

**The use of cost-effectiveness**

**by the National Institute for Health and Clinical Excellence (NICE):**

**No(t yet an) exemplar of a deliberative process**

**Author:**

Michael Schlander

**Original Paper** (*or paper on an issue of current controversy in medical ethics, at the editors' discretion*)

**Address of author with identification:**

Prof. Dr. med. Michael Schlander, M.B.A.

- (1) Institute for Innovation & Valuation in Health Care (INNOVAL<sup>HC</sup>);
- (2) University of Applied Economic Sciences Ludwigshafen (Germany);
- (3) Mannheim Medical Faculty, University of Heidelberg (Germany)

**Address** (1): P.O. Box 1107, D-65741 Eschborn, Germany

Tel.: +49-6023-929589; Fax: +49-6023-929591

E-Mail: [michael.schlander@innoval-hc.com](mailto:michael.schlander@innoval-hc.com)

**Key words:** National Institute for Health and Clinical Excellence (NICE) – National Health Service (NHS) – accountability for reasonableness (A4R) – cost-effectiveness – quality-adjusted life-year (QALY)

**Word count** (abstract): 247.

**Word count** (body of text, excluding title page, abstract, and references): 3,477.

## **Abstract**

Democratic societies find it difficult to reach consensus concerning principles for health care distribution in the face of resource constraints. At the same time the need for legitimacy of allocation decisions has been recognized. Against this background, the National Institute for Health and Clinical Excellence (NICE) aspires to meet the principles of procedural justice, specifically the conditions of *accountability for reasonableness* (A4R) as espoused by Daniels and Sabin, i.e., publicity, relevance, revisions and appeal, and enforcement. Although NICE has adopted a highly standardised approach and continuously publishes key documents on its website, its technology appraisal programme does not fulfil the *publicity* condition of A4R. Economic models are not made sufficiently transparent to enable public scrutiny, and decision criteria other than cost-effectiveness remain enigmatic. NICE's reliance on cost-utility analysis and "plausible" cost-per-QALY benchmarks further raises serious issues with regard to the *relevance* condition of A4R. This is illustrated by counterintuitive cost-per-QALY rankings that are difficult to justify using reflective equilibrium methods, and by the current debate surrounding expensive therapies for rare diseases ("orphan" treatments). In addition, an excessive focus on QALYs may stand in the way of exploiting the best available effectiveness evidence. The NICE mechanism for *revision and appeals* is also more restrictive than provided in A4R. As to the *enforcement* condition, no effective quality assurance processes are in place for technology assessments, and implementation of guidance remains imperfect. NICE, despite impressive efforts, appears to have a long way to go before meeting the conditions of A4R.

NICE's multiple technology appraisal process is broadly considered a role model for health technology assessments that include economic evaluation [1-3]. A review team of the World Health Organization (WHO) described key principles of the NICE approach as "the use of best available evidence in decision-making, transparency, consultation, inclusion of all key stakeholders, and responsiveness to change" [4]. They concluded that, "in all of these areas, it is clear that NICE is setting a new, international benchmark, for which it can and should be congratulated" [4]. Further, NICE has assumed a leading role internationally by fostering methodological advances such as the use of probabilistic sensitivity analyses (designed to capture decision uncertainty) and mixed treatment comparison techniques (in order to enable indirect comparisons of technologies in the absence of head-to-head studies).

### **The logic of cost-effectiveness**

The logic of cost-effectiveness, as adopted by NICE and in contrast to traditional cost-benefit analysis, does not represent an orthodox application of economic welfare theory [5-9]. The development of the cost-effectiveness framework was, instead, heavily influenced by decision analysts with operations research backgrounds, who were striving to transfer methods used to optimise the efficiency of manufacturing processes to the production of health [10].

NICE has chosen specifically to be prescriptive about the use of cost-utility analysis – a variant of cost-effectiveness analysis – as its reference case, with quality-adjusted life years (QALYs) as a universal and comprehensive measure of health-related outcomes [11]. Cost-utility analysis, however, is compatible with standard cost-

benefit analysis only under restrictive assumptions, including a constant (context-independent) willingness-to-pay for each QALY gained [12-14].

Although the use of QALYs is backed by a strong research agenda [15], important methodological issues still remain to be resolved. For example, different valuation techniques give rise to inconsistencies in utility values for similar health states, causing serious reliability problems [16-18]. Another well-known but unresolved issue concerns the difference between the utility of a health state *expected* by healthy persons (or for that purpose, a sample of the general population, as required by NICE for reference case analysis [11]) and the utility of this health state actually *experienced* by patients, often confounded by adaptation to disability and disease. This raises further concerns about the content validity of derived QALYs [19-24]. Yet a key motivation (beyond the integration in one index of multiple clinical *outcomes for defined patients*) for the use of QALYs has been the promise to allow meaningful comparisons across a wide range of *interventions* and, therefore, different *patient groups* [25,26]. Despite the obvious relationship of both subjects, however, examination of the usefulness of QALYs as a measure of health outcomes should not be logically confused with debate about interpersonal comparisons and the appropriateness of specific aggregation rules [27].

It is a fundamental and well-established principle of decision analysis that “the identification and structuring of objectives essentially frames the decision being addressed. It sets the stage for all that follows” [28]. To be relevant, analytic decision support relies on prior clarification of the values and objectives to be pursued [29].

