

THREE CHALLENGES TO ACHIEVING BETTER ANALYSIS FOR BETTER DECISIONS: GENERALISABILITY, COMPLEXITY AND THRESHOLDS. NOTES ON A CONFERENCE IN HONOUR OF BERNIE O'BRIEN.

Adrian Towse and Martin Buxton

Contents

Executive summary	1
The current use of economic evaluation in decision-making	3
Generalisability of results across geographical settings	5
The cost-effectiveness threshold and its relevance to local budgets	10
Conclusions	13
References	14

Obtaining a copy of the Conference Proceedings

The Conference Proceedings are published in the November 2006 issue of the journal *Pharmacoeconomics* (issue 24 (11)). The issue is available online at <http://pharmacoeconomics.adisonline.com> from the first week of November 2006. Single articles are available for purchase through the website. To order a print copy of the entire issue e-mail subs@nz.adis.com.

EXECUTIVE SUMMARY

In June 2006, a conference entitled *Better Analysis for Better Decisions: Bridging the Gap Between Economic Evaluation and Healthcare Decision-Making* was held at McMaster University in honour of the late Bernie O'Brien. The papers presented by leading health economists were reviews of the use of economic evaluation in the UK, Canada and USA, and more methodologically focussed contributions.

The reviews of the experience in the three countries suggest that economic analysis is playing an increasingly important role in health sector decision-making. But usage is patchy, rarely systematic and

¹ Many issues were covered in the conference and we would urge readers to look at the full proceedings. We focus on three issues which we thought to be recurring ones at the meeting, running through a number of papers and through the discussions.

explicit, reflecting both some irrational illogical resistance and some genuine concerns and problems with the current methods of analysis.

We identified three key issues¹ that recurred in the papers and in the discussion at the conference:

- generalisability: the extent to which cost-effectiveness analyses relevant and appropriate to one jurisdiction can be used in another;
- complexity: two related issues arise as economic evaluations become more complex:
 - credibility with decision makers and
 - the need to trade complexity against quantity to enable the analyses of more technologies within Health Technology Assessment (HTA) resource constraints;
- thresholds: the basis for, and validity of, thresholds values for the incremental cost-effectiveness ratio (e.g. cost per Quality Adjusted Life Year {QALY}) adopted by central decision-makers and their relevance at a local level.

The challenges posed by these issues result from success, i.e. the greater use of economic evaluation in decision making. Meeting them is fundamental to its continued growth in use.

Generalisability

Local decision makers want local information. For pharmaceutical companies a central problem is tension between a desire to have global development programmes and this desire of local decision-makers to have information about cost-effectiveness tailored to their own perceived needs and clinical settings. Sculpher and Drummond

(2006) focussed on the methodological element of this tension. They argued for a greater commonality of methodological requirements (and by implication of the data requirements to support them), recognising that some differences will remain because of unresolved methodological issues and “genuine” differences in what decision makers are trying to achieve and the institutional context in which they seek to accomplish it.

Manca and Willan (2006) focussed on the generalisability of trial-based economic evaluation. They challenged the notion that pooled clinical, let alone economic results from a multi-country trial can be taken to apply even to a country participating in the trial. They did not see it as appropriate for a decision maker to rely on a single country study for their country, or on the results from the centres within the trial that come from their country. Rather the challenge is to use all of the information available from centres wherever located but to test for “read across” relevance to countries inside and outside of the trial.

The papers and discussions suggested several actions to resolve the tension between the global and the local in an efficient way:

- seek greater clarity on what needs to vary methodologically across countries;
- reach agreement on the tests to be performed on and rules to be applied to the use of data and results from multi-centre trials within and outside of the countries for which estimates of cost-effectiveness are required;
- further develop and promulgate robust methods to combine different sources of evidence (evidence synthesis);
- apply value of information approaches to seek to resolve the issue of the extent to which local demands for additional work can be regarded as reasonable.

Complexity

Weinstein (2006) and Buxton (2006) both argued that there are tensions between investing the greater time and effort to ensure higher quality economic analyses which are of more *potential* value and:

- the risk of reducing their *actual* value because decision makers do not

² Following a negative appraisal by NICE of three drugs for the treatment of Multiple Sclerosis the UK Department of Health entered into risk sharing contracts with each of the manufacturers whereby the health outcomes achieved by patients would be measured and assessed against a £36,000 per QALY threshold with price adjustments if the threshold were exceeded. No formal appraisal of the Scheme has yet been published.

understand the methods and therefore do not trust the results. If so, Weinstein concluded, they are unlikely to be influential and change behaviour;

- the need for more analyses of a wider range of technologies within the resource constraints on appraisals (financial and/or the availability of time and of skilled researchers).

Several actions might reduce the potential dilemmas of trading off complexity against decision maker understanding and HTA resource constraints:

- development of a more contingent appraisal process where the level of analytical effort is more closely related to the nature of the decision-problem;
- use of Value of Information (VoI) methods to determine the relevant level of complexity in particular cases;
- use of VoI to determine which technologies to appraise;
- use of “MS type²” risk sharing schemes involving additional data collection with payment contingent on actual patient outcomes. This can allow adoption of technologies despite high levels of remaining uncertainty.

Thresholds

The estimation of incremental cost-effectiveness is only useful if decision-makers have a clear, and appropriate view of the maximum they are prepared to pay for health improvements. Buxton (2006) set out the evidence on NICE’s use of threshold values for a QALY. There is growing explicitness and evidence of reasonable consistency but is the threshold set at the right level? The key problem is local opportunity cost. Which health programmes or technologies will an efficient local decision-maker not invest in or have to give up in order to fund the technology recommended by the national HTA body? Is the national HTA body accurately reflecting this local opportunity cost in its threshold willingness to pay for health gain? Buxton suggested that NICE’s threshold may well have been too high.

Birch and Gafni (2006) argued that the threshold ICER is a fundamentally flawed concept, notably because:

- it ignores budget constraints;
- it assumes knowledge of the costs and benefits of technologies that might be displaced that we simply don’t have.

They argued for a programme budgeting approach. However, their information requirements seem to be similar in respect of technologies to be included in the programme budget as in ICER calculations. One problem they highlighted is however a very real one. Technologies are not perfectly divisible. Programme budgeting approaches can address this, although this information could be fed back into a revised ICER. The affordability /cost effectiveness distinction that is imposed on NICE is, they argued, not credible. Affordability is as much about opportunity cost as is cost-effectiveness. Their arguments make sense but are not inconsistent with the use of data to improve the estimation of the relevant ICER.

