. The burgeoning demand for economic evidence indicated by Box 1 suggests that established databases such as HEED and NHS EED will have a greater part to play in assembling the economic evidence base on which decisions are being made and will increasingly be made in future.

The objective of this briefing is to set out the types of studies included on HEED and to present an analysis of how these studies have changed over time, in terms of broad characteristics such as types of evaluation, disease areas covered and study design (similar to a previous review presented in Pritchard, 1998). Comparisons are made between two five year time periods, 1992–1996 and

1 INTRODUCTION

Health care decision makers are becoming increasingly concerned with obtaining value for money and therefore with the use of economic evidence, particularly as a criterion for the reimbursement of new pharmaceuticals. Box 1 indicates the range of countries in which economic considerations have been introduced into the decision making process together with some of the associated policy applications.

As the application of economic criteria has become more widespread, so sources of information on the cost-effectiveness of health care technologies have acquired a greater level of importance. One such source is HEED – the Health Economic Evaluations Database, which has contributed to the evidence base called upon in around a quarter of the first 75 technology appraisals completed by the National Institute for Clinical Excellence (NICE). In this context, HEED and NHS EED (see Box 2 for

Box 1

Use of economic evidence in health care decision making

Fourth hurdle i.e. reimbursement decisions: Australia, Baltic countries, Belgium, Canada (British Columbia, Ontario), Czech Republic, Denmark, Finland, France, Hungary, Netherlands, New Zealand, Norway, Portugal, Russia, Slovenia, Sweden, US managed care formularies

Pricing negotiations: Australia, France, Italy, New Zealand

Advice to the health services: England and Wales (NICE), Scotland, Wales (AWMSG)

Risk sharing arrangements: Australia, New Zealand, UK

* Statement of interest: Clive Pritchard is employed by the Office of Health Economics and is contracted for part of his time to HEED

Box 2

Comparison of HEED and NHS EED

FEATURES	HEED	NHS EED
Literature sources used	Electronic databases (Medline, Embase), grey literature, journal hand searches	Electronic databases (Current Contents, Medline, Embase, CINAHL), grey literature, journal hand searches
Inclusion criteria	Economic evaluations, cost-outcome descriptions, cost analyses, cost of illness studies	Economic evaluations
Full structured reports included	Applied studies, reviews of applied studies, older methodological studies	Applied studies
Bibliographic entries	More recent methodological studies, policy papers, editorials, letters	Cost, review and methodology records
Search fields	All fields, or 53 individual fields, including bibliographic details, abstract, countries applicable, data sources ICD-9, ATC codes, drug names, key results	All fields, bibliographic details, abstract, addition date, MeSH subject headings, HTA organisation
Target audience	Health economists in the public and private sector	NHS decision makers
Number of references	28,000 (14,500 fully reviewed), 300 references added per month (100 reviewed)	14,000 (3,700 structured abstracts), around 40–50 structured abstracts added per month

1997–2001. This briefing also presents a summary of a group of cost utility analyses which have employed modelling methods and discusses some of the key methodological issues involved in conducting this type of study. The briefing should be of interest to practitioners and consumers of economic evaluations.

Section 2 summarises the annual numbers of reviewed and bibliographic entries and the distribution of reviewed studies by study type. Subsequent sections focus on applied studies, the most common form of reviewed study, and the ones to which HEED's standard report format is best suited. Section 3 considers the distribution of applied studies by type of technology assessed, the coverage of HEED being in line with a broad definition of technology, such as that by Jonsson et al. (2002):

"the drugs, devices, medical, and surgical procedures used in health care, as well as measures for prevention and rehabilitation of disease, and the organisational and support systems in which health care is provided".

Section 3 also looks at the distribution of applied studies for all technologies by therapeutic area according to ICD-9 code and, for pharmaceutical studies, by ATC code. Section 4 summarises the distribution of applied studies by type of economic study, the study designs used in cost minimisation, cost effectiveness, cost utility and cost benefit analyses, and the methods used to value health outcomes in cost utility analyses and discusses the use of discounting and sensitivity analysis. Section 5 reports on sources of sponsorship for all applied studies and for pharmaceutical studies while section 6 reviews a sample of recent studies included in HEED and section 7 concludes.

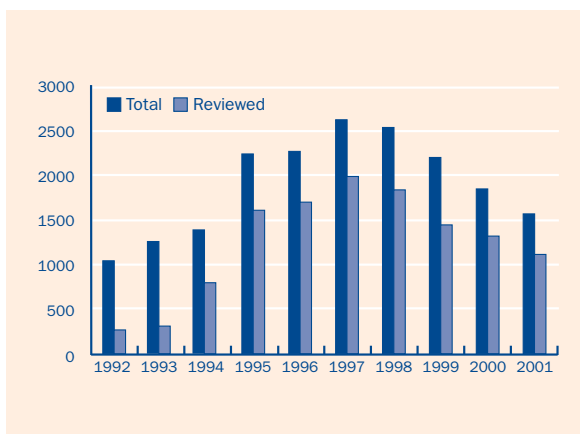
2 CONTENT OF THE DATABASE

HEED provides detailed reports on economic evaluations and other types of cost analysis of health care technologies. It is compiled primarily from monthly searches of Medline and Embase, coupled with hand searches of key journals, and trawls of the important grey literature. Reports compiled according to a standard format on eligible studies (reviewed studies) are supplemented with bibliographic references, as noted in Box 2. The overview presented here is based largely on searches of HEED conducted in late 2003. Box 3 lists the types of studies eligible for inclusion in HEED.

Figure 1 presents annual totals for references included in HEED over the ten year period from 1992. Those studies described as 'reviewed' have a detailed report compiled according to a standard template. Those references entered only in bibliographic form for this period are drawn from several sources. Firstly, studies other than applied studies or reviews of applied studies have, in recent years, been entered in bibliographic form only. The most numerous of these are methodological studies; with a few exceptions, those with a publication date of 1997 and later, have been affected. Secondly, non-English language articles are entered in bibliographic form only. Thirdly, references drawn from the reference lists of reviewed studies appear in bibliographic form. If these are subsequently reviewed, the bibliographic entry is removed (similarly, all studies which are about to be reviewed appear initially as a bibliographic entry which is later removed).

It is worth noting that HEED also contains the Battelle (Elixhauser et al., 1992) and Wellcome (Backhouse et al., 1992) bibliographies. Of these,

Figure 1 HEED entries by year of publication 1992–2001



only the second will feature in the numbers underlying Figure 1 since it is the only one containing references with a publication date as late as 1992. We can see that, overall, bibliographic entries exceed reviewed studies for the first two years of the period but reviewed

Box 3 Studies included in HEED

The following classifications are used: Applied Study, Review of Applied Studies, Methodological, Government/Public Policy, Editorial, Letter, Other.

Applied studies are those which provide some original analysis of the cost impact of an intervention or bring together information on costs and outcomes. This could be done by means of primary data collection or through the use of secondary data sources, either in a structured way (e.g. a decision analysis model) or unstructured way. Studies which consider both costs and outcomes include both economic evaluations and cost-outcome descriptions. An economic evaluation is the "comparative analysis of alternative courses of action in terms of both their costs and consequences" (Drummond et al., 1997).

Cost-outcome descriptions, on the other hand, do not provide an explicit comparison of interventions. Both types of study are classified according to the following types of evaluation:

- Cost minimisation analysis (CMA)
- Cost effectiveness analysis (CEA)
- Cost utility analysis (CUA)
- Cost benefit analysis (CBA)
- Cost consequences analysis (CCA)

The first four of these are defined in the standard way (Drummond et al., 1997). CCA is an additional category of study which presents data on costs and outcomes separately rather than in the form of a ratio and which will tend to capture most cost-outcome descriptions. HEED also includes cost analyses which compare the relative costs of alternative interventions and cost of illness studies which estimate the aggregate costs of a disease at a national, supra-national or local (e.g. state, provincial) level.

Figure 2 Distribution of reviewed studies by study type 1992–1996 (n=4910)

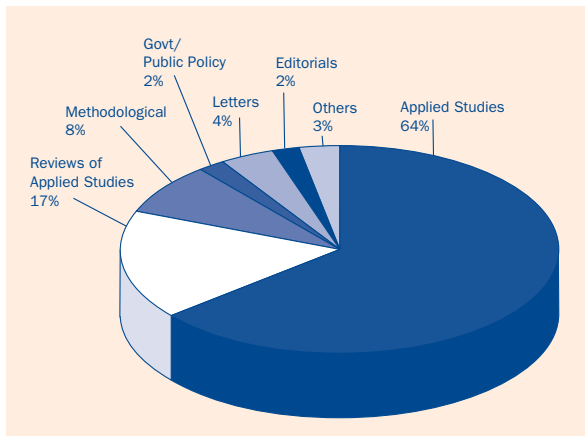
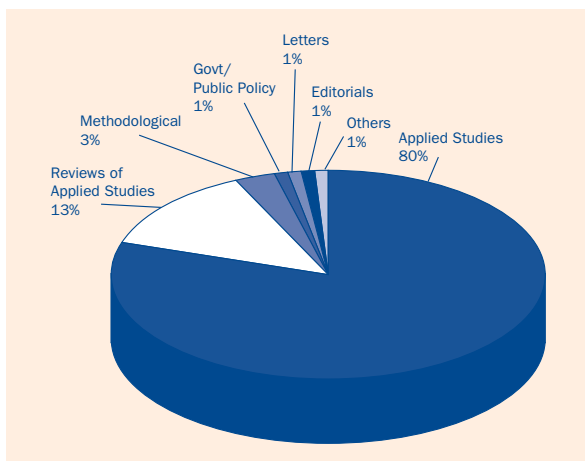


Figure 3 Distribution of reviewed studies by study type 1997–2001 (n=7917)



studies move ahead thereafter. Although, overall, nearly half of the 28,000 or so references on HEED are in bibliographic form, the balance is very much towards reviewed studies in later years.

