SCIENTIFIC CONTRIBUTION

NICE technology appraisals: working with multiple levels of uncertainty and the potential for bias

Patrick Brown · Michael Calnan

Published online: 25 December 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract One of the key roles of the English National Institute for Health and Clinical Excellence (NICE) is technology appraisal. This essentially involves evaluating the cost effectiveness of pharmaceutical products and other technologies for use within the National Health Service. Based on a content analysis of key documents which shed light on the nature of appraisals, this paper draws attention to the multiple layers of uncertainty and complexity which are latent within the appraisal process, and the often socially constructed mechanisms for tackling these. Epistemic assumptions, bounded rationality and more explicitly relational forms of managing knowledge are applied to this end. These findings are discussed in the context of the literature highlighting the inherently social process of regulation. A framework is developed which posits the various forms of uncertainty, and responses to these, as potential conduits of regulatory bias—in need of further research. That NICE’s authority is itself regulated by other actors within the regulatory regime, particularly the pharmaceutical industry, exposes it to the threat of regulatory capture. Following Lehoux, it is concluded that a more transparent and reflexive format for technological appraisals is necessary. This would enable a more robust, defensible form of decision-making and moreover enable NICE to preserve its legitimacy in the midst of pressures which threaten this.

Keywords Alzheimer’s · Complexity · Hope · Polycentric regimes · Regulation · Uncertainty

Introduction: the politicisation of uncertainty

In his seminal examination of the modern regulatory state in Britain, Moran (2003) describes the transition away from a stable governing compact based on ‘club rule’ to a period of hyper-innovation which occurred towards the end of the twentieth century. In this latter phase, a number of factors acting on the British system made it “uniquely pioneering” (Moran 2003, p. 155) in the “drive to subject areas of life not previously formally controlled to formal regulation with the aim of more synoptic legibility” (ibid., p. 153). The National Institute for Clinical Excellence1 emerged towards the end of this period of bureaucratic modernisation (thus fallings outside Moran’s assessment) in April 1999, and in a number of ways epitomises many of the features of the late-modern British regulatory state. NICE’s evidence-based approach, whether applied as clinical guidance or in the form of cost-containment recommendations, has elicited esteem from both sides of the Atlantic (Wall Street Journal 2009). Yet in this latter role, through making explicit the difficult decisions of healthcare rationing which are an inevitable part of a state-funded healthcare system, the institute faces legitimacy problems in England and Wales (Brown and Calnan 2010).

The implicit compact—between government, clinicians and public—by which the NHS initially functioned has been consistently weakened since the 1960s, and subjected to heightened strains post-1997 (Ham and Alberti 2002).

1 More recently renamed the National Institute of Healthcare and Clinical Excellence.
One of the central bases of this informal system of running the NHS was that the government determined the overall budget for the NHS whilst individual clinicians—GP gatekeepers and hospital consultants—decided healthcare priorities. The basic economic problem of healthcare—limited resources and ever-increasing demand—was in this way negotiated “through localized discretionary decisions of clinicians known as dilution” (Crinson 2004, p. 34). Yet because the legitimacy of any process of decision-making is always indirect and achieved in relation to other authorities (Habermas 1976, p. 101), a decline in the authority and esteem of medical professionals—at least in the eyes of policy-makers (Alaszewski 2002)—rendered this ad-hoc process of resource decisions increasingly untenable.

This perceived decline in professional authority, combined with other features of the NHS such as variations in care (Calman and Hine 1995) and spiralling costs across the service overall and for certain interventions in particular (Crinson 2004), led to the inauguration of a more formal means of regulating (limiting, assuring and adjudicating between) the availability of certain treatments as prescribed within the NHS. Yet a fundamental paradox surrounding NICE is that, in replacing the previous local/informal model of decision-making with a more bureaucratic regulatory apparatus, new layers of subjectivity and policy meddling have been introduced: “the new regulatory state, so often identified with the rise of neutral, non-majoritarian decision-making, has actually exposed hitherto ‘non-political’ domains to the power of elected politicians” (Moran 2003, p. 125). This politicisation is clearly apparent in interventions which have pre-empted (e.g Herceptin and beta-interferon) or legally contested (e.g donepezil) NICE recommendations on particular technologies.

A guiding assumption of this paper is that the authority of NICE—as a regulator—is itself regulated by sociopolitical and economic phenomena which in turn are strongly influenced by other actors (Brown and Calnan 2010), not least the mass media and the pharmaceutical industry. The regulatory regime within which NICE functions is therefore best understood as ‘polycentric’ (Black 2008); one where there are multiple state and non-state actors, obscured boundaries of power and knowledge, and which is “marked by fragmentation, complexity and interdependence between actors” (ibid., p. 1). So whilst a much more formal, centralised and (prima facie) transparent system of rationing has to some extent replaced the previous ‘club system’ of governance and resource decision-making (Moran 2003, p. 140), this new process is still highly and unavoidably relational, institutionally and socially embedded (Black 2008), and rife with uncertainty (Brown and Calnan 2010).