Pulling together the common strands of the Weinstein, Buxton and the Birch and Gafni contributions suggests the following actions make sense:

- more specific assessment of the opportunity cost in various localities to generate relevant ICERs for use in central decisions;
- more emphasis on evaluation of existing technologies that might be appropriately disinvested;
- greater use of programme budgeting approaches at a local level.

THE CURRENT USE OF ECONOMIC EVALUATION IN DECISION-MAKING

Review papers were presented on the use of economic evaluation in the UK (Buxton, 2006), Canada (Laupacis, 2006) and the USA (Neumann and Sullivan, 2006).

In the UK, the technology appraisal decisions of the National Institute for Health and Clinical Excellence (NICE) are dominated by cost-effectiveness considerations and the most controversial rejections by NICE have been on these grounds. NICE has an explicit £20,000-£30,000 per QALY cost-effectiveness threshold. However, there is controversy as to whether that threshold reasonably reflects the opportunity cost faced by local decision-makers within the NHS given their local fixed budgets. There is also a concern that NICE guidance has not been implemented consistently at a local level. There has also been criticism from various stakeholders that the NICE technology appraisal process takes too long. A shorter Single Technology Appraisal (STA) process, relying much more heavily on the economic evidence provided by the manufacturers, has now been introduced alongside the existing process (which has been renamed the Multiple Technologies Appraisal process). The STA process is similar to that used by the Scottish Medicines Consortium at the time of launch for all new pharmaceuticals. The effect of the STA process is unclear but it appears to shift the burden of proof so that the onus rests more clearly with the manufacturer to demonstrate that a drug will be cost-effective.

Buxton cautioned that the international prominence of NICE and transparency of its decisions should not lead people to infer that formal cost-effectiveness was consistently and extensively used throughout the UK health care system. There is little evidence of it informing local NHS decisions, and not even all of NICE's work (e.g. on clinical practice guidelines) is focussed on cost-effectiveness.

In Canada pharmaceutical assessment through the Canadian Common Drug Review (CDR) faces a number of challenges. Laupacis expressed concern about the validity of surrogate markers and of quality of life measures, in particular whether their use in clinical drug trials would help predict the likelihood of the drugs achieving clinically important (and cost-effective) outcomes in practice. He was also sceptical about the value of modelling, arguing that if sophisticated modelling is needed, it is because we don't know what we need to know. The "more frequent scenario is that the cost-effectiveness ratio varies .. and the reimbursement committee is left with genuine uncertainty about a drug's cost-effectiveness." In other words, additional research (not modelling) is required to help resolve the underlying uncertainties. However, he did not believe "that these models should be abandoned; just that we must be realistic about what we can expect them to provide." Laupacis also noted that:

- o the CDR did not have an explicit cost-effectiveness threshold and there had been limited public involvement in decision making to date. There was a tendency for priority setting bodies to deny that they are actually setting priorities. Increasing openness (by everyone) and inclusiveness is a necessary part of legitimate priority setting (i.e. getting decisions that are accepted and acted upon), although that won't make the decisions themselves any easier to make;
- o the Provinces did not necessarily follow the CDR outcome. It was unclear whether this was because local opportunity cost was different or if there was a "not invented here" syndrome;

Box 1

The Ontario Medical Advisory Secretariat (MAS) process

A comprehensive process for the assessment of new (non pharmaceutical) health technologies in Ontario was recently established. "Requests for new funding are directed through a department of the Ministry called the Medical Advisory Secretariat (MAS), who, using a systematic and evidenced-based approach, provides information on the safety, effectiveness, and cost-effectiveness of the technology to the Ontario Health Technology Advisory Committee (OHTAC)." This produces an estimate of cost-effectiveness and budget impact. MAS/OHTAC has undertaken evidenced-based analyses for over 50 new technologies.

"If OHTAC determines that the evidence for making recommendations is too uncertain, a request is made to reduce this uncertainty through the conduct of a 'real world' Ontario-based study" or "field evaluation". The PATH Reduction of Uncertainty through Field Evaluation (PRUFE) Iterative Evidenced-Based Framework is "an iterative process for collecting information, updating prior knowledge and providing new evidence back to the Ministry regarding the cost-effectiveness of new health technologies".

The field evaluations are considered necessary "if the health technology is likely to have an important impact on patient outcomes but where:

- The quality of evidence is conflicting or too questionable
- RCTs have focused only on a sub-group of patients likely to benefit from the technology, and there is concern that the technology would likely be diffused to other patient sub-groups with unproven effectiveness
- The number needed to treat and the unit cost of the technology are high
- Available follow-up data do not extend over long enough periods of time
- There are unresolved questions regarding patient safety
- The technology is likely to be highly disruptive in terms of budgetary impact."

"Field evaluations are distinct from the usual 'ideal' controlled trial with strict inclusion/exclusion criteria and fixed protocols for patient management and assessment. A distinction can be made between 'explanatory' trials (i.e. can the technology work?) and 'pragmatic' trials (i.e. does the technology work?). Of the seven field evaluations which have been started in Ontario since 2003, four have been pragmatic RCTs, two are cohort observational studies and one is a cohort registry study. These evaluations cover a broad range of disease areas and types of technologies."

- o there was little utilisation control (to see if drugs were used in line with CDR recommendations) and company reluctance to negotiate price-volume arrangements. There was no credible national approach to conditional listing or price negotiation.

For the US, the key message from Neuman and Sullivan was that the multiplicity of decision makers in the public and private sectors produces a complex environment with a variety of ways and contexts within which economic evaluation could be used. They

argued that there is increasing use of cost-effectiveness analysis (CEA) data in both sectors, much of which is not explicit but implicit. In the public sector, it is by stealth as there is no direct mandate to incorporate CEA into decision making. In the private sector, CEA is used within the constraints of each plan or provider's decision framework, but again often without explicit public acknowledgement. Neuman and Sullivan expected this incremental and experimental process to continue, albeit with an emphasis on regional initiatives and on issues around evidence synthesis, but with an increasing willingness to combine economic and clinical evidence, with

some explicit use of cost-effectiveness thresholds “in selected circumstances with public and private leaders willing to take the political risks.”

In addition to these three general reviews, two papers highlighted specific initiatives which may be pointers for the future. Claxton and Sculpher (2006) reported on some pilot projects from the UK on Value of Information approaches to prioritise research. We discuss these further in the “Complexity” section of this Briefing.

