Social Policy & Administration issn 0144–5596
Vol. 39, No. 5, October 2005, pp. 498–512

Health Policy and the Politics of Evidence

Timothy Milewa and Christine Barry

Abstract

National decisions on the drugs, treatments and medical devices that should be funded through public expenditure are a fundamental element of health policy. But despite a political emphasis upon evidence-based policy, the results of rigorous clinical trials and statistical modelling techniques rarely speak for themselves. So, does the pre-eminence traditionally accorded to quantitative data in the medical field underpin policy decisions on a consistent basis? Or are more subtle, less transparent characteristics of context and interaction evident in the shaping of attendant decisions? This article considers these questions by drawing on a study of decision-making in the National Institute for Health and Clinical Excellence (NICE)—an organization established by the British government in 1999 to decide whether selected health technologies should be made available throughout the National Health Service in England and Wales. In broad terms, the findings point to the primacy of arguments based on quantitatively oriented, experimentally derived data but also to a discursive hegemony of clinicians and health economists in mediating, including or debarring more qualitative, experientially based evidence. A more complex, dynamic understanding of policy governance in the field of health technology appraisal—founded on a discursive appropriation of the idea of the "common good"—goes some way to explaining the persistence of this hegemony despite an avowedly inclusive, plural approach to decision-making.

Keywords

Health technology assessment; Health technology appraisal; NICE; Health policy

Introduction

Over the past fifteen years, countries such as the United Kingdom, Canada and New Zealand have sought to institutionalize and systemize arrangements for health technology assessment (HTA). The broad aim of HTA is to provide policy-makers and clinicians with reliable, scientifically robust evidence on the clinical and cost effectiveness of particular drugs, medical appliances or therapies (Sackett *et al.* 1996: 71). But HTA and attendant strands of health

Address for correspondence: *T. Milewa, Sociology, School of Social Sciences and Law, Human Sciences Building, Brunel University, Uxbridge, UB8 3PH, UK. Email: timothy.milewa@brunel.ac.uk*

policy do not exist in a social or political vacuum. Since the early 1990s successive British governments have emphasized patient empowerment in negotiating treatment, patient choice in the location and timing of health care and notions of the "expert patient" as a source of authority and advice (Department of Health 2001a, 2001b, 2002, 2003). These trends, debates on health care rationing (of highly variable visibility and coherence at a policy level) and, more recently, a political commitment by the New Labour government to transparency and inclusiveness in public sector decision-making helped to fuel the creation in 1999 of the National Institute for Clinical Excellence (renamed National Institute for Health and Clinical Excellence in April 2005) (Mills 2000; Litva *et al.* 2002; Milewa 2004; Quennell 2003).

This organization is charged with the systematic and inclusive appraisal of evidence pertaining to selected drugs, therapies and medical devices ("health technologies") with a view to their routine admissibility or exclusion within the National Health Service in England and Wales. But the results of rigorous clinical trials and sensitive modelling techniques tell us little about how data on clinical and cost effectiveness are interpreted at the level of national health policy formulation. Does the pre-eminence traditionally accorded to quantitative data in the medical field underpin policy decisions on a consistent basis? Or—in a country where the vast majority of the population relies on publicly funded health care—do patient, public and political concerns expressed in the deliberative process shape attendant decisions on the admissibility of health technologies? This article considers these questions by drawing on a study of decision-making in the National Institute for Health and Clinical Excellence (NICE).

Evidence, Deliberation and Policy Guidance

An initial indication of the scope for interpretation and discretion in the road to policy guidance on the adoption or exclusion of health technologies can be seen in the nature of the process administered by NICE. Anyone, for example, can suggest that a particular health technology be appraised (although such proposals are screened by at least two committees and usually have to be sanctioned by the Department of Health or National Assembly for Wales). NICE then takes the lead in identifying those consultees or "stakeholders" seen to have an interest in the adoption or rejection of a technology chosen for appraisal. Consultees may include pharmaceutical companies, professional associations (such as the Royal College of General Practitioners), patient advocacy organizations, and local health care providers. Discussion at this stage is used to refine the scope of the appraisal in order to clarify the questions and issues that will be addressed. At this stage NICE commissions an independent Technical Assessment Group, often based in a university, to undertake a systematic review of evidence (clinical, economic and "rigorous" qualitative data) with regard to the technology in question. The latter's report is circulated for comment among consultees and, possibly, is further revised.

The key body in the process, the Appraisal Committee, then meets to consider a synthesis of all submissions and additional representations from invited (and sometimes non-invited) consultees (three "branches" of the

committee have since been established). Of the 39 members of the Appraisal Committee in post at the time of the research: six were health economists based in universities; five were medical statisticians; three were public health specialists in academic posts and 17 clinicians (three-quarters of whom also had substantive roles in academe). The remaining members included five managers or chief executives from within the NHS, two members of patient advocacy bodies and an individual from a pharmaceutical company whose professional role centred upon drug safety. Returning to the process, comments from this membership on the submissions and additional representations attending an appraisal are then synthesized by the group's Chair in conjunction with NICE personnel. A further meeting of the Appraisal Committee considers this synthesis and then issues a decision in the form of a "Final Appraisal Determination". Even this outcome can, under certain criteria, be challenged by referral to an Appeal Panel.