This latter, intractable feature of uncertainty attests to the limits of regulators in adjudicating between a range of competing concerns (Fuller 1978). Following a brief consideration of these competing interests and their relation to NICE decision-making, the thematic analysis which forms the core of this paper will be presented in describing a range of layers and forms of uncertainty which are more or less implicit within NICE appraisals (epistemic assumptions, bounded rationality, and relational forms of managing knowledge). These empirical findings then act as the basis of an extended discussion which develops a conceptual framework for future research; one concerned with the modes by which such uncertainties may form the basis of unforeseen and/or undesirable influence on, or capture of, NICE decisions. These influences may happen either more explicitly, where technical process may be legally challenged (Fuller 1978, pp. 387–388), or implicitly, where attempts to bridge over uncertainty privilege certain interests over others (Milewa 2008). A concluding section argues that while the problems of uncertainty are absorbed yet never fully solved by institutions such as NICE, regulatory formats which openly acknowledge and debate the multiple interests and values which are at stake (Lehoux 2006) may serve to heighten public legitimacy (Milewa 2008) and reduce the scope for technocratic challenges which subvert and skew effective decision-making.

Reconciling interests? Muddling through amidst exposed uncertainty

While the unknowables latent within bio-medical/pharmacological research and its application are far from novel, late-modern society is increasingly sensitive to such uncertainty (Taylor-Gooby 2000). Similarly the rationing typically assumed as inherent to public healthcare provision (c.f. Light and Hughes 2001) has become more explicit through public debate about the welfare state and the transition from local-informal to national-formal decisions invoked via NICE. In response to the augmented public scrutiny which has followed (as well as a legal ruling following the contesting of NICE’s decision regarding donepezil), the institute has made its appraisal format, calculations approach and cost-effectiveness threshold more transparent. Such accountability demands imposed on regulating agencies are far from unusual and reflect a wider, late-modern tendency—as referred to by Black (2008, pp. 10–11), following Power (2005)—where institutions are “turned ‘inside out’... the details of their internal decision-making structures and processes, including their incentive structures, audit and risk management processes, are seen as critically relevant to those outside them”. Thus the use of bureaucracy and proceduralism...
applied in attempting to ‘absorb uncertainty’ (Habermas 1976, p. 98) may actually serve to politicise decision-making and expose further uncertainty therein.

One explanation of this phenomenon is that the existence of NICE galvanises a ‘logic of collective action’ (Olson 1965) by which homogenous groups which stand to benefit from newly illuminated decisions are especially influential (see Meadowcroft 2008). The politicisation of decisions and demands for transparency by these groups highlight a range of contestable components within appraisals, pressurising NICE to change its approach and concede ground to these interests (Hedgecoe 2004, pp. 140–141). In contrast “more dispersed, heterogeneous groups for whom the costs of collective organization will be high” (Meadowcroft 2008, p. 434)—those who stand to lose from the disinvestment by PCTs in other services resulting from NICE recommended technologies—are much less active and therefore the uncertainty and arbitrariness around such disinvestment remains largely obscured.

With regard to this latter uncertainty, the blanket determining of what is affordable for the NHS as set out by NICE may result in recommendations against the prescribing of certain drugs which could otherwise be afforded by some local Primary Care Trusts (PCTs). Conversely NICE decisions may also compel other Trusts to forego other important services (undermining equity of access in certain areas) in order to be able to provide a particular product recommended by NICE (which PCTs are legally required to provide as a result). Which technologies are ‘disinvested’ in is of course very much at the discretion of local Trusts. The only way of preventing this enforced inequity/variation in care would be to: (a) ensure more equitable and efficient allocation and spending of money across the NHS; and moreover (b) to have a comprehensive threshold ruling and therefore cost-effectiveness appraisal of every aspect of healthcare treatment and technology provided across the entire NHS.

As Lindblom (1959, p. 79) makes apparent in his classic commentary on public administration: such a ‘root and branch’ review “would of course require a prodigious inquiry into values held by society and an equally prodigious set of calculations on how much of each value is equal to how much of each other value” (ibid. 1959:79). Thus as far as a fully rigorous assurance of equitable access to treatment is concerned, this “is of course impossible. It assumes intellectual capacities and sources of information that men [and NICE] simply do not possess” (ibid., p. 79). Given this impossibility of perfect decision-making (technically and normatively), Lindblom asserts that organisations such as NICE inevitably limit themselves to “the politically or legally possible—restrict their attention to relatively few values and relatively few alternative policies” (ibid., p. 79). Intrinsic to appraisals therefore are the profound assumptions of ‘constructed’ economic models (Light and Hughes 2001) and the “vacuum where the possible treatments for one condition are assessed relative to one another but without reference to alternative uses of those resources to treat other conditions” (Meadowcroft 2008).

This ‘muddling through’ (Lindblom 1959) of institutional decision-making is less benign than the term implies. Due to the multiple actors involved in polycentric regulatory regimes such as that of NICE’s technological appraisals, the influence upon judgements and recommendations can be described as “pluralistic, though grossly lopsided” (Lindblom 1979, p. 525; Fuller 1978). Not only may homogenous, well organised and resourced groups exert disproportionate influence (Meadowcroft 2008), but other interests may be ‘built in’ to the process. As Brown and Webster underline (2004, p. 181), new technologies “are saturated with talk of breakthroughs, advances, future visions and great leaps forward”—it is impossible for those making regulatory decisions to detach products from the compelling stories they are associated with. Because regulation and its means of coping with uncertainty remain profoundly relational (Black 2008), the potential influence of interests within these social processes requires significant investigation.