The Ontario Medical Advisory Secretariat (MAS) process for assessing non-drug health technologies was outlined by Goeree and Levy (2006) and is set out in some detail in Box 1. It involves commissioning additional field work through the PATH Reduction of Uncertainty through Field Evaluation (PRUFE). PRUFE considers:

- whether information from studies carried out elsewhere were likely to be sufficient to predict the cost-effectiveness of these interventions in routine use in Ontario;
- whether collecting additional data and further reducing uncertainty (beyond the

initial field evaluation) is likely to be cost-effective. Each HTA includes an assessment of the value of collecting additional information (e.g. designing a new trial or continuing to collect data to refine estimates from the current trial).

Goeree and Levy argued that PRUFE has been a successful process for reducing uncertainty and controlling health care expenditures while improving patient outcomes because it has led to resources being focussed on the use of technologies in cost-effective indications.

The country experiences suggest that economic analysis is playing an increasingly important role in decision making. But this increasing use is raising a number of questions. For us, three key issues emerged from the papers and the discussion at the conference: the generalisability of results from one setting to another; the extent to which greater complexity, particularly in modelling, is helpful to decision makers and an efficient use of scarce resources; and the appropriateness of cost-effectiveness thresholds and the understanding of the local opportunity cost. These themes are developed further in the following sections.

GENERALISABILITY OF RESULTS ACROSS GEOGRAPHICAL SETTINGS

Sculpher and Drummond (2006) considered whether it will ever be possible for the results of a specific analysis to be generalisable, that is to say relevant and applicable, to other jurisdictions. By this they meant not just whether costs and effects are the same in different places, but also if the decision-makers’ “objective functions” (e.g. to maximise health gain) and “constraints” (e.g. the relative magnitude and scope of their budgets) are the same and hence whether an analysis of costs and effects is likely to meet the requirements of decision-makers in more than one jurisdiction. They noted that:

- few decision-makers specify their objective function. NICE is an exception, focusing on health gain, although even here there are other factors. They commented that it “is clear from a case-history of decisions and other documentation that most systems are also concerned about the distribution of health gain, although none is willing to define trade-offs between efficiency and equity in such a way as to be useful for analysis.” There is, by implication, an ill defined “health plus” objective function.
- some systems are explicit that their budget constraint is also the required cost

perspective for all analyses undertaken to inform its decisions – this is true, for example, of the NICE guidelines. Some systems, such as the Netherlands, use a wider societal cost perspective, although they “typically require that system costs and wider costs are presented separately, implying that non-system costs may have limited direct impact in decisions”;

- there are often additional “political” constraints, but again, “these are rarely defined *ex ante* in a way that can inform the design and execution of an economic evaluation.”

Typically national agencies which require and use economic analysis to support decision-making publish guidelines stipulating preferred methods for economic evaluation. But these guidelines exhibit considerable variation. Sculpher and Drummond accepted that some variation between countries is inevitable. But how much of the existing variation is justified? They argued that:

- some of the variation can be justified by the differences between systems objectives and constraints;

- some relates to legitimate differences of opinion about appropriate analysis given current methodological uncertainty; but
- much is “difficult to justify and is inconsistent with the aims and objectives of the systems whose decision-makers the analyses are supposed to inform.”

For pharmaceutical companies, there is inevitably a tension between the aims of their global development programmes and the specific requirements of local decision-makers. This tension could be reduced. First, analysts and companies could address the methodological and practical challenges in making economic studies undertaken in the context of global development programmes applicable to many decision contexts. Second, decision-makers could collaborate to make methodological guidelines more consistent. Legitimate variation in the methodological and data needs of different systems will remain, but could be substantially reduced.

Drummond and Sculpher suggested that a first step “would be for each jurisdiction to develop a ‘reference case’, along the lines of that developed by NICE”, to stipulate a system’s requirements for study features such as the descriptive system for health, the use of monetary valuation versus non-monetary metrics, the source of preferences and method of elicitation. This should be required in its own right because of the need for consistency *over time within each jurisdiction*. The second step would be discussions between countries “to distinguish between:

- ‘justifiable’ differences where opinions differ on, say, the choice of health state descriptive system or valuation method, and
- differences that are not easy to justify because the approach suggested in one or more jurisdiction is inconsistent with the objectives and constraints of the system.”

However, although methods are clearly less consistent than they ought to be it may be that few of the differences really matter. Efforts should be concentrated on identifying those which undermine generalisability or require significant additional expenditure in a global development programme to meet local requirements. Such a process will inevitably raise issues as to the legitimacy of local decision-maker demands for data, or methods, that impose significant additional cost.

Manca and Willan (2006) focussed on the generalisability of data from economic analysis conducted as part of a clinical trial. They noted that “multinational trial-wide cost-effectiveness results may not be directly applicable to any one of the countries that participate in the clinical study”:

- “decision-makers are inherently country-specific and are more interested in results which are directly relevant to their own jurisdiction”;
- in many cases “the country of interest does not participate in the clinical trial”;
- “even when the country of interest is part of the original study, the presence of country-specific factors potentially affecting the geographical applicability of the overall study results (effectiveness, cost, and quality of life). ..This means that trial-wide results may not be informative for reimbursement decisions at country-level.”

Hence some form of additional modelling will typically be required to customise the results to the country of interest.

Manca and Willan argued that given “the time and effort required to complete a multinational trial-based CEA, it seems reasonable – from the viewpoint of both the industry and national/state governments – to support the use of methods which facilitate the ‘translation’ of cost-effectiveness data obtained from one country to make them applicable to another. The need to customise the economic study results to a specific jurisdiction is not purely academic, but stems from the decision-makers’ need for context-specific information. This raises two overarching methodological questions: ‘what methods are there to make cost-effectiveness estimates more country-specific?’, and ‘how can we account for factors that may affect the between-country generalisability of cost-effectiveness results?’” They identified on this basis two drivers affecting the type of analysis that can be done – the availability or not of “raw” Individual Patient Data as compared to aggregate results, and whether or not the country of interest had centres participating in the trial. This gave rise to several scenarios including:

Scenario 1. Individual Patient Data (IPD) are available, but the country of interest did not participate in the trial.

Analysts usually assume that the baseline clinical event risk is “location-specific, whilst the relative treatment effect is more generalisable”. While the former captures a range of country-specific factors (e.g. epidemiology, medical attitudes), the relative clinical effectiveness of the intervention is assumed not to differ greatly across countries. Trial-wide relative treatment effects (such as the relative reduction in risk of deaths, cardiac events, or side effects) are then applied to the reference (baseline) risk – i.e. the event rate without the treatment – for the country of interest. They gave the “well known example

of this methodology [as] the application of the West of Scotland Coronary Prevention Study (WOSCOPS) cost-effectiveness results" to a number of other countries (Caro et al, 2000.)

