NICE accountability for reasonableness: a qualitative study of its appraisal of treatments for attention-deficit/hyperactivity disorder (ADHD)

Michael Schlander

Institute for Innovation and Valuation in Health Care (InnoVal-HC), Eschborn, Germany; Department of Public Health, Social and Preventive Medicine, Mannheim Medical Faculty, University of Heidelberg, Germany; and University of Applied Economic Sciences Ludwigshafen, Germany

Address for correspondence: Prof. Dr. med. Michael Schlander, MBA., Institute for Innovation and Valuation in Health Care (InnoVal-HC), Rathausplatz 12–14, D-65760 Eschborn, Germany. Tel.: +49-6023-929589; Fax: +49-6023-929591; email: michael.schlander@innoval-hc.com

Key words: Accountability for reasonableness – Attention-deficit/hyperactivity disorder (ADHD) – Case study – Health Technology Assessment (HTA) – National Institute for Health and Clinical Excellence (NICE) – Publicity – Relevance

ABSTRACT

Objective: The UK National Institute for Health and Clinical Excellence (NICE) is widely regarded a role model for the implementation of Health Technology Assessments including cost-effectiveness evaluation. The aim of the present study was to explore the real-life robustness of the NICE technology appraisal process when addressing complex clinical problems, using the Accountability for Reasonableness (A4R) framework proposed by Daniels and Sabin as a reference.

Method: A qualitative case study of NICE Technology Appraisal No. 98, ‘Treatments for Attention-Deficit/Hyperactivity Disorder (ADHD)’, analyzing each step of the appraisal process.

Results: Scoping was narrower than that for corresponding clinical guidelines. Economic evaluation for assessment was primarily based on six short-term studies, was unable to differentiate compounds on grounds of effectiveness, and cost-effectiveness modeling suggested a clear recommendation driven by acquisition costs. After appraisal, all treatment options assessed were recommended within their licensed indications. With estimated costs per quality-adjusted life years (QALYs) compared to no treatment generally falling below £7000, NICE guidance specified that choice of drug should be primarily based on clinical considerations, followed by cost.

Conclusion: The appraisal process adhered to predefined timelines, which were sensibly adapted by NICE to changes in the environment. A4R criteria most pertinent to the case study were ‘publicity’ and ‘relevance’. The ‘publicity’ condition was greatly fulfilled, except for commercial-in-confidence data and economic model. ‘Relevance’ requires appraisals to reflect concerns for fairness and to be evidence-based; in that respect, principles and realization of the assessment deserve further scrutiny. Questions also remain regarding the ‘appeal’ and ‘enforcement’ conditions under A4R.

Introduction

Since its inception as the ‘National Institute for Clinical Excellence’ in April 1999, technology appraisals by the National Institute for Health and Clinical Excellence (NICE) have attracted international attention. Their high visibility has served to extend their influence beyond the Institute’s primary remit, notably (though not limited to) providing guidance to the National Health Service (NHS) of England and Wales. NICE
Table 1. Treatment options for ADHD in children and adolescents in the United Kingdom: overview of product availability and acquisition cost*