Abraham (1995) delineates a number of potential interest configurations, noting the corporate bias apparent through many of the regulatory mechanisms for drug safety and pricing (Abraham 2009). While NICE would seem to have been created in the public interest (optimising NHS provision) this stance should not be taken for granted. Regulatory regimes may be ‘captured’ by certain private interests or, more fundamentally, may be shaped by private interests in their inception and design. Corporatist and Marxist theories of regulation suggest, respectively, that states may share interests with industry or that public interests may be misconstrued as congruent with manufacturers. Linkages between NICE and the state, its clear tendency towards approving technologies (Abraham 2009, p. 111), and its partial reliance on industry data (Brown and Calnan 2010) underline the importance of investigating the avenues of uncertainty through which influence may occur.

The study

The central sections of this paper draw on a content analysis of a selection of NICE documents to explore and describe these avenues or forms of uncertainty. While the
analysis was informed by a wide range of documents made available on the NICE website—the content analysis presented here focuses around three in particular. Firstly, there is the Final Appraisal Determination (FAD) given for ‘Alzheimer’s disease—donepezil, galantamine, rivastigmine & memantine’ (NICE 2006). The sheer volume of documents made available relating to the decision-making process around this particular set of drugs created the need to narrow down and focus on a smaller amount of text. This document assimilates a wide range of submitted information and is analysed here as a means of understanding the modes by which NICE committees manage knowledge within the appraisal process. This appraisal has been chosen for its particular salience to the themes raised thus far: on the one hand it was an appraisal which took the relatively rare stance against recommending the technology, thus showing apparent resistance to influence; on the other, the appraisal decision was contested in the courts, with the case contesting uncertainty around the economic modelling and procedures applied by NICE (Dyer 2007).

On their own, FAD documents provide a rather disembodied and bureaucratically ‘sanitised’ account of the decision-making process. In order to develop a better understanding of the ‘backstage’ of decision-making—two further documents are made use of: the ‘Report of Threshold Workshop’ (NICE 2009) enables deeper insights into how the threshold is applied within NICE appraisal committees and moreover how it is viewed by key actors from within the institute and beyond; the transcript of a relatively candid interview given by Sir Ian Kennedy (2009), regarding his research into the way NICE deals with innovative health technologies, sheds light on the backstage salience of sectional interests and certain aspects of interpersonal interactions.

These three documents were initially coded in terms of cases of where uncertainty was indicated—either as a latent or more explicit phenomenon. These aspects of uncertainty were then re-coded in terms of the nature of the uncertainty as well as the ways in which this uncertainty was approached or ‘assumed away’ within the decision-making process. ‘Thick’ (or broad) codes relating to ‘epistemic assumptions’, ‘bounded rationality’, and ‘trust in procedures, norms and personalities’ were derived from this first stage of analysis. The material was then revisited and coded around these axial themes and consequently further sub-codes within these themes became apparent. The cases of uncertainty were then revisited and compared in terms of further clarifying the aptness of the codes and refining the conceptual framework (Neuman 2000). We now turn to an overview of the findings of this content analysis, using a range of excerpts from the documents by way of illustration and clarification.

Findings

Epistemic assumptions

System trust in empirical science

The most fundamental of assumptions on which the NICE appraisal process functions is that of the epistemological concreteness of ‘biomedicine’. Although a range of theoretical and empirical routes towards contesting this exist, ontological suppositions as to the existence of a physiological reality and corresponding notions of health and illness, alongside uncontested epistemologies where the ability of biomedicine to develop effective knowledge about this reality (e.g Alzheimer’s disease) and refine this over time, are utterly intrinsic to the functioning of NICE and remained completely uncontested within the documentation analysed. So whilst headings such as ‘Evidence and Interpretation’ (NICE 2006, p. 7) allude to the potential limitations of knowledge and its ability to represent ‘reality’, no substantive means of confronting these assumptions is apparent.

In many senses of course this recourse through the validity of science is entirely rational—indeed the section on ‘bounding rationality’ will note the inherent need for reducing complexity as a means to practical decision-making. NICE appraisal decisions are already incredibly intricate endeavours, without adding further layers of philosophical uncertainty. Hence confidence or ‘system trust’ in biomedicine—where its elemental limitations are not even considered—is an effective tool by which complexity is warded against (Luhmann 1979). So long as positive feedback is forthcoming in terms of treatment development, and in the absence of any glaring negative feedback, ‘system trust’ in the validity of biomedicine and its techniques for ‘knowing’ would seem to represent a useful foundation on which the appraisal process is built. Yet these necessary parameters and assumptions are nonetheless grounded in “notions of status and authority” which may “intervene to shape and truncate deliberation”, thus undermining the process and potentially meaning decisions are unduly influenced by certain more powerful interests (Milewa 2008, p. 361).

This picture begins to appear more problematic in the way that the scientific knowledge upon which decisions are made is rooted almost entirely in the randomised controlled trial. There exists a growing literature which makes apparent the limitations of RCTs—the key points of which are neatly summed up by Harrison (2009, p. 192; see also Busfield 2006) in his critique of ‘scientific bureaucratic medicine’: that evidence is constituted in terms of “somewhat narrow physiological parameters”, with a
corresponding ambivalence to the psycho-social, and where ‘evidence’ of interventions typically carried out in well resourced teaching hospitals in urban environments is then generalised across entire populations. Yet the way RCT findings are handled by NICE committees tends to overlook these short-comings:

The quality of the reporting and methods of the included published randomised controlled trials (RCTs) of the AChE inhibitors (donepezil, galantamine and rivastigmine) was mixed (NICE 2006, p. 7).