These junctures and deliberative spaces within the appraisal process mark NICE as an institutional terrain within which even the most "objective" procedural and methodological approaches to the appraisal of health technology are influenced by the aims, strategies and power of participating individuals and actor-groups. It is within this terrain that boundaries—denoting competence and appropriate knowledge with regard to sorting good evidence from bad—are drawn by scientists, bureaucrats, policy-makers and an assortment of lay interests in relation to specific health technologies. In theory, however, such boundaries and criteria should be quite straightforward. In terms of cost effectiveness, there exist a variety of approaches to the calculation and modelling of costs—approaches that may take into account factors such as life years gained, quality of life, patient compliance with treatment regimes and knowledge about patient preferences between forms of treatment (Dixon *et al.* 1997; Young *et al.* 2003). Corresponding evaluative criteria appear even clearer with regard to clinical effectiveness. Evidence-based medicine and the developing focus upon health technology assessment are routinely framed in terms of a positivist scientific tradition that, particularly through the work of the Cochrane Collaboration, is aligned with the ideal-typical "hierarchy of evidence".

In brief, the "gold standard", or "Level I", in the hierarchy reflects evidence based on the systematic review of multiple randomized controlled trials that nearly all report findings of statistical significance. Level II evidence points to a number of randomized controlled trials that tend—with some exceptions—to confirm or discount evidence for a health care technology to the point of statistical significance. The downward direction continues to Level III wherein evidence relates to findings from methodologically robust trials that are not necessarily randomized, data from cohort studies or case-controlled analytical studies. And, still further from the ideal, Level IV evidence is accrued from non-experimental research (in other words, the research design does not attempt to control for all extraneous factors that may influence the outcome of the research). The nadir, at least in terms of this mindset, is reached at Level V—evidence is based on studies that lack controls, such as findings that arise from case studies, descriptive studies or unstructured consultation with experts (Leys 2003: 219).

These ideal types do not accommodate real-world formal or informal attempts to integrate qualitative data or explicitly social or political priorities with quantitative data—an approach exemplified by the development of a Patient Involvement Unit within NICE designed to facilitate the representation of "lay" concerns in a manner appropriate to formal deliberation (National Institute for Clinical Excellence 2003). But the positivist experimental orientation of contemporary biomedicine is evident in findings that suggest that more qualitative or subjective forms of evidence are routinely downplayed by medical professionals and policy-makers (Fenton and Charsely 2000; Ray and Mayan 2001; Upshur 2001; Little *et al.* 2002). Related accounts have emphasized the discursive pre-eminence of science in general and medicine in particular in defining the procedures, characteristics and evaluative criteria by which evidence is judged to be sound or unsound and actors characterized as qualified or unqualified to offer informed comment. Indeed, it has been suggested that, before the mid-1990s, regulation concerned with health technologies was often premised on a view that greater transparency and consumer involvement would render rigorous scientific deliberation open to a deleterious political and public pressure (Abraham and Lewis 2002).

In a second respect, however, a growing body of work has focused upon the social, or interactional, aspects of judgements on health technologies within particular institutional settings, such as health care planning bodies and individual hospitals (Timmermans and Berg 2003; Jenkings and Barber 2004; Dobrow *et al.* 2004). So, turning to the National Institute for Health and Clinical Excellence, we were interested in the degree to which the appraisal of health technologies reflected these social or contextual factors. Do clinical and cost effectiveness data tend to speak for themselves in the formulation of policy guidance on the admissibility of health technologies, as the hierarchy of evidence might suggest? Or does the reality of professional and bureaucratic discretionary power and lobbying by stakeholders intrude to a discernible and significant extent upon the supposedly depoliticized adjudication on the adoption of health technologies?

Methods

We used two qualitative methods to address these questions. First, preparatory meetings with personnel from within NICE, a review of public-domain documents pertaining to the appraisal process and strands of the academic literature outlined above were used to develop and test a semi-structured interview guide. Topics centred upon the identification and recruitment of participants in the appraisal of health technologies, strategies and alliances exhibited by actor-groups, the handling of appeals against appraisal decisions and interaction within relevant meetings. This preparatory phase was also used as a starting point for the initiation of a purposive snowball sample ($n = 33$). The intention here was to speak to a selection of authoritative informants in relation to key stages in the health technology appraisal process. Informants included NICE employees ($n = 6$); members of the Appraisal Committee ($n = 7$); consultees and nominated experts at Appraisal Committee

meetings with a "professional" background in terms of clinical practice, the manufacture of health technologies or management of the NHS ($n = 9$) and consultees and nominated experts at Appraisal Committee meetings with a "lay" background in areas such as patient advocacy or the receipt of relevant medical care ($n = 7$). Other informants were based in the independent Technology Assessment Groups ($n = 2$), and the Department of Health and the NHS ($n = 2$). These interviews, lasting an average of 55 minutes each, were tape-recorded and transcribed.

