OHE Research Office of Health Economics

OHE Briefing

No 46 November 2008

NEW GUIDELINES FOR ECONOMIC EVALUATION IN GERMANY AND THE UNITED KINGDOM
Are we any closer to developing international standards?

Michael Drummond, *Centre for Health Economics, University of York, UK*
Frans Rutten, *Institute for Medical Technology Assessment, Erasmus University Medical Centre, The Netherlands*

Contents

1. Introduction 1
2. Guide to the methods of technology appraisal (NICE, 2008) 2
3. Methods for assessing the relation of benefits to costs in the German statutory health care system (IQWIG, 2008) 5
4. Towards the development of international standards 9
5. Conclusions 13
6. References 14

1. INTRODUCTION

The last twenty years has seen a substantial growth in the literature on economic evaluation in health care. In addition, since Australia made economic evaluation an important component of its decision-making process on the reimbursement of drugs in 1993, several jurisdictions have adopted economic evaluation as part of their formal decision-making procedures.

In situations where economic evaluation is formally adopted, it is customary for the decision-making authority (e.g. Ministry of Health or health technology assessment agency) to issue methodological guidelines for the conduct of studies. As more and more sets of guidelines have been published, researchers have compared and contrasted the methodological guidance prescribed by the various jurisdictions (Hjelmgren et al., 2001; Tarn and Smith, 2006).

The general conclusions of these assessments are that there are probably more similarities between the sets of guidelines than there are differences, especially among those guidelines that have been developed in the context of formal decision-making procedures. Nevertheless, Sculpher and Drummond (2006) point out that, despite the general similarities, there are important differences among guidelines in key elements of methodology, such as the choice of comparator (to the treatment of interest), or the ways in which uncertainty is to be handled. (See Tables 1 and 2)

Table I: Variability among guidelines on methods for health care economic evaluation in terms of choice of comparator (n = 27)

Recommended method for comparator selection	No. of guidelines
Most commonly used treatment	8
Existing, most effective or minimum practice	2
Existing or most effective treatment	1
Any treatment, provided that the choice is justified	1
Existing and no treatment	2
Most common, least costly, no treatment	1
Most common, least costly, most effective, no treatment	2
Most common, least costly, most effective treatment	1
Treatment most likely to be displaced	1
Most efficient, most effective treatment, plus the do nothing option	2
All relevant comparators	2
Most effective and no treatment	1
Not clear/specific	3

Source: Sculpher and Drummond, 2006

Table 2: Variability in guidelines on methods for health care economic evaluation in terms of the approach to sensitivity analysis (SA) N=27

Recommended for SA	No. of guidelines
Need to state approach and to justify it	3
Probabilistic sensitivity analysis (PSA)	3
One-way SA, multi-way SA	1
One-way SA, two-way SA	2
Multi-way SA (of most important parameters)	1
One-way SA, multi-way SA and PSA	5
One-way SA, multi-way SA and worst-best scenario analysis	1
One-way with tornado diagram	1
Not stated/no specific recommendation	10

Source: Sculpher and Drummond, 2006

Sculpher and Drummond (2006) argue that some of these differences could be considered legitimate, since they reflect differences in local value judgements (eg the choice of discount rate, though this is also subject of a methodological debate, Claxton et al., 2006, Gravelle et al., 2007). However, others could be questioned, in situations where there is considerable international agreement on how to tackle particular methodological issues (e.g. the need to characterize overall uncertainty in the estimates).

This raises the issue of whether it is possible to develop international standards for economic evaluation in health care. Such standards are likely to be more relevant as a growing number of jurisdictions request economic evaluations in support of their decisions since, even if a multi-national decision-making body for drug reimbursement (e.g. in the European Union) is some way off, it would greatly reduce the burden on those conducting studies if common methods could be agreed.

Although every set of methodological guidelines represents, in essence, what the authors feel should be the accepted standard, none has yet claimed to set an international standard. Perhaps the closest to this is the reference case proposed by the United States Public Health Service Panel (also referred to as the Washington Panel) in 1996 (Gold et al., 1996). This set of guidelines gained considerable international prominence, partly because of the expertise of the individuals involved and partly because of the comprehensive approach adopted.

Also, many authors considered them to be the standard to follow in submitting papers to the top journals in the USA. The concept of the reference case was widely embraced, since it prescribed a set of minimum requirements for studies (thereby encouraging standardization), without stifling methodological advances (e.g. using discrete choice analysis to value benefits). Provided that authors submitted an analysis in accordance with the reference case, they were encouraged to submit additional analyses using alternative methods.

In addition, most commentators supported the majority of the methods suggested by the Washington Panel, the major disagreement being over whether the gains from increased productivity could be assumed to be included in the estimate of the QALYs gained. Other researchers argued that, even if individuals did consider income when assessing the value to them of improved health, individual income may only have a weak link with production change, particularly in settings where individuals have protection against loss of income, or where they experience reduced productivity whilst remaining at work. Therefore, the best approach would be to estimate the value of improved health whilst asking individuals to ignore income effects and then to estimate productivity changes separately , for inclusion in the numerator of the cost-effectiveness ratio. (Brouwer et al., 1997a,b).

This briefing note examines two recent sets of methodological guidance issued by agencies in two major European countries (NICE, 2008; IQWiG, January 2008) and assesses whether analysts are moving closer to the development of international standards in economic evaluation and what it would take to achieve this aim.

2. GUIDE TO THE METHODS OF TECHNOLOGY APPRAISAL (NICE, 2008)

2.1 Historical development of NICE's methods guidance

This 2008 document outlines the third set of methodological guidelines to be used by NICE since its inception in 1999. The first set of guidelines was fairly general, building on a guidelines that had been issued by the Department of health several years previously. Two important features of the guidelines were, in fact, determined outside of NICE. First, the Department of Health insisted that, since NICE's role was to maximise the gain, in terms of improved health, from the National Health Service (NHS) budget, the appropriate perspective for costs was that

of the NHS and Personal Social Services (PSS). Secondly, the UK Treasury insisted that NICE use the current test discount rate. (At the time of the initial set of guidelines this was 6% per annum for costs and 1.5% per annum for health benefits.) Because there have been substantial developments in NICE's methods in recent years, the initial methods guidelines will not be discussed any further here.

The Guide to the Methods of Technology Appraisal issued in April 2004 (NICE, 2004) represented a major shift in emphasis. These guidelines were much more detailed, and prescriptive, than the earlier ones. However, they also embodied the concept of the reference case (discussed above) and therefore did not prevent analysts from undertaking other analyses, in addition to the required one, if they so wished.

2.2 The 2004 NICE reference case

The key elements of NICE's reference case were as follows;

Defining the decision problem: this involved a clear definition and justification of the technologies being compared and the relevant patient group(s). These elements were to be determined during NICE's 'scoping' of the appraisal.