The chart suggests that the number of studies in the literature rose to a peak and then declined. This can partly be explained by a tightening of the inclusion criteria in the last five years to exclude some types of study on which a full report would previously have been completed. Firstly, we

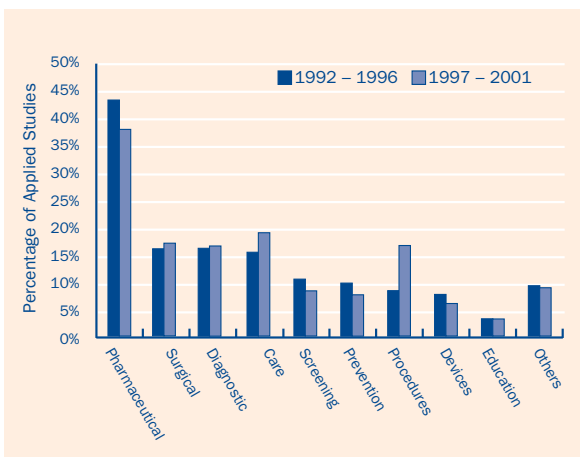
exclude applied studies which present costs simply as key items of resource use, rather than monetary values and, secondly, we discard reviews of applied studies if they cover only a few studies or provide insufficient detail of the studies included in the review. In these cases, an attempt is made to ensure that the relevant original studies are included on HEED. A second reason for the pattern in Figure 1 is that there are time lags in studies appearing in the literature. A comparison with the earlier briefing indicates that a year's entries continue to be supplemented with additional references for some time afterwards. For example, the previous analysis, based on early 1998 data, recorded a total of fewer than 2,000 references for 1995 and below 1,800 for 1996, compared with around 2,300 in each year currently. Some of this increase will be due to full reviews of studies which are picked up in later years but most is likely to be due to the addition of references from reviewed studies published in subsequent years. We may therefore expect significant numbers of references to be added to the later years of the chart.

Figures 2 and 3 show the distribution of reviewed studies according to the types of article noted in Box 3. The previous report showed applied studies increasing from less than half of studies in 1992 (42%) to over two thirds (71%) in 1996. The current average for the period 1992–96 of 64% compares with the average of 80% for the following five year period. This is due partly to the decision in recent years to focus on reviewing applied studies and reviews of applied studies, with other forms of study relegated to the status of bibliographic entries. However, there is not always a clear boundary between applied studies and, for example, methodological studies. A study which applies a novel technique of analysis to an existing set of data (perhaps from a clinical trial) could be considered an applied study but might equally be classified as a methodological study. Thus, some methodological studies will continue to be added in fully reviewed form.

3 TOPICS OF EVALUATION

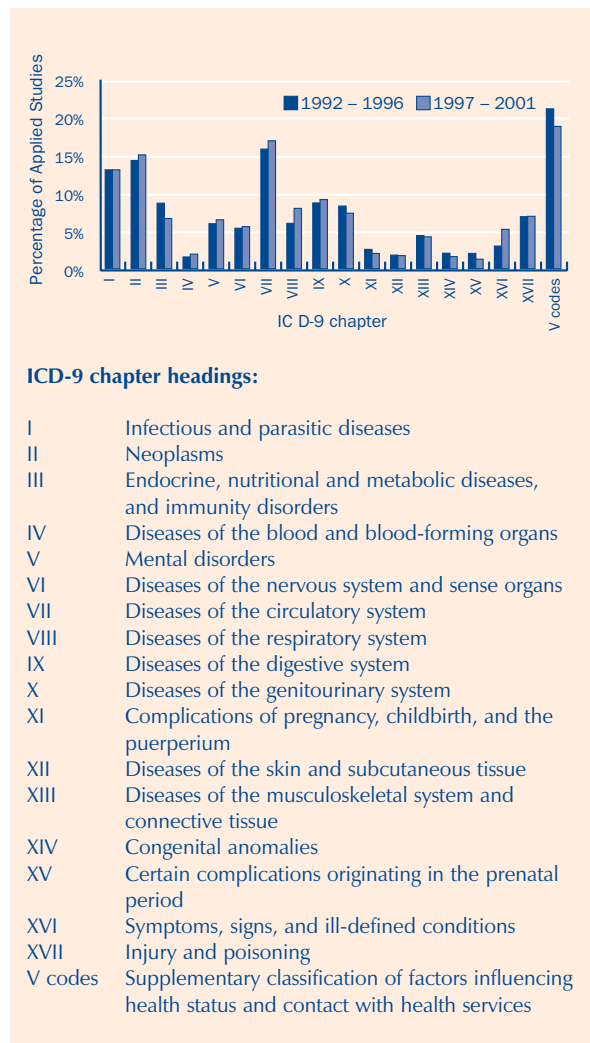
Figure 4 shows the distribution of applied studies according to type of technology assessed. The category of 'others' includes rehabilitation, dentistry, counselling, immunisation, radiotherapy and dialysis but excludes studies which were deemed to lie outside the pre-defined categories. In both time periods, these technologies were the six least investigated. The distributions are similar in both periods, but the relative importance of pharmaceuticals has continued to decline (it is worth noting that in 1992 pharmaceutical evaluations comprised one half of all applied studies). The two types of technology, care and procedures, which have shown the largest increase in their share of studies between the two time periods are perhaps the least precise. Care, for example, encompasses long term care and intensive care as well as packages of care, either in hospital or at home or elsewhere, which may not be well defined.

Figure 4 Distribution of applied studies according to type of technology



Figures 5 and 6 summarise the disease areas covered by HEED; firstly, by ICD-9 chapter for all applied studies and, secondly, by ATC chapter for applied studies which have included a pharmaceutical. The supplementary (V) codes of the ICD-9 coding system were not included in the previous briefing (Pritchard, 1998) but have been here; as the chart shows, these codes form the most commonly used of all the chapters. Their broad coverage (for example, immunisation, screening and pregnancy) means that they can be used in many instances to supplement the main disease classification. The distributions by ICD-9 chapter were similar for the two time periods, with the top five chapters in 1992–1996 also being the top five chapters in 1997–2001 (with the same

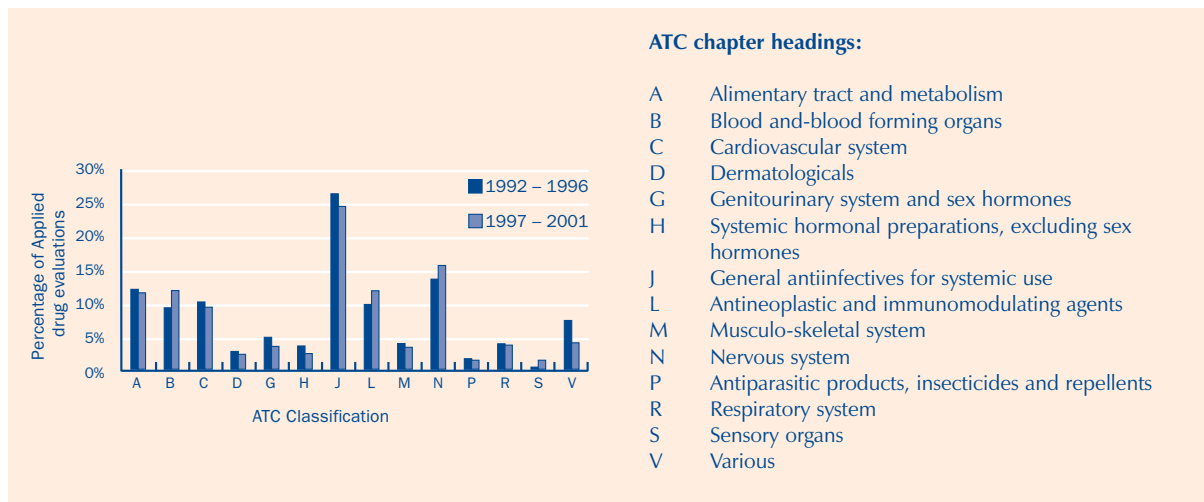
Figure 5 Distribution of applied studies by ICD-9 chapter



ranks). These were the supplementary (V) codes, diseases of the circulatory system, neoplasms, infectious and parasitic diseases, and diseases of the digestive system. The five least used chapters were the same in both time periods, but with different ranks. These were certain conditions originating in the perinatal period, congenital anomalies, diseases of the blood and blood-forming organs, diseases of the skin and subcutaneous tissue and complications of pregnancy, childbirth and the puerperium. Rankings of ICD-9 chapters varied by at most three places between the two time periods.