However, Manca and Willan noted that it is possible that the between-country variability in risk factors (that underlies differences in the baseline risk) could in turn lead to differences in the relative treatment effect. This could limit the generalisability of the relative treatment effect observed in the trial beyond participating countries.

Scenario 2. IPD are available and the country did participate in the trial.

Here there is a need to avoid a stark choice between 'pooling' (grouping data from all countries together) and 'splitting' (only using data relating to the country of interest). Pooled analysis ignores possible differences in relative treatment effects. They argued that "the presence of between country differences in the magnitude and sometimes in the direction of the relative treatment effect is a fundamental consideration which, when paired with international differences in factors affecting resource use and costs, makes the estimate of cost effectiveness for a particular country based on the trial wide relative treatment effect unreliable, even when individual [baseline] risk factors (and their distributions) are similar between countries."

They also noted that "Splitting the data, on the other hand, is impractical when the country of interest has recruited a limited number of patients compared with the rest of the countries in the trial....data collected from different countries (and patients) may share some degree of similarity and there may be advantages in trying to capture such similarities."

They recommended combining these approaches in a hierarchical model that allows country-specific estimates to be obtained by combining data from the country of interest with data from the other countries in the trial. Such estimates can be said "to be borrowing strength from each other."

Scenario 3. IPD are unavailable and the country did not participate in the trial.

The most challenging situation "is when IPD are unavailable and the country of interest did not participate in the trial. In this case the analyst has to rely on data published in the literature, and to assume that the relative treatment effect estimated from other countries is indeed generalisable to the country of interest." Methods similar to those for Scenario 1 can be

used, again "supplementing the evidence base with additional IPD specific to the country of interest" if available from elsewhere.

However, as discussed above, the relative treatment effect may be related to baseline risk – "for example, the higher the baseline risk, the lower the treatment effect – in which case the assumed independence between the two components of clinical effectiveness is not sustainable". It will be important to try to test for any relationship between the relative treatment effect (i.e. relative risk reduction) and the baseline risk rather than just assuming that baseline risks are not transferable internationally, but relative risk reductions are. They explained how this was done in the case of the NICE appraisal of GPAs (Sculpher et al, 2004).

Sculpher and Drummond supported Manca and Willan. They argued that the "methodological challenge is to use all available evidence whilst reflecting geographical variability". However, they saw a need for use of "evidence synthesis and decision modelling" because of the severe limitations of "trial-based economic evaluation". Manca and Willan noted "the perceived limitations of using evidence from a single trial to inform decision-making in a specific jurisdiction (country)." They agreed that "it is good practice to synthesise all the available evidence within a comprehensive decision-analytic model" but noted that "clinical trials will always play a crucial role in providing unbiased country-specific estimates of (clinical) treatment effects."

In this context, Sculpher and Drummond discussed the importance of IPD and how it can be combined and used to generalise – albeit outside of trial-based studies. They saw this as particularly valuable when IPD are available from one or more sources and can be used (wholly or partly) to estimate parameters for an economic model. When the IPD relate to a sample (e.g. in a trial) which includes a jurisdiction of interest, regression methods can be used to estimate baseline risks, treatment effects, costs or HRQL parameters for that jurisdiction.

It is clear that producing the best estimate of costs and effects in one jurisdiction, whilst drawing on data relating to other jurisdictions, is an important area for methodological development. Similarly, better techniques for combining IPD with aggregate data are needed, together with an understanding of the importance of using IPD as compared to more aggregated data. Further attention also needs to be paid to the weighting given to trial and non-trial based data in decision analytic modelling.

Unfortunately, such methodological development and thinking is likely to increase – at least in the short term

– the degree of methodological divergence between jurisdictions, and further increase the complexity of analysis. Analysts and companies face a challenge. Companies need to provide evidence for analyses relevant to specific decision-making contexts, within

the constraints of a global development programme for a new drug. Hence where possible they need to use studies which are ‘generic’ in terms of both data and methodology and hence generalisable or easily adaptable to a range of contexts.

COMPLEXITY

Weinstein (2006) and Buxton (2006) both worried about the growing complexity of economic evaluations, with Weinstein arguing in discussion that perhaps some economists were behaving rather like the medics they so often criticise for wanting to keep treating (analysing) until the marginal benefit is zero, rather than until the marginal benefit equals the marginal cost of treating (analysing).

Weinstein summed up his concern:

“The art and science of decision modelling in health care has become much more complex and even arcane since the early decision tree and state-transition models that it used to be possible for us pioneers to run on pocket calculators or to program ourselves in Fortran or Basic. With complexity comes loss of transparency, which is unfortunate. One of the values of a good model is that it can yield qualitative insights that cause a decision maker to say “Aha!” when a counterintuitive strategy proves to be optimal *and for an understandable reason*. A complex model that yields a result that is surprising but not understandable in simple terms is unlikely to change the behavior of decision makers until it is verified in a clinical trial or until it can be explained in logically understandable terms.”

His remarks echoed the comments of Laupacis about the use of economic modelling in the Canadian Common Drug Review Process.

Weinstein went on to discuss the ways in which more rigorous model validation could provide reassurance to decision-makers, but gave an example where data limitations could not be dealt with to the satisfaction of decision-makers by modelling alone:

“A notable example of how models have differed in their approaches to these types of assumptions based on short-term trials can be found in the literature on economic evaluation of drugs for multiple sclerosis.” Here “the issue is not survival or incidence of events, but the reductions in rates of disease progression and relapse, which affect the trajectories of health-related utility with and without treatment. Some of the most prominent models in the literature have obtained stunningly disparate results and cost-effectiveness ratios, largely driven by their disparate assumptions about the durability of treatment effects. The contingent reimbursement scheme for multiple sclerosis drugs in Britain is a unique policy response

to this situation, whereby payment will be retroactively adjusted in response to observed outcomes that could only be modeled from short-term trial data at the time the drugs were introduced.”