Perspective: the perspective on outcomes was all direct health effects whether for patients, or, where relevant, other individuals (principally carers). The perspective on costs remained that of the NHS and PSS, although in non-reference cost analyses, significant resource costs imposed outside the NHS could also be considered. These could include direct costs on patients or carers, or costs to other public sector organisations, but would not normally include productivity costs.

Time horizon: this should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Since many technologies have impacts on costs and outcomes over a patient's lifetime, NICE recognised that extrapolation, through modelling, was often necessary.

Synthesis of evidence on outcomes: this was to be achieved by a systematic review of all the relevant literature. Although it was recognised that estimates of relative treatment effect would be most valid if based on evidence from head-to-head randomised controlled trials (RCTs), it was acknowledged that indirect trial comparisons and non-RCT evidence may be required. Analysts were asked to consider the implications of selection bias when using these studies and to perform an analysis of uncertainty in their estimations.

Measurement of health benefit: the benefit measure of choice was the quality-adjusted life-year (QALY). Some of the restrictive assumptions of the QALY (e.g. constant proportional trade-off and additive independence between health states) were recognised. Analysts were asked to comment if these were considered inappropriate and to justify the use of alternative measures.

Description and valuation of health states: health states were to be described using a generic and validated classification system, for which UK population preference values were available, elicited using a choice-based method such as time trade-off or standard gamble. Although the EQ-5D was considered to be the most appropriate choice of instrument, NICE felt it was inappropriate to require the use of the EQ-5D to the exclusion of any other methods meeting its underlying criteria.

Discounting: because the Treasury's requirements on discounting had changed, in the 2004 guidelines both costs and benefits were to be discounted at an annual rate of 3.5%, with sensitivity analyses using 0% and 6% per annum, where this was thought to be important.

Dealing with uncertainty: for parameter uncertainty, the 2004 guidelines stated that probabilistic sensitivity analysis should be used, as this translated the imprecision in all input variables into a measure of decision uncertainty in the cost effectiveness of the options being compared.

2.3 Criticisms of the 2004 reference case

Although the 2004 guidelines were generally well-received, there have been criticisms of a number of aspects. First, the perspective on costs has been criticised as being too narrow. This issue arose in the judicial review of NICE's guidance against the use of alzheimer's drugs for people with mild disease, where some parties argued that caregivers' costs should have been considered in the primary analysis. Some commentators have also noted that, while the perspective on costs is fairly narrow for NICE's technology appraisals, the perspective for NICE's public health appraisals includes costs on all other public sector budgets (Drummond et.al., 2008a).

Secondly, NICE's reliance on QALYs has been criticised, particularly in relation to the use of a decision rule based on the incremental cost per QALY ratio. In a series of papers, Birch and Gafni (2003; 2007) have argued that, by using a threshold of (say) £30,000 per QALY as its base for accepting or rejecting new technologies, NICE is not adequately considering the true opportunity cost of these technologies to the NHS.

Thirdly, some analysts have questioned whether NICE's chosen methods are overly complicated and that it might be better for NICE to consider more technologies but to assess them in a less resource-intensive manner (Buxton, 2006). In particular, NICE's insistence on probabilistic sensitivity analysis has been singled out as a requirement that complicates the analysis but often offers few additional insights. In addition, some analysts have complained that the requirement for PSA limits their options in choice of model. This has resulted in a lively debate (Caro et al., 2007; Griffin et al., 2006; 2007).

2.4 The 2008 NICE guidelines

The guidelines published in June 2008 (NICE, 2008) are a development of the guidelines issued by NICE in April 2004 (NICE, 2004). As in the 2004 guidelines, the stated objective of NICE is to maximize the health gain from available resources (Section 5.2.8). As in the earlier guidelines, health gain is expressed in terms of quality-adjusted life-years (QALYs) and the efficiency of an intervention is assessed by its incremental cost per QALY gained, compared with relevant alternatives.

The perspective for costing in the primary analysis is National Health Service (NHS) and Personal Social Services (PSS) costs, as in the previous guidelines. It is specifically noted (Section 5.2.10) that productivity gains and losses should not feature in the base case analysis.

Although the QALY is clearly the main dimension of outcome to be considered, the guidelines recognize that various dimensions of patient experience could be important (Section 4.3.6) and that other characteristics of therapies (e.g. convenience) can be noted. However, there is an implied judgement that the QALY captures most of the important aspects of health gain.

The guidelines recognize that the randomized controlled trial (RCT) is the best way to determine relative treatment effect (Section 3.1.3), but also recognize its limitations (Section 3.2.3). Indeed, the 2008 guidelines contain an extensive discussion of the need for evidence synthesis, including the need to synthesize indirect and mixed treatment comparisons in situations where relevant head-to-head clinical trials do not exist. There is also a brief discussion of possible methods.

Given the interest in QALYs as the outcome measure, it is recognized that modelling will be necessary in most situations, since RCTs will not measure the most appropriate endpoints, or be conducted over a long-enough period.

The guidelines make it clear that the objective is to develop a reference case for economic evaluation that will facilitate decision-making across all disease areas (Section 5.2.1). Therefore, in formulating the decision problem all relevant comparators (to the treatment of interest) need to be considered and, with a generic outcome such as the QALY, studies in different disease areas can be compared.

Finally, the guidelines recognize that there is often considerable uncertainty surrounding the estimates from economic evaluations, either because of structural uncertainty in the methods used, or parameter uncertainty in the data inputs. There is also a discussion of how uncertainty can be reduced, but the guidelines stop short of recommending formal approaches, such as value of information analysis.

Although the 2008 guidelines are fairly similar to the 2004 guidelines, there are several key differences. First, it is acknowledged that, for some health care interventions, costs on other government budgets could be relevant. This change is due partly to the fact that, in its guidance on the economic evaluation of public health programmes, NICE had already recognized the relevance of a broader perspective. Therefore, it would be perverse to acknowledge the impacts, on criminal justice costs, of a programme to prevent substance abuse, yet to ignore these impacts in an evaluation of a drug treatment for heroin addiction. However, the 2008 guidelines require that these costs should be reported separately from the reference case and that the intention to submit such data should be identified in advance.

Secondly, whilst the QALY remains the outcome measure of choice, its methodological problems are again recognized (Section 5.2.12). Furthermore, it is recognized that, in making decisions about the suitability or otherwise of various health care treatments, social value judgements are required (Section 1.4.4). Therefore, the QALY, whilst being a reasonable measure of health gain, may not adequately capture all the relevant elements of social value. This means that treating all QALYs of being of equal value, no matter to whom they accrue, may not adequately reflect societal preferences (Section 5.12).