Then, to a great deal, applying the logic of cost-effectiveness to inform health care resource allocation decisions hinges on the assumption that “the principal objective of the National Health Service (NHS) ought to be to maximise the aggregate improvement in the health status of the whole community” [30,31]. While it appears trivial that health care services (should) produce health, it is by no means self-evident to make a quick leap from here to an assumed “principal objective” of collectively financed health care to simply maximise some construct (QALYs or else) of health-related consequences [32].

In fact, there is little if any evidence that an emphasis on maximisation (sometimes justified by an asserted “consensus in the literature” without specifying sources [10]) is shared by the general population [33]. On the contrary, there is a rapidly growing body of studies that collectively show that this assumption is “empirically flawed” [34-36]. Controversy revolves around (but is not limited to) a higher social priority for interventions when the severity of the patient’s condition increases, with life-saving interventions most highly valued (this is sometimes referred to as “the rule of rescue” [37-41]), and for people in so called double jeopardy (i.e., with more than one condition causing impairment) who have less QALYs to gain from successful interventions compared to otherwise healthy individuals [42-45]. To address these issues, there has been a call for more research into “*empirical ethics*” by leading health economists [46].

The QALY maximisation assumption is also critiqued from a *normative* perspective. Arguments prominently include the implied valuation of human life as a function of

health status, as opposed to viewing the value of human life as a dimension distinct from health, i.e., to assign individual life the same value independent of the presence of disorders and functional impairment [47,48]. Recently, the premise that “all people are equal regardless of their QALY score” and the presumption of potentially “disastrous effects [of denial of treatment for reasons of cost-effectiveness or, more precisely, the lack hereof] on the sense of personal worth and security” of afflicted patients [49] gave rise to a passionate debate in this journal [49-55].

In the absence of a gold standard against which to judge the criterion validity of the logic of cost-effectiveness, it has been proposed to use the so-called reflective equilibrium approach to examine the social acceptability of the resulting rankings of health care programmes [56-59]. Central to a reflective equilibrium approach is the claim that considered moral judgments about justice in particular cases carry weight [60]. The inconsistencies that can arise from the application of standard decision rules derived from the logic of cost-effectiveness are perhaps illustrated best using an example. Assuming the cost per QALY gained (incremental cost-effectiveness ratio, ICER) is, for example, ~£3,600 for sildenafil in erectile dysfunction [61], ~£7,000 for pharmacotherapy of children with attention deficit hyperactivity disorder [62,63], and >£120,000 for beta-interferons and glatiramer in multiple sclerosis [64], would this ranking reflect the comparative social desirability of these interventions? [65]

The issue of counterintuitive rankings is not a phenomenon encountered only in England and Wales but was a major obstacle faced by the protagonists of cost-effectiveness analysis for resource allocation in the Oregon Health Plan (OHP).

Under the OHP programme, for example, capping teeth for exposed pulp received a better ranking than an appendectomy for acute appendicitis. While some analysts correctly pointed out that capping a tooth for 150 patients (*not one!*) was ranked higher than an appendectomy for *one* person [66], others insisted that the ranking failed to reflect “the powerful human proclivity to rescue endangered life” [38]. Decomposing reflective equilibrium problems like those cited above reveals several questions, which are not addressed adequately by conventional cost-effectiveness analysis, including: (1) What priority should be given to the worst off – those with the most serious and/or immediate conditions? (2) When should small benefits to a large number of people outweigh large benefits to a small number of persons? [67] (3) How can the conflict between fair individual *chances* and best aggregated outcomes be resolved? [68].

It is conspicuous that reviews of the usefulness of such rankings (“QALY league tables”) by many health economists, while addressing a variety of technical issues in detail, have not given attention to the larger issue of the validity of the rankings [69,70]. Of course, the issue of counterintuitive rankings should not be confused with the problem of distorted human judgments due to “heuristics and biases” [71]. Moral intuitions in the sense of reflected values and beliefs – such as Rawls’ nonwelfarist account of primary social goods [56] and Sen’s appeal to a capabilities-based account [72,73] – cannot be invalidated simply on grounds of their incompatibility with competing normative claims [74]. It has been argued by some philosophers that there may exist an irreducible pluralism at the foundations of normative ethics [75].

### **Accountability for reasonableness**

Recognizing both the difficulty of democratic societies to achieve consensus on distributive principles for health care and the need for legitimacy of allocation decisions, Norman Daniels and James Sabin [76-79] proposed a framework for institutional decision-making, which they call “accountability for reasonableness” (A4R). In order to narrow the scope of controversy, A4R relies on “fair deliberative procedures that yield a range of acceptable answers” and consists of four conditions [79]:

**Publicity**, i.e., resource allocation decisions must be public, including the grounds for making them. Transparency should open decisions and their rationales for scrutiny by all affected, not just the members of the decision-making group.

**Relevance**, i.e., “the grounds for decisions must be ones that fair-minded people can agree are relevant to meeting health care needs fairly under reasonable resource constraints.” Arguments should rest on scientific *evidence*, though not necessarily a specific kind of evidence [77], and appeal to the notion of “fair equality of opportunity.” Although Daniels and Sabin acknowledge that *stakeholder participation* may improve deliberation about complicated matters, they believe it is neither a necessary nor a sufficient condition of A4R.

**Revisions and appeal**, i.e., there must be an institutional mechanism to engage a broader segment of society in the process, providing those affected by a decision to

reopen deliberation, and to offer decision-makers an option to revise funding decisions in light of further arguments.