Although the excerpt above makes a clear acknowledgement of the variable quality of RCTs, this qualification is done within understandings of what makes a good RCT, not in anyway calling into question the wider validity of RCTs in terms of their more inherent limitations:

Six RCTs reviewed by the Assessment Group showed a statistically significant improvement in cognition following treatment with donepezil compared with placebo, as assessed using the ADAS-cog scale. Higher doses of donepezil were associated with increasing benefit (NICE 2006, p. 7).

Similarly, the findings referred to in this excerpt are described in terms of the usage of the ‘Alzheimer Disease Assessment Scale’ tool for assessing cognitive functioning, with no recognition of the likely limitations of this and other scales (e.g Mini Mental State Examination) in reflecting either what is clinically meaningful (Thomas 2004; Wilkinson 2004) or, more importantly, what is relevant to patients’ day-to-day lives. Although the MMSE is often used due to being widely available, free and easily administered by non-medics, this cannot be equated with its efficacy in terms of accuracy (especially towards more severe diagnoses of dementia) or its construct validity in relation to lived experience. Once again these not insignificant limitations are not acknowledged. Practically speaking it would seem easier to assume away these imperfections rather than have no straightforward mechanism of comparison whatsoever.

System trust in publications

Within discussions of evidence in the report, there were frequent and clear distinctions made between: that evidence which was published; manufacturers’ data which remained unpublished; and review papers. An example of this distinction is visible in the following excerpt within the considerations of cost effectiveness:

Four RCTs assessed the effect of rivastigmine compared with placebo on the CIBIC-plus scale. In the two published RCTs, statistically significant mean improvements were recorded following treatment with rivastigmine in the high-dose—licensed —regimen only, compared with placebo… For the two trials, 16–20% of participants treated with placebo were judged to have responded versus 30–57% of those treated with rivastigmine. A statistically significant difference was found for the high-dose regimen only (NICE 2006, p. 17).

The clear distinction which is being made between published and non-published, as visible in the above excerpt, is typical of that which is made throughout the document—intimating a legitimacy attached to that which has been through the rigours of peer review in contrast to data which has not been published in academic journals:

Cost estimates in the model were taken from published UK data (NICE 2006, p. 27).

Implicit within this assessment are assumptions that published data is: (a) more robust in a scientific sense; and, stemming from this first assumption, (b) less prone to bias than that which is produced by manufacturers:

Twenty-one published economic evaluations of the three AChE inhibitors and memantine were available to the Appraisal Committee. All four manufacturers also submitted their own economic evaluations (NICE 2006, p. 24).

The highly rationalised presentation of evidence given in the FAD suggests a level of objectivism which is not concretely reiterated in the other two documents. Within the Threshold Workshop minutes, a professor who has chaired many appraisal committees describes the cost-effectiveness models as:

…detailed thought experiments with multiple variables open to manipulation (NICE 2009, p. 10).

In light of this comment, key issues which are not explicitly raised include the inherent tendency within the journal publications system towards publishing manufacturer sponsored trials (Horton 2004). A former editor of the British Medical Journal notes the myriad ways in which financial interests of journals make them more likely to publish studies funded by pharmaceutical companies as well as the highly selective manner in which drug companies can manage the ‘favourability’ of the evidence which reaches publication (Smith 2005). So whilst “readers [including NICE panels] see randomised controlled trials as one of the highest forms of evidence… studies funded by a company were four times more likely to have results favourable to the company than studies funded from other sources” (ibid., p. 138).
**Suspending doubts about the interests of the industry**

In a similar sense to confidence in science and RCTs more generally, the continued positive feedback, *prima facie*, from using hierarchies of knowledge (constituted around the soundness of published evidence) within the decision-making process ensures a ‘system trust’ in the publications process which facilitates inferences about the quality of data—entailing that concerns to the contrary (regarding validity and reliability) are ‘assumed away’. This mode of a ‘suspension’ of doubts over the possibility of negative or skewed outcomes is central to the process of trust (Möllering 2001a, b) and would seem to characterise aspects of NICE’s relationship with the pharmaceutical industry (Brown and Calnan 2010). Indeed the FAD account suggests a level of confidence in the validity of published evidence where doubts are not even entertained.

When it comes to more direct considerations of the pharmaceutical industry however, it is clear that significant doubts do exist:

I was quite taken by some of the submissions and certainly some the workshops by the ... you can almost taste a kind of state of undeclared war between Pharma and NICE and I was surprised, and what I’ve tried to say in the report is that we need some detente here, we need some mutual understanding (Kennedy 2009, p. 1).

The selective publication and ‘burying’ of unfavourable evidence by the industry (Smith 2005) is relatively well recognised—hence the clear conflict of interests between the objectives of NICE and those of the industry. Yet the bureaucratic process of the FAD appears, ostensibly at least, to leave this ‘culture’ of diverging interests to one side. Partly this would seem to be based on a confidence in the capabilities of the publishing system in overcoming such biases—one that has been seen here to be misplaced. But moreover there is pragmatic aspect to appraisals by which the evidence has to be taken (more or less) at face value, otherwise any kind of appraisal would be unworkable. It is to this expedient or *bounded* rationality, where certain levels of uncertainty and complexity are more explicitly warded against, that we now turn to in the next section.

**Bounding rationality**

**A ‘Science’ of muddling through?**

As was noted in the literature review above, given the impossibility of root and branch reviews, Lindblom (1959) suggests that organisations such as NICE tend to ‘muddle through’ the decision-making process in a pragmatic, as much as a scientific, mode. The work of Herbert Simon (1982) helps further clarify the nature of this pragmatic bureaucracy through the concept of ‘bounded rationality’. This notion can be partly characterised as pertaining to the benefits of ‘docility’ (Simon 1982, p. 202): That decision-making is streamlined by assuming certain ideas via instruction (from socialised-scientific norms, bureaucratic stipulation and the like), rather than empirical investigation and contestation of every single case or eventuality.