On the other hand, the new guidelines are much more specific about how QALYs should be estimated, by declaring a clear preference for the EQ-5D as the generic instrument to be used in the measurement of health-related quality of life in adults. In situations where EQ-5D data are not available, or are inappropriate for the condition or the effects of treatment, the valuation methods should be fully described and comparable to those used for the EQ-5D.

Finally, in one area, the characterization of uncertainty, there appears to be some backtracking from the previous guidelines. Namely, whereas the 2004 guidelines gave a fairly strong steer towards the use of probabilistic sensitivity analysis (PSA), the 2008 guidelines acknowledge that this may constrain the analyst's options in modelling. Therefore, while PSA remains the preferred approach for characterizing uncertainty, it is by no means compulsory.

3. METHODS FOR ASSESSING THE RELATION OF BENEFITS TO COSTS IN THE GERMAN STATUTORY HEALTH CARE SYSTEM (IQWiG, 2008)

3.1. Reimbursement policy in Germany and the restrictions imposed by IQWiG

As of April 1 2007, the German legislature stipulates that the Institute for Quality and Efficiency in Health Care (IQWiG) would be commissioned to carry out evaluations of the benefits and cost-benefit ratios of pharmaceuticals. The results provided by IQWiG will support the Central Federal Association of Health Insurance Funds in setting the ceiling price for specific drugs that cannot be included in a reference price group. The results may also be used to support the Federal Joint Committee in assessing the efficiency of medical interventions in general (see press statement of January 24 by IQWiG). In January 2008 a consultation document was produced by IQWiG in consultation with an international Expert Panel (IQWiG, January 2008).

As is mentioned in the preamble to the document, the specific requirements of German legislation (§35b SGB V) state that IQWiG should value the utility of interventions according to international standards, especially as these are established within health economics. But IQWiG's mandate to its Expert Panel imposed additional constraints, which are rather restrictive and are at considerable tension with the requirement to use methods according to international standards. The most important restrictions are:

(a) IQWiG should only address the determination of a ceiling price at which a superior health technology ***in a given therapeutic area*** *should be continued to be reimbursed (IQWiG, January 2008, page iv)*

This quite restrictive condition is in contrast with the NICE guidance where it is stated that analyses should facilitate decision-making across all disease areas (Section 5.2.1). According to the consultation document an important reason for focussing on a single therapeutic area is that Germany's health care system is not bound to a fixed national budget and therefore should not consider funding priorities across therapeutic areas. It is clear that this is an important deviation from common health economic methodology, where a common measure of benefit is sought and trade-offs are made across therapeutic areas and diseases. The rationale for the latter is, of course, that one would not like to spend $100,000 to gain a unit of health (e.g. a quality-adjusted life-year or QALY) in one therapeutic area, whilst denying an intervention in another therapeutic area with a cost per QALY of only $10,000 per QALY, just because one had failed to compare directly the interventions in the 2 disease areas under consideration.

Even if Germany is not bound to a fixed national budget, which is not obvious, it would be unwise to allow such inefficiencies by concentrating on one given therapeutic area at a time and not checking the consistency of decisions across therapeutic areas, as is done when a general measure like cost per QALY is used. To have different cost effectiveness thresholds for different therapeutic areas may also be judged inequitable, as patients in one disease area may have access to a health care service while patients in another may not, even if these services are equally cost effective and patients may be equally in need of these services.

Why this rather restrictive approach is chosen in this methodology paper remains unclear. One should acknowledge that this is the first attempt to give economic evaluation a formal role in German health policy. On the other hand it is unclear why so little attempt is made to learn from experience in other jurisdictions. Nowhere else are analysts restricted to one therapeutic area when considering the cost effectiveness of alternative strategies. The international group of experts was confronted with this restriction from the beginning and was not able to discuss relaxation of it. One can only speculate about the reasons behind this approach. However, in a recent editorial Jönsson (2008) suggests that 'economic evaluation in Germany is still not seen as a valuable instrument by decision makers'.

(b) The costs to be considered should only be those from the perspective of the community of citizens insured by statutory health insurance (IQWiG, January 2008, page v)

This would fall within the variation seen in the guidelines for pharmacoeconomics across national jurisdictions and is therefore in line with international standards. There is a large variation in perspectives mentioned in official guidelines, ranging from a narrow perspective suggesting that only direct costs borne by health care payers should be considered, to a societal perspective indicating that productivity

costs, costs borne by patients and costs of informal care givers may be included. It is fair to say that even in the latter situation health authorities tend to give more weight to the costs falling on the health care budget.

(c) The estimation of benefits should be according to standards of evidence based medicine (EBM), (IQWiG, January 2008, page vi)

This is, of course, completely acceptable, but the way it is interpreted in the methods section is rather stringent, as EBM seems to be restricted to results from RCTs. In the question and answer (Q&A) section of IQWiG's website it is stressed that well- founded evidence (from RCTs) should be separated from other evidence based on bad methods or professional opinion; but how evidence from other sources than RCTs should be valued is not stated.

Although most international guidelines stress the importance of high quality clinical evidence, experience from many settings shows that using results from RCTs is a necessary but insufficient approach for understanding the true value of health technologies. First, the available clinical trials often do not compare the relevant alternatives (for reimbursement decisions), are too short-term, or measure only a limited range of endpoints. Secondly, a recent review of (systematic) reviews of clinical trials undertaken for NICE in the UK has shown that these were often inadequate, because a pooled estimate of effectiveness could not be produced, or could only be produced for a restricted range of outcomes. This is why NICE has strongly supported economic modelling (Drummond et al., 2008b).

(d) IQWiG will assess only those technologies that have been demonstrated to be superior (IQWiG, January 2008, page vi)

Although this would prevent consideration of highly efficient new technologies (e.g. just a little less effective but much cheaper), other jurisdictions (e.g. the Netherlands) employ a similar restriction, as the decision making process is organised such that better or equal effectiveness has to be demonstrated first before cost-effectiveness is considered. In considering the cost effectiveness plane (Drummond et al., 2005), both the North East quadrant (new intervention is both more effective and more costly) and the South West quadrant (new intervention is less effective but also less costly) are equally interesting when exploring the relative efficiency of new technologies. However, it is commonly assumed that the line representing the societal threshold in the South West quadrant lies closer to the horizontal axis (costs) than the familiar threshold in the North East quadrant for new, more expensive technologies. In other words, relatively more resources must be freed to give up some benefit. There are few examples, however, of successful new technologies in the South West quadrant, and most policies tend to take the established standard of care as the point of departure, which prevents due consideration of such technologies. In drug prescribing the closest analogy is the use of an inexpensive generic product first-line, reserving more costly therapies for those patients who do not respond to initial therapy.

