

## Deliberative processes in decisions about health care technologies: combining different types of evidence, values, algorithms and people\*

**Professor Anthony J Culyer** *Professor of Economics, Department of Economics & Related Studies, University of York, Heslington, York YO10 5DD, UK. Ontario Research Chair in Health Policy & System Design, Department of Health Policy Management and Evaluation, 155 College Street, Toronto, M5T 3M6, Canada*

### Contents

1. Foreword	1
2. Introduction	1
3. Deliberative Processes and 'Evidence'	4
4. Deliberation Compared with Consulting or Commenting	7
5. Evidence Regarding Deliberative Processes	8
6. Under What Circumstances are Deliberative Processes Likely to be of Greatest Use?	9
7. The Effectiveness of Deliberation	10
8. NICE and Deliberative Processes	12
9. Comparators	13
10. How Does NICE do?	13
References	15
Recent OHE Publications	17
Compendium	18

there are a variety of 'structured' deliberative processes that may combine facilitation of some form with the use of decision weights generated both by the deliberative process and from outside.

This briefing follows the OHE lunchtime seminar held in September 2008, in which Professor Tony Culyer presented a series of conjectures about the circumstances under which deliberative processes are likely to be useful, and the characteristics that deliberative processes ought to possess in order for their successful application to health care decision making. A lively discussion was then introduced by Professor Jack Dowie, who put forward the view that the best collective decisions may be the product not of consensus and compromise, but of disagreement and contest. Jack also argued for the use of one form of Multi-Criteria Decision Analysis which used web-based software as a means of eliciting and aggregating inputs in an impersonal and efficient manner. A short comment by Jack, which mirrors his contribution to the seminar and provides the basis for much of Tony's discussion of algorithms, group polarisation and 'value of analysis analysis', is available online at [www.ohe.org](http://www.ohe.org).

## 1. FOREWORD

In the UK and elsewhere, choices about how to allocate limited health care resources are guided not only by cost-per-QALY calculations, but also by social value judgements such as those relating to equity and fairness. Factors such as uncertainty around key values, operational feasibility and stakeholder interests also need to be considered. The question of how one ought to combine these various inputs is central to the field of health care decision making. There is a continuum of approaches to answering this question. At one end sits the algorithm and formulaic methods of combining information, albeit with some attribution of a score usually required. At the other end sits a 'pure' deliberative process. In between

## 2. INTRODUCTION

*Parliament is not a congress of ambassadors for different and hostile interests ... but ... a deliberative assembly ... with one interest, that of the whole; where, not local purposes, not local prejudices ought to guide, but the general good...*

**Edmund Burke**

quoted in Kurland and Lerner (1987) pp. 391-2

This is as good a statement as can probably be found on the general political case for deliberation. It is notable for its emphasis on deliberation, or on at least

\*I am grateful for the comments of Richard Barker, Martin Buxton, Roger Chafe, Jack Dowie, Chris McCabe, Dave Parkin, Katherine Payne, Larry Phillips, James Raftery and Adrian Towse and participants at a seminar held at the OHE in November 2008. I have not followed the advice of every one of them but have nonetheless benefitted from it in every case.

the kind of discourse in a deliberative assembly of the parliamentary sort, as a means of suppressing the arbitrary and subjective self-interest of its participants – a practical man’s practical equivalent of a Rawlsian theoretical person’s theoretical notion of ‘the veil of ignorance’ (Rawls 1971) – a means of achieving a disinterested state of mind. Burke sees the deliberative assembly as a helpful way by which people of goodwill might be able to restrain their more selfish concerns in pursuit of a wider, or deeper, idea of the social good – one that is not simply the sum of the preferences and prejudices (admirable or not, well-informed or not, representative or not, based on mature reflection or not) of those participating in the debate.

More philosophically, another possible (Kantian) progenitor might be Habermas (e.g. 1990) who initiated the idea of ‘discourse ethics’ through which both factual and normative argument and communication seems best done through the deliberation of rational agents: through discourse the ‘unforced force’ of the better argument prevails.

For some that might be enough. There are doubtless those for whom deliberation and all that it implies is inherently a good thing. Some would seem to think of it as a veritable cornucopia of political good things. For example, Tali Mendelberg has written (citing numerous political authorities as he goes along):

*If it is appropriately empathic, egalitarian, open-minded, and reason-centered, deliberation is expected to produce a variety of positive democratic outcomes .... Citizens will become more engaged and active in civic affairs .... Tolerance for opposing points of view will increase .... Citizens will improve their understanding of their own preferences and be able to justify those preferences with better arguments ... . People in conflict will set aside their adversarial, win-lose approach and understand that their fate is linked with the fate of the other, that although their social identities conflict they "are tied to each other in a common recognition of their interdependence"... . Faith in the democratic process will be enhanced as people who deliberate become empowered and feel that their government truly is "of the people" ... . Political decisions will become more considered and informed by relevant reasons and evidence ... . The community's social capital will increase as people bring deliberation to their civic activities ... . The legitimacy of the constitutional order will grow because people have a say in and an understanding of that order ....*

**Mendelberg** (2002) p. 153

It is notable that just about all of these good things flow from the process *itself* rather than its outcome. That being the case, and assuming the consequences

claimed being also empirically the case, then it would seem to follow that there are consequences here that would be lost if one were to substitute for a deliberative process some other more hierarchical, or more bureaucratic, process – or a more algorithmic process, as some analysts of ‘value of analysis analysis’ have suggested (e.g. Dowie n.d. a, b). It is, of course, an empirical question as to whether deliberation as a process really has these advantages (and whatever disadvantages it must doubtless also have) – and perhaps it is a question that is worth trying to answer seriously rather than relying on mere assertion, though my impression is that mere assertion (and wishful thinking) are for the most part all that we have. However, that is not the focus of this paper and it is notable that advocates of undoubtedly useful methods, like Multicriteria Decision Analysis (MCDA), have tended to move away from seeing them as stand-alone algorithms in favour of their being decision and process *aids*:

*The advantage of ... modern MCDA in general, is not limited ... to its technical ability to aiding in the construction of single value functions and the weighting of criteria. It should be seen in a much broader socio-technical scope. MCDA has evolved from a mechanism to rank alternatives, to a structured approach to organise factual and judgmental information and to argue logically about the pros-and-cons of the alternative courses of action. The ultimate goal...is therefore to facilitate the process of learning about the problem and the alternatives, by enabling people to think about their values and preferences from several points of view.*

**Bana e Costa et al.** (2004)

I shall record no curmudgeonly dissent regarding the desirability of any of these promised Mendelbergian things, nor from the perhaps optimistic prediction that they are all the fruits of deliberative processes. Nor am I concerned, being myself a confirmed consequentialist in most things, with the highly consequentialist flavour of this case for deliberation. The essence of my point is, instead, the conjecture that a rather strong case can be made for deliberative processes on grounds that *they lead to better decisions*, regardless of whatever other desirable political outcomes may be their consequences. That raises a number of questions like: ‘what is a better decision?’ and ‘how would you recognise one if you saw one?’ I shall try to address these questions later.

### **BOX 1 CONSEQUENTIALISM**

A philosophical doctrine according to which the moral rightness of thoughts and deeds is to be judged in terms of the moral rightness of their probable or necessary consequences.

## BOX 2 ALGORITHM

A mathematical procedure or formula for solving a problem in a sequential fashion, with each step depending on the outcome of the previous one. Named, in corrupted Latin form, after the great scholar Mohammed ibn-Musa al-Khwarizmi (780-850) who was born in the ancient Persian empire (in the territory now called Uzbekistan) and spent most of his working life in Baghdad, then arguably the greatest intellectual centre in the world.