<table>
<thead>
<tr>
<th>Trade name</th>
<th>Manufacturer</th>
<th>Active ingredient</th>
<th>Formulation</th>
<th>Abbreviation</th>
<th>Cost/daily dose† (£)</th>
<th>Assumed total daily dose‡ (mg/day)</th>
<th>Daily dosage schedule‡</th>
<th>Authorization§</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dexedrine</td>
<td>UCB Pharma Ltd. Slough</td>
<td>Dexamphetamine sulphate</td>
<td>Tablets (5 mg)</td>
<td>DEX</td>
<td>0.43</td>
<td>20</td>
<td>2 (-3) times</td>
<td>≤ 2000</td>
</tr>
<tr>
<td>Ritalin</td>
<td>Cephalon UK Ltd. Guildford, Surrey</td>
<td>Methylphenidate hydrochloride</td>
<td>Immediate-release tablets (10 mg)</td>
<td>MPH-IR</td>
<td>0.56</td>
<td>30</td>
<td>3 (2–4) times</td>
<td>≤ 2000</td>
</tr>
<tr>
<td>Equasym</td>
<td>UCB Pharma Ltd. Slough</td>
<td>Methylphenidate hydrochloride</td>
<td>Immediate-release tablets (5, 10, 20 mg)</td>
<td>MPH-IR</td>
<td>0.53</td>
<td>30</td>
<td>3 (2–4) times</td>
<td>≤ 2000</td>
</tr>
<tr>
<td>Concerta XL</td>
<td>Janssen-Cilag Ltd. High Wycombe, Bucks</td>
<td>Methylphenidate hydrochloride</td>
<td>Modified-release tablets (18, 36 mg)</td>
<td>MPH-MR12</td>
<td>1.23</td>
<td>36</td>
<td>1 time</td>
<td>2002 (Feb, 19)</td>
</tr>
<tr>
<td>Strattera</td>
<td>Eli Lilly and Co Ltd. Basingstoke, Hampshire</td>
<td>Atomoxetine hydrochloride</td>
<td>Hard capsules (10, 18, 25, 40, 60 mg)</td>
<td>ATX</td>
<td>1.95**</td>
<td>Irrelevant (flat pricing)</td>
<td>1 (–2) times</td>
<td>2004 (May, 27)</td>
</tr>
<tr>
<td>Equasym XL</td>
<td>UCB Pharma Ltd. Slough</td>
<td>Methylphenidate hydrochloride</td>
<td>Modified-release capsules (10, 20, 30 mg)</td>
<td>MPH-MR08</td>
<td>1.17</td>
<td>30</td>
<td>1 time</td>
<td>2005 (Feb, 11)</td>
</tr>
</tbody>
</table>

*Note that exact licensed indications and ages differ between products
†NHS acquisition costs (net prices, excluding VAT and not accounting for negotiated procurement discounts in some settings), taken from British National Formulary 51, March 2006; note that individual doses and thus costs may vary
‡Average assumptions underlying ‘cost per daily dose’, not to be confused with treatment recommendations; doses need to be optimally titrated for each individual patient; see for instance Wilens and Dodson
§First authorization in UK, from electronic Medicines Compendium, available online at http://emc.medicines.org.uk [Last accessed August 12, 2005]
**Due to its flat pricing policy, the cost per daily dose of ATX increases to £3.80 when administered twice daily
ADHD = attention-deficit/hyperactivity disorder
is frequently being perceived as a role model for the implementation of cost-effectiveness analysis (CEA), to inform decisions about the rational allocation of health care resources in an environment of economic limitations. This has led to a call by some economists to expand the NICE approach internationally. The European High-Level Group on Innovation and Provision of Medicines (G-10) engaged in debate about creating a ‘Euro-NICE’, although it recognized that pricing and reimbursement structures for medicines fall within the competence of the member states. Only in July 2006, the German grand coalition government agreed on the outline of a new health care reform, expanding the mission of the German Institute for Quality and Efficiency in Health Care (‘Institut fuer Qualitat und Wirtschaftlichkeit im Gesundheitswesen’ [IQWiG]) to include ‘cost-benefit evaluations’ of pharmaceutical products. These evaluations should follow ‘international standards’ with explicit reference to NICE. Also, in the United States, a comparable debate has been stimulated by the recent introduction of the Medicare Part D prescription drug benefit, and the creation of one or more new institutes has been proposed to provide advice to Medicare on cost-effectiveness when determining the coverage of new medical interventions. Not surprisingly, in a field as ideologically charged and as generously subsidized as health care, the approach adopted by NICE has not been without controversy. Nevertheless, the processes and transparency utilized by NICE have been widely regarded as exemplary and it has been asserted that ‘NICE demonstrates the potential of a new organization with a specific mandate to consider cost-effectiveness’. Furthermore, NICE recently updated its methods guidance for technology assessment, (among other aspects) endorsing probabilistic sensitivity analyses, thus assuming a leadership role in this important area. Indeed, the traditionally cautious and, for that matter, initially skeptical editors of the British Medical Journal even proclaimed ‘the triumph of NICE’. They suggested ‘NICE may prove to be one of Britain’s greatest cultural exports, along with Shakespeare, Newtonian physics, the Beatles, Harry Potter, and the Teletubbies’.