In a second respect, meetings within NICE were selected for observation with a view to encompassing specific stages in the appraisal process. The first such meeting was a briefing seminar organized by NICE for consultees—such as patient advocacy groups and manufacturers—in relation to the planned appraisal of cannabinoids (drugs based on cannabis) as a treatment for the symptoms of multiple sclerosis. Two meetings of the Appraisal Committee were also observed. One session centred upon a drug treatment for aggressive non-Hodgkin's lymphoma; the other covered a drug treatment for rheumatoid arthritis and endometrial ablation (surgical techniques for the treatment of heavy menstrual bleeding). The fourth meeting observed was that of the Appeal Panel, convened in relation to a decision not (on two occasions) to approve a specific growth hormone. Particular attention was given to interaction within these meetings and their management (the time accorded to different participants, the way in which contributions were managed by the Chair, and spatial arrangements). Data capture in the briefing seminar and two meetings of the Appraisal Committee centred on contemporaneous note-taking, post-meeting discussions with selected participants and examination of written submissions. The Appeal Panel was recorded, as a condition of access, by means of non-contemporaneous notes after its observation, study of corresponding minutes, examination of written submissions and non-standard *post hoc* discussion with selected participants. Notes pertaining to each meeting were collated and supplemented on the same day or on the next working day.

In terms of the analysis, interview transcripts were coded with reference to an iterative, incremental coding frame based in the first instance on the research themes in the original schedule and attendant indicators. A similar approach was adopted with regard to analysis of notes made at the meetings. Penultimate paper-based codings were then digitized, revised further and interrogated with the use of *QSR NVivo* software. Final interpretations were informed by annotated telephone discussions, correspondence and documentation and discussion between the two investigators. We focus here upon influences, beyond clinical and cost-effectiveness data alone, seen by informants to act upon decision-making in the appraisal process.

The Framing and Prioritization of Evidence

Perhaps not surprisingly, the framing and prioritization of evidence in relation to different health technologies reflected a concern with more quantitative, experimentally derived data over partial or qualitative submissions. This was not, however, straightforward. The key decision-making

body, the Appraisal Committee, could be seen—on a highly variable basis—to interpret and act upon such data in the light of two particular factors. These centred on the views of committee members on the "appropriate" presentation of evidence by stakeholders and variation in the tendency to take into account additional "considerations" (in addition to formal submissions).

Specifically with regard to perceptions of appropriate evidence, 12 informants made detailed reference to a specific distaste on the part of the Appraisal Committee for emotive or "political" presentations:

> *"Groups such as [a charity concerned with a fatal degenerative disease] just give you the entire textbook about the charity and the condition they deal with and why it is absolutely vital that you ignore the cost-effectiveness data (because, morally, that is what is expected of you). They might as well have torn up the paper and thrown it in the bin rather than post it to us."* (Appraisal Committee member, ap13:10,11)

> *"If there are high quality randomized controlled trials in a particular area of symptomology (like pain, stiffness, spasm) then that would drive what the appraisal says. And what will not drive the appraisal would be 'And when I use the cannabis drugs I feel much better but I don't know how to explain it' and there are no trials showing that."* (Appraisal Committee member, ap21:4)

Another nine informants took the view that the committee favoured a synthesis of quantitative evidence supplemented by personal testimonies. This was particularly the case where a representation complemented the Appraisal Committee's interpretation of quantitatively oriented data:

> *"I think the whole committee realized that—in order to make good judgements that did balance the clinical, the cost effectiveness (in other words, the trials)—real clinical experience . . . is very important . . . We had someone come from one of the groups as a patient representative. And, with the [quantitative evidence] we had, it was an extremely persuasive explanation as to why . . . people with hearing impairments found these new digital devices so valuable."* (Appraisal Committee member, ap19:7,9)

A second emphasis centred upon recourse to "considerations"—factors additional to formally presented evidence that might sway the Appraisal Committee. Ten informants (including three experts/consultees invited to participate in particular appraisals) mentioned a variety of possible considerations. These included the Appraisal Committee's thoughts on the degree to which a new technology constituted a significant innovation that should be supported (despite less than decisive evidence in its favour); awareness that evidence relating to some drugs is obscured by their use beyond the terms of their licence and the durability of evidence (technological or medical advances might alter a specific evidence-base markedly). Another potentially significant factor related to committee decisions on the admissibility of late or unsolicited evidence:

"Sometimes [unregistered consultees] send us unsolicited comments . . . I think the policy is not to treat them as, sort of, like formal consultees. But we read them, definitely. And if there's any sort-of, like, valid issue, that would be raised in some way during the [Appraisal Committee meeting]." (NICE employee, n26:10)

This point and the other data accentuate the issue of discretion on the part of "insiders" within NICE. But, as the findings below indicate, those arriving from outside (in pursuit of particular appraisal outcomes rather than administration of the process) could exercise some degree of influence beyond the simple presentation of evidence.