Buxton’s concern was the opportunity cost of unnecessarily complex analysis. He argued that in the UK, within the constraints on analytical resources and time, there was an imbalance between the very high quality but very limited quantity of economic evaluations commissioned by NICE and the lack of evidence of any kind on the cost-effectiveness of most technologies used by the NHS. The need was for “reasonable evidence on the cost-effectiveness of numerous new technologies and of the plethora of existing technologies that they might replace within the fixed health service budget.” But for many interventions no economic analysis was available. He observed that:

“Academics are good at raising the bar on the gold-standard, but we are collectively much less helpful in suggesting, what is the minimum level of evidence and analysis that is likely to be enough in many or most circumstances to lead to a reasonable decision. An example of what we might have to see as a rarely used luxury might be probabilistic sensitivity analysis (PSA). It is hard to deny the value of PSA in identifying priorities for future research, but it is not necessary if the only real question is what should we recommend on the basis of existing evidence. We need to be willing to fit the analysis to the context and situation. At present problematic technologies, with patchy and contradictory evidence, which hinge on long-term modelling, and for which the ICER is likely to be close to the threshold, face the same NICE analytical process as ‘no-brainers’.”

Buxton suggested that what is needed is to better align these assessment and appraisal activities into “a much more contingent process, where an initial review, rather like that of the SMC [Scottish Medicines Consortium], is undertaken at launch and then a decision is made as to what level of additional scrutiny and appraisal is needed. In that way the heavy gun resources that NICE has employed to date could be devoted to those technologies, and at a time in the development of the evidence relating to them, where there would be most benefit.”

Claxton and Sculpher argued that it was not sensible for health care payers and HTA bodies to make

Box 2

Is Complexity Good or Bad?

We argue that there are tensions between investing the greater time and effort to ensure higher quality economic analyses which are of more *potential* value and:

- the risk of reducing their *actual* value because decision makers do not understand the methods and therefore do not trust the results;
- the need for *more analyses of a wider range of technologies* within the appraisal resource constraints, i.e. the returns to investing additional appraisal resource in a particular technology diminish.

Reviewers of the draft of our paper have pointed out that complexity, in and of itself, is neither good nor bad. If increased complexity leads to a better decision making process and better decisions then it is a good thing. Complex problems usually require complex methods and approaches. For example, we have seen substantial changes and increasing complexity in the use of statistical methods around clinical data over the past few decades.

As discussed at the Conference and reflected in this Briefing, complexity has arisen from trying to make decisions more evidenced-based. Three key drivers have been:

- understanding how best to use and combine evidence from different sources (for example experimental and observational data) and settings (the generalisability issue);
- improving the comprehensiveness with which uncertainty is handled;
- trying to estimate the costs and benefits of investing in research (notably through the use of Value of Information approaches to model the impact of additional evidence on uncertainty.)

We accept that greater complexity of analysis should improve the quality of decision making in relation to particular problems. That does not mean however that it will or that it is overall an efficient use of resources when opportunity cost is taken into account. Trade-offs remain. These need to be acknowledged and debated if we are to be sure we are achieving “Better Analysis for Better Decisions”.

decisions about what research to commission without information on the likely cost-effectiveness of that research in terms of its ability to add to our knowledge (by reducing our uncertainty about the value of a technology.) Not only could research get commissioned that was of little value but research that could “in some circumstances offer greater benefits than the decision to adopt a technology” may often not take place. There were two questions:

- can “value of information” (Vol) methods provide realistic estimates of the potential value of a piece of research in adding to knowledge? They reported on two opportunities to apply Vol methods to directly inform policy decisions about research priorities in the UK – a pilot study for the UK National Co-ordinating Centre for Health Technology Assessment (NCCCTA) and a pilot study for the NICE.;
- can institutional boundaries and regulatory hurdles be overcome so that cost-effective research does get commissioned? Claxton and Sculpher note that institutions (in the UK and

elsewhere) with the remit for making adoption and reimbursement decisions “are often separated from those responsible for prioritising and commissioning research. The former often come under closer public scrutiny than the latter and have, therefore, been more willing to adopt transparent and explicit approaches to their decisions.”

The NCCHTA pilot study (with 4 case studies) was to consider whether value of information methods might help in “identifying research priorities and commissioning research for the NHS Health Technology Assessment (HTA) programme. The NCCHTA has no role in issuing guidance on the adoption and reimbursement of health technologies and the research topics that it considers come from a variety of sources including a web-based general call for suggestions, special interest groups, as well research recommendations made by NICE.”

The purpose of the NICE pilot study (with 6 case studies) was to:

- investigate:

- o whether existing evidence is sufficient to support the use of a technology;
 - o the appropriate length of time until reconsideration of the guidance;
 - o the nature of any needs for further research.
- “establish the feasibility and requirements of routine value of information analysis in addition to the existing reference case for submissions and Technology Assessment Reports.”

NICE currently separated decisions about additional research from recommendations on the use/non use of technologies. Suggestions are made informally by the Appraisal Committee (which formulates guidance) as part of a Technology Appraisal but formally by the Research and Development Committee.

“The results for the NCCHTA were considered by the panels which prioritise suggested research topics and also by the Commissioning Board which commissions research from the prioritised topics.The results of the NICE pilot were presented to the recently formed Research and Development committee who were considering which methods of prioritisation to adopt.” In both cases “the formal analysis failed to have a significant impact on the decisions taken”. Although “the NICE Research and Development committee believed that value of information analysis should be considered and developed for future prioritisation decisions, they decided to adopt a subjective scoring system in the short term.” In the case of the NCCHTA “the decisions made by the panels and commissioning board did not seem to be informed by the results of the analysis”.

Claxton and Sculpher identified “a number of reasons why the pilots did not have the impact on decision making that could have been hoped for:

- most of those responsible for research prioritisation at NCCHTA were not familiar with cost-effectiveness analysis, most were unfamiliar with decision modelling, and almost all had not been presented with probabilistic analysis and the type of evidence synthesis required for the

analysis. This was also true at NICE, since the Research and Development committee is separate from the Appraisals Committee which formulate guidance and is very familiar with the methods. As a result there was a reluctance to accept and base decisions on such unfamiliar methods.

- as well as a general lack of familiarity with more formal methods of evaluation, there was some questioning of the relevance and quality of some included studies as well as some of the structural assumptions...it could have been possible to explore the impact of these alternative views of the evidence and structural assumptions if there had been an iterative process between the decision makers and the analysts. However, in both pilots, this was not possible.”

Claxton and Sculpher argued, in effect, that the issue of the complexity of the analysis and decision maker lack of familiarity could be overcome. (And indeed Goeree and Levy showed how Vol could be successfully incorporated into a research commissioning exercise). The real problem “may be the separation of research prioritisation and commissioning decisions from adoption and reimbursement. Those responsible for research prioritisation and commissioning do not face the same pressure or scrutiny [as those making adoption and reimbursement decisions] and maybe able to maintain processes which are less explicit and transparent.”