In stark contrast to the deliberative process stands the algorithm, to which reference has already been made. An algorithm is a systematic mathematical process sequentially linking various strands in a decision problem to an outcome. A good example of an algorithm for present purposes is the EQ-5D version of the QALY, the quality-adjusted life-year, which combines a set of pre-defined characteristics of good health, measurable at a variety of intensities, and weighted in a pre-set fashion in order to measure a health outcome such as the difference between a patient's health when taking a particular medicine and their health when not taking it or when taking a different medicine. You can make the algorithm as

## BOX 3 ALGORITHMS VERSUS DELIBERATIVE PROCESSES IN NICE

### Algorithm

NICE recommends the use of Quality-Adjusted Life-Years (QALYs) as the main outcome measure in the economic appraisals that are presented to its multi-disciplinary and multi-professional appraisals committee. The particular form of QALY recommended is the EQ-5D, which is an algorithm embodying various health state characteristics (five) measured on a three-point scale and added together. The guidance explicitly states *an additional QALY should receive the same weight regardless of the other characteristics of the individuals receiving the health benefit.* (NICE 2008 para 5.12.1)

### Deliberation

NICE has also referred an aspect of this algorithm to its Citizens Council, a form of consensus group to engage in a deliberative approach regarding the weighting (if any) to be given to older people. Their recommendations to the NICE board included, somewhat ambiguously, *Overall, the majority of us on the Citizens Council [22 of us] felt very strongly that no judgement should be made about being more generous to certain age groups because of the social roles those age groups tend to fulfil.* (NICE 2004, p.16)

complicated as you like – at least in principle – by adding characteristics, refining intensities, changing the weights, including probabilities and uncertainty, discounting future health changes, and so on, and every element of the algorithm can even be moderated by the results of consultative engagement with patients, say for their values, and doctors, say for their beliefs about the transitional probabilities. The process remains, however, mechanical, unidirectional and, if used without interaction between decision makers, not conducive to learning. Rather than enabling the exercise of judgment about the merits and interpretation of evidence it conceals extensive conclusions that have already been reached. These may (as with EQ-5D) have been reached in earlier and deliberative stages of preparation for a decision to be reached by deliberation but the nature of dispute resolution, the character of value judgments, the extent of agreement about them, the adequacy of the information base available, and so on, all become subsumed in the algorithmic solution. (See Box 3 for the way NICE uses the EQ-5D algorithm.) The use of algorithms is thus likely to be perceived as impenetrable to those not involved in the decision making process but who may nonetheless have significant stakes in its outcome. Even when an algorithm has had deliberation embodied in some stages of its construction it is not itself deliberative nor is it necessarily designed to be particularly useful in deliberative processes. By being possibly deliberative in its construction, but not a self-evidently satisfactory adequate substitute for a deliberative process in the making of actual decisions, any algorithm lacks particular virtues which we shall explore and its effective use requires there to be sufficient expertise within the decision group for its members as a whole to have confidence that no unacceptable short cuts have been taken. The same may be said about the use of computerised models to simulate decision making processes. In their review of facilitation in theory and practice, Phillips and Phillips (1993) concluded:

*... while computers are good at storing, retrieving, manipulating and communicating information, they cannot exercise judgement. The facilitator and members of the group must perform that function: formulating problems, identifying key issues, considering risk and uncertainty about the future, forming preferences, making judgements of subjective value, establishing goals and objectives, and assessing trade-offs among objectives.* (p. 548)

I should add that I am not against computers. Nor am I against algorithms. Indeed I once helped to invent a precursor of the QALY (Culyer et al. 1971). But I am suspicious of algorithms that substitute for an interactive deliberative process and I am suspicious of

them used alone and without a decision making context. I have always followed my mentor, Alan Williams, in believing that the role of an algorithm is to be an aid to thought, not a substitute for it:

*I take the objective of CBA [another, rather complicated, algorithm] to be to assist choice (not to make choice, nor to justify past choice, nor yet to delay matters so that some previously chosen course of action has a greater chance of adoption ...*

**Williams** (1972) p. 201 (square brackets added)

Similar thoughts have been expressed by most of the major proponents of cost-effectiveness analysis as well (for example, Weinstein and Stason 1977).

A good way of ‘assisting choice’ is, however, almost certainly likely to embody algorithms in deliberative processes – or at least to embody those likely to ‘assist choice’. Again this is an empirical matter that would be worth testing. The interesting story of how Lawrence Phillips came to see the crucial interplay between decision analytical methods that use algorithms and group processes, which led to his concept of the ‘decision conference’, is told in Phillips (2007).

Simplicity in the elements of a decision or in the decision rule is not a sufficient ground for supposing that deliberation is unnecessary, just as complexity is not a sufficient ground for supposing that it is necessary. For example, the Committee for the Evaluation of Drugs in Ontario has one simple and rather strictly-applied decision rule as its criterion for recommending drugs for the publicly funded Medicare system there. It uses a threshold incremental cost-effectiveness ratio of \$50k Canadian per QALY. It is an example of a single and simple criterion – but underneath it are lots of complexities that might well warrant the use of deliberative processes to assess claims and make judgments about them. So a *simple* rule may still warrant a deliberative process. Conversely, it is easy to think of occasions where complexity could be directly addressed through deliberation. For example, the use of multi-attribute decision analysis (MADA) could be enhanced by having it placed in the context of a deliberative process rather than substituted for it, as when the MADA approach is itself to be examined as a part of the decision process, or when decision makers need to challenge one another’s beliefs about the reliability of the empirical data they have been given and learn through discussion, or when they need to explore with systematic reviewers precisely how bias may have crept into the data extraction process, or when transparency to non-participating stakeholders is deemed to be critical to success.

But this is to jump ahead. Let me end this introduction with a caution. *I am not advocating the indiscriminate use of deliberative processes. They are costly and may not only be inappropriate but, even if they were thought to be appropriate, they may not be worth their cost. Indeed, I am not advocating anything in this paper apart from good decision making and, as we shall see, a more empirical approach to finding out ‘what works’ from amongst various possible decision processes.* I hope to show that there can be a reasonable expectation that deliberation *can* in principle enhance the quality of decisions and that it is possible to make coherent conjectures about the circumstances when this is likely to be so. But ultimately these are empirical matters and here, alas, as is so common in all realms of public policy, the current ratio of evidence to speculation, assertion and ‘guesstimation’ is woefully low.

---

### 3. DELIBERATIVE PROCESSES AND ‘EVIDENCE’

I am strongly in favour of having relevant and timely evidence to inform decisions. Deliberation is nearly always required in considering ‘evidence’. One thing about which one might deliberate indeed concerns what it is that is to be regarded as evidence. Evidence can be classified in a variety of ways, as illustrated in a recent systematic review (Lomas et al. 2005) and summarised in Box 4. Three general approaches emerged from that study. One group of authors focused on categories according to the *method of collection used* for the evidence, for example, whether it was experimental or survey. Another focused on the *general purpose* to which the evidence would contribute, such as identifying a problem or measuring the effectiveness of an intervention. The third emphasised *source*, usually distinguishing research by professional researchers from unsystematic forms of evidence such as ‘clinical experience’. Given the potentially diverse elements of the practical contexts in which evidence is commonly used, it is not surprising that a multiplicity of methods might be needed to create scientific evidence for even a few of them. As one source noted after describing ethical, judicial, expert, practical, and theoretical ‘dimensions’ to providing guidance, *these dimensions tend to operate within different frameworks that seek to answer different questions in different ways, based on different evaluative criteria* (Buetow and Kenealy 2000).

It is not inherent in the concept that deliberation be face-to-face, for the possibilities of effective interaction through, for example, electronic conferencing like chat rooms or blogs are very numerous. Their relative effectiveness as ways of making good decisions or even as ways of effective communication, however, has yet, so far as I know, to be tested.