***Enforcement*** entails some form of regulation to make sure that the first three conditions are met.

### **NICE's use of cost-effectiveness as an exemplar of a deliberative process?**

Seeking to combine legitimacy and pragmatism, and realizing that utilitarianism “has next to nothing to offer in eradicating health inequalities” [50], NICE put aside questions of whether matters of content can be resolved solely with a reference to “due process” [80] and explicitly subscribed to the principles of accountability for reasonableness (A4R) [50,79]. At the same time, NICE reaffirmed its preference for cost-utility analysis with QALYs “as its principal (though not only) measure of health gain” [50].

Qualitative research may serve to illuminate the performance of NICE in relation to A4R. A preliminary case study of a recent NICE Technology Appraisal (No. 98 [63]; see [www.nice.org.uk](http://www.nice.org.uk)) focused on the processes adopted by NICE. The case study, which was largely in agreement with the positive findings of the WHO review [4], confirmed the high (albeit not perfect) level of transparency, predictability, and the participatory nature of the NICE approach [81]. However, the analysis also indicated a need for further in-depth inquiry. A subsequent, more comprehensive review focusing on the technology assessment report informing NICE Technology Appraisal No. 98 [62,63] did not confirm the expected robustness of the NICE evaluation

process, revealing a striking number of limitations and anomalies [82]. Collectively these left the assessment open to critique regarding *all* essential components of a technology review question, namely the population studied, the choice of interventions, the clinical and economic criteria used, as well as the study designs and selection criteria [82-84]. Furthermore, the structure of the economic model itself was found to be prone to distortion and bias in various ways. An unsettling number of consistency problems were identified within the assessment report [82,83]. As a consequence, the assessment did not consider fully the best available evidence and was unable to identify any differences in clinical effectiveness between the treatment options evaluated [82,83].

A number of underlying problems were suggested as causing the observed limitations in the assessment, including, notably, an insufficient integration of clinical and economic perspectives; a high level of standardisation demanding the problem fit a preconceived solution approach, including (but not limited to) the use of QALYs as effectiveness measure; and, somewhat surprisingly, issues related to the technical quality of the assessment itself [83].

Significant gaps are observed when the NICE technology appraisal *process* is compared to the conditions of accountability for reasonableness:

**Publicity.** The overall process was well structured and followed well-defined timelines with predictable opportunities for (some) stakeholders to provide input; key documents were continuously published at the NICE website. Major limitations of

transparency were related to the use of commercial-in-confidence information (a situation on which NICE has taken action meanwhile), the economic model developed by the assessment group, and decision-making criteria beyond cost-effectiveness used by the appraisal committee. For example, the detailed health state vignettes used to elicit utility estimates were not published with Technology Appraisal No. 98 [62,63]. Some of the company submissions may have been biased, as the results submitted tended to favour their respective products. Subsequent in-depth review drawing on peer-reviewed publications (which had not been in the public domain at the time of assessment) identified specific sources of distortion, contributing to these inconsistencies [82,83].

Even more importantly, NICE designate economic models as “proprietary”. This insulates a major component of their technology assessments from public scrutiny and does not meet established standards of good economic modelling practice [85,86].

Read-only copies of models are provided to consultees and commentators only upon their request in writing, with the *caveat* that these stakeholders “must not publish the model wholly or in part” [87] and are not permitted to “re-run the model with alternative assumptions or inputs” [87]. This practice also prevents academic debate and, therefore, is not conducive to the further development of health economic evaluation methods.

As admitted by NICE (cf. above), quasi-utilitarian maximisation of QALY gains irrespective of their distribution does not provide for a sufficient basis for health care resource allocation in tune with social preferences. Thus, it is a further critical

transparency issue that decision criteria other than cost-effectiveness have not (yet) been codified by NICE [4,83]. Official statements by NICE have remained vague [88,89]. Appraisal committee meeting minutes are hardly informative [81,83], despite reasonable expectations created by NICE's own process description, claiming, "the minutes provide an accurate record of its proceedings and discussions and also inform the public of the matters discussed at the meeting" [87].

**Relevance.** In the absence of codified criteria for fairness and with its heavy (albeit not exclusive) reliance on cost-effectiveness benchmarks, the NICE approach may be characterised as an "efficiency-first" strategy, with "efficiency" defined according to the logic of cost-effectiveness [11,36,87]. The NICE priority for "efficiency" is demonstrated by its expectation that adopting a cost-per-QALY threshold, even if interpreted somewhat flexibly [1,88,89], will "maintain *consistency* across the many different types of health care technologies that NICE appraises" [1]. It has been argued by observers that this approach in practice will result in the marginalization of other factors "as outside of NICE's terms of reference" [90]. It seems indeed unlikely that the current approach will enable to adequately capture social preferences for health care provision.

The current debate surrounding the cost-effectiveness of expensive drugs to treat patients with rare disorders ("orphan drugs") illustrates this issue. Given the high fixed (i.e., volume-independent) and low variable cost structure of the pharmaceutical industry [91,92], applying the logic of cost-effectiveness would inevitably deprive these patients of *any chance* to receive effective treatment [93-96]: The costs per

QALY gained for these treatments often exceed £100,000 at current NHS acquisition prices. It is therefore impossible to justify NHS coverage of these treatments using NICE's cost-effectiveness benchmark of "a most plausible" (maximum) cost per QALY in the range of £20,000 to £30,000 [88,89,97-99]. However, a majority of the members of the NICE Citizens Council (cf. below) believed that "the National Health Service should consider paying premium prices for drugs to treat patients with very rare diseases," reasoning *inter alia* that this approach avoids breaching "the human rights of individuals [afflicted with rare disorders]" and "helps shape a more humane society" [100]. Economists agree that the shadow price of (or willingness-to-pay for) a QALY may well depend on the budgetary impact of the intervention under consideration [101]. Empirical data further suggest that the public "places a very high value on giving everyone a *chance* at receiving scarce resources," even if that is associated with a significant loss of efficiency in terms of maximising aggregated outcomes [102,103].