“Epistemic assumptions” section noted the relatively ‘unaware’ aspects of docility which are implicit within the NICE appraisal process—those which are taken for granted. However NICE is also highly active in designing and applying its decision-making boundaries in a much more purposive sense. Indeed the very existence of a ‘threshold workshop’ (NICE 2009) is evidence of one clear means of erecting a boundary—around which the notion of cost-effectiveness can be considered—and the level of organisational reflexivity (self-confrontation) which exists towards this:

[The workshop’s] focus was on exploring whether there is a need for the ‘threshold’ to be amended, and if so what methods would be available, and what would need to be done by NICE and by others for any possible improvements to be realised (NICE 2009, p. 2).

**Bounding through procedures**

One clearly apparent basis for bounding rationality was through the creation of procedures. These are seen as an “objectively more robust” (Moran 2003, p. 29) means of dealing with risk and uncertainty as well as a means of creating short-cuts across complexity. Even after this complexity reducing process, the procedures described in the FAD are vastly intricate and deal with a potentially endless list of variables and permutations. From the starting point of the considerations of the cost effectiveness of donepezil, uncertainty was already clearly apparent:

In five (of 11) studies donepezil was found to be cost saving (NICE 2006, p. 24).

From here a great number of boundaries are described as being negotiated, adjusted and re-established within the process of determining the value-for-money of the drug and refining such an assessment. These range from the way a specific model for taking account of caring costs was calibrated:
The predictive risk equation for full-time care of the AHEAD³ model was used unchanged, while an annual mortality rate of 11.2% replaced the risk equation for mortality used in AHEAD (NICE 2006, p. 29);

to questions over the extent to which the benefits (of the effects of the drug) for carers of those prescribed the various drugs ought to be accounted for:

Comments received during consultation highlighted the positive impact that treatment with AChE inhibitors had on the quality of life of carers. However, quantitative evidence on the impact of AChE inhibitors on carer benefits in the form of utilities is lacking (NICE2006, p. 39).

…the effect of the drug would be to delay progression of the condition, in which case the carer would still be faced at some time in the future with the same difficulties caused by disease progression. Exceptions could be if the person did not progress to later and more difficult stages of the disease within 5 years or because of death (NICE 2006, p. 39).

In each of the three excerpts quoted here the unavoidable and yet relatively arbitrary nature of bounding decision-making is very much in evidence. In the first quote (p. 29), the choice of which model to use and the extent to which this choice should then be adjusted is far from straightforward. One would assume that the imposition of an annual mortality rate (in place of that framed by the model) is based on compelling evidence for the figure applied (11.2%), although the relevance of this is likely to be linked to the decision to model the cost-effectiveness over a 5 year period.

The use of 5 years, as opposed to 4.3 or 8 (for example), would seem to be an arbitrary rule of thumb—as also applied in the latter excerpt as a basis of deciding what should be seen as a ‘reasonable’ adjustment to make to the cost-effectiveness model in the context of a lack of quantitative evidence.

Bounding by experts

As already stated, the utility of proceduralism in the regulatory process is partly to reduce complexity but moreover in terms of its ‘legibility’—the outward visibility of rule-based conduct as a means of asserting legitimacy. In this context the manifold modifications to the models by which NICE appraisals are reached, as evident within the FAD, would prima facie appear to threaten the robustness (bureaucratically and politically) of this process. The legitimacy of the FAD however is consistently reaffirmed by appeal to the authority of expertise—usually in the guise of the independent academic ‘Assessment Group’. Often the legibility of their decisions is itself underlined through justification:

…this study was excluded by the Assessment Group because the study population was not described as patients with mild to moderately severe Alzheimer’s disease by any definition and the MMSE scores fell outside the range of 10–26 (NICE 2006, p. 20).

Yet in many other cases there is no reasoning given, with the expertise of the group apparently sufficient to explain the decision reached:

The Assessment Group reran each of the manufacturer’s economic models using its preferred assumptions (NICE 2006, p. 24).

Of course the constitution and attribution of ‘expertise’ is itself a means of bounding the rationality of evidence assimilation. Which academic centres are considered as potential ‘Assessment Group’ candidates and how choices are made between candidates is far from self-evident. Assumptions that publications or prestigious universities/research-centres are an unproblematic indicator of ability and impartiality are highly criticisable. And whilst the lack of a more effective alternative means of selecting experts may be a good reason for the current format, the presentation of expert decisions to committees nonetheless erases (or bounds against) the level of subjectivity involved in theory selection and evidence appraisal (Brown and Webster 2004). The sheer volume, complexity and contestability of evidence which is required to be processed make the use of heuristics (more or less informed rules-of-thumb) intrinsic to the process. The inevitable subjectivity of the appraisal committee is candidly acknowledged in the excerpt below, yet any similar perspectives upon decisions taken within the Assessment Group remain invisible within the FAD:

But there is also a “gut” component to decision making, and opportunity cost decisions can be as much influenced by this as by processes of reasoning. The committees’ gut feelings also have to take account of the people unrepresented in the decision process: those whose NHS gains will be displaced (NICE 2009, p. 10).