There was also a close correspondence between the relative importance of the 14 ATC chapters in the distribution of pharmaceutical evaluations between the two time periods, as Figure 6 shows. The chapters ranked one and two in importance

Figure 6 Distribution of applied drug evaluations by ATC chapter



for the period 1992–1996, general anti-infectives for systemic use and nervous system, were ranked one and two in 1997–2001. This shows that the increase in the importance of nervous system drugs as a subject for research observed between 1992 and 1996 has been maintained. As with the ICD-9 chapters, the five least researched ATC chapters were the same in both time periods (with two

chapters swapping ranks). These were (from the least frequent in 1997–2001): antiparasitic products, insecticides and repellents, sensory organs, dermatologicals, systemic hormonal preparations, excluding sex hormones and musculo-skeletal system. Rankings did not vary by more than two places between time periods.

4 METHODS OF EVALUATION

Table 1 shows the distribution of applied studies by type of evaluation for the two time periods considered above. Since studies can be classified according to more than one type of evaluation, percentages sum to more than 100. Little change in the distribution has been observed, with cost consequences analysis being the largest category, followed by cost effectiveness analysis and cost analysis. Those studies with shares of all applied studies in single figures range from the least useful for informing resource allocation decisions, namely cost of illness studies, to potentially the most useful, cost utility analyses (CUAs) and cost benefit analyses (CBAs). CUAs accounted for 6.6% of studies and CBAs, despite advances in the available methods, only 0.9%, for the period 1997 to 2001. This compares with 9.3% and 1.4% of studies on the NHS Economic Evaluation Database (NHS EED) for all records up to February 2000 (Nixon et al., 2000).

The NHS EED classification ‘cost effectiveness analysis’ encompasses three separate sub-categories of study. Firstly, it includes cost effectiveness analyses (CEAs) as conventionally defined (and as classified by HEED), namely

studies which report their results in terms of cost per natural unit of effect, such as cost per symptom-free day. The other two sub-categories are cost minimisation analyses (CMAs), in which alternative technologies are taken to be of equal effectiveness and cost consequences analyses (CCAs), in which costs and health outcomes are presented separately. Table 1 compares the breakdown of these three types of study for HEED and NHS EED, respectively, the label CEA being applied to the first of the three NHS EED categories. The main difference is that HEED has around twice the proportion of CCAs compared with NHS EED; this discrepancy is likely to be explained, at least in part, by the inclusion of cost-outcome descriptions in HEED but not NHS EED.

Figures 7 to 10 explore the way in which cost minimisation, cost effectiveness, cost utility and cost benefit analyses have been designed. We can be reasonably confident that these studies fulfil all the criteria of an economic evaluation, although we acknowledge the problematic status of cost minimisation analysis. In HEED, this is a term loosely applied to studies which find no significant

Table 1 Applied studies by type of evaluation

Types of economic evaluations – % of applied studies			
	1992–1996 (n = 3,118)	1997–2001 (n = 6,392)	NHS EED to Feb 2000 (n = 1,718)
Cost Minimisation Analysis (CMA)	9	6	–
Cost Effectiveness Analysis (CEA)	24	23	–
Cost Utility Analysis (CUA)	5	7	9.3
Cost Benefit Analysis (CBA)	1	1	1.4
Cost Consequences Analysis (CCA)	50	45	–
Cost of Illness	6	7	–
Cost Analysis	24	23	–
% of CMA, CEA, CCA:	(n = 2,529)	(n = 4,651)	(n ≈ 1,500)
CMA	12	9	16
CEA	29	31	55
CCA	61	61	29

difference between the outcomes associated with alternative interventions, rather than a formal demonstration of equivalence. Attaching the cost minimisation label to a study does not imply that the analysis has been performed in the most appropriate way. As Briggs and O'Brien (2001) have pointed out, estimation of an incremental cost-effectiveness ratio (ICER), with exploration of uncertainty around the ICER, may well be a preferable approach in these cases.

The pairing of Figures 7 and 8 compares evaluations of any technology with drug evaluations for the period 1992–1996, while the pairing of Figures 9 and 10 makes the same comparison for the subsequent five year period 1997–2001. Observational studies are defined as studies which have 'observational data' alone recorded for the data sources 'probability of main clinical events', 'quantities of resource use' and 'outcomes'. Similarly, evaluations alongside randomised trials have 'randomised trial' alone recorded for these key data sources. The studies selected in this way would typically be prospective economic evaluations conducted alongside a single observational study or randomised trial.

Modelling studies have been identified by selecting those studies which have 'Modelling' recorded for the 'quantities of resource use' and 'outcomes' field, without restriction on the other types of data which could be recorded (outcomes might, for example, be modelled on the basis of data from randomised trials and from epidemiological data sources). We excluded 'probabilities of main clinical events', since here modelling refers to the

generation of probabilities by statistical modelling, for example the estimation of a survival function (meta-analysis is a separate category). In comparison, in the context of outcomes and quantities of resource use, it refers to the kind of modelling with which we are mainly concerned in economic evaluation. That is, it concerns the estimation of values of these parameters by the use of techniques which synthesise data from different sources, typically through decision analysis models. Once we had defined observational studies, randomised trials and modelling studies, the 'other' category was simply a residual made up of studies not captured by the other three groups.

It is worth noting the similarity between the two time periods in the proportions of economic evaluations accounted for by observational studies and analyses conducted alongside randomised trials. From 1992–1996 to 1997–2001, the absolute and relative shares of observational studies and randomised trials as proportions of all studies changed little for pharmaceuticals or for all technologies. In both cases, randomised trials accounted for 17% of pharmaceutical evaluations which were about as likely as studies of health care technologies generally to be performed in this way, and about half as likely as studies of all technologies to be assessed by means of observational studies. Part of the difference in proportions of studies based on observational studies was compensated for by an increased likelihood of drug evaluations (which account for about half of all studies in both time periods) to be based on modelling. In both 1992–1996 and 1997–2001, the proportion of all economic

evaluations in these categories being conducted alongside a single randomised or non-randomised study was about 40%. This compares with a figure of nearly 60% for studies on NHS EED conducted on the basis of a single effectiveness study. However, our definition of a single-study evaluation identified studies in which resource use had also been assessed in a single type of study and in which none of the main data sources had been drawn from the literature. We may note that, if cost consequences analyses are added to the other categories of evaluation, then the proportion of studies conducted alongside a single randomized trial or observational study rises to over 60%.

The most notable feature of Figures 7 to 10 is the increase in the share of modelling studies in the totals for all technologies and pharmaceutical evaluations, in the former case from around one

quarter to around one third and, in the latter case, from 32% to 42%. While the point is often made that economic evaluations are increasingly being performed alongside randomised trials, the share of this type of study changed little between the two time periods presented here. The absolute numbers of studies conducted alongside randomised trials has certainly increased and randomised trials are perhaps more likely to contain an economic component than was the case in the past, but this review of HEED indicates that, of the study designs considered, analysts are increasingly turning their attention to modelling approaches as a means of estimating cost-effectiveness.

A further category of studies we considered consisted of analyses which applied modelling to extrapolate from the authors' own clinical trial, rather than being used to combine inputs from

Figure 7 Design of applied CMA, CEA, CUA, CBA – all technologies 1992–1996 (n=1144)

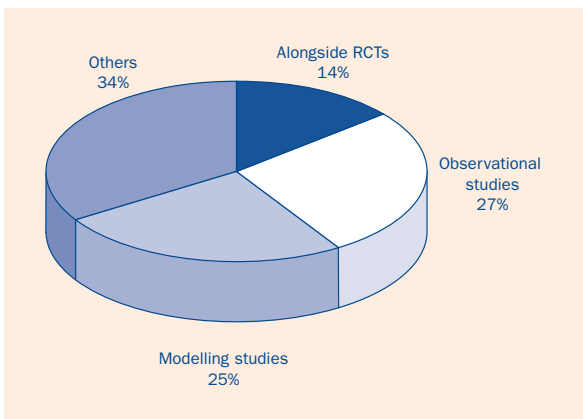


Figure 8 Design of applied CMA, CEA, CUA, CBA – drug evaluations 1992–1996 (n=578)