The example of orphan drugs sheds light on the end of a continuum, not a distinct well-defined category. From an economic perspective, this example illustrates the role of budgetary impact (as the opportunity costs of programmes depend crucially on their size [104]) in reimbursement decision-making – a role that NICE has repeatedly denied taking into consideration [1,88], despite at least some indications to the contrary [89]. While the position taken by NICE appears questionable on both theoretical [7,105] and pragmatic [106,107] grounds, it is evident that explicit

recognition of budgetary impact would have fatal implications for any attempt to interpret the logic of cost-effectiveness in a *normative* way [7,35,108].<sup>i</sup>

Daniels and Sabin reasoned that fair-minded people “should accept many kinds of evidence and reasons as relevant”, including “scientific evidence about effectiveness and safety” [77]. An overly narrow focus on data which are thought to enable the computation of QALYs may result in the exclusion of relevant information and thus contribute to a situation where technology assessments, adhering to NICE reference case provisions, fail to use the best available clinical evidence [109]. The analysis of Technology Appraisal No. 98 illustrates that this is not merely a theoretical concern [82,83]. The occurrence of such situations is in stark contrast to claims by leading NICE representatives that its “guidance is based on the best available evidence” [88].

***Revisions and appeal.*** NICE provisions for appeal are more restrictive than those provided for by A4R. Appeals are narrowly limited to specific grounds and do not permit the debate to reopen [110]. New evidence or simply disagreement with an appraisal will “almost certainly” not be accepted [110]. Although understandable from a pragmatic perspective, these limitations are not adequately compensated for by opportunities for (invited) consultees and commentators to provide inputs during the

---

<sup>i</sup> A *league table algorithm* selecting programmes in ascending order (i.e., those with the lowest ICERs first) until available resources are exhausted might *in theory*, under restrictive assumptions, satisfy the need to consider opportunity costs. In practice, however, ICERs are not available for all competing programmes – hence this approach is not feasible. (Even if it were feasible as a method, it would still be impractical, as its implementation would imply a permanently changing threshold, with programmes around the cut-off line to be added to or excluded from coverage in an ongoing process. The *ICER threshold rule* is therefore adopted in practice but cannot satisfy the requirement to consider opportunity costs [104].

appraisal process. Only relatively short windows of opportunity are provided, with a massive amount of data to be reviewed under limited transparency [81,83,87].

***Enforcement.*** There is no indication that NICE has implemented an effective quality assurance system for its technology assessments. Again, the case of Technology Appraisal No. 98 suggests that this is not simply a theoretical observation [82,83]. This issue is exaggerated by the limited transparency of economic models, as discussed earlier. Independent analysts concluded that “absolute transparency of reporting is needed” to address the problem of poor methods in economic evaluations [111]. Conventional peer-review processes are unlikely to be up to the task of assessing the quality of economic evaluation models [85,112,113]. Design of effective quality assurance systems must take into account these challenges.

Further, following Hasman and Holm [80], proper enforcement of appraisal-based decisions should be implemented to ensure that reasoning is “decisive in priority setting and not merely a theoretical exercise”. Although NICE and the NHS have made substantial efforts to improve actual *implementation* of guidance, significant gaps remain in this area as well [114,115]. It has been suggested that guidance may be “more likely to be adopted when there is strong professional support, a stable and convincing evidence base” and that “guidance needs to be clear and reflect the clinical context” [114]. These conditions were arguably not fulfilled in the case of Technology Appraisal No. 98 [82,83].

NICE has established a Citizens Council to provide input on the “topics it wants the council to discuss” [116] and to ensure that its “value judgments resonate broadly with the public” [88], while maintaining that its guidance “is based on clinical and cost-effectiveness evidence” [116]. The Citizens Council has shown some concern for considerations of social justice, but has broadly endorsed NICE’s approach, concluding that “cost-utility analysis is necessary but should not be the sole basis for decisions on cost-effectiveness” [117,118]. What is unclear is whether the Citizens Council was confronted with the issue of counterintuitive cost-per-QALY rankings such as those cited above, i.e., with the logic that the benefit of providing ten people with a utility gain of 0.1 for the rest of their life (corresponding to sildenafil treatment for men with erectile dysfunction [61]) is indeed considered equivalent to saving the life of a single (otherwise healthy) person. With respect to NICE’s attempts to ensure stakeholder input in general, and to its Citizens Council in particular, it seems worth mentioning that Daniels and Sabin, with explicit reference to advisory bodies and commissions, believe that the absence of a democratic representational procedure “constitutes a decisive objection to claiming that consumer participation contributes to legitimacy” [79]. They further refute the idea “that organizationally based deliberation can substitute for broader democratic processes” [79]. Although selection of Citizens Council members might be described as only “symbolic representation” of the public, NICE should be acknowledged for their efforts to achieve a diverse cross-section of the population [119].

In conclusion, there are good reasons to be suitably impressed by the attempts of NICE to ensure rigorous systematic reviews, objective economic evaluation,

stakeholder participation, and transparency of process as well as value judgments in their assessment. This notwithstanding, NICE is still in its infancy [120], and there remains a long way to go until it will have met its stated objective [50,118] to fulfil the conditions of accountability for reasonableness.