(e) Transferability of economic evaluations to Germany is allowed when adjustments are made for local conditions

Again, this is in line with other international guidelines and is not contentious. The document specifies that the analysis should consider local conditions relating to epidemiology, availability of resources, access to health care, clinical practice, reimbursement of providers, and organisational structures. In the guidelines of other jurisdictions one also finds statements about the transferability of treatment effects and health utilities, which are not mentioned in the IQWiG document.

These five restrictions, especially the first one, form an impediment for performing proper economic assessments in Germany. Therefore it is disappointing that the reasons for imposing these restrictions are not thoroughly discussed in the IQWiG document, especially as they do not appear to be derived from the legal context. In a thorough analysis of the German legislative framework and its implications for the methodology of cost benefit assessment, von der Schulenburg et al. (2007) determine a minimum catalogue of methods and criteria that meet the legal requirements in Germany. Their most relevant statements regarding what this catalogue should consist of are:

- as benefit dimensions the legislature explicitly mentions the quantity and quality of life: this suggests the use of patient reported outcomes and comprehensive measures like QALYs, rather than staying away from such concepts;

- all available evidence meeting certain requirements regarding perspective, study form, databases, calculation of costs and benefits and modelling should be taken into account (including non-RCT evidence) and these requirements are further specified by the authors using international standards;

- there is no requirement from the German legislative context that allocation decisions should only be considered within a therapeutic area.

The resulting catalogue of von der Schulenburg et al. (2007), which they claim to be appropriate in the German context, resembles the toolkit described in most international textbooks and therefore deviates considerably from IQWiG's methodological recommendations as discussed here.

In conclusion, it remains unclear and very much disputed by many German health economists (see Health Economics Committee of the Social Policy Association, 2008 – a position statement signed by 29 German health economists), why these restrictions are required given the German context. As is stated in the methodology paper (IQWiG, January 2008, page iv), there is room for interpretation of what is needed given this context and IQWiG itself has imposed additional constraints. As suggested above in discussing restriction a) and, to a lesser extent, restriction c) the arguments given in the methodology paper are not convincing.

3.2. The main impact of the restrictions on the methodology of IQWiG: the efficiency frontier approach

The methodology section begins by saying that none of the existing methods for economic evaluation is universally accepted and therefore cannot be used for ceiling price assessments in Germany. This suggests that cost-utility analysis, which is recommended as the reference case in most textbooks and in various national guidelines, is also not to be used in the German context. Whilst it is true that there are some differences among the various national and international guidelines, there is quite considerable agreement on the general approach. For example, we observed above that the reference case proposed by the Washington Panel (see Table 3, Gold et al., 1996) included QALYs and that these guidelines have been widely followed in the literature.

Nevertheless, costs and benefits still need to be compared in a given therapeutic area to arrive at a ceiling price for that area. In Section 2.3.1. of the IQWiG document it is suggested that 'accepted clinical measures' should be used as measure of benefit, the advantage being their familiarity to clinicians and their availability from clinical trials. There are various problems with this, as also described in the document. Some of the problems relate to the fact that the measures may not be cardinal (see below), that they may only provide a benefit measure in one dimension, or that the relation between the surrogate endpoint and the final outcome to the patient may not be stable across interventions (or over time). Though these problems are acknowledged, no solutions are provided. Of course, in international studies QALYs have been constructed to overcome these problems, but QALYs are definitely not recommended and not even mentioned in the IQWiG document (except in footnote 1 on page v).

More information on the aversion towards using QALYs can be found in the Q&A section of IQWiG's website. There it is suggested that using QALYs would not represent the values of the majority of the general population and would therefore be very risky. There are no references provided to support these statements. Also, the possible use of 'willingness to pay' is not mentioned, a cardinal measure that may be determined by using discrete choice analysis.

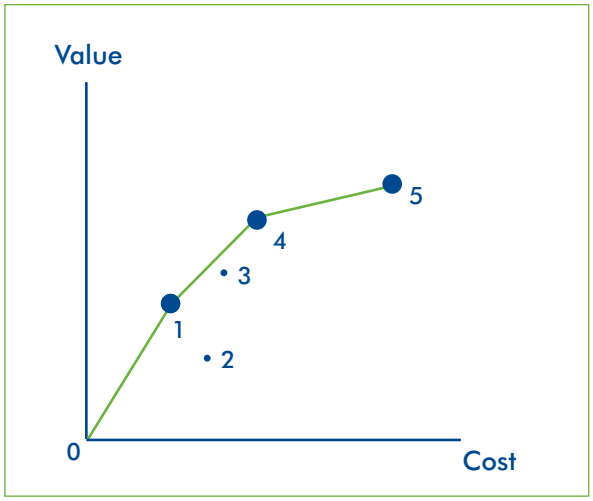
To compare costs and benefits in a particular therapeutic area the document suggests constructing a diagram with costs on the X-axis and 'value' on the Y-axis and then to plot the existing therapies in this therapeutic area as points on the graph (see Figure 1). By using arguments of dominance (intervention 2) or extended dominance (intervention 3) the most efficient interventions can be selected (in this case 1, 4 and 5) and these together form an efficiency frontier. The information value of this efficiency frontier graph depends on several factors:

- the extent to which the measure of value captures the overall benefit to the patient; indeed if this is not the case different efficiency frontiers may apply;
- the extent to which the measure of value is cardinal. If benefits are plotted against costs (in cardinal units) it is imperative that one can infer from the graph that for x more costs y more benefits can be produced (thus benefits on a cardinal scale). As the document itself points out many clinical measures used in trials do not have cardinal properties like functional scales of ADL (going from level 8 to level 9 may be less valuable than going from level 4 to level 5), but also intermediary clinical outcome measures, or the proportion of patients reaching certain targets, may lack these preferred properties;
- whether the information on value and costs for a specific intervention is up to date. This not only relates to the date of the study from which the data are derived but also whether the data are updated with information about costs and benefits in actual practice;
- whether all relevant interventions are depicted, even those for which no cost effectiveness information (with the chosen measure of value) is available in the literature;
- whether the positions on the 'efficiency frontier' really represent efficient decision making in the past. It is not at all clear what one can learn from

past decisions especially as these were made in a time when systematic consideration of efficiency was not common.

Each of these factors may cast serious doubts on the value of this frontier to the decision maker.

Figure 1: The efficiency frontier based on past decisions about five interventions



It is also very difficult to collect the data required for constructing the efficiency frontier for each therapeutic area. In quite a few situations new studies would be needed to provide the data for constructing this frontier, especially as old interventions have often not been assessed in terms of their costs and benefits. There may be some analogy here with the discussion about the WHO approach on generalized cost effectiveness, where it has proven difficult to determine the counterfactual of the null set of related interventions (Murray et al., 2000). Often one needs to reconstruct the past on the basis of incomplete data. Even if economic evaluations have been performed in the past, considerable adjustment and updating would be needed to make the plotted costs and values representative of the actual outcomes in current practice. In addition, foreign studies using cost per QALY as reported outcomes may be of little use. In sum, considerable effort would be needed to plot the efficiency frontier for each therapeutic area; indeed one may expect this effort to be much greater than would be required if IQWiG were allowed to use the same methodology as in other jurisdictions and hence be able to draw on cost effectiveness studies performed abroad.