#### **BOX 4 CATEGORIES OF EVIDENCE**

##### **Defined by method of collection, discipline or theoretical framework:**

- observational, experimental, quasi-experimental, extrapolated, survey, experiential, administrative
- quantitative, qualitative, economic, ethical/philosophical
- narrative review, systematic review, meta-analysis
- legal, epidemiological, clinical
- clinical epidemiology, decision science, expected utility theory

##### **Defined by general purpose:**

- problem identification, description or scoping
- cost-containment, efficacy, effectiveness, cost-effectiveness, implementability
- cultural, leadership, measurement
- philosophical-normative, practical-operational
- academically driven by discipline (clinical, biostatistics, economics, sociology,...)

##### **Defined by source:**

- primary research data, secondary data (meta analyses etc.) administrative data,
- clinical experience
- patient/carer experience
- political necessity
- local managerial experience
- professional (scientific, theoretical, practical, expert, judicial, ethical)

A deliberative process entails the careful, deliberate consideration and discussion of the advantages and disadvantages of various options and an important element in this 'consideration and discussion' is the weighing up of the 'evidence'. Deliberative processes are mechanisms for *both* eliciting and combining evidence:

*a more fundamental means by which the public can influence the generation of data and the derivation of the policy options as well as discussing acceptable decisions, thus, taking account of public as well as expert knowledge.*

Petts (2004) p. 115

Petts thus implies that there are (at least) two forms of knowledge, one possessed by the experts and the other by the public; the evidence provided by the latter, and often by practising professionals too, is often generated through the deliberative process itself, whereas 'expert' scientific evidence is generally gleaned from professional reports in learned journals.

When people in the clinical, management, or policy world are asked what they consider to be 'evidence', they tend to think of a medley of scientifically verifiable and locally idiosyncratic types of information – what Lomas et al. call 'colloquial' interpretations – drawing on a wide range of experiences and using a broad definition of 'evidence'. Thus, clinical effectiveness data compete with expert assertion, cost-benefit calculations are balanced against political acceptability, and public or patient attitude data are combined with the recollection of recent personal encounters with strong personalities. The evidence-based decision making 'movement' has, however, engendered for many of them a greater regard for the more scientific forms of evidence than would have been usual twenty years ago and there is an increasing tendency to 'dress up' the conclusions of a decision making process in the language of science (Jenkins and Barber 2004).

The research community's view of evidence, both in clinical subjects and the social sciences, tends to be restricted to information generated through a prescribed set of processes and procedures recognised as scientific (e.g. Scott-Findlay and Pollock 2004, Atkins et al. 2005, Whitehead et al. 2004, Norheim 2002). In this case both scientific tradition and more modern infusions from the philosophy of science determine what is evidence, which can be summarised as knowledge that is explicit (that is, codified and propositional); systematic (that is, uses transparent and explicit methods) and replicable (that is, it can be tested whether others following the same methods with the same samples will arrive at the same results).

Hence, when evidence is defined colloquially, its inclusion is determined through tests of local and professional relevance. When it is defined scientifically, its legitimacy for inclusion as scientific evidence is determined by the methodology used to generate it.

Jonathan Lomas with myself and other colleagues (2005) suggested further that it is helpful to think of three forms in which evidence may come. At a basic level, the general notion of evidence concerns actual or asserted facts (a fact is defined as a 'thing certainly known to have occurred or be true' in the OED) intended for use in support of a conclusion. As we have seen, most decision makers view evidence colloquially and eclectically, as anything that increases their degree of belief in a 'fact'. They define it by its resonance with experience and relevance to the kinds of decisions they have to make. This is the first form: colloquial evidence. The second and third forms are provided by scientists. Scientists' views on the role of evidence divide into

those who emphasise context-free universal truths (identified closely with ‘evidence-based medicine’) and those who emphasise a context-sensitive role for evidence in a particular decision process (identified more with the applied social sciences). The appropriate methods for obtaining scientific evidence about context factors are not the same as those for obtaining evidence related to the testing for the validity of bioscientific hypotheses but this makes such evidence no less ‘scientific’ from a methodological perspective even though the research designs may be very different. Hypothesis testing is common to both, as is the control of ‘confounding’ variables. But both the phenomena hypothesised about and the method required to do the testing differ. Thus, whereas the gold standard procedure for controlling for confounding variables in clinical sciences might be a form of prospective randomised trial, the gold standard in assessing the resource consequences over time is more likely to be a retrospective multivariate econometric study. Scientific evidence on context must be more than merely medical and can embrace attitudes, implementation, organisational capacity, forecasting, economics/finance and ethics. Not all will always be relevant but some will always be relevant (given the context). Colloquial evidence will typically embrace the resources likely to be available, expert and professional opinion on a matter, political judgment, values, habits and traditions, lobbyists and pressure groups, and the particular pragmatics and contingencies of a situation. In health care decisions, all three kinds of evidence are more or less constantly in play.

*In practice, the range of ideas about what constitutes ‘evidence’ is likely to be very wide. For ... improvements in public health decision-making to occur, the systematic evaluation of research on potential interventions and the contextual factors, such as acceptability to stakeholders and implementation constraints, need to be considered together.*

**Petticrew et al.** (2004)

Moreover, participants come with widely different perceptions about ‘evidence’. As Walshe and Rundall have pointed out:

*The clinical culture is highly professionalised, with a formal body of knowledge which is shared by all members of the profession and which acts as frame of reference for intraprofessional dialogue and debate ... In contrast, health care managers are a highly diverse group drawn from different professional and disciplinary backgrounds, and they often lack even a shared language or terminology with which to describe and discuss what they do.*

**Walshe and Rundall** (2001)

These three different forms of evidence – colloquial, context-free scientific and context-sensitive scientific evidence – will not combine of themselves to produce health system guidance. Combining and interpreting them requires a process and the most suitable process may be deliberative. Regardless of which of the three types of evidence one is considering, any suitable process needs to address a common set of complexities. The following list is unlikely to be all-inclusive:

- all evidence needs to be interpreted
- its relevance needs to be assessed
- its quality needs to be assessed
- its applicability in the current context, as compared with that in which it was generated or collected, needs to be assessed
- its completeness needs to be assessed
- qualitative evidence needs to be weighed alongside quantitative
- any technical controversy over its standing needs to be settled
- the precision of estimates of effectiveness needs to be assessed
- the robustness of the results need to be tested by sensitivity analyses
- the evidence, of whatever kind, needs to be infiltrated by values to determine priorities, ‘worthwhileness’ and to specify what ought to be done and by whom.

Facts do not ‘speak for themselves’ and any single piece of evidence, whether of the scientific or colloquial type, is rarely complete enough to enable guidance to be created without further evidence and assessment. To be useful a deliberative process must therefore facilitate the combination and interpretation of the population of evidence (however defined) for the purpose intended and enable those engaged in it to explain why they decided as they did.

Maintaining a common understanding of what constitutes evidence is likely to become increasingly difficult as further interest groups or stakeholders are added in any guidance-producing process. Conversely, the more homogeneous the group in terms of professional background and level of responsibility, the less tension and disagreement is likely to exist on what constitutes permissible evidence. However, it seems unlikely that the object ought ever to be to maximise the homogeneity merely for the sake of achieving a ‘common understanding’. If a ‘common understanding’ can be reached it ought to be reached with an appropriate degree of heterogeneity (for example, one in which the full

range of stakeholder interests is represented) and, if it cannot be reached, then the differences and the reasons for them are worth facing up to explicitly, not obscured through selection bias. There are also other dangers, as we shall see, in having too homogeneous a membership of a decision making body.