## References

1. Pearson SD, Rawlins MD. Quality, innovation, and value for money. NICE and the British National Health Service. *Journal of the American Medical Association (JAMA)* 2005; 294 (20): 2618-2622.
2. Neumann, PJ, Rosen, AB, Weinstein MC. Medicare and cost-effectiveness analysis. *New England Journal of Medicine* 2005; 353 (14): 1516-15.
3. Culyer AJ. NICE's use of cost effectiveness as an exemplar of a deliberative process. *Health Economics, Policy and Law* 2006; 1 (3): 299-318.
4. Hill S, Garattini S, van Loenhout J, et al. *Technology Appraisal Programme of the National Institute for Clinical Excellence. A review by WHO*. Copenhagen: WHO Regional Office for Europe, June-July 2003.
5. Birch S, Gafni A. Cost effectiveness ratios: in a league of their own. *Health Policy* 1994; 28: 133-141.
6. Tsuchiya A, Williams A. Welfare economics and economic evaluation. In: Drummond M., McGuire A (eds) *Economic evaluation in health care: merging theory with practice*. Oxford: Oxford University Press, 2001: 22-45.
7. Donaldson C, Currie W, Mitton C. Cost-effectiveness analysis in health care: contraindications. *British Medical Journal* 2002; 325: 891-894.
8. Birch S, Donaldson C. Valuing the benefits and costs of health care programmes: where's the 'extra' in extra-welfarism? *Social Science & Medicine* 2003; 56 (5): 1121-1133.
9. Birch S, Gafni A. Information created to evade reality (ICER): things we should not look to for answers. *Pharmacoeconomics* 2006; 24 (11): 1121-1131.
10. Torrance GW. Utility measurement in healthcare: the things I never got to. *Pharmacoeconomics* 2006; 24 (11): 1069-1078.
11. National Institute for Clinical Excellence (NICE): *Guide to the Methods of Technology Appraisal (reference N0515)*. London: NICE, April 2004.
12. Ubel PA, Richardson J, Prades JL. Life-saving treatments and disabilities: are all QALYs created equal? *International Journal of Technology Assessment in Health Care* 1999; 15: 738-748.

13. Bala M, Zarkin GA, Mauskopf JA. Conditions for the near equivalence of cost-effectiveness and cost-benefit analyses. *Value in Health* 2002; 5 (4): 338-346.
14. Brent RJ. *Cost-Benefit Analysis and Health Care Evaluations*. Cheltenham: Edward Elgar, 2003.
15. Gold MR, Stevenson D, Fryback DG. HALYs and QALYs and DALYs, Oh My: similarities and differences in summary measures of population health. *Annual Review of Public Health* 2002; 23: 115-134.
16. Read JL, Quinn RG, Berwick DM, et al. Preferences for health outcomes: comparisons of assessment methods. *Medical Decision Making* 1984; 4: 315-329.
17. Hornberger JC, Redelmeier DA, Peterson J. Variability among methods to assess patients' well-being and consequent effect on a cost-effectiveness analysis. *Journal of Clinical Epidemiology* 1992; 5: 505-512.
18. Garber AM. Advances in cost-effectiveness analysis of health interventions. In: Culyer AJ, Newhouse JP (eds) *Handbook of Health Economics*, Vol. 1A. Amsterdam: Elsevier, 2000: 181-221.
19. Boyd NF, Sutherland HJ, Heasman ZK, Cummings BJ Whose values for decision-making? *Medical Decision Making* 1990; 10: 58-67.
20. Richardson J. Cost utility analysis; what should be measured? *Social Science & Medicine* 1994; 39 (1): 7-21.
21. Kahneman D, Wakker PP, Sarin R. Back to Bentham? Explorations of experienced utility. *The Quarterly Journal of Economics* 1997; 112: 375-405.
22. Menzel P, Dolan P, Richardson J, Olsen JA. The role of adaptation to disability and disease in health state valuation: a preliminary normative analysis. *Social Science & Medicine* 2002; 55: 2149-2158.
23. Ubel PA, Loewenstein G, Jepsen C. Whose quality of life? A commentary exploring discrepancies between health state evaluations of patients and the general public. *Quality of Life Research* 2003; 12 (6): 599-607.
24. Damschroder LJ, Zikmund-Fisher BJ, Ubel PA. The impact of considering adaptation in health state valuation. *Social Science & Medicine* 2005; 61 (2): 267-277.
25. Torrance GW. Measurement of health state utilities for economic appraisal. A review. *Journal of Health Economics* 1986; 5: 1-30.
26. Dolan P, Cookson R. A qualitative study of the extent to which health gain matters when choosing between groups of patients. *Health Policy* 2000; 51: 19-30.
27. Dolan P. Utilitarianism and the measurement and aggregation of quality-adjusted life-years. *Health Care Analysis* 2001; 9 (1): 65-76.
28. Keeney RL, Raiffa H. *Decisions with multiple objectives: preferences and value tradeoffs*. Cambridge: Cambridge University Press, 1993.