Influencing decisions from outside

This influence on the part of external stakeholders (such as representatives of patient advocacy bodies) appeared to centre on four discursive and strategic repertoires. The latter encompassed the production of "new" evidence: the timing and accentuation of evidence: the building of alliances and, finally, lobbying. In terms of the first such strategy, 15 informants emphasized the conduct or commissioning of independent research by stakeholders in order to produce new evidence for inclusion in the appraisal or to challenge existing findings—a questioning, in effect, of the degree to which preparatory work by NICE and technical assessment groups had covered relevant material. As 11 consultees/experts ("lay" and "professional" in background) indicated, this was particularly the case with regard to patient advocacy groups, pharmaceutical companies and professional associations:

"I'd put together a small survey questionnaire [on growth hormone treatments] that we sent to all the . . . people on our database . . . We got an amazing response rate. I can't remember what it was now—it was sort of like in the nineties (the response rate) . . . Our last submission was basically [built] on the results of that." (Lay expert/consultee, ep30–4)

"There is very little really good work done on the health economics of rheumatic diseases. So we asked an expert, as best as we could get, from another country (I think it was Sweden) to actually criticize the NICE's own commissioned report—and they did so. So we actually got an expert in the area to review their work." (Professional expert/consultee, emp7–3)

Comment was also made on the different resources available to stakeholders in terms of commissioning independent research. A handful of informants noted that pharmaceutical firms would sometimes withhold findings from their own research if disclosure in the appraisal process might prove commercially disadvantageous:

"Some of the manufacturers, it would appear . . . don't like all this 'commercial in confidence' stuff—it's very, very difficult . . . How can we possibly make public and transparent judgements on things unless [manufacturers] are allowing [their] data to be

scrutinized by others? That's the very essence of peer review." (Appraisal Committee member, ap19–12)

In a second respect, the timing and accentuation of evidence or related data was mentioned by 14 informants (distributed fairly evenly across categories of interviewee). Particular topics included the approaches adopted by patient advocacy groups with regard to the use and timing of "unscientific" personal testimonies. Speaking in relation to an appraisal of cannabinoids, one lay expert/consultee observed:

"The [Appraisal Committee] might say 'No, we're not interested in this, we're only interested in quantitative data'. But we are constantly trying and pressing organizations to think about the impact [of multiple sclerosis] upon people's whole lives (which you can only generally get from telling their stories frankly, rather than just clinical measures). I think in this particular circumstance, [patient testimonies] would probably come into place if the [appraisal] decision was a tough one to take. If it's an easy one to take, [patient testimonies are] fairly irrelevant [to the Appraisal Committee]." (Lay expert/consultee, ep16–14)

The third factor accentuated by informants was that of the building of alliances by stakeholders. Five informants (three of whom were lay or professional experts/consultees) focused on the impact, actual or potential, of cooperation or co-ordination between different stakeholders. Most emphasis was placed on the development of links between stakeholders but reference was also made to learning from the experience of other consultees and the borrowing of "successful" presentational strategies. One informant, for example, a lay expert, described the relatively systematic dialogue and coordination that preceded a submission to the appraisal of a drug for the treatment of rheumatoid arthritis:

"We've had quite a number of tele-conferences with a variety of people listed as stakeholders. There were only four of us who were the expert witnesses who actually attended the appraisal process. But there were probably a group of 12 or 15 people involved in the tele-conferences . . . We wanted to make sure that we were covering every aspect. We all want this drug to be licensed." (Lay expert/consultee, ep27–5)

Interestingly, no direct mention was made of the "sponsorship" of patient advocacy groups by health technology manufacturers—a phenomenon that has excited some comment in recent years (Herxheimer 2003). But discussion of lobbying by stakeholders did encompass some claims of "unofficial" representations—in other arenas—to individuals involved in the appraisal process as well as mention of external levers of influence. In this latter respect one example of external lobbying was reflected in a specially brokered arrangement for the provision of a drug for the treatment of people suffering from multiple sclerosis. As this illustrates, the appraisal process is not hermetically sealed and it can benefit interested parties to lobby the government directly:

"In the end [a drug for the treatment of multiple sclerosis] was turned into a special project by the Secretary of State for Health purely because of the pressure that was put on the government by the [patient advocacy body for people with multiple sclerosis] . . . And there are also a few clinicians who have built up an empire treating MS who were backing up the [patient advocacy body for people with multiple sclerosis]. But if you look at the treatment . . . in terms of its cost-effectiveness, it's one of the least cost-effective drugs there is." (Appraisal Committee member, ap12:16)

Overall, though, this emphasis upon pro-active stakeholders in the appraisal process did not appear to challenge the general dominance of the clinical-economic discourse of cost and clinical effectiveness over that of personal, subjective testimonies. We saw, for example, an explicit institutional preference for the "appropriate" presentation of evidence by stakeholders and—usually—a low but variable significance attached to patient testimonies. But it was also clear that the appraisal process was not immune to lobbying.