A positive effect of considering the frontier is that it would allow the identification of inefficient strategies (i.e. those not on the frontier, such as interventions 2 and 3 in Figure 1) that may still be implemented in practice. But an active policy may be required to discourage such strategies, which is often lacking in most jurisdictions.

3.3. Using the efficiency frontier for decisions about ceiling prices

Once the frontier has been constructed the next important question is how to use the information for decision making. If one believes that the efficiency frontier represents the relative efficiency of interventions that exist, or have been considered in the past in the therapeutic area of interest, then clearly interventions with a value higher than intervention 5 (Figure 1) and with a cost lower than intervention 5 are acceptable at the prevailing price (that is, they dominate previous interventions). In the same way interventions with higher costs but lower value than intervention 5 are not acceptable (that is, they are dominated by intervention 5). Of course, most new interventions will be positioned North East from the position of intervention 5, providing additional value at higher costs. For this highly relevant area it is stated in the IQWiQ document (page 42) that there 'cannot be a firm decision rule for health technologies in this zone'. However, a number of options that may be considered by decision makers are outlined. One may use a rather strict rule by extrapolating the frontier from intervention 5 using the steepest slope of the efficiency frontier (0-1) and allowing only interventions above that position, or being more permissive by using the least steep slope (4-5). One may also use the average slope by extrapolating from the origin through point 5. How to use multiple frontiers in cases where there are multiple relevant benefits, which cannot be aggregated in some way, is even more problematic and is not even discussed.

On page 45 IQWiG suggests that the frontier at least provides some framework for decisions about ceiling prices and also states that there is 'no conceptual foundation for alternative approaches that do not directly project the efficiency frontier'. This statement, of course, is untenable as the obvious approach would be to use a common threshold for cost-effectiveness, as is done in so many jurisdictions. The use of a common threshold would, of course, also promote consistency of decision making across therapeutic areas. In this way we avoid the situation where a rather cost ineffective intervention may be approved considering the efficiency frontier, if it is lucky to be situated in a disease area where there exist only rather inefficient programmes (adding inefficiencies to inefficiencies).

Furthermore, imagine two therapeutic areas having the same relevant measure of outcome, e.g. increased survival. Because of historically differing frontiers it would be possible to pay more for increased survival in one disease area than the other. The common threshold may be very differently positioned in Figure 1 than any of the proposed extrapolations of the efficiency frontier.

In sum, much effort goes into the construction of the efficiency frontier but the directions on how to use that information provide little concrete guidance to the decision maker.

3.4. Other comments on the IQWiG methodology paper

Uncertainty
When decision makers are presented with efficiency data they also need to be informed about the reliability of such data. Recently Claxton (2007) suggested that decision makers may use lower (i.e. tougher) cost per QALY thresholds in cases where the uncertainty about the reported cost effectiveness ratio is larger. Methods for dealing with uncertainty are well addressed in the international health economic literature and established methods are available for constructing confidence intervals for cost effectiveness ratios. In one way it is surprising that this topic is not addressed in the methodology paper. But on the other hand it is not at all obvious how to present the efficiency frontier framework with due consideration of uncertainty.

First, it may be very difficult to reconstruct the uncertainty around the positions of interventions in the past. But even if this information were available, it would translate into an area of possible positions of maximum efficiency rather than an efficiency frontier. This would significantly add to the complexity of an already rather complex framework and further hamper proper inference from the analysis. If multiple efficiency frontiers have to be constructed because benefits in different dimensions need to be considered, the analysis becomes cumbersome.

Costing
Most of the comments on costing in the IQWiG methodology paper seem relatively straightforward. Rather curious, however, is the remark on page 49 is that 'a clearer categorisation (of costs), unfortunately not often used in economic evaluations, would be into 'insured', referring to those the payer covers and 'not-insured', referring to those borne by others regardless of what goods and services they are paying for.' Indeed in textbooks a clear distinction is made between the resources deployed for medical interventions (the costs) and the way of paying for these resources, emphasizing that only the former have to be determined in the context of an economic evaluation. As IQWiG also allows costs not covered by insurance (see IQWiG's section 3.1.2.), this alternative categorisation is not very useful in the context of IQWiG's own methodological recommendations.

Another comment is that the recommendation that productivity costs or benefits should not be treated as a cost or savings, but included on the benefit side (page 52). Though this is consistent with previous recommendations by the Washington Panel (Gold et al; 1996), it is not done in actual research practice and has several disadvantages (Brouwer et al., 1997a,b).

3.5. Overall assessment

By imposing the restriction to consider the efficiency of resource allocation only within a therapeutic area and not across therapeutic areas, IQWiG has manoeuvred itself into a difficult position. This restriction makes it impossible to conduct economic evaluations to international standards and only allows the presentation of information which is of limited value to the decision maker and gives little guidance on how to decide on the introduction and pricing of medical technologies. Furthermore, by not considering the relative efficiency of interventions across different therapeutic areas it runs the risk of allowing clearly inefficient technologies or rejecting clearly efficient technologies. Finally, constructing the efficiency frontiers for each therapeutic area will consume many resources, only a small part of which would be needed to conduct a standard economic analysis, especially as available information on cost effectiveness from studies abroad can be used.

3.6 Revised methods document

IQWiG produced a revised version of its methods document in October 2008 (Version 1.1). The new document provides some further clarification of why particular methods were chosen and includes a fuller critique of alternative approaches, especially those based on QALYs. However, the basic methods remain unchanged and hence the comments made here still apply. Nevertheless, individuals wishing to gain a full understanding of IQWiG's methods should consult the new document, along with the supporting papers on modelling, uncertainty and cost estimation (IQWiG, October 2008).

4. TOWARDS THE DEVELOPMENT OF INTERNATIONAL STANDARDS

The recent guidelines from NICE and IQWiG could not be more different, despite being issued within months of each other. Does this mean that the development of international standards for economic evaluation is a distant prospect?

Before undertaking a detailed analysis of the issues it is important to acknowledge that the IQWiG guidelines, despite being the most recent, represent an outlier. If one represented all the existing international guidelines on a spectrum, with the NICE guidelines at one end and the IQWiG guidelines at the other, most impartial observers would say that the weight of international opinion is in favour of NICE's methods.

Nevertheless, much can be learned from a more detailed examination of the issues. Namely, how is it that two jurisdictions facing the same problems can arrive at such different proposals at (roughly) the same time?