29. Keeney R. *Value-Focused Thinking*. Cambridge, MA, Harvard University Press, 1992.
30. Culyer AJ. The rationing debate: maximising the health of the whole community – the case for. *British Medical Journal* 1997; 314: 667-669.
31. Weinstein MC, Stason WB. Foundations of cost-effectiveness analysis for health and medical practices. *New England Journal of Medicine* 1977; 296: 716-721.
32. Harris J. The rationing debate: maximising the health of the whole community. The case against: what the principal objective of the NHS should really be. *British Medical Journal* 1997; 314: 669
33. Coast J. Is economic evaluation in touch with society's health values? *British Medical Journal* 2004; 329: 1233-1236.
34. Dolan P, Shaw R, Tsuchiya A, Williams A. QALY maximisation and people's preferences: a methodological review of the literature. *Health Economics* 2005; 14 (2): 197-208.
35. Schlander M.: Kosteneffektivität und Ressourcenallokation: Gibt es einen normativen Anspruch der Gesundheitsökonomie? In: Kick, HA, Taupitz J (eds) *Gesundheitswesen zwischen Wirtschaftlichkeit und Menschlichkeit*. Münster: LIT-Verlag, 2005: 37-112.
36. Richardson J, McKie J. Economic evaluation of services for a National Health Scheme: The case for a fairness-based framework. *Journal of Health Economics* 2007; 26 (4): 785-799.
37. Jonsen AR. Bentham in a box: technology assessment and health care allocation. *Law, Medicine and Health Care* 1986; 14: 172-174.
38. Hadorn DC. Setting health priorities in Oregon. Cost-effectiveness meets the rule of rescue. *Journal of the American Medical Association (JAMA)* 1991; 265 (17): 2218-2225.
39. Nord E. *Cost-value analysis in health care: making sense out of QALYs*. Cambridge: Cambridge University Press, 1999.
40. Ubel PA. *Pricing life: why it's time for health care rationing*. Cambridge, MA, London: The MIT Press, 2000.
41. McKie J, Richardson J. The rule of rescue. *Social Science & Medicine* 2003; 56 (12): 2407-2419.
42. Singer P, McKie J, Kuhse H, Richardson J. Double jeopardy and the use of QALYs in health care allocation. *Journal of Medical Ethics* 1995; 21 (3): 144-150.
43. Harris J. Double jeopardy and the veil of ignorance: a reply. *Journal of Medical Ethics* 1995; 21 (3): 151-157.
44. McKie R, Kuhse H, Richardson J, Singer P. Double jeopardy, the equal value of lives and the veil of ignorance: a rejoinder to Harris. *Journal of Medical Ethics* 1996; 22 (4): 195-196.

45. Menzel P, Gold MR, Nord E, et al. Toward a broader view of values in cost-effectiveness analysis of health. *Hastings Center Report* 1999; 29 (3): 7-15.
46. Richardson J, McKie J. Empiricism, ethics and orthodox economic theory: what is the appropriate basis for decision-making in the health sector? *Social Science & Medicine* 2005; 60: 265-275.
47. Harris J. QALYfying the value of life. *Journal of Medical Ethics* 1987; 13 (3): 117-12.
48. Arnesen T, Nord E. The value of DALY life: problems with ethics and validity of disability adjusted life years. *British Medical Journal* 1999; 319: 1423-1425.
49. Harris J. It's not NICE to discriminate. *Journal of Medical Ethics* 2005; 31: 373-375.
50. Rawlins M, Dillon A. NICE discrimination. *Journal of Medical Ethics* 2005; 31: 683-684.
51. Harris J. Nice and not so nice. *Journal of Medical Ethics* 2005; 31: 685-688.
52. Claxton K, Culyer AJ. Wickedness or folly? The ethics of NICE's decisions. *Journal of Medical Ethics* 2006; 32: 373-377.
53. Harris J. NICE is not cost effective. *Journal of Medical Ethics* 2006; 32: 378-380.
54. Claxton K, Culyer AJ. Rights, responsibilities and NICE: a rejoinder to Harris. *Journal of Medical Ethics* 2007; 33: 462-464.
55. Harris J. NICE rejoinder. *Journal of Medical Ethics* 2007; 33: 467.
56. Rawls J. *A theory of justice*. Cambridge: Harvard University Press, 1971.
57. Daniels N. *Reading Rawls*. Oxford: Blackwell, 1975.
58. Daniels N. Wide reflective equilibrium and theory acceptance in ethics. *Journal of Philosophy* 1979; 76: 256-282.
59. Nord E. Methods for quality adjustment of life years. *Social Science & Medicine* 1992; 34: 559-569.
60. Tersman F. *Reflective equilibrium: an essay in moral epistemology*. Stockholm: Almqvist & Wiksell International, 1993.
61. Stolk EA, Busschbach JJ, Caffa M, et al. Cost utility analysis of sildenafil compared with papaverine-phentolamine injections. *British Medical Journal* 2000; 320: 1165-1168.
62. King S, Riemsma R, Drummond M, et al. *A systematic review of the clinical and cost-effectiveness of methylphenidate hydrochloride, dexamfetamine sulphate and atomoxetine for attention deficit hyperactivity disorder (ADHD) in children and adolescents*. York: December 2004.
63. National Institute for Health and Clinical Excellence (NICE). *Methylphenidate, atomoxetine and dexamfetamine for attention deficit hyperactivity disorder*