³ The ‘Assessment of Health Economics in Alzheimer’s Disease’ model.
⁴ Acetylcholinesterase inhibitors, of which donepezil (Aricept) and other other drugs being appraised within this FAD are examples, act to impede the breakdown-process of the cholinesterase enzyme which is associated with Alzheimer’s disease.
Relational forms of managing knowledge: social norms and trust

Norms of deference as a means towards ‘verifying’ expertise

Sections “Epistemic assumptions” and “Bounding rationality” have already made clear that the social is significantly influential on the ostensibly scientific process of NICE technology appraisals. In many ways it is quite appropriate that decisions which have significant social impact—on patients and the wider population—are socially orientated. However, much of the influence of the social that has been noted thus far has been found to be unintentional, docile and/or implicit. The utility of these assumptions or modes of bounded rationality are undeniable—given the need to reduce the complexity which would otherwise overwhelm the entire appraisal system. So long as these means of influence—such as socialised assumptions about the innate validity of science, the reliability of published data, or arbitrary choices justified by faith in experts—remain implicit then their potential for warping the process and outcomes of decision-making will rest unimpeded.

Norms of deference towards authority and personality have already been indicated within our earlier discussion of the use of experts. The language of acceptance towards the expert authority of the Assessment Group, as visible within Kennedy’s (2009) account of his own ‘expert’ role within NICE, displays a taken-for-grantedness of the scientific expertise of the Assessment Group, as visible within the use of experts. The language of acceptance towards the have already been indicated within our earlier discussion of the social norms and obligations acting on those who make representations towards the committee.

Whether Kennedy was indeed the right expert for this role is impossible to verify—what is clear is that he was not exceedingly familiar with NICE procedures prior to his appointment:

I was operating at the plane of a kind of political and philosophical approach on the one hand and also then on the other trying to work out what a FAD meant (Kennedy 2009: 1).

This makes evident how the choice of experts can bear greatly on the outcomes of expert decisions. Cooksey’s background, alongside that of Kennedy, will inevitably have coloured their prioritisation of certain issues ahead of others—especially where irreconcilable values (cost control in the NHS versus pharmaceutical industry profitability) have to be compared.

Trust as a means of ‘verifying’ motives

The importance of Kennedy’s track-record for his appointment seems almost certain, with the positive outcomes from his previous appointments making it easy to assume—or trust—in his competence and willingness to carry out an effective role for NICE. Kennedy’s renown, and the visibility of his role, make likely that norms and expectations of conduct are perceived as exerting a compelling force on him to fulfil his professional obligations and carry out a robust study (Möllering 2005). Similar assumptions are likely to be made by the appraisals committee in terms of the social norms and obligations acting on those who make representations towards the committee.
How these assumptions are applied towards actors with different interests—as clinician, patient or industry representatives—cannot be investigated through the data assessed here though would make a valuable focus of future research.

In spite of the cynicism and antipathy, as set out above, which exists between NICE and pharmaceutical manufacturers, this was not explicit within the bureaucratic narrative of the FAD. Although the existence of conflicting interests could be read between the lines of aspects of the report:

The Committee carefully examined the cost-effectiveness models provided by the Assessment Group and the manufacturers, and it noted the substantial differences in cost-effectiveness estimates between the manufacturers’ models and those of the Assessment Group (NICE 2006, p. 37).

Given the clear conflict of interests which exists between the goal of the committee and the accounts of pharmaceutical manufacturers, the matter of how evidence is received and absorbed or discounted comes to be of vital significance for the robustness of the whole appraisal process. On the one hand manufacturers’ data is visible throughout the FAD and would seem to be necessary in their understandings of the nature and efficacy of interventions. On the other hand a certain degree of scepticism is received and absorbed or discounted comes to be of vital significance for the robustness of the whole appraisal process. On the one hand manufacturers’ data is visible throughout the FAD and would seem to be necessary in their understandings of the nature and efficacy of interventions. On the other hand a certain degree of scepticism could be read between the lines of aspects of the report:

The Committee noted that the Assessment Group considered that the manufacturers’ cost-effectiveness calculations needed to be treated with considerable caution because: optimistic assumptions on estimates of mortality and costs were used (NICE 2006, p. 37).

The means by which this ‘considerable caution’ is applied, how this highly contestable evidence bears upon the perceived robustness of other evidence provided by manufacturers, and the nature of the tension between a compromised trust in manufacturers and the pragmatic need to work with their representatives—all emerge as important issues of trust within this context. If trust is conceptualised as the belief that the trustee will put the interests of the trustees (the appraisal process) first and has no agenda to the contrary (Williams 2007)—then the relation of the committee (and others within NICE) to representatives from the industry, and the evidence they present, are problematic and worthy of exploration.

**Value-attachment as a result of cultural representations of illness and the NHS**

A further mode by which social factors may influence the decision-making within NICE appraisals is through the value-laden means of understanding illness and NHS institutions as driven by dominant cultural representations of these within society. In the case of the former, the task of comparing radically different technologies for use in divergent illness types is a very difficult one. Although the QALY-based process of NICE appraisals seeks to overcome this, a level of pragmatism around the cost-per-QALY threshold, as influenced by value-orientations, would seem unavoidable:

In practice, decision making involves a great deal of what Professor Stevens described as “pragmatic thresholdry”. Factors that might influence are many and various. Cancer tends to be viewed more sympathetically, as are children, or patients nearing the end of their lives or who are particularly disadvantaged by their condition (NICE 2009, p. 11).