Little, however, is known about the real-life robustness of the NICE approach when addressing complex clinical problems. To explore this important question, a qualitative study was conducted using the example of a recent technology appraisal of treatments for attention-deficit/hyperactivity disorder (ADHD) in children and adolescents. The decision to analyze this particular case was motivated by the convergence of a personal scientific interest in the areas of pharmaceutical market regulation and ADHD, and by the observation that the scope published in August 2003 excluded psychosocial interventions, which represent a therapeutic mainstay in most European countries.

The aim of the present study was to elucidate the challenges faced by decision-makers when addressing particularly difficult clinical problems. The study does not intend to invalidate all NICE technology appraisals conducted to date; rather it provides insights based on qualitative research, which is not an alternative but a complement to quantitative work. It shall enable the readers to ‘reach the parts other methods cannot reach’, i.e., to explore complex areas not amenable to quantitative research, in particular in health service organization and policy.

### Attention-deficit/hyperactivity disorder

Broadly, ADHD is characterized by a ‘persistent pattern of inattention and/or hyperactivity/impulsivity that is more frequent and severe than is typically observed in individuals at a comparable level of development’. Though economic studies of ADHD are still in their infancy – and thus far have been published predominantly in the United States – it is already clear that this disorder is associated with a substantial economic burden. The case of the appraisal of ADHD interventions by NICE may be of particular interest for a number of reasons, all of which illustrate its relevance as well as its complexity.

First, even though the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV)-based prevalence estimates vary widely (in 8 to 10-year-old children from 3.9% up to 19.8% in one study) and depend on the population studied and diagnostic criteria used, ADHD is believed to represent the most common psychiatric disorder in children and adolescents.

Second, ADHD is associated with high rates of coexisting (comorbid) conditions. Externalizing signs such as oppositional defiant disorder and conduct disorder occur in 50–60% of all children with ADHD, and internalizing mental health problems, notably, anxiety and depression, in 12–26%. A wide range of other psychiatric, neurological and somatic disorders may be associated with or superimposed on signs and symptoms of ADHD.

Third, the complexity of the situation is exacerbated by international differences in commonly accepted diagnostic criteria (notably DSM-IV in North America, and the more restrictive International Statistical Classification of Diseases and Related Health Problems (ICD-10) criteria that are preferred in Europe), giving a 1.7% prevalence for ‘hyperkinetic disorder’ (HKD) in boys in the United Kingdom. There also exist international differences of standards of care (as exemplified, for instance, by lower prescription...
rates of psychostimulants in Europe compared to the United States; see below). This creates issues related to the transferability of the results of US studies to a European context.

Fourth, accrued data on ADHD is suggestive of an apparent increasing prevalence. However, this needs to be viewed in the context of raised awareness among parents, educators and health professionals of the adverse effects of behavioral and learning problems in children. This raised awareness is, perhaps, reflected by a striking increase in the number of prescriptions for psychostimulant in younger people in the US during the 1990s, and, with some time lag compared to the US, in a number of European countries, including the United Kingdom during the last decade. In England, prescription cost analysis data from the Prescription Pricing Authority show an increase in the number of prescription items of methylphenidate hydrochloride from 126,600 in 1998 to 389,200 in 2005.

In turn, this has, fifth, aroused much controversy and emotive debate both among professionals, as well as the general public, not only about the possibility of overuse of psychotropic medications in children, but also about the validity of ADHD as a distinct disease entity.