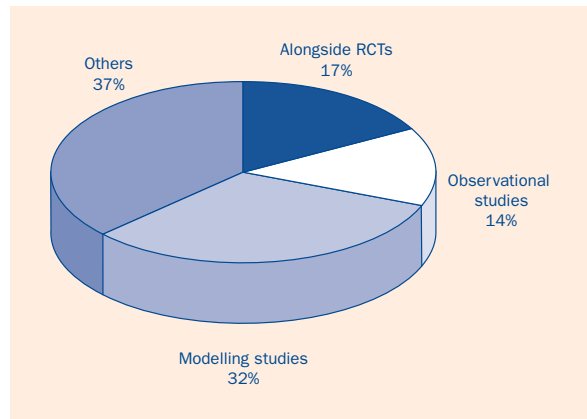


Figure 9 Design of applied CMA, CEA, CUA, CBA – all technologies 1997–2001 (n=2142)

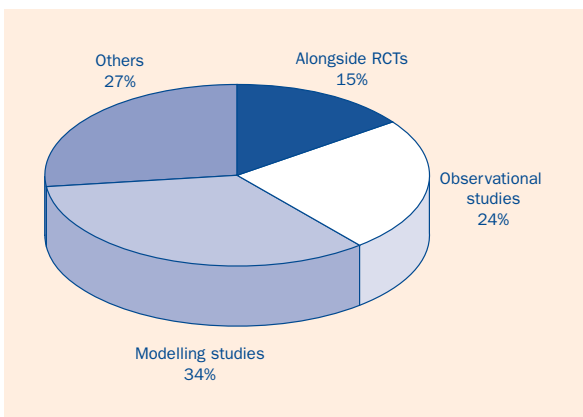
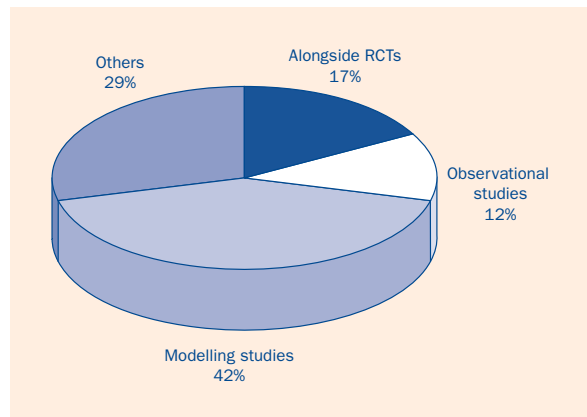


Figure 10 Design of applied CMA, CEA, CUA, CBA – drug evaluations 1997–2001 (n=1053)



trials reported in the literature. However, across all technologies, these studies represented less than 1% of all the four types of evaluation we considered in either time period and were therefore not presented separately. If modelling was recorded for both quantities of resource use and outcomes, then these studies would have been captured by the modelling category, but if only one was subject to modelling (normally outcomes), then it would have fallen into the 'other' category.

The increased use of modelling appears to have taken place at the expense of the 'other' category, which includes studies drawing data from more than one source (or from the literature) but without synthesising data on costs and outcomes in a systematic way through modelling. Therefore, we can conclude that when analysts attempt to estimate the costs and effects of health care interventions by combining data from different sources, they are more likely to do so through formal modelling techniques now than in the past. This emphasises the importance of establishing principles of good practice for these types of studies, an issue addressed by a number of practitioners, and one to which we will return in our review of a sample of recent studies which appear on HEED.

Narrowing the sample down further to cost utility analyses (CUAs, of which there were 147 in 1992–1996 and 421 in 1997–2001), the quality adjusted life year (QALY) was the overwhelming choice of outcome measure for cost utility studies in HEED (over 90% in both time periods); only one study was recorded as using the Healthy Years Equivalent (HYE) in each of the two time periods considered. 16 studies categorised as cost utility analyses used the disability adjusted life year (DALY) in the period 1997–2001 and one between 1992 and 1996. No studies were recorded as using the saved young life equivalent (SAVE).

It was unusual to find cost utility analyses declaring that an established utility measure had been used to weight life years. Only 16% of studies for the period 1992–1996, and 15% of studies in the period 1997–2001 were recorded as doing so. The measures considered applicable were the five multiattribute utility scales discussed by Sculpher and O'Brien (2000), plus the more recently developed SF-6D (Brazier et al., 1998). The latter was used in a single 2001 study. Of the other five, EQ-5D/EuroQol was the most popular, being used in 50% of studies for which one of the six

measures was recorded in the later time period (38% in the earlier period). The Health Utilities Index (HUI) was used in 26% (29%) of studies and the Quality of Well Being (QWB) scale in 24% (17%). Use of the Rosser index declined from 33% to 5% of studies between the two time periods.

In some cases where the utility scale was specified, values were simply drawn from the existing literature. There were also many studies in which the investigators had recourse to values from the literature, but the underlying utility scale was not specified; overall 45% of cost utility analyses in the period 1997–2001 are recorded as drawing health state values from previously published sources. This is in broad agreement with the results of Gerard et al. (2000) who found that, of a sample of 43 cost utility analyses published in 1996, over half had failed to carry out or had only partially carried out reporting of how QALYs were obtained and a similar percentage had inadequately reported the source of QALY valuations.

The earlier briefing looked at discounting practices for the period 1992–1996 for cost minimisation, cost effectiveness, cost utility and cost benefit analyses (discounting of both costs and benefits being potentially relevant for these studies). Over this period as a whole, 21% of studies discounted costs and 17% of studies applied a positive discount rate to benefits. Table 2 provides corresponding figures for the years 1997–2001. Over this later period, the proportion of studies using discounting has increased for both costs and benefits, but with benefits lagging behind. At the same time, the proportion of these types of studies using life years, lives saved or QALYs gained (outcome measures which might be considered amenable to discounting) has increased from 27% to 33% of studies.

Although, without undertaking a more in-depth review of studies, we cannot identify the time horizon employed, studies in which costs were discounted but benefits were left undiscounted suggest themselves as being worthy of further investigation. A brief inspection of HEED entries for cost utility analyses between 1997 and 2001 with discounting applied to costs but not benefits revealed no clear reason why authors had adopted this approach but authors may have been influenced in some cases by continuing methodological debates. The controversy surrounding the appropriateness of discounting benefits is discussed by Lipscomb et al. (1996).

Table 2 Use of discounting

Discounting – Applied CMA, CEA, CUA and CBA (% using)					
	1997	1998	1999	2000	2001
Costs Discounted	27%	25%	32%	36%	36%
Benefits Discounted	22%	20%	25%	27%	25%

For those studies which did use discounting, it is worth noting that the vast majority of studies in the period 1992–1996 applied a discount rate of 5% (74% for costs and 80% for benefits), whereas fewer than half of studies in the period 1997–2001 did so (43% and 42%). In comparison, 51% of studies which discounted costs used a rate of 3% in the later period as did 57% of those which discounted benefits, compared with 11% and 12% of the corresponding studies in the earlier period. This may well be a reflection of methodological developments, specifically the publication of the Panel on Cost-Effectiveness in Health and Medicine report which recommended a common discount rate of 3% for costs and benefits (Lipscomb et al., 1996). We should bear in mind, however, that health technology assessment bodies vary in their preferred discount rates. For example, NICE has recommended rates of 6% for costs and 1.5% for benefits (revised to 3.5% for both in the 2003 review of appraisal methods). The Australian Pharmaceutical Benefits Advisory Committee, on the other hand, prefers a common rate of 5% for costs and benefits to be adopted in submissions.

Debate over the conduct of sensitivity analysis centres not on whether it should be conducted but on how well it is performed. Taking the same types of evaluation as discussed in relation to discounting, 43% of studies used sensitivity analysis in the period 1992–1996 and 41% reported the results in quantitative terms, rising to 56% and 55% for the period 1997–2001. These figures were approximately double the percentages of all applied studies making allowance for uncertainty in this way. However, even the improved figures for the later time period show poorer results than

those found by Briggs and Sculpher (1995) for the use of sensitivity analysis by a sample of studies published in 1992 (36% used no sensitivity analysis). These findings perhaps suggest that the type of study selected is important in the extent to which sensitivity analysis is used. Indeed, Briggs and Sculpher (1995) themselves report a range of results from different studies.

Simply recording whether or not this type of analysis was used and reported does not, however, indicate whether sufficient allowance was made for uncertainty. For example, a one-way sensitivity analysis will not establish the degree of variability of the results to changes in several variables simultaneously. Even a multi-way sensitivity analysis is unlikely to elucidate fully the responsiveness of the results to changes in cost and outcome variables. On the other hand, analysis of trial-based studies may exploit the properties of the data to undertake a statistical analysis of uncertainty. This will enable the analyst fully to explore the extent of uncertainty associated with the parameters of the study while sensitivity analysis remains a useful way of assessing the impact of, for example, altering the discount rate or using different analytic methods. In modelling studies, all the uncertainty around parameter estimates in the model can be evaluated by attaching probability distributions to the model's inputs and employing a process which samples from these distributions. The use of this approach, usually termed probabilistic sensitivity analysis, is NICE's recommended method of allowing for uncertainty in economic models submitted by manufacturers and sponsors (NICE, 2003).