Colonizing the Management of Uncertainty

We began by positing two possible tendencies with regard to deliberation within the National Institute for Health and Clinical Excellence on the admissibility of health technologies. In one respect, it was suggested that clinical and cost-effectiveness data might tend, effectively, to speak for themselves in influencing decisions. Alternatively, we speculated that such decisions may be demonstrably influenced by the actors and actor-groups in such a way that the idea of a wholly "objective" deliberative process comes into question. In broad terms, the findings point to the primacy of arguments based on quantitatively oriented, experimentally derived data but also to a discursive hegemony of clinicians and health economists in mediating, including or debarring more qualitative, experientially based evidence. This conclusion should be tempered by the investigation's limited scale and by the fact that the very specific nature of the appraisal process within NICE precludes direct comparison with similar arrangements abroad (usually concerned more with the technical aspects of health technology assessment than deliberative appraisal). With these qualifications in mind, how can the findings be interpreted?

At one level the findings can be taken at face value. They reflect a highly structured approach to deliberation and governance in which "rules about rules" ensure that each health technology under consideration by NICE is considered "consciously, collectively and comprehensively"—an aspect of what has been called a "Type I" form of governance (Marks and Hooghe 2004: 28). The inclusion of relevant stakeholders, in addition to a review of available data, in health technology appraisals provides a therapeutic and experiential context for technical details. Such context is not, however, intended to supplant the primacy of data derived from rigorous research. Accordingly, the apparent dominance of a clinical-economic discourse is self-explanatory. Those whose professional training and experience centre upon disciplines and methodologies apposite to clinical trials or health economics can speak with greater authority than the untrained when interpreting data

in order to inform policy. This dominance is reinforced on an ongoing basis by publicly acknowledged standards of professional training and qualifications, shared ideas of methodological rigour, common systems of peer reviews and collegial norms of appropriate professional practice and discourse (Gieryn 1995: 393, 399; Salter 2001).

This essentialist notion of health technology appraisal is persuasive but only to a limited degree. The procedural focus within NICE upon information meetings, solicited and unsolicited submissions from consultees, interrogation by the Appraisal Committee of those called to give evidence, and even a "court of appeal" within the system suggests a clearly delineated deliberative sphere. But, as some of the interviewees indicate, a focus upon health technology appraisal within NICE and social relations therein only tells part of the story. Informants referred, for example, to independent research by patient advocacy groups and health technology manufacturers, lobbying by stakeholders of the government and the influence of professional medical associations. There is thus no reason to suppose that any particular "public sphere of the political realm" (in this case the appraisal process administered by NICE) is closed in nature (Habermas 1992: 231). In relation, for example, to HIV/AIDS in the United States, the "Treatment Action Group"—among a number of such bodies—has attempted to blur the distinction between lay and expert knowledge in the development, evaluation and adoption of health technologies. The group was central to a review of all relevant research projects (costing $1.3 billion) funded by the National Institutes of Health in the 1990s. And in 1994 the Treatment Action Group used a high-profile campaign to persuade a large drugs company, Roche, and government regulators to double the number of people involved in a trial of a new drug, *Saquinavir* (Harrington 1997: 283).

This pattern is by no means exceptional—advocacy groups concerned with other conditions, such as multiple sclerosis and breast cancer, continue to focus upon the research, development, testing, acceptance and diffusion of relevant health technologies (Wood 2000; Brown *et al.* 2004). The governance of health technology appraisal can thus be seen as far less prescribed than previously implied. Indeed, we might think in terms of a "Type II" form of governance—a potential plethora of deliberative and decision-making spaces beyond those afforded by NICE "may be created, deleted or adjusted through inter-jurisdictional competition" (Marks and Hooghe 2004: 28).

But the very co-existence of the prescribed, rule-bound, Type I governance of health technology appraisal and the far looser Type II pluralistic scenario suggests a need for some consideration of how the two forms are melded. In other words, we need to consider the factors that privilege some discursive spaces over others (with attendant discursive asymmetries between participants). One of the more obvious factors in this respect centres upon the framing and representation of particular deliberative spaces. The appraisal process within and through NICE has been framed in terms of transparency, inclusion and reasoned deliberation. This representation has been reinforced by a requirement that National Health Service providers in England and Wales make resources available to comply with guidance issued by NICE—provision within a health care system designed to provide comprehensive health care for all, regardless of individual ability to pay. NICE is thus portrayed as an institution

charged with one aspect of a commitment to a form of collective justice founded on the notion of an *a priori* equality among health service users/citizens (Marshall 1963). Health technology appraisal administered by NICE is thus seen as one means of pursuing—as far as practical—a "common good".