In short, the decision making process ought to provide a way through which the preferences of participants can be transformed rather than merely aggregated; a process that allows participants to change their minds; that is a means through which three kinds of evidence can be assessed and combined – colloquial (e.g. from professional experience, case-studies, other gossip), context-free science with high internal validity (such as evidence from randomised clinical trials), context-specific science with high external validity (such as evidence from cost-effectiveness analyses, most budget impact analyses); and that enables what people bring to the deliberation to count (such as their own values, experience, attitudes to risk and degrees of understanding and knowledge).

---

## 4. DELIBERATION COMPARED WITH CONSULTING OR COMMENTING

Deliberative processes often include but are not the same as consultation or comment. A famous example of consultation is the Oregon experiment to help determine which clinical procedures ought to be included in that state's Medicare program. It was not a deliberative process, but a process of consultation in which there were 47 community meetings, 12 public hearings, and 54 panel meetings for health care providers. All the data thereby gathered was fed into a committee (the Oregon Health Services Commission) for prioritisation of procedures (Garland 1992). Thus many were *consulted* prior to the decision but relatively few *participated* in its making. The Commission itself doubtless engaged in much deliberation but the participation of all those people who were consulted was not part of the decision-making.

Nor are opportunities to comment the same as deliberation. NICE provides opportunities for people to comment on technologies that are under appraisal, alongside consultation and deliberation. The public in general might be invited to comment (say, via a website) and some individuals or organisations may receive specific invitations. Like consultation, commenting can be a part of a

deliberative process but it is not to be equated with one. Neither consulting nor commenting involves mutual deliberation – there is limited interchange, there is restricted participation – and neither is an arrangement for the actual taking of decisions, whereas deliberative processes can be. These are what make deliberative processes different.

One approach that embraces the whole range of comment, consultation, and deliberative participation is the 'Cooperative Discourse Model' of Renn (1999). This entails the elicitation of values and criteria from stakeholder groups, the provision of policy options by expert groups, and the evaluation and design of policies by randomly selected citizens. This was a model that seems to have been used to good effect by the UK Committee on Radioactive Waste Management, which is an independent committee established by the UK Government in November 2003 to develop recommendations for the long term management of higher level radioactive wastes, and which faced a classic set of issues of science and of value. Its terms of reference explicitly required that the review

*be carried out in an open, transparent and inclusive manner, ... must engage members of the UK public, and provide them with the opportunity to express their views. Other key stakeholder groups with interests in radioactive waste management ... [had also] to be provided with opportunity to participate. The objective of the review [was] to arrive at recommendations which can inspire public confidence and [were] practicable in securing the long term safety of the UK's radioactive wastes. It must therefore listen to what people say during the course of its work, and address the concerns that they raise.*

**CoRWM** (2006)

The use of the Cooperative Discourse Model seems to have been a success – at least as judged by the criterion that the client knows best. The Government's response to the report included this:

*The reflection of a wide range of viewpoints, and a basis in sound science is key to providing recommendations which inspire public confidence for managing the wastes in the long term, providing protection for people and the environment. The open and transparent manner in which CoRWM has conducted its business has been ground breaking. Accordingly Government welcomes CoRWM's report and believes it provides a sound basis for moving forward. Most recommendations can be acted on immediately; others require us to undertake more work.*

**UK Government** (2006)

### **BOX 5 DELIBERATION THAT IS EMBODIED IN THE PRODUCTION OF EVIDENCE**

- summarised scientific evidence on (context-free) efficacy in the form of narrative reviews, systematic reviews or meta-analyses, each of which will have involved a lot of 'judgment'. Reviewers often engage in deliberative processes to reach agreement on (e.g.) search terms, search engines, inclusion-exclusion criteria, quality criteria
- other scientific evidence about contexts with controlled social science experiments, which in turn have to be summarised and synthesised for decision makers
- artificial evidence such as that from a model extrapolating beyond experimental time periods (such models typically have very different outcomes according to their provenance (e.g. industry or academia))
- the summarising of elicited colloquial evidence through consultative processes including public meetings and the hearing of special witnesses, as well as directly from those participating in the deliberative process.

Deliberation need not be a late stage process. In most cases, the production of evidence itself will have embodied deliberative processes as, for example, in scientific discussions of the design of a research project, clinical trial or systematic review. Box 5 indicates how deliberation may be embodied in a cascade of evidence production. This cascade creates a very fertile soil for deliberative digging. The typical scientific evidence on (context-free) efficacy is summarised in the form of narrative reviews, systematic reviews or meta-analyses (each of which will themselves have involved a lot of 'judgment') and each of which in itself will have often embodied mini-deliberative processes. So there are deliberative processes within deliberative processes. Other scientific evidence may relate to relevant contexts with controlled social science experiments, which in turn may also be summarised and synthesised and raise similar issues of judgment for deliberation. 'Artificial' evidence, such as evidence from economic models that extrapolate beyond experimental time periods, is particularly suited to deliberation, as is the evidence that comes up through colloquial processes like public meetings, hearings from special witnesses and survey material.

Within some of these embodied further deliberative processes (for example, in systematic reviews) there is scope for the involvement of 'lay' people with

scientists. I remember doing a systematic review with some colleagues on a review of the effectiveness of management systems for health and safety in workplaces. It was a very complex literature of generally low quality (as judged by Cochrane standards). At one stage we thought we had got all the relevant search terms and thought we had done a really good job. We then tried the list of search terms out on the clients in a joint workshop, a process that *doubled* the length of the list. Bringing the clients in was thus actually *scientifically* useful.

No evidence is totally authoritative; it all involves judgments by people in its creation, assembly and presentation. Some of the judgments are technical and scientific (was the most efficient estimating procedure used?). Some are scientific but also interpretive (are the trial results applicable in another setting?). Some are scientific and judgmental (were the scientists at risk of bias from their funding sources?). Some have the character of social value judgments (was the outcome measure an appropriate indicator of health?). Moreover, these are all questions about which it is perfectly possible for both scientifically trained and 'lay' people to disagree amongst themselves.

---

## **5. EVIDENCE REGARDING DELIBERATIVE PROCESSES**

### **BOX 6 SIX STANDARDS OF QUALITY DECISIONS**

- An appropriate frame (scoping the question(s) to be addressed)
- Creative, doable alternatives from which to choose
- Meaningful, reliable information
- Clear values and trade-offs
- Logically correct reasoning
- Commitment to action

**(Matheson & Matheson 1998, ch. 3)**

Most of the evidence relating to answering the 'does it work?' question of deliberative processes is qualitative, impressionistic, casuistical and, in short, 'colloquial'. Done well as a process, it might plausibly satisfy the six standards for quality decisions devised by Matheson and Matheson (1998) (see Box 6). Their methods, however, do not enable us to get at the ultimate criterion of success, which might best be specified as a decision deemed to be the best possible under the circumstances by the responsible authority

to whom the decision makers are accountable. At the highest level of decision making, in the case of a private company, that might be its shareholders. In the case of the NHS it is the Secretary of State for Health. In other cases, the appropriate 'deemers' are those judged to be the key stakeholders or an accountable 'authority'. The evidence we ultimately seek, if this approach is considered the right one, is thus ultimately subjective, probably qualitative, and probably best discovered only in a Darwinian fashion – decision processes that work best in the environment of which they are a part will be *those that survive*. The test is thus the test of history. That does not mean that the evidence for or against the effectiveness of deliberative processes is not explicit, systematic and replicable, as was discussed earlier, though it seems likely that it would be sufficiently deficient in each regard so as rarely to deliver a 'knock-out' verdict. Meanwhile, I shall offer some conjectures that are at least in principle testable, though they certainly fall short of this ultimate judgment (see Culyer and Lomas 2006). The ultimate approach entails taking an 'output' or 'outcome' effectiveness view of success (Schilling et al. 2007) which is bedevilled both by problems of measurement (is the effect 'large' or 'small' and relative to what?) and of attribution (can any observed effects be fairly attributed to the process?). An alternative is 'process effectiveness' which may offer as good a way of assessing the effectiveness of deliberative processes as we are likely to have for some time: did the process itself work well, even if we cannot know whether the decisions it generated were the truly 'right' ones?.