5 SPONSORSHIP OF STUDIES

Figures 11 and 12 show the contribution of different types of sponsoring organisation for, respectively, all applied studies and pharmaceutical evaluations only. The figures are percentages of studies with a declared sponsor, with more than one type of sponsoring organisation being possible for each study. The two time periods show a similar picture whether all technologies or pharmaceuticals only are considered. In the former case, government and publicly funded policy making bodies are the most important source of funds, being involved in around 40% of studies,

followed by the pharmaceutical industry, at around one third, research councils/universities (20%–25%) and charities (around 15%). In the case of pharmaceuticals, as we would expect, the pharmaceutical industry is the most important sponsor at 57% of studies in 1997–2001, with government and publicly funded policy making bodies providing support to only half as many studies. Again, research councils/universities and charities are the only other funding organisations providing support to 10% or more of studies.

Figure 11 Sponsors of applied studies – all technologies

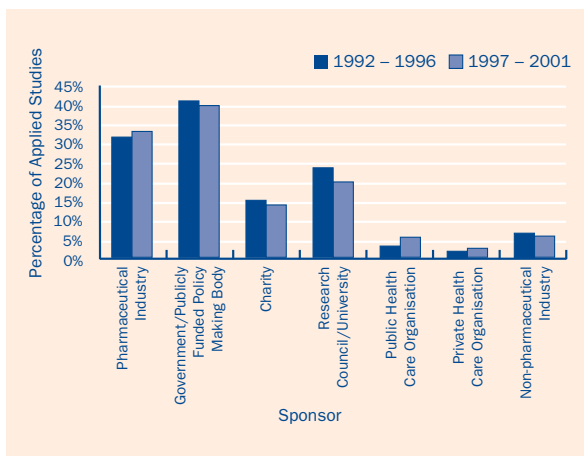
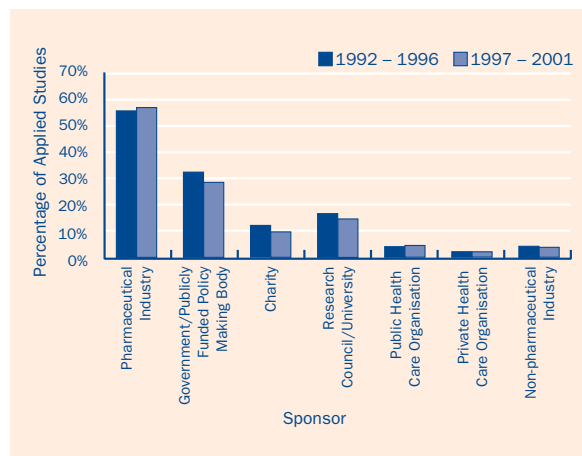


Figure 12 Sponsors of applied studies – pharmaceuticals



6 REVIEW OF A SAMPLE OF STUDIES

We selected a sample of studies from HEED for more detailed examination. Our selection criteria were to some extent arbitrarily determined to yield a sample of manageable size. However, we took account of the following factors. Firstly, we felt it appropriate to focus on cost utility studies since they are potentially of greatest interest to decision making bodies and are actively promoted by NICE as the preferred method of compiling submissions. Secondly, we restricted our search to modelling studies partly because of their increased share of studies and partly because of their importance in submissions to health technology assessment bodies. Not only does modelling form the basis of the methods proposed by NICE but it is widely accepted as a useful tool of analysis in submissions made elsewhere, for example in Ontario, Canada and the Netherlands (to be fully operational from

2005). Thirdly, we wanted to select studies which might illustrate recent methodological developments. Focusing on more than one year would have provided too unwieldy a sample; we reduced the numbers to a manageable level by selecting, from the year 2002, studies which had evaluated a pharmaceutical technology.

Thirty-three studies were initially identified, of which two were excluded, one because the authors set out to estimate QALYs but did not present the results in this form (despite utility weights being reported), the other because the authors estimated morbidity avoided days rather than survival adjusted for health quality. Of the remaining 31 studies, which are listed in the HEED sample references at the end of this briefing, two were published by technology assessment

organizations and the remainder were published in peer reviewed journals. In seven cases, publication was in economics or technology assessment journals while the other 22 studies were published in medical journals. Contrary to expectations from the figures presented in section 5, the pharmaceutical industry was a minority sponsor of these studies. A company or companies provided exclusively industry sponsorship for four studies and joint sponsorship of one further study, with the Janssen Research Foundation sponsoring another two studies. In comparison, public sector and charitable organisations provided sponsorship of 10 studies in our study sample (one jointly with industry).

The boxes accompanying this section highlight some of the overall features of the sample while here we attempt to make some overall quality assessments, without undertaking a formal rating exercise for individual studies. Given the likely increasing importance of modelling in the conduct of economic evaluations, it is worth noting the interest in establishing measures of the quality of cost-effectiveness models by, for example, Sculpher et al. (2000) and McCabe and Dixon (2000) and of the reporting of modelling studies by Nuijten et al. (1998).

These authors go somewhat beyond the established tools for assessing quality, such as the BMJ checklist (Drummond and Jefferson, 1996). This is a checklist measure which has been used (with some refinements) by, for example, Gerard et al. (2000), in assessing the quality of cost utility studies and has been referred to by Sculpher et al. (2000) as an instrument which might be complemented (rather than replaced) by a modelling-specific tool. Indeed, a number of the weaknesses in the sample of studies considered here would be picked up in items of the BMJ checklist, for example, justification of the model used and the parameters on which it is based, or justification of the choice of variables for sensitivity analysis.

Nevertheless, there is a case for a separate quality assessment of modelling studies since the results are crucially dependent on the structure of and inputs into the model. Sculpher et al. (2000) firstly make a plea for transparency of methods; "assessing the quality of a cost-effectiveness model requires the analyst to have adequately described their methods and to be able to provide clear, honest and transparent justification for the numerous components of their approach". As the discussion in Box 4 indicates, a majority of

the studies in our sample would fall foul of this requirement since they did not go so far as to explain the structure of the model used. Neither was it clear in most cases how the literature had been searched nor how studies had been selected to provide model inputs. In this regard, Sculpher et al. (2000) recommend that, as a minimum, analysts should search those data sources which require a relatively low investment of researcher time.

Where it is necessary to use expert opinion, Sculpher et al. (2000) suggest elements of good practice, including a statement of how experts were selected and what they were asked. While a number of studies referred to the use of expert opinion, none provided such details. Nuijten (1998) emphasises the importance of study selection criteria, arguing that the reader should be able to evaluate possible sources of bias in the data used. He reinforces the point made by Buxton et al. (1997) that the use even of a meta-analysis as a source of data carries the risk of bias by the inclusion or exclusion of particular studies.

With regard to model inputs, Nuijten (1998) recommends full disclosure of sources, including a description of the search strategy and of the data sources used and model design. This would extend to details of study design, such as the period of follow-up for patients and drop-outs and an assessment of the strengths and weaknesses of the selected data sources. McCabe and Dixon (2000) feel that these recommendations provide useful guidance but caution against a full-blown evidence-based medicine approach. However, it is clear that the studies reviewed here fall some way short of this level of transparency. Their usefulness would have been much improved by observing some basic tenets of good practice such as explaining the model structure, how data inputs were identified and what sources were used for this purpose. Authors should also describe how transition probabilities were derived from the source trials and other studies, and report key probabilities of clinical events and associated costs.

Few researchers revealed how the literature sources which provided the inputs to their model were identified. Only two stated the databases which had been searched and one reported that a "systematic search of electronic reference databases and abstract listings" had been performed. None gave an indication of the search terms which had been used or inclusion and exclusion criteria. Neither did authors describe the methods by which the data retrieved had been combined to generate probability inputs for the model, although

Box 4

Sources of data

Studies routinely used a range of sources in the literature to provide the relevant inputs which enabled the investigators to estimate quality adjusted survival. These could include clinical trials, observational studies and routinely collected data on general population mortality rates. Only two studies modelled treatment effects on the basis of a single clinical trial. Studies frequently provided a list of input data and their associated sources in tabular form (21 studies, plus two which tabulated resource inputs only); however, in only 14 studies did authors include a diagrammatic representation or explanation of the structure of their model.

Utility weights were drawn from the literature in the overwhelming majority of cases. It appeared that in only three studies were utility weights specifically generated for the analysis (not being based on previous studies or authors' judgement). One used a panel of experts, one conducted a time trade-off among a group of five patients, while another set of researchers drew their utilities from a parallel study of health related quality of life in lung transplantation using EuroQol. In most cases,

studies which extracted utilities from the literature did not specify whether they were based on a particular multiattribute utility instrument. Of those which did, two referred to EQ-5D and two to the Health Utilities Index.

Most studies restricted themselves to a consideration of direct health care costs, although a third (10 studies) also included productivity costs (the value of lost work time). For both cost categories, the literature, as in the case of clinical and quality of life data, was heavily utilised. The costing aspect of the analysis seemed generally to be the area where researchers used the flexibility afforded by the model to adapt their estimates to the 'real world' decision making context with which they were presented. Thus, cohort studies, cross-sectional studies and patient records were called upon to provide estimates of resource use observed in clinical practice, rather than in the confines of a clinical trial (although some studies referred to trial data as a source of resource use estimates). Researchers also sought out expert opinion, a source which is inevitably used in modelling studies to a greater or lesser extent, if only to construct the clinical pathways represented in the model.

one group of investigators stated that their model's parameters had mostly been derived from the weighted-averages of relevant studies.