Socialised norms and the cultural resonance of certain concepts of illness such as cancer—in terms of a valiant battle (Sontag 1978)—or Alzheimer’s means that the very use of these linguistic utterances “is capable of making present that which is spatially, temporally and socially absent from the ‘here and now’” (Berger and Luckmann 1966, p. 174).

In contrast, the services which are being displaced by favourable opinions from appraisal committees are highly nebulous and distant. Although such ‘at-risk’ services are not necessarily neglected, the lack of more tangible evidence for their effectiveness and a corresponding reliance on the ‘gut feelings’ of committee members as cited above (NICE 2009, p. 10), makes it probable that their remoteness within the social process of appraisals leads to their neglect. The FAD records that:

The Committee was also mindful of the need to ensure that its advice took account of the efficient use of NHS/PSS resources (NICE 2006, p. 36).

Yet the abstract notion of resource allocation—compared with the visible and interactive presence of sufferers of the condition under consideration—would seem to make it probable that the emotional and moral impact of the illness experience of certain diseases (particularly those with significant cultural resonance), and the potency of hope attached to its treatment, have an important impact on committee members (Brown 2011). Even when leaving notions of affect to one side, the gap between the threshold applied by NICE and the ‘shadow’ threshold (the price of QALY that a PCT cannot quite afford) will vary from PCT to PCT (depending on funding, efficiency, and variations in need) and hence there is no way of knowing which interventions will be displaced. Hence although NICE’s threshold should represent its best guess about the measure of the shadow price (NICE 2009, p. 16),
in practice the committee has no way of knowing which treatments are being displaced. This high level of uncertainty amplifies the need for committee members to resort to a ‘gut feeling’ approach.

Discussion: a framework for investigating the impact of uncertainty

The findings presented above suggest the existence of multiple layers of uncertainty. Yet thus far, specific perceptions regarding the format of uncertainty, as to how unknowns are considered and pursued (or not), have yet to be considered fully. The social scientific literature pertaining to decision-making within contexts of risk and uncertainty offers a number of theoretical insights into these formats which are decidedly relevant to NICE appraisals. Knight (1921—see also Langlois and Cosgel 1993: 460) differentiates between approaches to uncertainty which: seek to categorise and probabilistically calculate in their response to uncertainty (‘calculable’ risk); distinguish between potential outcomes in the face of uncertainty, but where numeric calculation of their likelihood is seen as impossible (‘categorisable uncertainty’); recognise uncertainty in a more vague and nebulous manner (‘known uncertainty’). As considered above, socially-constructed assumptions (Schutz 1972) within prevailing systems (Luhmann 1979) influence the type of approach and application of knowledge forms within each system.

While ostensibly NICE decisions are purely calculative, the ‘gut component’ referred to highlights the extent to which ‘categorisable uncertainty’ and even mere ‘known uncertainty’ are inherent within decisions. Responses within these latter categories may include applying experiential or tacit knowledge (Lam 2000) and affect-based heuristics (Finucane et al. 2000). Such mechanisms—rules-of-thumb, intuition, and trust—are common and effective modes of decision-making amidst uncertainty (Knight 1921; Gigerenzer 2007; Zinn 2008). Trust is an especially vital means of bridging the unknowable (Möllering 2001a, b) through gauging the reliability of individuals. Luhmann (1979) notes moreover how continuous positive feedback from abstract systems of knowledge or institutions may similarly lead to ‘system trust’. Enduring positive feedback may even result in ‘confidence’ (ibid), where potential fallibility or uncertainty is no longer even envisaged: thus a fourth format—‘unknown uncertainty’—emerges. The following matrix represents one possible basis for future research, where different layers of uncertainty may be investigated in light of particular (or multiple) formats for tackling these:

<table>
<thead>
<tr>
<th>Layers/ forms of Uncertainty</th>
<th>Potential approaches/responses to various layers of uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calculating risk</td>
<td>Categorising uncertainty</td>
</tr>
<tr>
<td>Epistemic</td>
<td>Procedural</td>
</tr>
</tbody>
</table>

By its scientific approach, NICE has developed sophisticated techniques for modelling cost-effectiveness. Yet these calculating approaches—in their technocratic nature—are open to contestation and dispute via legal challenge. So while NICE was able to extend its access to (and recalculate) patient data, thus reducing the problem of suspending doubts in industry information, Eisai (donepezil manufacturer) and Pfizer (the distributor) disputed the appraisal decision as “procedurally flawed and irrational” (Dyer 2007, p. 1337). In spite of findings that donepezil and other acetylcholinesterase inhibitors were far from sufficiently cost-effective for mild stage Alzheimer’s, the decision was nonetheless exposed to legal contestation and therefore widespread media debate.5

Lehoux et al. (2009) suggest that novel health innovations are typically contested through three different types of argument: scientific, clinical and social. These three aspects are congruent with the layers of uncertainty noted above and indeed all three emerged within the contestation of this particular appraisal decision. Deeper assumptions about the clinical relevance and consistency of the MMSE, emerging in the form of its procedural application in the appraisal, were at the centre of manufacturers’ (and the Alzheimer’s Society’s) objections presented at the High Court. Further concerns about the economic modelling were also challenged in this way. The scientific-technocratic (procedural uncertainty) were challenged legally, while more practical-clinical and social arguments were raised in the public sphere (mass media)—for example the difficulty in telling patients they were unable to be prescribed a drug until their condition was more developed.