4.1. Defining the decision problem

Probably the greatest difference between the approaches adopted by NICE and IQWiG is that of whether the analyses are intended to inform decisions across all disease areas, or within a particular disease area. Ideally, the analysis should be able to assist the decision-maker in allocating the resources at his or her disposal. Therefore, the choice of approach should be based on institutional reasons, as opposed to methodological criteria.

Whereas it is possible to conceive of situations where, for example, a decision-maker may only have jurisdiction over (say) therapies for cancer, this is not the way most health ministries, or health insurers, are structured. Assuming that this is not the case in Germany, it is important that the Joint Federal Committee (G-BA) makes decisions that optimize the overall use of health care resources. It is thus surprising that IQWiG has followed the approach outlined. The approach followed by NICE would make more sense in the majority of health care systems, since decisions on the allocation of resources are based on a consideration of expenditures in all therapeutic areas.

Other key elements in the definition of the decision problem include the specification of the comparators and the viewpoint adopted. In developing international standards one must surely consider all relevant comparators in the setting concerned, as stipulated by NICE. However, it was mentioned earlier that some jurisdictions specify a narrower range (see Table 1).

It is difficult to assess the implications, on choice of comparators, of IQWiG's chosen approach. In principle, all relevant comparators can be included in the construction of the efficiency frontier, but in practice they may not. However, this would be a deficiency in execution of the method, rather than an inherent deficiency in the method itself. Of course, an important implication of the efficiency frontier method is that those treatment options that are not on the frontier should not be used at all.

In most jurisdictions, decisions tend to focus on the allocation of resources to new health technologies (the incremental approach), rather than addressing inefficiencies in the current treatment mix. The reasons for this are that it is difficult to challenge current practice, which is supported by so many professionals, or that information on the cost-effectiveness of the current treatment mix against the 'no treatment' option may not be available or hard to produce. IQWiG's recommendation to construct efficiency frontiers for separate disease categories may be potentially useful for gaining insight into the average efficiency of treatment programmes within a disease category. To consider the relative efficiency of existing programmes in one disease area against another, however, it would be necessary to use a common measure for benefit such as the QALY, which IQWiG does not support. It is somewhat ironic that this benefit of the efficiency frontier approach (which indeed may be the only one), is not thought to be relevant in the framework of IQWiG where allocation of resources across disease areas cannot be considered.

Both NICE and IQWiG restrict the range of costs to be considered, at least in the base case analysis. On the international level, there are few guidelines that advocate a true societal perspective, although some do allow consideration of productivity changes, or costs falling on the patient and family. In an international guideline for economic evaluation it would probably make sense to adopt a broad, societal perspective and then to allow decision-makers in particular jurisdictions to consider a narrower range of costs if desired.

4.2. The role of evidence-based medicine

Both the IQWiG and NICE guidelines recognize the importance of good quality clinical data in conducting economic evaluations. In addition, both agencies subscribe to the accepted hierarchy in the quality of clinical studies. For example, the highest quality evidence on relative treatment effect is obtained from systematic overviews of the relevant randomized clinical trials. Lower levels of evidence, in declining levels of quality are: (a) individual RCTs; (b) non-randomized comparisons; and (c) observational studies, such as registries and clinical case series.

Therefore, as an essential component of their evaluations, both IQWiG and NICE undertake systematic reviews of the available RCTs. However this is where the similarities end. Since IQWiG plans

to construct efficiency frontiers using the most important indicator(s) of clinical success in each clinical area, the estimates from the systematic review are directly useable.

On the other hand, since NICE requires the health gain to be estimated in QALYs, the systematic review becomes just one input, albeit an essential one, to the evidence considered. For example, observational studies may be required to help convert an intermediate outcome (e.g. change in HbA1c or LDL cholesterol) to a final outcome (e.g. change in coronary heart disease events or survival). In addition, estimation of the QALYs gained will require information on the effect the therapy has on the patient's quality of life.

In some clinical areas it is possible that much of the relevant data for assessment of QALYs can be obtained directly from the systematic review (e.g. trials of drugs to treat various forms of cancer, that often measure both survival and quality of life). However, in the majority of clinical areas this may not be the case and the measures summarized in the systematic review may be insufficient, either because they are intermediate, or because they relate only to one component of the clinical value of the new drug.

One example of this comes from IQWiG's evaluation of the rapid-acting insulin analogues (lispro, glulisine) in the treatment of diabetes. A systematic review identified 1,017 studies, of which only nine met the criteria for inclusion. IQWiG decided to focus on HbA1c since this is a good predictor of events relating to diabetes, is unambiguously measured and is included in all studies. The conclusion was that the newer products were equivalent to standard insulin in their control of Hba1c over the period studied in the trials.

However, the main benefits of the newer products, and hence the justification for any price premium, lie in their reduction of hyperglycaemic episodes. These can be troubling to the patient and, in some cases may lead to some costs due to hospital emergency room visits. For this reason, the newer products are recommended for a subset of the patient population in some countries (e.g. by NICE in the UK). On the other hand, based on the IQWiG review, the Joint Federal Committee clustered the insulin analogues with standard insulin, at a low reference price level. The consequence of this would normally be that patients requiring the products would face a high co-payment unless the manufacturers reduced their price to a level close to the reference price. In this case, the outcome was that the manufacturers reduced prices to the health insurers (sickness funds) so that they would be willing to make insulin analogue available.

It should be noted that the IQWiG methods document does acknowledge that, if the relevant clinical trials contained the appropriate data, the clinical outcome measure could be in QALYs. However, the widespread inclusion in clinical trials of measures like the EQ-5D is some way off. In addition, this does not deal with the problems of extrapolation of benefits beyond the period of the clinical trial.

Therefore, NICE is committed to modelling, using outcome data from RCTs, observational studies and free-standing studies on quality of life. Thus, in respect of the role of evidence-based medicine, IQWiG and NICE are at opposite ends of the spectrum from, on the one hand, being considered the only relevant approach to, on the other hand, being a necessary but insufficient input to the decision.

In developing international guidelines, it is hard to imagine an approach relying only on data from systematic reviews of RCTs. Indeed, most of the existing national guidelines for economic evaluation allow some modelling, although few foresee a broad role. This is because the decision-makers in most jurisdictions have residual concerns about economic modelling, because of the assumptions involved. For this reason, initiatives to improve the methodological standards of models are essential to the growth of this approach (Weinstein et al., 2003; Phillips et al., 2004).

4.3 Assessment of therapeutic and social value

NICE favours QALYs as a measure of health gain, whereas IQWiG prefers to use measures of therapeutic value relating to each individual clinical area. The advantages and disadvantages of these two approaches are self-evident. The QALY, being a generic measure of health gain, enables us to make comparisons of interventions across the whole of health care. On the other hand it requires a number of assumptions.