---

## 6. UNDER WHAT CIRCUMSTANCES ARE DELIBERATIVE PROCESSES LIKELY TO BE OF GREATEST USE?

There is some evidence with a bearing on an answer to this question but much of what I have to suggest will still be my own conjecture. Recall that evidence is always ambiguous. Some of the problems posed by evidence that might be resolvable through deliberation include situations where:

- evidence from more than one expert discipline is involved
- evidence from more than one profession is involved
- some stakeholders' interests are threatened by evidence
- there are technical disputes to resolve

- evidence is scientifically controversial
- evidence is incomplete
- evidence is lacking
- evidence gathered in one context such as a controlled trial in a large teaching hospital is to be applied in another such as a district hospital or a primary care unit
- issues of outcome, benefits and costs go beyond the conventional boundaries (of concept and end-point) of medical research design
- there is substantial uncertainty about key values
- there are risks (quantified or unquantified) to patients that need to be assessed and weighed
- there are risks (e.g. of malpractice suits) to professionals that need to be assessed and weighed
- there are other social and personal values not taken into account in the scientific evidence
- there are issues of equity and fairness of treatment (e.g. of patients similar in many respects but differing in their capacity to benefit)
- there are issues of implementability and operational feasibility
- there are issues of short term financial feasibility
- there are reasons to suppose that implementation may seriously destabilise local strategies and priorities
- wide professional 'ownership' is desired
- public credibility is desired
- political 'trust' is involved (e.g. no unpleasant surprises for ministers; help on how to handle unwelcome or embarrassing evidence).

When there is evidence from more than one expert discipline issues can easily arise about language. 'Cost' and 'outcome' are unlikely to mean the same to a clinician, a sociologist or an economist. Confusion may arise through failing to distinguish between statistical, clinical and policy significance. Views about the relative virtues of cross-sectional and time-series data are not shared. Bayesians and frequentists do not always see eye to eye. Equilibrium gets confused with equipoise. There are a lot of conventions that are manifestly different as between disciplines and these can easily become barriers to communication. Many such issues can be resolved only by talking – and, moreover, by frequent engagements of a deliberative character.

Feeling threatened is something that is dangerous, not only for the person threatened but also for the whole decision making process. A deliberative process can be one in which people's interests are exposed and the character of the risks to which they are exposed

assessed. That in itself may be sufficient protection, for example, through enabling those affected to take preliminary steps to minimise adverse impacts, or for further analysis of the size of the threat, and any more extensive protection or compensation that might be warranted, to become an issue for discussion.

Deliberation is likely to be useful when there are technical disputes to resolve in connection with evidence. These are endemic and non-trivial. Some relate to the evidence itself, some to its generation and some to the methodology used to summarise it.

Several characteristics of evidence that suggest the beneficial use of deliberation arise from complexity. Examples include issues concerning outcomes, benefits or costs, any or all which might go well beyond someone's conventional boundaries – of concept (for example, when the principal beneficiary is a family member rather than the patient), of metric (biological measures of outcome like blood pressure in comparison to the social consequences of such indicators), of end point (end of trial versus remainder of expected life) and in lots of other ways too.

Uncertainty is all-pervading, both that which is formally measured through conventions about statistical significance (for example, less precision in an estimate is usually indicated by a larger standard error) and that which is qualitatively expressed, for example, via a Likert scale of 'more' or 'less' likeliness. There can be uncertainty about the right methodology (Should benefits be discounted by the same factor as costs? Was the sample large enough to make statements with confidence about the experience of subgroups of patients? Was the measurement of other social and personal values not normally taken into clinical account appropriate? Ought such effects be taken into account at all?) In politically controversial decisions it may be helpful for 'the minister' to be able to explain in Parliament and to the public that there has been extensive consultation, much deliberation, full consideration of expert opinion and the ample weighing of the values of those most affected by the decision. At a minimum, the case becomes easier to make that the decision was not arbitrary and its rationale becomes communicable. This will take on specific significance if the decision is an unpopular one. *Both* the process *and* its outcome help to legitimise the decision. Some situations may require special handling: it is, for example, not unknown for ministers to have previously committed themselves to a view that the evidence now suggests to have been wrong or to have unwisely anticipated the outcome of the deliberation – and got it wrong.

---

## 7. THE EFFECTIVENESS OF DELIBERATION

The ultimate product of a deliberative process remains a decision, guidance or recommendation shaped by judgment about evidence, in a context that is both ethical – it inevitably involves value judgments as well as other kinds of judgment such as whether the science is any good – and practical. So guidance generated, for example, by NICE deliberative processes actually affects what citizens are going to get from the NHS; it actually affects what it is that manufacturers of all kinds are going to be able to sell to the NHS. It evidently has real consequences for both consumers and producers.

A deliberative process in itself guarantees nothing but I conjecture that, in the main, it increases the probability that decisions will be good ones. By giving empirical content to 'good ones' we might be able to test this claim. Were the resultant judgments reached using deliberation more comprehensively evidence-informed than other decisions, were they better matched to the context of their application, more efficiently implementable and more widely acceptable to those affected by it, whether positively or negatively?

Are these the right tests or are there others that should be applied?

One way of tackling a question like this is to ask what we might reasonably expect of a deliberative process. Here are some conjectures. They all embody the idea of having a comparator process, a counterfactual, which appears to be a feature of evaluation that, while commonplace in the received ways of conducting technology appraisals in health care, is scarcely mentioned at all in the literature (such as it is) on appraising decision processes. A deliberative process is, I conjecture, more likely than a non-deliberative process – without specifying what that other process might be – to:

- generate guidance that is both consistent with the context-free scientific evidence and its reasonable interpretation in particular contexts
- identify relevant clinical, social and political contexts for interpreting context-free scientific evidence, simply by virtue of the fact that people who can represent those sorts of views and who can interpret the scientific evidence on external validity are there at the table
- command a wide credibility in professional circles and beyond simply because professionals whom they respect are there at the table

- result in a quality and power of residual opposition that is low. This is self-evidently important for NICE because NICE decisions nearly always offend or hurt somebody. The prediction is that there will be less hurt, less offence and therefore less opposition with deliberation than without it
- result in less alienation. If the process is one whose design was actually shaped by everybody with a stake in its outcome, so that they actually become parties to its design and committed to the nature of the process, you are much less likely to be alienated by its outcome. After all, it was a process that you helped to design and even approved rather than some other arbitrary process that somebody else invented and thrust upon you. You may well be able to live with the consequences of deliberation even if on occasion the approved process produces results that are not your preferred ones
- generate guidance whose implementation will be speedy
- identify impediments to the implementation of guidance and to find solutions to those impediments – ways of leaping over or going around them
- identify knowledge gaps that might be resolved by further enquiry and research.
- the quality of the scientific research (both kinds) available at the start of the process and subsequently is high
- the quality of the colloquial evidence is high (for example, comes from reputable and respected sources)
- good quality systematic reviews and meta-analyses are available
- research into the public's views on contextual and other 'non-scientific' matters of relevance has been gathered and made available. Some of this may be directly represented on the decision-making committee, in other cases it may be better discovered through research or opinion surveys
- the rationale for the decision is subsequently written up and made available to all stakeholders
- the process is self-critical and points for improvement noted for the future
- the support staffing, whether in house or out-sourced, has been good in terms of both quantity and quality.

Other characteristics leading to success are about the process itself and are capable also of being turned into testable propositions. Here are some further conjectures. Success (again in any of the aspects discussed above) is more likely if:

It seems to me time to get empirical about these things and attempt, through both qualitative and quantitative methods, to test these conjectures.