Thus it would seem that different layers of uncertainty may be addressed through different formats, but moreover may be prone to different modes of influence. NICE’s open recourse to procedures and expertise in the face of uncertainty means that these formats are most open to challenge—through procedural means (i.e legal process) and by contesting expertise. Conversely, due to the way that epistemic and relational uncertainty is largely veiled by NICE behind calculative approaches, contestation of these uncertainties (and influence through this) is accordingly

5 NICE revised its stance towards a more favourable position in 2011.
much less detectable. Yet this lack of visibility should not be misinterpreted as a lack of influence, rather such influence may well occur ‘behind the backs’ of decision-makers: first, due to the impossibility of fact/value distinctions and the corresponding influence of the ‘dynamics of hope and expectation’ upon ‘objective’ experts (Brown and Webster 2004); and second, through the epistemic assumptions on which the whole system is based (Milewa 2008).

Epistemic frameworks are significant in that, as with legal challenges to technical procedures, innate barriers exist to certain interests and not to others. Hence the preponderance of the bio-medical alongside particular economic models privileges those interests which function within these paradigms—the ‘health-industrial complex’ (Meadowcroft 2008; Milewa 2008). Such a scientific-bureaucratic framework facilitates the rationalising of instrumental tendencies such as economic and research progress, while the communication which considers values, interests and moral outcomes becomes increasingly separated off and its refinement impeded (Habermas 1987). Whereas single decisions may be deemed rational or not in terms of their realising a particular ‘end’, rationality over time is a function of an institution’s selection of ‘ends and values’, its ‘consistency over time’ and ‘self-understanding’ (Brubaker 1984, p. 94). This separation is thus of serious concern because NICE appraisals have a profound impact on patients and society.

NICE’s legitimacy is developed within such a broader context, yet its tendency towards the calculative, in defending its technical-instrumental rationality in the face of contestation, leads to the neglect of this communicative rationality. It is in this sense that more overt contestations, and more implicit tendencies towards certain forms of rationality, are very much linked. Legal contestations and more general criticism of NICE procedures are not only effectual on the particular decision in question but more fundamentally have resulted in reform of NICE’s appraisal practice and other assumptions (Hedgecoe 2004). Fuller’s (1978) notion of neglected parties is salient here, as is Olson’s (1965) consideration of the ‘logic of collective action’, in that interests which are poorly resourced and/or overly diffuse are more likely to become neglected by NICE. Contrast this with the concerted and well resourced effort to review NICE as a possible fourth hurdle to industry innovation (Kennedy 2009). In the short term this may mean that the very procedures NICE apply in bounding rationality may be shaped by certain outside interests (and not others), in the longer term this may engender the forming of more profound assumptions—such as that of the need for NICE to further concern itself with incentives for industry ‘research and development’ as well as immediate NHS resource efficiency.

Conclusion: considering the influence of uncertainty

This paper began by denoting the creation of NICE as an attempt to inaugurate a systematic, neutral and evidence-based regulation of the use of expensive pharmaceutical products by the NHS. Yet this attempt to absorb uncertainty and promote objectivity has paradoxically resulted in an increasingly politicised rationing and amplified awareness of decision-making uncertainty. The core of the paper described three main layers of uncertainty within appraisals—epistemic, procedural and social. The latter part of the paper set out a framework for considering the extent to which these layers of uncertainty and the various formats of responses to them may act as avenues through which outside interests, for example the pharmaceutical industry, may exert influence which skews the regulatory process.

This potential bias within the appraisals process may be more overt—through the ability of certain interests (and not others) to contest procedures—or via more subtle systemic tendencies. These two modes of influence have been seen to be linked. So while NICE is clearly aware and seemingly cynical (Kennedy 2009) to the motives of the pharmaceutical industry, this is not sufficient to protect appraisals within a polycentric regime. Pressures exerted by the industry indirectly (through the media and the state) may lead to the regulation of the regulator (Black 2008; Brown and Calnan 2010) while built in biases engender neglected parties (Fuller 1978). “Accordingly, the regulation of health technology … should be seen as a distributed collective process of which interdependencies and the need to make persuasive claims are key features” (Lehoux 2006, pp. 157–158).

The existence of unknowables is unavoidable and future research is necessary to explore their management as a potentially significant source of regulatory bias. Such research would undergird the proposals of Lehoux (2006, pp. 170–186) who presents a framework outlining how biased tendencies amongst uncertainty could be made explicit and dealt with publicly. Her five principles include making values, the various norms of scientific networks, and the interests of the private sector much more explicit. This would enable a more ‘reflective’ regulatory practice enlivened through ‘socio-political’ debate and would “make civil society a pivotal locus of deliberations” (p. 186). Involving civil society more generally and flexibly (Milewa 2008), rather than specific vested interest groups, alongside a broadening of the debate beyond the scientific bureaucratic, may be one step towards overcoming built in biases. This would enhance legitimacy by acknowledging and involving more diffuse interests. So long as NICE decisions hinge on the technocratic, legal challenge is always a lingering threat. By broadening the format of decisions (Milewa 2008), and highlighting
potential losses through disinvestment (Meadowcroft 2008), legitimacy would be enhanced and the potential for the asymmetric influence of interests reduced. Yet the first step along this road must be more developed understandings of how these biases become manifest through the uncertainties outlined here.

Acknowledgments The authors are grateful to the anonymous reviewers, special edition editors, and Dr Richard Brown for their highly constructive comments in assisting the development of this paper.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References


Dyer, C. 2007. NICE’s decision on dementia drugs was “irrational.” High Court is told. BMJ 334: 1337.