The biggest concern about IQWiG's approach is whether it can even help us to interpret therapeutic value within a given clinical field. It probably works best in fields where there is one unambiguously superior measure of outcome. If this is not the case it may be necessary to construct multiple efficiency frontiers and to make trade-offs among them. The methodology for this still needs to be developed, but is likely to be complex. The most important limitation is likely to be that systematic reviews may not be available to produce summary estimates of a range of clinical outcomes. (However, this may impede the calculation of QALYs also. See Drummond et al,. 2008b).

The methodological problems with QALYs have been widely discussed and will not be repeated here (Drummond et al., 2005, Chapter 6). However, it is clear that QALY estimates depend on the methodology used to obtain them. Even the various generic instruments (e.g. EQ-5D, HUI, SF-6D) generate different estimates of the QALYs gained. Also, it is known that many of the axioms upon which QALYs are based (including that the utility of a health state is independent of the time spent in the state) do not hold.

Therefore, the QALY is clearly a measure of convenience rather than a measure of choice. That is, it provides a reasonably informative, weighted index of several of the key components of health gain, that many decision-makers consider a pragmatic solution to the valuation problem. However, it is not, in any sense, a 'true' measure of health gain, since important value judgements are involved in defining the relevant dimensions of health and in making the trade-offs among them.

One obvious recommendation for international standards in economic evaluation is that the important consequences of the alternative therapies should be identified and measured in a cost-consequences analysis. This step, which is often overlooked, would ensure that all the relevant dimensions of clinical outcome are addressed and would add transparency to the calculation of QALYs.

As mentioned above, the assumption is that QALYs gained are applicable across all areas of health care and, under the standard methods, they can be simply added. That is, 'a QALY is a QALY is a QALY' no matter to whom it accrues. Economists have had concerns about the simple addition of QALYs for some time (e.g. Williams, 2001; Nord, 2001). However, more recently, a growing body of literature has suggested that the community might not be indifferent to whether a QALY is given to someone in a very poor health state (as compared with someone who is in near full-health), or to someone who is elderly (as compared with someone who is young). These factors have led the authorities in the UK to commission more research into how QALYs are valued by the general public (Dolan et al., 2008; Donaldson et al., 2008).

At the present time, the international standard would probably suggest that a cost-consequences analysis be performed and the health gain (in QALYs) estimated. Then decision-makers may introduce other factors alongside the evidence of cost-effectiveness. These could include items such as the seriousness of the health condition, the availability of alternative therapies, the need for equity of access to care, and so on.

4.4 Assessment of value for money

Once the estimates of cost and benefit have been made, the decision-maker needs to make an assessment of value for money. In the NICE process, this is called 'appraisal' (that is, the application of decision-making criteria), as opposed to 'assessment' (that is, the scientific analysis of evidence).

Under the NICE approach, where the evidence is summarized as an estimate of the incremental cost per QALY of the new technology compared to the old, the debate centres on the threshold willingness-to-pay. That is, what level of incremental cost per QALY is deemed reasonable? As Rawlins and Culyer (2004) point out, this is a societal value judgement to which there is no easy answer. At best there may be a relevant range of 'acceptable' cost per QALY, within which it is worth debating whether or not the new technology should be adopted.

The notion of a threshold value of the incremental cost-effectiveness ratio has been extensively debated (Birch and Gafni, 2003; 2007). Those involved in NICE's decision-making describe this as a deliberative decision-making process (Culyer et al., 2007), where the decision-maker is 'searching for the threshold'. Certainly, it is envisaged that many of the factors referred to above (e.g. equity of access to health care) are important inputs to the decision.

As mentioned previously, some research has been commissioned in the UK (Donaldson et al., 2008; Dolan et al., 2008) to answer questions such as: (i) what is the maximum amount members of the community are willing to pay for a QALY; and (ii) does the community's willingness to pay for a QALY differ by factors such as the age of the recipient or the severity of their condition. However, a recent editorial has likened this to a 'search for the Holy Grail' (Brouwer et al., 2008), in that there is not, and never will be, a simple answer to these questions.

In contrast, IQWiG's approach is to use prior funding decisions, as expressed by points on the efficiency frontier, as possible upper and lower bounds on the willingness to pay (i.e. ceiling price) for the new drug. To be fair, the IQWiG document does not lay down any firm rules, stating that this is all up for debate. However, it is necessary to give some guidance to the decision-maker, beyond the suggestion that any price premium for the new drug should be between zero and infinity!

Therefore, suggestions are made to the extent that the price premium should be limited to the level of the previous one (i.e. extending the last slope on the efficiency frontier), or limited to the level of that paid for the greatest advance seen in the given clinical

area (i.e. the steepest slope on the existing frontier), or some kind of average of past payment for results. Of course, all this assumes that there is one efficiency frontier for each clinical area; if there were multiple frontiers, the decision-making criteria would be even more complex.

In judging the two contrasting approaches against international standards, it is clear that there are a multitude of approaches. The QALY appears to be most accepted in Belgium, Canada, Hungary, The Netherlands, Sweden and the UK). We are not aware of any other jurisdiction adopting IQWiG's approach, although the Pharmaceutical Benefits Advisory Committee (PBAC) in Australia comes close by insisting on a within-trial analysis, using the clinical outcome of choice, prior to any modelling to estimate life-years or QALYs gained.

Therefore, it would not be appropriate to argue that NICE's approach is right and IQWiG's approach is wrong (or vice versa). However, the two approaches can be viewed from a practical perspective. At least the NICE approach provides a fairly comprehensive framework for the decision-maker to appreciate, and to debate, the main issues. On the other hand, it is not clear how IQWiG's approach helps the decision-maker assess the real value of a new technology. Key elements of value may not be considered (because they are not included in systematic reviews) and there is no consistency across different clinical areas.

5. CONCLUSIONS

So what would international standards for economic evaluation look like, given our current state of knowledge? Our understanding of these is set out in Table 3.

One important point is that, given the uncertainties about some elements of the methods and the fact that these incorporate both technical and value judgements, it is difficult to divorce the methods themselves from the decision-making process that accompanies them. The NICE guidelines make some comments about this (NICE, 2008, Section 6). Namely, the manufacturer and other stakeholders are important participants in the process and have the opportunity to comment on the scoping of the analysis, the interim assessment and the final appraisal. They then have the opportunity to appeal if they are dissatisfied with NICE's final guidance.

In addition, the NICE guidelines make it clear that the data on cost effectiveness will be considered within a deliberative decision-making process, in which other factors will be taken into account. Also, it is clear that there is a strong emphasis on transparency, in both the process and the decision.