What characteristics of a deliberative process are likely to lead to success? If relevant characteristics can be identified, they could be conjectured to entail less muddle when 'muddling through' (Lindblom 1959). The following conjectures draw on CHSRF (2006) and ongoing work by my colleagues Mark Dobrow and Roger Chafe. Success (in any of the aspects discussed above) is more likely if:

- chairing skills are high
- when different from the chair, the enabling skills of a (preferably trained) facilitator are high
- the process has clarity and openness
- the time-lines are reasonable – not too long as there is always pressure for decisions to be made quickly; on the other hand the shorter the time-line the less opportunity there is for consultation, discussion, deliberation, changing one's mind and learning
- colloquial evidence is used to challenge context-free evidence, set contexts and plug gaps in the science
- there is time for study, thought, discussion, reflection
- there are opportunities for all interested parties to comment
- there is scope for members to request information and take oral evidence to supplement written evidence when they have doubts about something
- there is an opportunity to appeal if the process is suspected of being flawed or the guidance that comes out appears unreasonable
- there are opportunities to hold in camera discussions to encourage free expression concerning sensitive matters, though otherwise making things as open and transparent as possible.
- the questions related to research evidence to be answered by the process have been clearly articulated
- the scoping of the questions to be answered has been thorough and has involved members of the deliberating group
- relevant comparators have been identified
- the decision group has a clear mandate
- if the nature of the decision is that it forms a recommendation or an input into a more ultimate decision process, the target 'clients' be identified and the 'knowledge transfer' mechanism be resourced and in place

There is an evident tension in this final conjecture. In the early days of an organisation, in camera sessions might be used more frequently than in its maturer days because at least some members might feel intimidated by the presence of a public, afraid of unpleasantness downstream should their support for a decision lead to an unwanted outcome or simply to avoid looking indecisive because they have changed their mind about something. Others may play to the crowd. Members have to learn to be Burke-like – to realise that they are not representing sectional interests but are there to represent the interests of the ‘general good’.

A final set of conjectures relates to membership of the decision making group. Success (again in any of the aspects discussed above) is more likely if:

- representative expertise in the relevant scientific (clinical and other) evidence exists amongst the panellists
- there is a representative breadth of colloquial sources of evidence (for example, respected people from the major professional communities with topic-specific interests and experience, including senior management)
- the deliberating membership is heterogeneous – a diversity of interests and opinions exists amongst them
- members are willing to share values openly
- stakeholder consultation has been inclusive with ample opportunities for all affected parties to be heard.

On one of these at least there is a fairly extensive literature (amusingly discussed in Surowiecki 2004). The sociological phenomenon of ‘group polarisation’ occurs when deliberating groups tend to move to increasingly extreme views. Although an artificially constructed consensus is not the purpose of deliberative processes, neither is an artificially induced extremism. Group polarisation can arise through a number of (imperfectly understood) mechanisms, as when members do not wish to stand out against what they perceive to be a trend in the opinions expressed, possibly reinforced by the forceful behaviour of strong personalities who manage to talk first and who might create strong framing effects for subsequent discussion. Of course, it is precisely this phenomenon that effective chairing and facilitation are designed to prevent.

---

## 8. NICE AND DELIBERATIVE PROCESSES

Are NICE processes deliberative, particularly those to do with technology assessment? In practice they

embody in varying degrees all of the following: extensive consultation, commenting, and deliberation. In the Appraisals Committee there is a lot of interchange that enables (though it may not always be used) a full deliberation and consideration of all the relevant matters: even things that the Secretary of State has said are *not* relevant, like costs that do not fall on social services or the health service, and which the methodology guidance singles out as explicit departures from the Reference Case.

Scientific judgment is usually about an effect (positive or negative), its size, the ways in which it can be achieved, for whom, for how long, its opportunity cost and so on. Value judgments tend to be in a different territory but they might be about, for example, how worthwhile a technology is, how defensible the tough bits of the decision are, how tolerant of uncertainty the committee ought to be, how inter-personal comparisons ought to be made or were made between potential beneficiaries, whether the QALY was a good tracker of the relative health benefits of the interventions that were compared.

My impression is that, aside from the issue of whether the benefits of late stage cancer treatments should be treated differently from other benefits, nearly all the deliberation relates to the evidence – and largely the scientific evidence – and is rarely about how the doubts regarding the evidence (all three kinds) are reconciled or combined, and the reporting is virtually zero. Subsequent decision-makers therefore cannot learn from previous ones and the outside world does not really know how and how well they addressed these issues.

Regarding the values that informed (or at least ought to have informed) the decision-making, the making of them explicitly and certainly the making of them public is something that one might reasonably expect of a deliberative process. But, needless to say, it is scarcely possible to make public the fruits of a deliberation that did not take place. As I implied in my earlier remarks, the NICE Appraisals Committee is hardly ever concerned with reporting how the dilemmas, the conflicts and the doubts were resolved, traded off and so on. And some dilemmas may have simply been buried. That is perfectly understandable but nonetheless unfortunate on lots of grounds – but perhaps chiefly because it denies the opportunity for everybody first to understand what has gone on and second to learn from it. The chief student of NICE’s guidance decisions ought, of course, to be the Appraisals Committee itself, through a kind of cumulative casuistical process.

---

## 9. COMPARATORS

As with any evaluative exercise, one seeking to evaluate the effectiveness of deliberative processes would need to specify appropriate counterfactuals or comparators: so the research question is whether these processes are effective *relative to* some other decision making process. In any given context, this could be the status quo process and might well be some much less costly alternative. For example, the Ontario system for evaluating cancer treatments is much less costly than NICE's Single Technology Appraisal procedure (which is itself less costly than NICE's Multiple Technology Appraisal process), relying basically on industry-supplied evidence which they may ask an expert to look at and review, but which is much less thorough in nearly all of the respects that have been mentioned. It sounds like a Mini compared to a pro-deliberationist's ideal Rolls Royce but that does not imply that it is more cost-effective. Low cost does not equate to high efficiency here, as elsewhere.

It is more than merely interesting to use less costly comparators because the costs of running decision-making and guidance-issuing organisations like NICE can be very high and there are always questions about whether the process is itself cost-effective. Such questions are not unreasonable ones to ask but they are hard ones to answer. One of the ways one might set about answering them would be to get very explicit about what the alternatives are and then trying to mount as controlled a comparative study as is possible. Whether NICE would relish the prospect of running parallel appraisals of the same technology but using processes that differed in critical respects is, of course, a moot point. In Ontario, however, the University of Toronto Priority Setting in Health Care Research Group has created an 'experimental' Citizens' Council that might in principle conduct such comparative experiments.

---

## 10. HOW DOES NICE DO?

Can NICE do it better? What does NICE do well at the moment? There are several things that one can say on behalf of NICE: it certainly commands international authority. It seems highly likely that NICE methods are destined to become a kind of world template for other jurisdictions to adapt for their own purposes. NICE has assured lots of stakeholder participation, both in the creation and establishment of NICE itself, but also in its governance, its methodologies, its procedures, and its processes. NICE scores highly on all these criteria. NICE believes that its processes are in line with 'accountability for reasonableness' (Daniels and Sabin

1998) and I think they are probably right. It is also much less rigid in using the 'threshold rule' than many, like the committee in Ontario. Although a judgment in terms of the 'ultimate' criteria for making the 'right' decisions may not be possible, less demanding tests, of the sort suggested (orally) to me by Lawrence Phillips are encouraging: I would conjecture that most members of NICE's Appraisals Committees would share the view that its processes get "everyone pulling in the same direction, ... a shared understanding of the issues, ... [and] a sense of common purpose and commitment to a way forward" and that a reason for this is that the processes in question are deliberative. That is a conjecture that could easily be evidence-based (or evidence-refuted).