At the heart of the problem, argued Claxton and Sculpher, was a failure to recognise that unless adoption and research decisions are made at the same time and on the same basis then technologies may be adopted with insufficient evidence. Early diffusion can “damage the prospects for gathering the evidence needed to inform future clinical practice...the opportunity cost of adopting a technology without considering whether further research is needed ...is the value of information which maybe forgone. In some circumstances this may be greater than the net benefits offered by the technology itself.” There is progress however. In the UK “NICE is now able to identify up to 3 research priorities annually for funding by the NHS Health Technology Assessment programme annually.”

THE COST-EFFECTIVENESS THRESHOLD AND ITS RELEVANCE TO LOCAL BUDGETS

Buxton’s (Buxton, 2006) review of the use of economic evaluation in the UK emphasised the general consistency of application by NICE of an explicit £20,000-£30,000 cost per QALY threshold with a few often cited exceptions. Buxton cited a

recent paper (Raftery, 2006) which reviewed 86 NICE technology appraisal guidances covering 117 specific technology/patient group combinations. Key points for Buxton were:

- of these 117, NICE gives an unrestricted 'yes' to 27 (23%), 'yes with minor restrictions' to 30 (26%), 'yes with major restrictions' to 38 (32%) and NICE has said no to 22 (19%);
- the highest cost per QALY it had accepted remains that for riluzole for motor neurone disease where the cited estimate of cost per QALY was £34,000-£44,000;
- in the recent decision on trastuzumab for advanced breast cancer NICE cited the company's estimate of £37,500 per QALY but noted that it considered their estimate to be unduly pessimistic;
- "The Committee's landmark appraisal of interferon beta and glatiramer acetate for multiple sclerosis, which took two years to complete and for which 3338 documents are listed on the NICE website, finally deemed it not cost-effective at an estimated incremental cost per QALY of £35,000-£104,000 (with a mean of £70,000 cost per QALY). The Department of Health then intervened with a risk-sharing scheme that effectively accepts a maximum threshold cost per QALY of £36,000".

Buxton noted that the "nature of the threshold value for NICE of the ICER – the maximum value it is prepared to pay for a QALY – has emerged only slowly. For a considerable period NICE steadfastly, denied the existence of any such threshold value.... NICE's revised methodological guidance on technology appraisal published in May 2004 was more transparent than hitherto and cited two values: £20,000 and £30,000. The guidance suggests that, for technologies with an ICER below £20,000, this fact alone should generally be sufficient to ensure their acceptance. Above that level, there needs to be other factors favouring the technology, and above £30,000 these additional arguments have to be pretty strong! A subsequent paper by Rawlins and Culyer elucidated the logic further, and stressed the reasons why NICE rejects the idea of a single, absolute threshold."

Buxton set out the two approaches for determining an appropriate threshold value for a QALY:

- the threshold is based on some concept of social valuation: "the value that the public, or politicians on their behalf, place on an additional QALY". This would require some process of preference elicitation using willingness to pay or associated techniques. "The result would be that the derived value then effectively determines how much should be spent on

health, with cost-effectiveness analysis used to identify any opportunities that exist to generate incremental QALYs for less than that threshold. In principle then, the cost of the resultant set of cost-effective technologies sets the health-care budget";

- the threshold is chosen so that it results "in the most efficient use of a pre-determined health budget. In this case, the threshold has to be set to ensure that, at the margin, adopted technologies have a better cost-effectiveness ratio than the cost-effectiveness ratios of any technologies that are not adopted or which have to be disinvested in order to free resources."

NICE has adopted the latter approach. But Buxton found "at least circumstantial evidence that, even during the past period of considerable growth in the real value of health care spending, NICE's threshold may well have been too high. Too high a threshold, combined with the political difficulty of disinvestment, could explain and to a degree justify why many local health services have found it difficult, or have been unwilling, to fully implement NICE guidance."

Buxton argued that economists need to recognise the reality of pressures facing local decision makers: "If as economists, we would like to see what, in terms of cost-effectiveness, is a more rational allocation of spending at a local level, we need to be actively arguing for less perverse budgetary arrangements rather than passively criticising local behaviour.... NICE, and similar central bodies in other countries, has the great luxury that it does not itself have to balance these competing and inconsistent managerial messages in the recommendations it makes. Local managers have no such luxury."

A possible approach is set out in more detail in Culyer et al (forthcoming). They note that NICE has no authority to determine the health budget and therefore it has to find the threshold value that ensures that what is recommended does not displace something that would have been more cost-effective. They describe this process as 'searching' for a threshold rather than 'setting' one.

By implication, NICE has to seek out evidence as to what is the cost-effectiveness of the local things that NICE recommendations currently displace. It also has to find cost-ineffective technologies to recommend for disinvestment. Disinvestment raises the question of whether threshold values for cost-effectiveness in health care should reflect the observed 'kink' in consumers' values with willingness to accept values being greater than willingness to pay? If such a 'kink' is the political reality, but is not

recognised by NICE, then local managers will find NICE disinvestment decisions even more difficult to act on than if a realistic threshold (in terms of public acceptance) is used.

If the NHS cannot identify relevant thresholds and cannot act to recommend disinvestment from existing cost-ineffective technologies then Buxton concludes that “NICE may have simply to accept the very much less satisfactory position, advocated by some (Cookson et al, 2001, Maynard et al, 2004) of being responsible for setting priorities within a distinct and predetermined ‘new technology’ budget.”

In the UK cost-effectiveness and a concept of what is an acceptable ICER is playing a significant role in some difficult, and very high profile, decisions. But Birch and Gafni complained that it is a bastardised form of cost-effectiveness analysis, which has long ceased to produce values that can be relied on to enable society to maximise economic-welfare. In a typically forthright paper Birch and Gafni pursued the same concerns as Buxton about the local opportunity cost of NICE decisions and but concluded that the fundamental problem is the use of the ICER in this situation is potentially misleading:

“We show that the ICER represents an attempt to provide comparative information on what are non-comparable (and therefore unreal) options and hence is irrelevant to, and evades the reality of, the decision-maker’s problem.”