- (ADHD) in children and adolescents. *Review of Technology Appraisal 13*. London: NICE, March 2006.
64. National Institute for Clinical Excellence (NICE). *Technology Appraisal Guidance No. 32: Multiple sclerosis – beta-interferon and glatiramer acetate*. London: NICE, January 2002.
  65. McGregor M. Cost-utility analysis: use QALYs only with great caution. *Canadian Medical Association Journal (CMAJ)* 2003; 168 (4): 433-434.
  66. Eddy DM. Oregon's methods. Did cost-effectiveness analysis fail? *Journal of the American Medical Association (JAMA)* 1991; 266 (15): 2135-2141.
  67. Mortimer D. The value of thinly spread QALYs. *Pharmacoeconomics* 2006; 24 (9): 845-853.
  68. Brock DW. Ethical issues in the use of cost effectiveness analysis for the prioritisation of health care resources. In: Anand S, Peter F, Sen A (eds) *Public Health, Ethics, and Equity*. Oxford: Oxford University Press, 2006: 201-223.
  69. Drummond M, Torrance G, Mason J. Cost-effectiveness league tables: more harm than good? *Social Science & Medicine* 1993; 37 (1): 33-40.
  70. Mauskopf J, Rutten F, Schonfeld W. Cost-effectiveness league tables: valuable guidance for decision makers? *Pharmacoeconomics* 2003; 21 (14): 991-1000.
  71. Gilovich T, Griffin D, Kahneman D (eds): *Heuristics and biases: the psychology of intuitive judgment*. Cambridge: Cambridge University Press, 2002.
  72. Sen AK *Inequality reexamined*. Cambridge: Harvard University Press, 1992.
  73. Sen AK. *Development as freedom*. New York: Alfred A. Knopf, 1999.
  74. Kagan S. *Normative ethics*. Boulder, CO: Westview Press, 1998.
  75. Nagel T. The fragmentation of value. In: Nagel T. *Mortal questions*. Cambridge: Cambridge University Press, 1979, 128-142.
  76. Daniels N, Sabin J. Limits to health care: fair procedures, democratic deliberation, and the legitimacy problem for insurers. *Philosophy and Public Affairs* 1997; 26 (4): 303-350.
  77. Daniels N, Sabin J. The ethics of accountability in managed care reform. *Health Affairs* 1998; 17: 50-64.
  78. Daniels N. Justice, health, and healthcare. *American Journal of Bioethics* 2001; 1 (2): 2-16.
  79. Daniels N, Sabin JE. *Setting limits fairly – can we learn to share medical resources?* Oxford: Oxford University Press, 2002.
  80. Hasman A, Holm S. Accountability for reasonableness: opening the black box of process. *Health Care Analysis* 2005; 14 (4): 261-273.
  81. Schlander M. NICE accountability for reasonableness: a qualitative study of its appraisal of treatments for attention-deficit/hyperactivity disorder (ADHD). *Current Medical Research and Opinion* 2007; 23 (1): 207-222.

82. Schlander M. Is NICE infallible? A qualitative study of its assessment of treatments for attention-deficit/hyperactivity disorder (ADHD). *Current Medical Research and Opinion* 2007; *in press*.
83. Schlander M. *Health technology assessments by the National Institute for Health and Clinical Excellence: a qualitative study*. New York: Springer, 2007 (*in press*).
84. Centre for Reviews and Dissemination (CRD): *CRD Report No. 4: Undertaking systematic reviews of research on effectiveness: CRD's guidance for those carrying out systematic reviews*. York: NHS Centre for Reviews and Dissemination, University of York: 2nd edition, March 2001.
85. Brennan A, Akehurst R. Modelling in health economic evaluation: What is its place? What is its value? *Pharmacoeconomics* 2000; 17 (5): 445-459.
86. Philips Z, Ginnelly L, Sculpher M, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technology Assessment* 2004; 8 (36): 1-174.
87. National Institute for Clinical Excellence (NICE) *Guide to the Technology Appraisal Process (reference N0514)*. London: NICE, April 2004.
88. Rawlins MD, Culyer AJ. National Institute of Clinical Excellence and its value judgments. *British Medical Journal* 2004; 329: 224-227.
89. Dakin HA, Devlin NJ, Odeyemi IAO. "Yes", "no" or "yes, but"? Multinomial modelling of NICE decision-making. *Health Policy* 2006; 77: 352-367.
90. Redwood H. *The use of cost-effectiveness analysis of medicines in the British National Health Service: Lessons for the United States*. (Expertise supported by the Pharmaceutical Research and Manufacturers of America, PhRMA.) Felixstowe, Suffolk: April 2006.
91. DiMasi JA, Hansen RW, Grabowski HG, et al. Cost of innovation in the pharmaceutical industry. *Journal of Health Economics* 2001; 10: 107-142.
92. Médecins Sans Frontières. *Fatal imbalance. The crisis in research and development for drugs for neglected diseases*. Geneva: Médecins Sans Frontières Access to Essential Medicines Campaign and the Drugs for Neglected Diseases Working Group, 2001.
93. McCabe C, Claxton K, Tsuchiya A. Orphan drugs and the NHS: should we value rarity? *British Medical Journal* 2005; 331: 1016-1019.
94. McCabe C, Tsuchiya A, Claxton K, Raftery J. Orphans drugs revisited. *QJM – Monthly Journal of the Association of Physicians* 2006; 99 (5): 341-345.
95. Hughes D. Rationing of drugs for rare diseases. *Pharmacoeconomics* 2006; 24 (4): 315-316.
96. Hughes DA, Tunnage B, Yeo ST. Drugs for exceptionally rare diseases: do they deserve special status for funding? *QJM – Monthly Journal of the Association of Physicians* 2005; 98 (11): 829-836.