The reality of finite resources in the face of continuing demand for health technologies remains. But the deliberative and pluralistic methodology of appraisal is seen to link diverse stakeholders in a common enterprise with a shared objective—balancing the needs of taxpayers and patients by ensuring, as far as possible, that health technologies used in a collectively funded health care system have been considered for their clinical and cost effectiveness. Decision-making in the name of the common good by Appraisal Committee members, in the wake of suitable consultation with invited stakeholders, is thus "something that emerges from an initial commitment to the group. It is a 'group rationality' of people that develop a like mind as a consequence of their discussions and deliberations" (Tenbensel 2002: 185). Subsequent decisions may be contestable or controversial but they are seen to arise from a system based upon explicit procedures and consensus or compromise among those deemed to know best.

This line of argument is far from uncontentious. Reliance on decision-making by experts runs the risk of conflating "the ideas of deliberation and the common good by assuming that deliberation must be deliberation about the common good" (Fraser 1992: 130). Deliberative arenas founded on notions of determining the optimum (or even least worst) outcome for the common good—expressed here in the form of rationally determining which health technologies should be available in a collectively funded health service—assume that participants can set aside personal or partisan interests and proceed to engage in debate on behalf of wider society. This assumption—even before decisions on health technologies are made, accepted or contested—is in itself contentious. And perhaps even more uncertainty arises if the basic components of analyses in health economics—such as maximum utility, cost effectiveness and quality adjusted life years—are linked upon an *a priori* basis to ideas of the common good. Were such a link to be drawn, the very disciplinary "language" that underpins much of the appraisal process—regardless of its applications—would be privileged over alternative discursive forms.

But if we set aside this uncertainty and accept the "common good" thesis, the appraisal process will still depend for its viability on two factors. The first centres upon prevailing attitudes towards experts and expertise within the state or in those institutions operating on its behalf. For a large part of the postwar period sociological approaches to expertise suggested that the development of a highly complex state administration has facilitated the rise of interests based upon the possession of technological, economic and bureaucratic expertise—an expertise seen as necessary, if not broadly beneficial, to society as a whole. But the supposed rise of a far more atomized and fragmented society, together with greater disenchantment with comprehensive state intervention, has, it is often claimed, been reflected in an increasing willingness on the part of interest groups (and politicians) to challenge such authority (Pickstone 2000; Tovey *et al.* 2001). This challenge can manifest itself in direct conflict or, instead, stimulate reforms in the governance or

oversight of particular policy domains to render them more inclusive or transparent (Chandler 2001). In the case of health technology appraisal, debates upon implicit or explicit health care rationing that emerged in the 1980s plus the continuing electoral importance of the British health service perhaps stimulated a preventative deflection of such challenges through an emphasis upon inclusion and transparent deliberation in decision-making within NICE. The basis of this strategy does not reside, however, in direct deference to an electorate or even in a neo-pluralist system through which rival parties seek to persuade their peers that they should, at least temporarily, control decision-making levers (Held 1996: 99, 217). Instead, through the pivotal role of NICE and the Appraisal Committee in mediating views from invited stakeholders in relation to specific health technologies, the state appears to be operating a highly focused form of "episodic" quasi-corporatism designed to generate informed consensus ("quasi" by virtue of the fact that stakeholders are invited to participate primarily on the basis of their perceived expertise rather than formal representative roles).

But in a second, more nuanced regard, the common good thesis can be refined further if we concede that NICE, by its very nature, challenges and reframes the distinction made at the beginning of this article—the contrast between an "objective" decision-making process based upon "facts" and the substantive/moral (or interest-oriented) principles with which partisan actors might seek to infuse the process. The NICE appraisal system appears to operate according to dispassionate procedures that admit few substantive principles beyond a fundamental notion of the common good (founded on access to health care on the basis of citizenship and a reasoned, evidence-based approach to decisions that might not meet the needs or wants of certain citizens). Yet as Gutmann and Thompson (2000) argue, the dispassionate or process-oriented approach to decision-making—in a democratic polity and in relation to something as fundamental as comprehensive, publicly funded health care—depends upon more than the foundational value of reciprocity and trust between actors (necessary to any idea of elite governance in the name of the common good). In other words, something more than fairness and objectivity is required. In the case of NICE this involves an acceptance that appraisal decisions are necessarily provisional and thus ultimately contestable. The positivist paradigm (seen earlier to privilege experimentally derived data) necessarily admits the possibility of new evidence (Gutmann and Thompson 2000: 14). Intelligible propositions regarding the efficacy and cost effectiveness of health technologies are only seen as viable because there exists an awareness of circumstances and data, however unlikely, that would act to nullify their veracity. This evidence may take the form of new or better data from recently reported clinical trials; *post hoc* critiques of experimental methodology (perhaps several years after original findings were reported); innovation in health technologies such as drugs and prosthetics and challenges to current interpretations of data.