Their argument is that:

- “... the comparison of ICERs (or IECRs) as a basis of making investment decisions aimed at maximizing health (or corn flakes) gains is only valid where the interventions being compared have identical total costs (i.e., the alternatives are truly interchangeable). But if the two alternatives have the same cost there is no need to calculate a ratio – common sense is again sufficient to determine that the intervention with the greater effectiveness represents the more efficient use of resources.”
- Remarkably little attention has been given in the literature to how the particular ICER thresholds are selected and more importantly how they relate to the opportunity cost considerations of the constrained maximization problem facing the decision-maker. Thus it is not simply that calculating the ratio is unnecessary or unhelpful, in some situations it lead to interventions being adopted that prevent

the maximization of health gain from available resources. Claxton et al (2006) note that the maximization of health benefits from available resources requires information on the shadow price of the decision maker’s budget constraint, which represents the marginal opportunity cost of available resources. This information is not available – witness the emerging discussion of NICE as a “threshold searcher”.

- “The origin of the ICER as a tool for determining whether an intervention represents an efficient use of available resources is found in Weinstein and Zeckhauser (1973). Using assumptions of perfect divisibility and constant returns to scale for all interventions, they show that total health benefits are maximized” by
 - o Ranking all interventions “from the lowest to the highest ICER and selected in descending order until available resources are exhausted (the *league table approach*)”, or
 - o “Specification of the ‘critical ratio’, λ , given by the opportunity cost of the resources at the margin, and implementation of all interventions with an ICER less than or equal to (the *threshold ICER approach*).”

Information on the incremental costs and effects of all current and potential interventions is required for both approaches. But “in the real world information for health care decision makers is incomplete. Consequently complete rankings of interventions cannot be produced and hence λ cannot be determined”. Birch and Gafni argued that “the marginal opportunity cost of resources, λ , depends crucially on, inter alia, the quantity of available resources (i.e., the size of the budget). Hence, communities with the same health care needs but different budgets will have different values of λ against which to judge the efficiency of interventions.” In practice, λ is determined arbitrarily and without any explanation of how application of the λ value leads to the efficient use of available resources.

- In summary, the ICER represents “Information Created to Evade the decision-maker’s Reality”.

Birch and Gafni noted that they had previously “presented integer programming solutions to dealing with the problems of indivisibilities and non-constant returns to scale in the context of a constrained maximization problem....These mathematical programming techniques require information on the costs and effects of all current and potential new interventions, together with the resources available for investment. Although these data requirements may be difficult to satisfy, they reflect the complex nature of the decision makers’ problem.”

This may seem a lot of information given that they criticise those using ICERs for lacking the information they need. However, for Birch and Gafni “the use, instead, of a restricted league table of ICER values or an arbitrarily determined ICER threshold might offer an intuitive short-cut to resolving the decision-maker’s problem but as Williams noted “..reality is horrendously complicated...the more complex the reality is, the more dangerous it is to rely on intuitive short-cuts rather than careful analysis” (Williams, 2004).

Birch and Gafni also criticised the approach (Rawlins and Culyer, 2004) of separating cost-effectiveness from cost as “matters of efficiency cannot be separated from matters of affordability. Because

money represents only command over resources, value for money is determined in relation to what it can purchase. Hence whether a particular intervention represents ‘value for money’ is determined by what is forgone in order to ‘afford it’. As Williams (Williams, 2004) notes, if affordability could be separated from efficiency there would be no need for a threshold.”

In effect Birch and Gafni argued for the need to look at budget impact and the opportunity cost of using that budget on the favoured technology. This is where Buxton also wanted to take NICE – to a better understanding of what would in practice be displaced by the use of a technology – and indeed what lies behind the notion of NICE as a “threshold seeker”. They approached the issue from radically different perspectives – Birch and Gafni rejecting on theoretical grounds the use of ICERs (essentially bypassing them to get to a measure of opportunity cost) and Buxton looking for further research as to what is happening in the NHS to get a better proxy for the ICER that NICE should use. Both implicitly rejected the use of a social valuation of the QALY (what would the public be willing-to-pay for), preferring approaches grounded in the reality of fixed budget constraints, (by implication, budgets in aggregate that are below public willingness-to-pay sums).

CONCLUSIONS

Overall the Conference marked a significant step in the world of health economic evaluation, where the discussion moved forward from the broad rhetoric of why it should be used to the more sensitive recognition that it is now being used, and that its very use is posing challenges that must now be met if the promise it offers is to be delivered. Three clear issues emerged:

- generalisability and the extent to which it is possible and efficient, to make economic evaluations relevant to specific jurisdictions within the constraints of global development programmes and multi-national clinical trials;
- increasing analytical complexity and the need to balance this against the needs, and abilities, of decision-makers to

understand and act on economic studies and the need for more studies of a wide range of technologies within reasonable constraints on analytical capacity;

- the threshold value for the incremental cost-effectiveness ratio and the need to ensure that this does lead to an efficient use of constrained health care resources.

The various speakers and participants would each, no doubt, place a different emphasis or priority on the next steps. We set out some possible ways forward in Box 3 overleaf. Certainly the conference offered no easy solutions, but it delivered a lively debate and constructive discussion and the shared recognition that these and other problems need to be addressed.

Box 3

Resolving these issues

Generalisability

The central problem is the tension between global development programmes and the desire of local decision-makers to have tailored information about cost-effectiveness. Actions arising in order to resolve tension between the global and the local in an efficient way could include:

- greater clarity on what needs to vary methodologically across countries;
- agreement on the tests to be performed and rules to be applied to the use of data and results from multi-centre trials within and outside of the countries for which estimates of cost-effectiveness are required;
- further development and promulgation of robust methods to combine different sources of evidence (evidence synthesis);
- application of Value of Information (VoI) approaches to seek to resolve the issue of the extent to which local demands for additional work can be regarded as reasonable.

Complexity

Several actions might reduce the potential dilemma of trading off complexity with decision maker understanding and Health Technology Assessment (HTA) resource constraints:

- development of a more contingent appraisal process where the level of analytical effort is more closely related to the nature of the decision-problem;
- use of VoI approaches to determine the relevant level of complexity in particular cases;
- use of VoI approaches to determine which technologies to appraise;
- use of risk sharing schemes to allow the adoption of technologies despite high levels of remaining uncertainty. It will be important to structure such schemes so that they collect data that does resolve key uncertainties. Use could in some circumstances reduce the ability to obtain data.

Thresholds

The central problem is local opportunity cost. Which health programmes and / or technologies in practice will or should an efficient local decision-maker give up or not invest in order to provide the technology recommended by the national HTA body? The way forward could include:

- more specific assessment of the opportunity cost in various localities of central decisions to generate relevant ICERs for these decisions;
- more emphasis on evaluation of existing technologies that might be appropriately disinvested;
- greater use of programme budgeting approaches at a local level.