NICE has not allowed methodological uncertainties, particularly in economics, to impede its progress. It has not merely forged ahead as though the issues did not exist but it has identified them, held many workshops to thrash the issues out, fed them as possible research issues into the NHS R&D Programme, and then carried on as best as it can and in full knowledge of the additional sources of uncertainty that such unsettled issues introduce. That seems to me to be exemplary – and highly deliberative.

NICE has certainly got through a lot of work and it has certainly got ministers off hooks.

What does it not do so well? It has not been very good at specifying explicitly what the other criteria might be that could be used in conjunction with the ICER threshold and how they might be balanced against cost per QALY. This is an active territory for enquiry. NICE is not all that good in supporting members in *ways of thinking* about non-threshold related criteria. The Secretary of State says NICE has to take innovation into account – how ought an Appraisals Committee do that in the context of a technology appraisal? I can think of quite a lot of bad ways in which it could do it – but NICE has not actually formulated any policy here as far as I am aware. What does NICE understand by equity? NICE is commanded to deal equitably, but how is that done in the context of economic appraisal of a technology? What does that ineffable term 'clinical significance' mean and how, if at all, ought it to be embodied in appraisals? What is the characteristic of an orphan drug (or an orphan anything), if there is one, that requires it to have a special set of appraisal criteria distinct from other interventions? It is quite clear to me that NICE does not regard these questions as unimportant and safely to be ignored – and, indeed, they are topics for the Citizens Council – but it is only beginning to come up with answers to some of them.

One danger in not leading on these issues is that other, not very thoughtful, presumptions may become established that NICE would not otherwise have adopted. It might, for example, have been railroaded into an acceptance of the so-called 'rule of rescue' by assigning a higher value to some people's health and length of life than to others', and for no clear and persuasive reasons (Cookson et al. 2008); or into assigning a higher value to an extension of life for those near death, or those with a specific disease (like cancer), than to the others – children, Mums and Dads, ... Everyman – who will not have the care they would otherwise have merited and solely on the grounds that they can effectively manipulate the media and stampede politicians into ad hoc decisions. It is striking that NICE's latest advice to its Appraisal Committee members, on the treatment of patients with terminal disease and who are near the ends of their lives, frames the matter for discussion without resorting to any algorithmic solution. They should assume:

*that the extended survival period is experienced at the full quality of life anticipated for a healthy individual of the same age, and [consider] the magnitude of the additional weight that would need to be assigned to the QALY benefits in this patient group for the cost-effectiveness of the technology to fall within the current threshold range.*

**NICE** (2009)

This sounds as though NICE's approach to issues of this sort might be casuistical – by offering a 'way to think' about an issue, it may encourage deliberations whose product over time may be a set of 'cases' which might at some later stage become consolidated into more specific guidance or serve as a basis for adjusting the future framing of members' deliberations. In the absence of consensus amongst ethicists and others who have thought systematically about such ethical issues, this casuistical approach may prove the best way forward.

I do not think NICE is very good at weighing qualitative factors explicitly. In fact, NICE is not alone – none of us is very good at doing it! But NICE ought to be better than it is. Nor is it very good at explaining recommendations of technologies with ICERs above the £20k threshold (there have been very few where the ICER has been over £30k). There is quite a lot of confusion outside NICE (and possibly within it) about the meaning of the threshold range of £20-30k.

What is next? Here are some more conjectures: some things that I think NICE ought to do. It ought to:

- explicitly use qualitative support methods
- be explicit about the use of casuistical methods, record the reasons for departures from the Reference Case and appraise the overall situation once sufficient cases are deemed to have been accumulated
- develop a database of 'typical' (imagined or real) scenarios which deal with some of the conflicted issues, with further recommended factors to consider such as 'here is how to think about this particular issue', 'what kind of utility is this?'
- ask NHS R&D to commission qualitative research into some of these issues, for example research into some of the issues that confront commissioning authorities, as they cope with the sequelae of NICE guidance.

Deliberation is not about establishing consensus. There is a lot to be said, however, for discovering whether there is or is not consensus. Sometimes it is important to understand that there is no consensus, in which case the Appraisals Committee has to make its recommendations with that in mind. It is quite important for public policy to discover whether there is any consensus about how equity ought to be used in health technology assessments. It would be likewise very useful to discover whether there is any consensus about the character of the anonymous individuals who are denied care, or receive less than otherwise, as a result of NICE guidance leading resources in other directions. Are they *just* as intrinsically meritorious as those who are often vehemently represented by special interest groups, as seems to be implied by most NICE decisions, or are they intrinsically *less* meritorious, as the special pleaders imply? We know that these people are anonymous but invisibility ought not to be sloppily taken to imply unworthiness or less need. Here indeed is a wonderful topic for deliberation!

---

## REFERENCES

- D Atkins, J Siegel, J Slutsky. Making policy when the evidence is in dispute. *Health Affairs*, 2005. 24: 102-113.
- C A Bana e Costa, P A da Silva, F N Correia. Multicriteria evaluation of flood control measures: the case of Ribeira do Livramento. *Water Resources Management*, 2004. 18: 263-283.
- S Buetow, T Kenealy. Evidence-based medicine: the need for a new definition. *Journal of Evaluation in Clinical Practice*, 2000. 6: 85-92.
- Canadian Health Services Research Foundation. Weighing up the evidence: making evidence-informed guidance accurate, achievable, and acceptable. CHSRF: Ottawa, 2006.
- Committee on Radioactive Waste Management. Managing our Radioactive Waste Safely: CORWM's Recommendations to the Government. Committee on Radioactive Waste Management: London, 2006.
- R Cookson, C McCabe, A Tsuchiya. Public healthcare resource allocation and the Rule of Rescue. *Journal of Medical Ethics*, 2008. 34: 540-544.
- A J Culyer, J Lomas. Deliberative processes and evidence-informed decision-making in health care – do they work and how might we know? *Evidence and Policy*, 2006. 2: 357-371.
- A J Culyer, A Williams, R J Lavers. Social indicators: health. *Social Trends*, 1971. 2: 31-42.
- N Daniels, J Sabin. The Ethics of Accountability in Managed Care Reform. *Health Affairs*, 1998. 17: 50-65.
- J Dowie. The future of HTA is MCDA. Knol. <http://knol.google.com/k/jack-dowie/the-future-of-hta-is-mcda/2rrcx2o0qjpsa/14#>.
- J Dowie. Decision Resource-Effectiveness Analysis: towards Value of Analysis Analysis. Knol. <http://knol.google.com/k/jack-dowie/decision-resource-effectiveness-analysis/2rrcx2o0qjpsa/15#>.
- M J Garland. Rationing in public: Oregon's priority-setting methodology. In: M A Strosberg, J M Wiener, R Baker, I A Fein (eds.) *Rationing America's medicalcare: the Oregon plan and beyond*. Brookings Institution: Washington DC, 1992. 37-59.
- J Habermas. Discourse Ethics: Notes on Philosophical Justification. In: J Habermas. *Moral Consciousness and Communicative Action*. MIT Press: Cambridge Mass., 1990. 43-115.
- K N Jenkins, N Barber. What constitutes evidence in hospital new drug decision making? *Social Science and Medicine*, 2004. 58: 1757-1766.
- P Kurland, R Lerner (eds.) *The Founders' Constitution*. University of Chicago Press: Chicago, 1987.
- C E Lindblom. The science of 'muddling through'. *Public Administration Review*, 1959. 19: 79-88.
- J Lomas, A J Culyer, C McCutcheon, L McAuley, S Law. *Conceptualizing and Combining Evidence for Health System Guidance*. Canadian Health Services Research Foundation: Ottawa, 2005.
- D Matheson, J Matheson. *The Smart Organization: Creating Value through Strategic R&D*. Harvard Business School Press: Boston MA, 1998.
- T Mendelberg. The deliberative citizen. In: M X Delli Carpini, L Huddy, R Y Shapiro (eds.) *Political decision making, deliberation and participation*. *Research in Micropolitics Vol. 6*. Elsevier: Amsterdam, 2002.
- National Institute for Health and Clinical Excellence. *NICE Citizens Council Report on Age*. NICE: London, 2004.
- National Institute for Health and Clinical Excellence. *Guide to the Methods of Technology Appraisal*. NICE: London, 2008.
- National Institute for Health and Clinical Excellence. *Appraising Life-extending, End of Life Treatments*. NICE: London, 2009.
- O F Norheim. The role of evidence in health policy making: a normative perspective. *Health Care Analysis*, 2002. 10: 309-317.
- M Petticrew, M Whitehead, S J Macintyre, H Graham, M Egan. Evidence for public health policy on inequalities: 1: the reality according to policymakers. *Journal of Epidemiology and Community Health*, 2004. 58: 811-816.
- J Petts. Barriers to participation and deliberation in risk decisions: evidence from waste management. *Journal of Risk Research*, 2004. 2: 115-133.
- L D Phillips. Decision conferencing. In: W Edwards, R F Miles, D von Winterfeldt (eds.) *Advances in decision analysis: from foundations to applications*. Cambridge University Press: Cambridge, 2007. 375-399.
- L D Phillips, M C Phillips. Facilitated work groups: theory and practice. *Journal of the Operational Research Society*, 1993. 44: 533-549.