In other words, "professional" influence over the appraisal of health technologies rests as much on systems and procedures designed to manage uncertainty as it does on the ability to declare a body of evidence decisive. In this light—although our data revealed occasional dissatisfaction among informants

with the appraisal process and some disappointment with decisions made by the Appraisal Committee—attacks on the legitimacy or viability of the system in place were non-existent. This suggests, initially at least, that the "common good" representation of deliberative governance in this field holds sway. The simultaneous procedural emphasis upon inclusion and transparency does, however, point to a selective and pervasive (but possibly quite fragile) divorce between elements of form and function within bodies such as NICE. The perceived legitimacy of the appraisal system among the diverse actor-groups we encountered may rest on an inclusive, pluralistic form of Type II participation *in tandem* with an organizational function based upon a more rarefied, exclusive, clinical-economic discourse oriented to Type I governance. In other words, arrangements for broadly drawn participation and consultation are not necessarily synonymous with mechanisms of control but can co-exist in the same deliberative space. The appraisal of health technologies in this particular context thus reflects two characteristics of pluralistic deliberation—access by stakeholders from outside and the opportunity for these stakeholders to express their views. But a third characteristic, joint decision-making, is limited by the *a priori* hegemony of a form of discourse associated primarily with specific actor-groups—in this case clinicians, positivist-oriented medical researchers and health economists.

Conclusion

This article focused upon the way in which data on the clinical and cost effectiveness of drugs, therapies and medical devices is interpreted within a system of appraisal designed to offer clear policy guidance on the adoption or exclusion of health technologies in the British National Health Service. Perhaps not surprisingly, interviews with key informants and observation of meetings administered by NICE suggest that a discourse founded on ideas of clinical and cost-effectiveness data tended to take precedence over more subjective, experientially based perspectives. But a focus on the detail of decision-making runs the risk of overlooking the more fundamental issue of how this key aspect of health policy governance is framed and represented to the citizenry, who collectively fund the appraisal of health technologies and who may, individually, experience the profound impact of attendant decisions. A more complex understanding of policy governance in the field of health technology appraisal—founded on a discursive appropriation of the "common good"—goes some way to explaining a clinical-economic discursive hegemony in an avowedly inclusive, plural approach to decision-making. In short, power in the sifting and evaluation of evidence does not centre on the ability to dispel uncertainty. Instead, influence accrues to those who can cast themselves as best equipped to manage uncertainty.

Acknowledgements

This article is based in part on research funded by Brunel University. Cooperation extended by the National Institute for Health and Clinical Excellence, interviewees and their organizations is acknowledged with gratitude,

as are referee comments. The interpretations expressed in this article are those of the authors and should not be taken to represent the opinions of the National Institute for Health and Clinical Excellence or the funding body.

References

Abraham, J. and Lewis, G. (2002), Citizenship, medical expertise and the capitalist regulatory state in Europe, *Sociology*, 36, 1: 67–88.

Brown, P., Zavestoski, S., McCormick, S., Mayer, B., Morello-Frosch, R. and Altman, R. (2004), Embodied health movements: new approaches to social movements in health, *Sociology of Health and Illness*, 26, 1: 50–80.

Chandler, D. (2001), Active citizens and the therapeutic state: the role of democratic participation in local government reform, *Policy and Politics*, 29, 1: 3–14.

Department of Health (2001a), *The Expert Patient: a New Approach to Chronic Disease Management for the 21st Century*, London: Stationery Office.

Department of Health (2001b), *Extending Choice for Patients: A Discussion Document*, London: Department of Health.

Department of Health (2002), *Extending Choice for Patients: Information and Advice on Establishing the Heart Surgery Scheme: Draft for Stakeholder Consultation*, London: Department of Health.

Department of Health (2003), *Building on the Best: Choice, Responsiveness and Equity in the NHS*, London: Department of Health.

Dixon, P., Gravelle, H., Carr-Hill, R. and Posnett, J. (1997), *Patient Movements and Patient Choice* (report), York: York Health Economics Consortium, University of York.

Dobrow, M., Goel, V. and Upshur, R. (2004), Evidence-based health policy: context and utilizations, *Social Science and Medicine*, 58: 207–17.

Fenton, S. and Charsley, K. (2000), Epidemiology and sociology as incommensurate games: accounts from the study of health and ethnicity, *Health*, 44, 4: 403–25.