97. Freemantle N, Bloor K, Eastaugh J. A fair innings for NICE? *Pharmacoeconomics* 2002; 20 (6): 389-391.
98. Gericke CA, Riesberg A, Busse R Ethical issues in funding orphan drug research and development. *Journal of Medical Ethics* 2005; 31: 164-168.
99. Connock M, Juarez-Garcia A, Frew E et al. A systematic review of the clinical effectiveness and cost-effectiveness of enzyme replacement therapies for Fabry's disease and mucopolysaccharidosis type 1. *Health Technology Assessment* 2006; 10 (20): 1-113.
100. NICE Citizens Council. *NICE Citizens Council Report: Ultra Orphan Drugs*. London: National Institute for Health and Clinical Excellence, November 2004.
101. Drummond M. Economic evaluation and decision-makers. In Jones AM (ed) *The Elgar Companion to Health Economics*. Cheltenham: Edward Elgar, 2006: 537-545.
102. Ubel PA, Loewenstein G. Distributing scarce livers: the moral reasoning of the general public. *Social Science & Medicine* 1996; 42 (7): 1049-1055.
103. Ubel PA, Loewenstein G. Public perceptions of the importance of prognosis in allocating transplantable livers to children. *Medical Decision Making* 1996; 16 (3): 234-241.
104. Birch S, Gafni A. Decision rules in economic evaluation. In Jones AM (ed) *The Elgar Companion to Health Economics*. Cheltenham: Edward Elgar, 2006: 492-502.
105. Sendi P, Gafni A, Birch S. Opportunity costs and uncertainty in the economic evaluation of health care interventions. *Health Economics* 2002; 11 (1): 23-31.
106. Audit Commission. *Managing the financial implications of NICE guidance*. (September 08, 2005) Available online at: <http://www.audit-commission.gov.uk/reports/national-report.asp?categoryid=&prodid=cc53ddfe-42c8-49c7-bb53-9f6485262718&fromreportsanddata=national-report&page=index.asp&area=hplink>. Last accessed December 20, 2005.
107. McGregor M. What decision-makers want and what they have been getting. *Value in Health* 2006; 9 (3): 181-185.
108. Schlander M. Zur Logik der Kosteneffektivität. *Deutsches Ärzteblatt* 2003; 100: 2110-2111.
109. Schlander M. *Lost in translation? Over-reliance on QALYs may lead to neglect of relevant evidence*. Paper presented at the 6th World Congress on Health Economics, International Health Economics Association (iHEA), Copenhagen July 8-11, 2007. Available at: [www.michaelschlander.com/presentations.htm](http://www.michaelschlander.com/presentations.htm).
110. National Institute for Clinical Excellence (NICE) *Technology Appraisal Process: Guidance for Appellants (reference N0520)*. London: NICE, April 2004.

111. Jefferson T, Demicheli V, Vale M. Quality of systematic reviews of economic evaluations in health care. *Journal of the American Medical Association (JAMA)* 2002; 287 (21): 2809-2812.
112. Reinhardt UE. Making economic evaluations respectable. *Social Science & Medicine* 1997; 45 (4): 555-562.
113. Hill SR, Mitchell AS, Henry DA. Problems with the interpretation of pharmacoeconomic analyses: a review of submissions to the Australian Pharmaceutical Benefits Scheme. *Journal of the American Medical Association (JAMA)* 2000; 283 (16): 2116-2121.
114. Sheldon TA, Cullum N, Dawson D, et al. What's the evidence that NICE guidance has been implemented? Results from a national evaluation using time series analysis, audit of patients' notes, and interviews. *British Medical Journal* 2004; 329: 999-1006.
115. Freemantle N. Is NICE delivering the goods? *British Medical Journal* 2004; 329: 1003-1004.
116. National Institute for Health and Clinical Excellence (NICE): *About the citizens council*. Available online at: [www.nice.org.uk/page.aspx?o=113692](http://www.nice.org.uk/page.aspx?o=113692). Last accessed January 7, 2007.
117. National Institute of Health and Clinical Excellence (NICE): *Social Value Judgements. Principles for the development of NICE's guidance. Post-consultation draft*. London: NICE, September 21, 2005.
118. National Institute of Health and Clinical Excellence (NICE): *Social Value Judgements. Principles for the development of NICE's guidance*. London: NICE, December 08, 2005.
119. Davies C, Wetherell M, Barnett E, Seymour-Smith S. Opening the box. Evaluating the Citizens Council of NICE. Report prepared for the National Co-ordinating Centre for Research Methodology, NHS Research and Development Programme. Milton Keynes: The Open University, March 2005.
120. Williams A. *What could be nicer than NICE? Annual Lecture 2004*. London: Office of Health Economics, 2004.

*Competing Interest: None Declared.*

The corresponding author has the right to grant on behalf of all authors and does grant on behalf of all authors an exclusive license on a worldwide basis to the BMJ Publishing Group Ltd and its Licensees to permit this article (if accepted) to be published in JME editions and any other BMJ PGL products to exploit all subsidiary rights, as set out in our license (<http://jme.bmjournals.com/ifora/licence.dtl>).

At the same time, the corresponding author requests permission to use some of this material in his book "*Health Technology Assessments by the National Institute for*

*Health and Clinical Excellence (NICE): A Qualitative Study*” to be published with Springer Science + Business Media. Permission is requested to use this material with word rights in all languages, in this and any future editions, revisions, adaptations or translations, and in any other form including, without limitation, mechanical & electronic storage on any carrier. Springer is a signatory of the STM guidelines agreement.