In general, authors did not state why a model had been used or why a particular type of model as opposed to any other had been employed. An exception was Karnon and Brown (2002) who explicitly based their preference for a discrete event simulation (DES) model on its ability to provide a "more flexible and accurate representation of the available data" than a Markov model. Potentially, there are a number of influences on the choice of model (Siebert, 2003) and in our sample of studies usually a number of factors seemed relevant. But in one model based on a single clinical trial the sole purpose of the model seemed to be to allow data collected in the trial to be combined with other data not collected in the trial.

In addition to the validity of the inputs, a number of commentators refer to the concept of validity in relation to the outputs of a model, generally interpreted as the model's performance against real world data from primary studies outside the

model. Few studies in our sample considered this aspect of the analysis, one simply stating that the model had undergone extensive verification. Of the two studies which explained their method, Kuehne et al. (2002) made an assessment of their model's predictive validity by comparing its predictions of cirrhosis and mortality risk for an HIV-infected cohort of patients with mild chronic hepatitis C infection against data from published studies not used to provide inputs for the model. The most detailed discussion of validation was provided by Karnon and Brown (2002) who reviewed a number of literature sources for comparison with their model. They used a pooled estimate of survival from two studies which did not provide data for the model as the basis of their validation estimates of costs and effects. They also conducted a number of verification tests on the internal consistency of the model.

An aspect of study validity often discussed in the context of economic evaluations based on randomised trials is that of external validity, or generalisability from one context to another. In the case of a modelling study, the objective is

Box 5 Modelling

The most common modelling approach, used by a majority of studies in this sample (18 studies), was a Markov model. One study used a discrete event simulation (DES), while the others simply stated that they had used a decision analytic, decision tree or simulation model or provided no description of the modelling approach used. Five studies made use of existing models; in four cases the model was developed for a particular disease or treatment while the fifth used a more generic disease-focused model. This could have been prompted by a desire to transfer results from one setting to another or a belief that there was little scope for improving on a model already developed. The desire to extrapolate from or adapt existing data for the real world setting could also be discerned in this group of studies. Some authors mentioned this factor explicitly, while the common use of observational studies and patient registries for resource use data suggest that this was frequently a consideration in the choice of study design.

Extrapolation over time appeared to be the most important reason for using a model, with only four studies adopting a time horizon of a year or less. One study, despite adopting a short time horizon of six months, nevertheless used modelling as a means of projecting

beyond the results given by trial data. Inclusion of data on clinical practice patterns and expert opinion allowed the analysis to be extended beyond the eight week time frame of the clinical trials. In contrast, one study was based on a single randomised trial but did not extrapolate beyond the time horizon of the trial. Of those studies adopting a time horizon of more than a year, the shortest period examined was two years; 14 studies took a whole of life time horizon, or an approximation to it, for example, Lee et al. (2002) ran their model until only 5% of their cohort remained alive after hospital discharge.

The objective of evaluating the cost-effectiveness of a range of different treatment options where direct comparisons are unavailable was a factor in 11 studies. This could be for the purposes of considering as few as two alternatives, as in the study by Grima et al. (2002), on the basis of, as they comment, "the lack of head-to-head trials" in the patient group being investigated. Secondly, there were studies such as that by Malin et al. (2002), who compared more treatment strategies than it seemed likely would be directly evaluated in clinical trials. Thirdly, there were studies which evaluated what Siebert (2003) refers to as "fine-tune technologies". For example, Buti et al. (2002) considered a variety of strategies in the management of hepatitis C, involving different approaches to duration of treatment and diagnostic testing.

often to obtain results applicable to a specific context, perhaps using the results from multi-national trials as a component of the input data. Authors of modelling studies are more or less explicit about the particular decision-making context to which the results are intended to apply and make more or less concerted efforts to generate results relevant to a given setting. However, most analysts, while trawling the international literature for data on probabilities of clinical events to enter into the model, typically identify sources of resource utilisation and unit cost data relating to their own local or national setting. The analyst can also generate estimates more relevant to the usual clinical practice or 'real world' setting than those produced in a clinical trial.

In addition, as Siebert (2003) notes, a modelling approach can be used to transfer evidence from one location to another by adapting a single model to a variety of contexts. Adoption of an existing model was a method used in several studies in our sample (Box 5). However, there was little discussion in the sample of the decision-making context to which the analysis was intended to be applicable, or of which data sources might give the study more or less generalisability. For example, as Nuijten et al. (1998) point out, some databases providing inputs to the model may be specific to a particular type of provider. Similar comments apply to the choice of alternative treatments, since a model may focus on the normally recommended reference point of usual clinical practice which may vary from setting to setting. On the other hand, the ability of a model to incorporate a wide range of comparators may

mean that it yields results of interest to decision-makers in a variety of settings. In general, the reason for the authors' particular approach to these factors was unclear.

Box 6 discusses the ways in which studies made allowance for uncertainty. As Sculpher et al. (2000) point out, "part of the rationale for the use of models is that they enable decisions to be reached through the systematic handling of uncertainty". Few studies, however, exploited this potential in their models to the full. Twenty-one studies restricted themselves to one-way sensitivity analysis, with little discussion of the choice of parameters or their ranges to be varied in the sensitivity analysis. Of the other ten, half employed Monte Carlo simulation as a means of recognising the stochastic nature of parameter inputs; in one case this took the form of a first order simulation. As Briggs (2000) points out, this approach does not allow the parameter values to vary and provides a range of estimates of costs and effects analogous to those which might be derived from individual patients in a trial. He notes that this is of limited relevance for decision makers concerned with uncertainty around mean costs and effects. The appropriate method for probabilistic sensitivity analysis is to explore second order uncertainty (Briggs, 2000) by allocating distributions to parameter inputs. It is worth reporting, in each study in which this was done, the preference for the cost-effectiveness acceptability curve as a means of presenting the results, perhaps an indication of the tractability or the ease of interpretation of this approach compared with, for example, confidence intervals.

There are therefore a number of aspects on which the studies considered here exhibit weaknesses relative to criteria used to assess economic evaluations generally and elements of good practice proposed for modelling studies. An alternative way of attempting to assess quality, and the starting point taken by Nuijten (1998), would be to refer to nationally produced guidelines or those developed to support particular decision making processes. Although Nuijten points out the shortcomings of many existing sets of guidelines, the use of modelling in the provision of data to reimbursement and other authorities is likely to become increasingly important. Particularly noticeable in its emphasis on modelling is the guidance developed by NICE in 2003 for submissions compiled by manufacturers and sponsors of technologies being appraised. The 'reference case' approach, while containing

Box 6 Allowance for uncertainty

While all studies made some allowance for uncertainty, only ten studies went beyond simple one-way sensitivity analysis and in three of these investigators restricted themselves to two-way or multivariate sensitivity analyses. One study was deemed to have exceeded a simple one-way sensitivity analysis by virtue of presenting three sets of cost-effectiveness results, each based on a different trial used to provide effectiveness estimates. One study undertook a first-order Monte Carlo simulation, in which each individual member of a hypothetical cohort of patients is followed through the model one at a time. Four studies analysed second-order uncertainty; in each case, results were presented in the form of cost-effectiveness acceptability curves. One other study calculated a confidence interval, with the distributions of incremental costs and QALYs being estimated using a bootstrap approach. No study estimated net benefits (a potentially more tractable statistic for estimating confidence intervals) or undertook value of information analysis to identify those variables contributing the greatest uncertainty to the cost-effectiveness estimates.

context-specific elements such as the appropriate discount rate, could be seen as the basis for a template for good modelling practice, incorporating as it does recent methodological developments such as probabilistic sensitivity analysis rarely seen in the studies considered here.

One technique proposed by the guidance and not used in any of our studies, namely value of information (VOI) analysis, can help to address issues relating to data quality by indicating whether further collection of data on particular variables would be worthwhile. For example, we have seen that utility weights were estimated in our sample in a relatively crude way but more sophisticated data collection may not be justified if this parameter is not a significant element in the overall uncertainty surrounding the results. By exploring the contribution of particular parameters to overall uncertainty, VOI will not help to inform the current treatment decision but can inform the decision as to whether and for what variables further research is warranted to inform future decisions.

7 CONCLUSIONS

Our analysis of the studies contained in HEED shows that some trends apparent in the five years 1992–1996 have continued in 1997–2001, for example the reduced share of pharmaceutical evaluations as a proportion of applied studies. Meanwhile, other comparisons show little change in the characteristics of studies, for example the distribution of studies by disease area and of pharmaceutical evaluations by ATC chapter.