Fraser, N. (1992), Rethinking the public sphere: a contribution to the critique of actually existing democracy. In C. Calhoun (ed.), *Habermas and the Public Sphere*, Cambridge, MA: MIT Press.

Gieryn, T. (1995), Boundaries of science. In S. Jasanoff, G. Markle, J. Petersen and T. Pinch (eds), *Handbook of Science and Technology Studies*, Thousand Oaks, CA: Sage.

Gutmann, A. and Thompson, D. (2000), Deliberative democracy beyond process. Conference paper at *Deliberating about Deliberative Democracy*, Austin, Texas, 4–6 February.

Habermas, J. (1992), *The Structural Transformation of the Public Sphere: an Inquiry into a Category of Bourgeois Society*, Cambridge: Polity Press.

Harrington, M. (1997), Some transition in the history of AIDS treatment activism. In J. Oppenheimer and H. Reckitt (eds), *Acting on AIDS: Sex, Drugs and Politics*, London: Serpent's Tail.

Held, D. (1996), *Models of Democracy* (2nd edn), Cambridge: Polity Press.

Herxheimer, A. (2003), Relationships between the pharmaceutical industry and patients' organizations, *British Medical Journal*, 326, 7400: 1208–10.

Jenkings, K. and Barber, N. (2004), What constitutes evidence in hospital drug decision making? *Social Science and Medicine*, 58: 1757–66.

Leys, M. (2003), Health care policy: qualitative evidence and health technology assessment, *Health Policy*, 65: 217–26.

Little, M., Jordens, C. F. C., Paul, K., Sayers, E., Cruikshank, J. A., Stegeman, J. and Montgomery, K. (2002), Discourse in different voices: reconciling N = 1 and N = many, *Social Science and Medicine*, 55: 1079–87.

Litva, A., Coast, J., Donovan, J., Eyles, J., Shepherd, M., Tacchi, J., Abelson, S. and Morgan, K. (2002), "The public is too subjective": public involvement at different levels of health care decision-making, *Social Science and Medicine*, 54: 1825–37.
Marks, G. and Hooghe, L. (2004), Contrasting visions of multi-level governance. In I. Bache and M. Flinders (eds), *Multi-Level Governance*, Oxford: Oxford University Press.
Marshall, T. H. (1963), Citizenship and social class. In T. H. Marshall (ed.), *Sociology at the Crossroads and Other Essays*, London: Heinemann [orig. pub. 1950].
Milewa, T. (2004), Local participatory democracy in Britain's health service: innovation or fragmentation of a universal citizenship? *Social Policy & Administration*, 38, 3: 240–52.
Mills, F. (2000), *Patient Groups and the Global Pharmaceutical Industry: the Growing Importance of Working Directly with the Consumer*, London: Urch Publishing.
National Institute for Clinical Excellence (2003), *Patient Involvement Unit Newsletter*, no. 1 (December), London: National Institute for Clinical Excellence.
Pickstone, J. (2000), *Production, Community and Consumption: the Political Economy of Twentieth Century Medicine*. In R. Cooter and J. Pickstone (eds), *Twentieth Century Medicine*, Boston: Harvard Academic.
Quennell, P. (2003), Getting a word in edgeways? Patient group participation in the appraisal process in the National Institute for Clinical Excellence, *Clinical Governance*, 8, 1: 39–45.
Ray, L. and Mayan, M. (2001), Who decides what counts as evidence? In J. Morse, J. Swanson and A. Kuzel (eds), *The Nature of Qualitative Evidence*, Thousand Oaks, CA: Sage.
Sackett, D., Rosenberg, W., Gray, M., Haynes, R. and Richardson, W. (1996), Evidence based medicine: what it is and what it isn't, *British Medical Journal*, 312: 71–2.
Salter, B. (2001), Who rules? The new politics of medical regulation, *Social Science and Medicine*, 52: 871–83.
Tenbensel, T. (2002), Interpreting public input into priority-setting: the role of mediating institutions, *Health Policy*, 62: 173–94.
Timmermans, S. and Berg, M. (2003), *The Gold Standard: the Challenge of Evidence-based Medicine and Standardization in Health Care*, Philadelphia, PA: Temple University Press.
Tovey, P., Atkin, K. and Milewa, T. (2001), The individual and primary care: service user, reflexive choice maker and collective actor, *Critical Public Health*, 11, 2: 153–66.
Upshur, R. (2001), The status of qualitative research as evidence. In J. Morse, J. Swanson and A. Kuzel (eds), *The Nature of Qualitative Evidence*, Thousand Oaks, CA: Sage.
Wood, B. (2000), *Patient Power? The Politics of Patients' Associations in Britain and America*, Buckingham: Open University Press.
Young, T., Neuberger, J., Longworth, L., Ratcliffe, J. and Buxton, M. (2003), Survival gain after transplantation for patients with alcoholic liver disease: a comparison across models and centres, *Transplantation*, 76, 10: 1479–86.