ABOUT THE AUTHORS

Adrian Towse is the Director of the Office of Health Economics. He is a Visiting Professor in the Department of Economics and Related Studies at the University of York and a Visiting Researcher at the Department of Public Health and Primary Care at the University of Oxford.

Martin Buxton is Professor of Health Economics at Brunel University and Director of the Health Economics Research Group (HERG) which he founded in 1980. For 3 years until 2002 he was a member of the Appraisal Committee of the NICE and continues to be involved in various of its activities.

REFERENCES

The conference proceedings publication Papers referred to in this Briefing

- Buxton M (2006). Economic evaluation and decision making in the United Kingdom. *PharmacoEconomics* 2006; 24 (11): 1133-42
- Birch S, Gafni A (2006). Information Created to Evade Reality (ICER): Things we should not look to for answers. *PharmacoEconomics* 2006; 24 (11): 1121-31
- Claxton K, Sculpher M (2006). Using Value of Information Analysis to Prioritise Health Research: Some Lessons from Recent UK Experience. *PharmacoEconomics* 2006; 24 (11): 1055-68
- Goeree R, Levin L. Building bridges between academic research and policy formulation: the PRUFE framework – an integral part of Ontario’s evidence-based HTPA process. *PharmacoEconomics* 2006; 24 (11): 1143-1156
- Laupacis A (2006). Economic Evaluations in the Canadian Common Drug Review. *PharmacoEconomics* 2006; 24 (11): 1157-1162
- Manca A, Willan A (2006). “Lost in translation”: accounting for between-country differences in the analysis of multinational cost-effectiveness data. *PharmacoEconomics* 2006; 24 (11): 1101-19
- Neuman P, Sullivan S (2006). Economic Evaluation in the U.S. What’s the Missing Link? *PharmacoEconomics* 2006; 24 (11): 1163-1168
- Sculpher M, Drummond M (2006). Analysis Sans Frontières: Can we Ever Make Economic Evaluations Generalisable Across Jurisdictions? *PharmacoEconomics* 2006; 24 (11): 1087-99
- Weinstein M (2006). Recent Developments in Decision-Analytic Modelling for Economic Evaluation. *PharmacoEconomics* 2006; 24 (11): 1043-53

Additional papers included in the conference proceedings volume.

- Briggs A, Levy A (2006). *PharmacoEconomics & Pharmacoepidemiology: Curious bedfellows or a match made in heaven?* *PharmacoEconomics* 2006; 24 (11): 1079-86
- Torrance G (2006). *Utility Measurement in Health Care: The Things I Never Got To.* *PharmacoEconomics* 2006; 24 (11): 1069-78

Other references

- Briggs A et al (2004). *Unfinished Symphony: A tribute to the life and career of Bernie O'Brien (1959-2004)*, *Medical Decision Making*, vol 24, pp 538-44
- Caro J et al, (2000). Are the WOSCOPS clinical and economic findings generalizable to other populations? A case study for Belgium. *Acta Cardiologica*, 55(4): 239-246
- Claxton K et al (2006). Discounting and cost-effectiveness in NICE – stepping back to sort out a confusion. *Health Economics*, 15: 1-4
- Cookson R et al (2001). Wrong SIGN, NICE mess: is national guidance distorting allocation of resources? *British Medical Journal*; 323:743-5
- Culyer A et al (forthcoming). Searching for a threshold, not setting one: the role of the National Institute of Health and Clinical Excellence. *Journal of Health Services Research and Policy*
- Maynard A et al (2004). Challenges for the National Institute for Clinical Excellence. *British Medical Journal*, 329: 227-229
- Rafferty J (2006). Review of NICE’s recommendations, 1999-2005. *British Medical Journal*, 332: 1266-8
- Rawlins M, Culyer A (2004). National Institute for Clinical Excellence and its value judgements. *British Medical Journal*, 329: 224-227
- Sculpher M et al (2004). Generalisability in Economic Evaluation Studies in Health Care. a Review and Case-Studies. *Health Technology Assessment*, 8(49):1-206
- Weinstein M, Zeckhauser R (1973). Critical ratios and efficient allocation. *Journal of Public Economics*, 2: 147-157
- Williams A (2004). *What could be nicer than NICE?* Office of Health Economics, London



In June 2006, a conference entitled *Better Analysis for Better Decisions: Bridging the Gap Between Economic Evaluation and Healthcare Decision-Making* was held at McMaster University in honour of Bernie O'Brien. Bernie, until his untimely death in February 2004, was Professor in the Department of Clinical Epidemiology and Biostatistics and Director of the Programme for Appraisal of Technology in Health (PATH) funded by the Ontario Ministry of Health and Longterm Care. Prior to his move to McMaster, Bernie worked with Martin Buxton at Brunel University in the Health

Economics Research Group where he took his PhD. He was part-funded for his PhD by the Office of Health Economics and wrote a number of publications for the OHE. For a comprehensive review of Bernie's academic career please read Briggs et al (2004).

The conference brought together a group of leading health economists, other analysts and health policy-makers, many of them people with whom Bernie had worked. There was lively debate between three schools of thought – methodologists, practical decision makers and iconoclasts. Many participants would, correctly, regard themselves as being in more than one of these groups. This Briefing summarises the main themes that in our view emerged from the papers and highlights the issues they raise – topics that need to be addressed if economic evaluation is to continue to improve decision-making in health care.

The papers are published in the November 2006 issue of *PharmacoEconomics*. Full references for the papers and details as to how to get hold of the journal issue are included in this Briefing.

About the OHE

The Office of Health Economics (OHE) 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, consultancy and commissioned research.

Peer Review

All OHE publications have been reviewed by members of its Editorial Board and, where appropriate, other clinical or technical experts independent of the authors. The current membership of the Editorial Board is as follows:

Professor Christopher Bulpitt – *Royal Postgraduate Medical School, Hammersmith Hospital*
Professor Martin Buxton – *Health Economics Research Group, Brunel University*
Professor Tony Culyer (Chair) – *Institute for Work and Health, Ontario University, Canada*
Dr Jennifer Dixon – *The King's Fund*
Professor Hugh Gravelle – *Centre for Health Economics, University of York*
Mr Geoffrey Hulme – *Director, Public Finance Foundation*
Professor Carol Propper – *Department of Economics, University of Bristol*
Professor Bonnie Sibbald – *National Primary Care Research and Development Centre, University of Manchester*
Mr Nicholas Wells
Professor Peter Zweifel – *Socioeconomic Institute, University of Zurich*

Details of all OHE publications, and how to order them, can be found on the OHE website at www.ohe.org



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