Perhaps the most noteworthy finding of the current review relates to the design of applied studies, an issue not fully explored previously. The comparison between the two five-year time periods shows that randomised trials are maintaining (but not increasing) their importance as a basis for conducting economic evaluations, while modelling studies are increasing in importance as a proportion of all studies. This seems to be at the expense of studies which draw on a number of different data sources but combine them only in an informal way rather than using formal modelling.

Restricting the sample to cost minimisation, cost effectiveness, cost utility and cost benefit analyses (which should satisfy the criteria for full economic evaluations) highlights this trend. The proportion of studies on HEED which derived estimates of cost-effectiveness using more than one data source (or data drawn from the literature), rather than using a single primary study, was around 60% in both the 1992–1996 and 1997–2001 time periods. However, in the later period, modelling was used in a majority of these studies whereas it was used in only a minority of studies in the earlier period. Therefore, the need of analysts to draw on a variety of data sources is increasingly being met by decision analysis techniques which combine data in a structured way. Given that these techniques are well established and accepted as methods of estimating cost-effectiveness, their use is likely to increase further in importance in future. This is particularly so given that decision makers are increasingly asking for information on long term outcomes for health care technologies, especially drugs, early in their life cycle and in many cases at the time of launch.

Decision analytic models offer a number of advantages over clinical trials, such as the ability to provide projections of disease progression over long periods of time. In many of the modelling studies we reviewed, even if it was not the prime motivation for using a model, authors took advantage of this facility, frequently extending the

analysis to the whole of life. Researchers also exploited the capabilities of decision analysis models generally and Markov models in particular (used by over half of the studies considered) to generate estimates of costs and QALYs for a range of possible treatment options or which would be applicable to a local decision-making context.

What studies did not generally do was to exploit the ability of models to explore and quantify the uncertainty around the estimates of costs and QALYs. Most studies allowed for uncertainty simply by undertaking sensitivity analysis (usually one variable at a time). Four studies attached probability distributions to the parameters of the model and performed probabilistic sensitivity analysis to elucidate all the uncertainty contained in the model.

Other weaknesses were that authors did not generally explain how the literature providing the inputs of the model had been identified, nor how the model inputs were derived; fewer than half explained or illustrated how their model was constructed.

Thus, although we did not formally assess the quality of studies, the growth in the numbers of modelling studies did not appear to be accompanied by a wide recognition of some of the methodological and transparency issues discussed in the literature. While some of the shortcomings observed in this sample of studies would be identified by general assessments of the quality of economic studies, e.g. by using the BMJ checklist, there is a strong case for additional guidance and quality checks specifically targeted at modelling studies. The salient features of these types of studies would differ from, for example, those relevant to economic evaluations alongside clinical trials, where statistical issues around cost distributions and methods of dealing with missing and censored data would be important. For the particular group of studies considered here, some of the main weaknesses were that authors frequently did not:

- report their search strategy for the data on which the model was to be based;
- describe the search strategy to identify relevant data;
- justify their choice of input parameters;
- explain or justify the model structure;
- explain how transition probabilities were derived from the underlying clinical trial and other data;

- report key probabilities or costs for key clinical events;
- justify the parameters and data ranges used in sensitivity analysis.

This is by no means an exhaustive list but simply highlights some of the issues which, if they had been addressed in our sample, would have rendered studies more useful for decision makers and given a clearer indication of the appropriateness of the methods used.

While transparency is to be encouraged, there is no one set of methods on which there is general agreement. Some areas of methodological controversy raised by the studies we looked at are:

- the use of probabilistic sensitivity analysis. This was one of the most hotly debated topics during NICE's 2003 review of methods and may become more widely used given that NICE has declared a preference for this type of analysis. Further discussion of this issue in the methods literature can be expected;
- the appropriateness of different models for different purposes. In our sample of studies, Markov models were favoured by most investigators but the pros and cons of these models compared with, for example, discrete event simulation, have not been fully explored;
- validation of models. Methods both for internal and external validation are another aspect of the modelling process not generally addressed and for which guidance would be useful;
- the impact of different model specifications. Most of the studies reviewed here used a model specifically developed by the authors, with only a few adopting an existing model, but they may well be capable of comparison with models developed to address similar questions. Comparisons with existing models could help to identify areas of disagreement

between analysts but equally could generate some consensus around the appropriateness of methods in different therapeutic areas;

- the importance of different data sources to the uncertainty surrounding the results of models. For example, whether greater effort is justified to obtain information on data sources, such as utility weights, which tend to be estimated in a relatively crude fashion. Value of information analysis can help here but this tends to go hand in hand with probabilistic sensitivity analysis, which is itself controversial.

While guidance on good methodology in modelling studies is a matter of some importance, the impact of the proposed methodological standards on the quality of studies is unclear. Nevertheless, despite areas of controversy with regard to methods, the development of a limited set of key criteria may be possible. In this respect, there is a role for methodological standards to be maintained by policy makers given that much of the cost-effectiveness evidence they will appraise is likely to be generated through the use of modelling techniques.

The guidance produced by NICE for those making submissions to its technology appraisals programme can be seen as an encapsulation of current good practice, with the flexibility to absorb developments in methods. Although some aspects of the guidance are context-specific, it might form the basis for a minimum set of standards or prioritised list of requirements which may be more readily applied than guidance which attempts to specify every detail a potential user would like. It remains to be seen whether this will result in cost utility analyses produced to higher standards than the studies considered here and a more widespread adoption of methodological advances emerging in the literature in future years.

REFERENCES

Backhouse M E, Backhouse R J, Edey S A (1992). *Economic evaluation bibliography*. Health Economics 1(suppl.):1-236.

Buxton M J, Drummond M F, van Hout B A, Prince R L, Sheldon T A, Szucs T, Vray M (1997). *Modelling in economic evaluation: an unavoidable fact of life*. Health Economics 6:217-227.

Brazier J, Usherwood T, Harper R, Thomas K (1998). *Deriving a preference-based single index from the UK SF-36 Health Survey*. Journal of Clinical Epidemiology 51(11):1115-1128.

Briggs A H (2000). *Handling uncertainty in cost-effectiveness models*. Pharmacoeconomics 17(5):479-500.

- Briggs A H, O'Brien B J (2001). *The death of cost-minimization analysis?* Health Economics 10:179-184.
- Briggs A, Sculpher M (1995). *Sensitivity analysis in economic evaluation: a review of published studies.* Health Economics 4:355-371.
- Drummond M F, Jefferson T O (1996). *Guidelines for authors and peer reviewers of economic submissions to the BMJ.* The BMJ Economic Evaluation Working Party. BMJ 313(7052):275-83.
- Drummond M F, O'Brien B, Stoddart G L, Torrance G W (1997). *Methods for the economic evaluation of health care programmes.* New York: Oxford University Press, 2nd edition.
- Elixhauser A, Luce B R, Taylor W R, Reblando J (1992). *Health care CBA/CEA: an update on the growth and composition of the literature.* Medical Care 31(7):JS1-JS11.
- Gerard K, Seymour J, Smoker I (2000). *A tool to improve quality of reporting published economic analyses.* International Journal of Technology Assessment in Health Care 16(1):100-110.
- Jonsson E, Banta D H, Henshall C, Sampietro-Colom L (2002). *Summary report of the ECHTA/ECAHI project.* International Journal of Technology Assessment in Health Care 18(2): 218-237.
- Lipscomb J, Weinstein M C, Torrance G W (1996). Time preference. In Gold M E, Siegel J E, Russell L B, Weinstein M C (eds). *Cost-effectiveness in health and medicine.* New York: Oxford University Press.
- McCabe C, Dixon S (2000). *Testing the validity of cost-effectiveness models.* Pharmacoeconomics 17(5):501-513.
- NICE (2003). *Guide to the methods of technology appraisal.* London: National Institute of Clinical Excellence.
- Nixon J, Stoykova B, Glanville J, Christie J, Drummond M, Kleijnen J (2000). *The UK NHS Economic Evaluation Database.* International Journal of Technology Assessment in Health Care 16(3):731-742.
- Nuijten M J C (1998). *The selection of data sources for use in modelling studies.* Pharmacoeconomics 13(3):305-316.
- Nuijten M J C, Pronk M H, Brorens M J A, Hekster Y A, Lockfeer J H M, de Smet P A G M, Bonsel G, van der Kuy A (1998). *Reporting format or economic evaluation. Part II: focus on modelling studies.* Pharmacoeconomics 14(3):259-268.
- Pritchard C (1998). *Trends in economic evaluation.* London: Office of Health Economics.
- Sculpher M, Fenwick E, Claxton K (2000). *Assessing quality in decision analytic cost-effectiveness models: a suggested framework and example of application.* Pharmacoeconomics 17(5):461-477.
- Sculpher M J, O'Brien B J (2000). *Income effects of reduced health and health effects of reduced income: implications for health-state valuation.* Medical Decision Making 20(2):207-15.
- Siebert U (2003). *When should decision-analytic modeling be used in the economic evaluation of health care?* European Journal of Health Economics 4:143-150.