The IQWiG guidelines say nothing about decision-making processes, although these considerations may feature in other documentation from the Institute. Further specification of these processes would be a useful contribution, since international standards for economic evaluation need to consider processes as well as methods.

Table 3: Possible International Reference Case

• Study perspective	Health/social care, other government budgets, family costs and productivity costs
• Comparators	All relevant comparators
• Source of effectiveness data	Synthesis of data from trials and observational studies
• Role of modelling	Essential
• Main economic outcome	QALYs or DALYs*
• Source of utilities	Generic measure
• Characterising uncertainty	One-way sensitivity analysis and summary approach (e.g. PSA)

Note: * Disability-adjusted life years

6. REFERENCES

Birch S, Gafni A. NICE methodological guidelines and decision making in the National Health Service in England and Wales. Pharmacoeconomics 2003;21:149-57.

Birch S, Gafni A. Economists' dream or nightmare? Maximising health gains from available resources using the NICE guidelines. Health Economics Politics and Law (2007) 2:193-202.

Brouwer WBF, Koopmanschap MA, Rutten FFH. Productivity costs measurement through quality of life? A response to the recommendation of the Washington Panel. Health Economics 6 (1997a), 253-259

Brouwer WBF, Koopmanschap MA, Rutten FFH. Productivity costs in cost-effectiveness analysis: numerator or denominator: a further discussion. Health Economics 6 (1997b), 511-514

Brouwer W, van Exel J, Baker R, Donaldson C. The new myth: the social value of a QALY. Pharmacoeconomics (2008) 26: 1-4

Buxton MJ. Economic evaluation and decision making in the UK. Pharmacoeconomics 2006;24(11): 1133-1142

Caro JJ, Getsios D, Moller J. Regarding probabilistic analysis and computationally expensive models; necessary and required? Value in Health 2007;10(4):317-8.

Claxton K. OFT, VBP: QED? Health Economics 16 (2007), 545-558

Claxton K, Sculpher M, Cuyler A, McCabe C, Briggs A, Akehurst R, Buxton M, Brazier J. Discounting and cost-effectiveness in NICE-stepping back to sort out a confusion. Health Economics 15 (2006), 1-4

Culyer AJ, McCabe C, Claxton K, et al. Searching for a threshold, not setting one: the role of the National Institute for Health and Clinical Excellence. Journal of health Services Research and Policy (2007) 12:56-58

Dolan P, Tsuchiya A, Brazier S et al. The relative social value of health gains to different beneficiaries. NIHR Methodology Panel Project. URL http://www.hta.ac.uk/project/1577.asp [Accessed 15 May 2008]

Donaldson C et al. What is the value to society of a QALY? NIHR Methodology Panel Project . URL http://www.hta.ac.uk/project/1578.asp [Accessed 15 May 2008]

Drummond MF, Ferguson BA, Weatherly HL. Is it time to broaden the perspective for economic evaluation? BMJ In press (2008a)

Drummond MF, Iglesias CP, Cooper NJ. Systematic reviews and economic evaluations conducted for the National Institute for Health and Clinical Excellence in the United Kingdom: A game of two halves? International Journal of Technology Assessment in Health Care 24 (2008b), 1-6

Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes. Oxford, Oxford University Press, 2005

Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-effectiveness in Health and Medicine. Oxford University Press, New York, 1996

Gravelle H, Brouwer W, Niessen L, Postma M, Rutten F. Discounting in economic evaluations: stepping forwards towards optimal decision rules. Health Economics 16 (2007), 307-317

Griffin S, Claxton K, Hawkins N, Sculpher M. Probabilistic analysis and computationally expensive models: necessary and required. Value in Health 2006;9(4): 244-252

Griffin S, Claxton K, Hawkins N, Sculpher M. Response to Caro et al. Value in Health 2007;10(4): 319

Health Economics Committee of the Social Policy Association (Ausschuss für Gesundheitsökonomie im Verein für Socialpolitik) (2008) available from: http://www.uni-bielefeld.de/gesundhw/ag5/aktuelles/IQWiG-Stellungnahme-Ausschuss.pdf [Accessed 6 October 2008]

Hjelmgren J, Berggren F, Andersson F. Health economic guidelines: similarities, differences and some implications. Value in Health (2001) 4: 143-57

IQWiG. Methods for assessment of the relation of benefits to costs in the German statutory health care system: version 1.0 (for consultation). Cologne, 24 January 2008

IQWiG. Methods for assessment of the relation of benefits to costs in the German statutory health care system: version 1.1. Cologne, 9 October 2008

Jönsson B. IQWiG: an opportunity lost? European Journal of Health Economics 9 (2008), 205-207

Murray CJL, Evans DB, Acharya A, Baltussen RMPM. Development of WHO guidelines on generalized cost-effectiveness analysis. Health Economics 9 (2000), 235-251

National Institute for Clinical Excellence. Guide to the methods of technology appraisal. London, NICE, April 2004.

National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal. London, NICE, June 2008

Nord E. The desirability of a condition versus the well being and worth of a person. Health Economics 2001; 10: 579-81

Philips Z, Ginnelly L, Sculpher Philips Z, Ginnelly L, Sculpher M, Claxton K et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. Health Technology Assessment (2004) 8(36): 1-172
Rawlins MD, Culyer AJ. National Institute for Clinical Excellence and its value judgements. BMJ (2004) 329: 224-7
Sculpher MJ, Drummond MF. Analysis sans frontieres: can we ever make economic evaluations generalisable across jurisdictions? Pharmacoeconomics (2006) 24(11): 1087-99

Tarn TY, Smith MD. Pharmacoeconomic guidelines around the world. ISPOR Connections (2004) 10(4): 5-12 [online]. Available from URL: http://www.ispor.org/news/index_new.asp [Accessed 10 October 2006]

Von der Schulenburg J-M Graf, Vauth C, Mittendorf T, Greiner W. Methods for determining cost-benefit ratios for pharmaceuticals in Germany. European Journal of Helath Economics 8 (2007), suppl 1: S5-S31

Weinstein MC, O'Brien BJ, Hornberger J et al. Principles of good practice for decision-analytic modelling in health-care evaluation: Report of the ISPOR Task Force on Good Research Practices – Modeling studies. Value in Health (2003) 6: 9-17

Williams AH. The 'fair innings' argument deserves a fairer hearing! Comments by Alan Williams on Nord and Johannesson. Health Economics 2001; 10: 583-5

As with all OHE publications, this briefing was peer reviewed by its Editorial Board and by other experts in the field and is intended to be a contribution to research and to public policy making. It does not represent the views of the OHE.

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**.

Office of Health Economics
12 Whitehall London SW1A 2DY
Telephone: +44 (0)20 7747 8850
Facsimile: +44 (0)20 7747 8851
www.ohe.org