J Rawls. *A Theory of Justice*. Harvard University Press: Cambridge Mass., 1971.

O Renn. *A Model for an Analytic-Deliberative Process in Risk Management*. *Environmental Science and Technology*, 1999. 33: 3049-3055.

L S Robson, J Clarke, K Cullen, A Bielecky, C Severin, P Bigelow, E Irvin, A J Culyer, Q Mahood. *The Effectiveness of Occupational Health and Safety Management System Interventions: A Systematic Review*. *Safety Science*, 2007. 45: 329-353.

M S Schilling, N Oeser, C Schaub. *How effective are decision analyses? Assessing decision process and group alignment effects*. *Decision Analysis*, 2007. 4: 227-242.

S Scott-Findlay, C Pollock. *Evidence, research, knowledge: a call for conceptual clarity*. *Worldviews on Evidence-Based Nursing*, 2004. 1: 92-97.

J Surowiecki. *The Wisdom of Crowds: Why the Many are Smarter than the Few*. Abacus: London, 2004.

UK Government. *Response to the Report and Recommendations from the Committee on Radioactive Waste Management (CoRWM) by the UK Government and the Devolved Administrations*. DEFRA: London, 2006.

K Walshe, T G Rundall. *Evidence-based management: from theory to practice in health care*. *Milbank Quarterly*, 2001. 79: 429-457.

M C Weinstein, W B Stason. *Foundations of cost-effectiveness analysis for health and medical practices*. *New England Journal of Medicine*, 1977. 296: 716-721.

M Whitehead, M Petticrew, H Graham, S J Macintyre, C Bambra, M Egan. *Evidence for public health policy on inequalities: 2: assembling the evidence jigsaw*. *Journal of Epidemiology and Community Health*, 2004. 58: 817-821.

A Williams. *Cost-benefit analysis: bastard science? And/or insidious poison in the body politick?* *Journal of Public Economics*, 1972. 1: 199-225.

---

## RECENT OHE PUBLICATIONS

All OHE Publications are peer reviewed by members of it's OHE Editorial Board:

<b>Prof A Culyer (Chairman)</b>	Institute for Work and Health, Canada and University of York
<b>Prof C Bulpitt</b>	Imperial College School of Medicine, Hammersmith Hospital
<b>Prof M Buxton</b>	Health Economics Research Group, Brunel University
<b>Dr J Dixon</b>	Nuffield Trust
<b>Prof H Gravelle</b>	Centre for Health Economics, York University
<b>Mr G Hulme</b>	Director, Public Finance Foundation
<b>Prof C Propper</b>	Dept. of Economics, Bristol University
<b>Prof P Zweifel</b>	Socioeconomic Institute, Zurich University

### **American Exceptionalism and American Health Care: Implications for the US Debate on Cost-Effectiveness Analysis**

Peter J Neumann  
March 2009  
Price £7.50

### **New Guidelines for Economic Evaluation in Germany and the United Kingdom: Are we any closer to developing international standards?**

Michael Drummond, Frans Rutten  
November 2008  
Price £7.50

### **The Market for Biosimilars: Evolution and Policy Options**

Deven Chauhan, Adrian Towse, Jorge Mestre-Ferrandiz  
October 2008  
Price £7.50

### **QALYs versus Experience: A Perspective from Experimental Economics**

OHE 14th Annual Lecture 2007  
Professor Daniel Kahneman  
August 2008  
Price £10.00

### **Health Care Services and the Single European Market**

Diane Dawson, Lyndsay Mountford  
September 2008  
Price £7.50

### **Challenges and Opportunities for Improving Benefit-Risk Assessment of Pharmaceuticals from an Economic Perspective**

James Cross, Louis Garrison  
August 2008  
Price £7.50

### **The New Global Economics of Vaccines: will the Scientific Potential be Realised?**

OHE 13th Annual Lecture 2006  
Jean Stephenne  
October 2007  
Price £10.00

### **Saving Lives, Buying Time: Economics of Malaria Drugs in an Age of Resistance?**

OHE 12th Annual Lecture 2005  
Professor Kenneth Arrow, Sir Richard Peto  
November 2006  
Price £10.00

Details of all OHE publications are available from our website [www.ohe.org](http://www.ohe.org).

---

## COMPENDIUM

### OHE Compendium 20th Edition 2009

#### The Essential Data Source for Understanding the UK Health Sector

The annual OHE Compendium of Health Statistics is the one stop statistical information source on population and health trends in the UK, and the changing finances and structure of the NHS.

The Compendium is compiled independently by the Office of Health Economics, drawing data together from a variety of scattered sources. The disparate data, some spanning more than 50 years, is consolidated into tables and graphs which are quicker and easier to use than most other information sources.

The Compendium contains nearly 300 easy to use, high quality tables and charts, with full commentary and explanatory notes provided.

The Compendium is available in both print and on-line versions, as a convenient desk top reference.

#### OHE Compendium online

On-line access provides a powerful text search facility, and enables charts, tables and graphs to be downloaded directly into Microsoft Excel spreadsheets or Microsoft Powerpoint presentations. With the online Compendium:

- Instantly access information anytime, anywhere
- Analyse, search and download charts, tables and graphs directly into reports or presentations.

Alternatively have the latest vital statistics at your fingertips with a copy of the reference book on your desk. With:

- New index and easy to use content pages
- Simple to use tables of statistics
- Charts clearly showing comparable trends
- Summaries of key trends in health and UK health care provision

#### Prices

The price for a hard copy of the 20th Edition of the Compendium for Health Statistics is £399.00. Special rates are available for charitable and public sector organisations\* upon proof of status at £89.00.

The price for a single user licence fee to the online Compendium of Health Statistics 2009, 20th Edition, including a hard copy of the Compendium is £799 + VAT. Special rates are available for charitable and public sector organisations\* upon proof of status at £199.00 + VAT. Please contact Claire Devaney at OHE for further details on + 44 (0)207 747 8855.

Site licences for multiple users of the on-line Compendium are also available. Please contact Radcliffe Publishing for further details on + 44 (0)1235 528820.

\*Charitable and public sector organisations only can purchase a single-user licence at this special rate. The publisher reserves the right to decide whether this price applies.



---

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 or of its funding body the ABPI.

---

### 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](http://www.ohe.org).