APPENDIX 1: HEED SAMPLE REFERENCES

- Anyanwu A C, McGuire A, Rogers C A, Murday A J (2002). *An economic evaluation of lung transplantation.* The Journal of Thoracic and Cardiovascular Surgery 123(3):411-420.
- Barry M, Heerey A (2002). *Cost effectiveness of statins for the secondary prevention of coronary heart disease in Ireland.* Irish Medical Journal 95(5):133-135.
- Bos J M, Rümke H C, Welte R, Postma M J, Jager J C (2002). *Health economics of a hexavalent meningococcal outer-membrane vesicle vaccine in children: potential impact of introduction in the Dutch vaccination program.* Vaccine 20:202-207.
- Buti M, Casado M A, Fosbrook L, Esteban R (2002). *Which is the most cost-effective combination therapy strategy using interferon α -2b plus ribavirin for naïve patients with chronic hepatitis C?* Clinical Drug Investigation 22(1):31-39.

- Caro J J, Salas M, Ward A, Getsios D, Mehnert A (2002). *Economic analysis of galantamine, a cholinesterase inhibitor, in the treatment of patients with mild to moderate Alzheimer's disease in the Netherlands*. *Dementia and Geriatric Cognitive Disorders* 14:84-89.
- Chodick G, Lerman Y, Wood F, Aloni H, Peled T, Ashkenazi S (2002). *Cost-utility analysis of hepatitis A prevention among health-care workers in Israel*. *Journal of Occupational and Environmental Medicine* 44(2):109-115.
- Crowley S, Tognarini D, Desmond P, Lees M, Saal G (2002). *Introduction of lamivudine for the treatment of chronic hepatitis B: expected clinical and economic outcomes based on 4-year clinical trial data*. *Journal of Gastroenterology and Hepatology* 17:153-164.
- François C, Sintonen H, Toumi M (2002). *Introduction of escitalopram, a new SSRI in Finland: comparison of cost-effectiveness between the other SSRIs and SNRI for the treatment of depression and estimation of the budgetary impact*. *Journal of Medical Economics* 5:91-107.
- Gaspoz J M, Coxson P G, Goldman P A, Williams L W, Kuntz K M, Hunink M G M, Goldman L (2002). *Cost effectiveness of aspirin, clopidogrel, or both for secondary prevention of coronary heart disease*. *New England Journal of Medicine* 346(23):1800-1806.
- Glick H A, Orzol S M, Tooley J F, Remme W J, Sasayama S, Pitt B (2002). *Economic evaluation of the randomized aldactone evaluation study (RALES): treatment of patients with severe heart failure*. *Cardiovascular Drugs and Therapy* 16:53-59.
- Goldie S J, Kaplan J E, Losina E, Weinstein M C, Paltiel A D, Seage III G R, Craven D E, Kimmel A D, Zhang H, Cohen C J, Freedberg K A (2002). *Prophylaxis for human immunodeficiency virus-related pneumocystis carinii pneumonia. Using simulation modeling to inform clinical guidelines*. *Archives of Internal Medicine* 162:921-928.
- Grima D T, Burge R T, Becker D L, Tosteson A N A (2002). *Short-term cost-effectiveness of bisphosphonate therapies for postmenopausal osteoporotic women at high risk of fracture*. *P and T* 27(9):448-455.
- Hershman D, Sundararajan V, Jacobson J S, Heitjan D F, Neugut A I, Grann V R (2002). *Outcomes of tamoxifen chemoprevention for breast cancer in very high-risk women: a cost-effectiveness analysis*. *Journal of Clinical Oncology* 20(1):9-16.
- Ikedo S, Yamada Y, Ikegami N (2002). *Economic evaluation of donepezil treatment for Alzheimer's disease in Japan*. *Dementia and Geriatric Cognitive Disorders* 13:33-39.
- Karnon J, Brown J (2002). *Tamoxifen plus chemotherapy versus tamoxifen alone as adjuvant therapies for node-positive postmenopausal women with early breast cancer: a stochastic economic evaluation*. *Pharmacoeconomics* 20(2):119-137.
- Khan K, Muennig P, Behta M, Zivin J G (2002). *Global drug-resistance patterns and the management of latent tuberculosis infection in immigrants to the United States*. *New England Journal of Medicine* 347(23):1850-1859.
- Kobelt G, Jönsson L, Miltenburger C, Jönsson B (2002). *Cost-utility analysis of interferon beta-1b in secondary progressive multiple sclerosis using natural history disease data*. *International Journal of Technology Assessment in Health Care* 18(1):127-138.
- Kobelt G, Lindgren P, Young A (2002). *Modelling the costs and effects of leflunomide in rheumatoid arthritis*. *European Journal of Health Economics* 3:180-187.
- Kuehne F C, Bethe U, Freedberg K, Goldie S J (2002). *Treatment for hepatitis C virus in human immunodeficiency virus-infected patients: clinical benefits and cost-effectiveness*. *Archives of Internal Medicine* 162:2545-2556.
- Lee H, Manns B, Doig C J, Johnson D, Donaldson C (2002). *A cost-effectiveness analysis of recombinant human activated protein C for severe sepsis in the intensive care unit*. *Working Paper 02 – 06*. Edmonton: Institute of Health Economics.
- Lehmann K, Radbruch L, Gockel H-H, Neighbors D, Nuyts G (2002). *Costs of opioid therapy for chronic nonmalignant pain in Germany: an economic model comparing transdermal fentanyl (durogesic) with controlled-release morphine*. *European Journal of Health Economics* 3:111-119.
- Malin J L, Keeler E, Wang C, Brook R (2002). *Using cost-effectiveness analysis to define a breast cancer benefits package for the uninsured*. *Breast Cancer Research and Treatment* 74:143-53.

Marshall J K, Blackhouse G, Goeree R, Brazier N, Irvine E J, Faulkner L, Dipchand C, O'Brien B J (2002). *Clinical and economic assessment: infliximab for the treatment of Crohn's disease. Technology Overview 8*. Ottawa: Canadian Coordinating Office for Health Technology Assessment.

O'Connor J B, Singer M E, Imperiale T F, Vaezi M F, Richter J E (2002). *The cost-effectiveness of treatment strategies for achalasia*. *Digestive Diseases and Sciences* 47(7):1516-1525.

Palmer C S, Nuijten M J C, Schmier J K, Subedi P, Snyder E H (2002). *Cost effectiveness of treatment of Parkinson's disease with entacapone in the United States*. *PharmacoEconomics* 20(9):617-628.

Pepper P V, Owens D K (2002). *Cost-effectiveness of the pneumococcal vaccine in healthy younger adults*. *Medical Decision Making* 22(Suppl): S45-S57.

Smith K J, Roberts M S (2002). *Cost-effectiveness of newer treatment strategies for influenza*. *American Journal of Medicine* 113:300-307.

Stein K, Rosenberg W, Wong J (2002). *Cost effectiveness of combination therapy for hepatitis C: a decision analytic model*. *Gut* 50:253-258.

Stevenson M, Beard S, Finn A, Brennan A (2002). *Estimating the potential health gain and cost consequences of introducing a pre-school DTPa pertussis booster into the UK child vaccination schedule*. *Vaccine* 20:1778-1786.

Vidal-Trécan G M, Stahl J E, Durand-Zaleski I (2002). *Managing toxic thyroid adenoma: a cost-effectiveness analysis*. *European Journal of Endocrinology* 146:283-294.

Zhou F, Bisgard K M, Yusuf H R, Deuson R R, Bath S K, Murphy T V (2002). *Impact of universal Haemophilus influenzae type b vaccination starting at 2 months of age in the United States: an economic analysis*. *Pediatrics* 110(4):653-661.

About the Office of Health Economics

The Office of Health Economics was founded in 1962. Its terms of reference are to:

- commission and undertake research on the economics of health and health care;
- collect and analyse health and health care data from the UK and other countries;
- disseminate the results of this work and stimulate discussion of them and their policy implications.

The OHE is supported by an annual grant from the Association of the British Pharmaceutical Industry and by revenue from sales of its publications, consulting services and commissioned research.

The research and editorial independence of the OHE is ensured by its Policy Board. This OHE publication has been peer reviewed by members of its Editorial Board. Further information about the OHE, including details of these boards, can be found at www.ohe.org.

About HEED – the Health Economic Evaluations Database

HEED is a subscription service and makes use of these subscriptions to finance all database operations. HEED receives no additional financial support from any company or organisation.

HEED has been developed as a joint initiative between the Office of Health Economics (OHE) and the International Federation of Pharmaceutical Manufacturers' Associations (IFPMA) and contains information on studies of cost-effectiveness and other forms of economic evaluation of medicines and other treatments and medical interventions.

A comprehensive description of HEED and the subscription rates can be found at www.ohe-heed.com.



Office of Health Economics
12 Whitehall London SW1A 2DY
Telephone: +44 (0)20 7930 9203
Facsimile: +44 (0)20 7747 1419
www.ohe.org

Price: £7.50

© Office of Health Economics