In addition, sixth, the substantial variety of instruments to measure clinical outcomes, symptom relief, and health related quality of life across clinical studies agitates existing difficulties in determining utility values to calculate quality-adjusted life years (QALYs) in patients with a diagnosis of ADHD.

Seventh, the therapeutic armamentarium for ADHD, comprising two principal options – behavioral treatment and medication management – has been expanded by the emergence of new treatment options, notably extended- or modified-release preparations of methylphenidate (MPH-MR08, MPH-MR12) that eliminate the need for a mid-day dose, and atomoxetine (ATX), a non-stimulant compound, both of which have the potential to profoundly change the therapeutic landscape.

Eighth, the resultant expectation of changes in service provision derived from the combined influences of increased awareness and more frequent diagnoses, growing acceptance of pharmacotherapy in the light of new clinical studies, and the availability of novel medication options that command higher unit costs compared to previously available options, are likely to have important budgetary impacts (Table 1).

Finally, ADHD is associated with a substantial cost of illness. Health care costs for individuals with ADHD have been reported as twice those for individuals without the disorder. Parents and other family members of patients have also been found to have about 60% more medical claims than matched controls, and data strongly suggest that a child’s ADHD places a substantial economic burden on parents and other family members, including a negative impact on parents’ absenteeism from work and productivity. The economic impact of ADHD is further exacerbated by its frequent persistence into adulthood, thus constituting a chronic condition, and by serious long-term sequelae that have been linked to the disorder. These sequelae include poor driving abilities, higher risks of accidents and injuries, increased rates of tobacco, alcohol and other substance use disorders, more frequent antisocial behaviors and encounters with the criminal justice system across the lifespan, as well as relatively poor educational outcomes and lower-ranking occupational positions than controls.

Each and all of the foregoing issues raises questions about how technology appraisal processes adopted by NICE can accommodate these clinical complexities. The present paper will focus on process-related observations, whereas the underlying technology assessment as well as a discussion of potential implications for international health care policy makers will be subject of subsequent, separate papers.

**Accountability for reasonableness**

The present analysis will be guided by a framework developed by Daniels and Sabin who have argued that the legitimacy of controversial limit-setting decisions in public health care systems hinges on a fair institutional decision process. In order to narrow the scope of controversy, they have proposed principles of ‘accountability for reasonableness’ (A4R), which ‘fair-minded people’ should accept based on the idea that there exists a core set of reasons, that all center on fairness, on which there will be no disagreement. A key element of fair process under A4R involves transparency about the grounds for decisions (the ‘publicity’ condition, opening decisions and their rationales for scrutiny by all affected, not just the members of the decision-making group). Second, the ‘relevance’ condition imposes an important constraint on arguments, because they are required to rest on scientific evidence, though not necessarily a specific kind of, and to appeal to the notion of ‘fair equality of opportunity’. Although Daniels and Sabin acknowledge that stakeholder participation may improve deliberation about complicated matters, they believe it is neither a necessary nor a sufficient condition of A4R. They advocate, however, an ‘appeals’ component as an institutional mechanism to engage a broader segment of society in the process, providing those affected by a decision an opportunity to reopen deliberation, and to offer decision-makers an option to...
Objective

The objective of the present report is to analyze the real-life performance and robustness of NICE technology appraisal processes, using the case study of ADHD treatments because these represent a particularly challenging field for economic analysis (see Introduction). Further the present report provides the context for a separate detailed analysis of the technology assessment report underlying the appraisal and for a discussion of implications for international policy makers. Given the policy relevance of NICE guidance in England and Wales, as well as in light of suggestions that NICE ‘is conquering the world’, one should expect that, when applied in practice, the approach adopted by NICE consistently meets the highest quality standards. Thus, successful accommodation of the complexities of ADHD treatment would be reassuring, whereas any problems observed would be of significant interest given the high profile of NICE and might stimulate debate about the adoption (or modification) of NICE-like approaches in other jurisdictional settings.

Methods

A qualitative case study was carried out of NICE Technology Appraisal No. 98, ‘Methylphenidate, atomoxetine and dexamfetamine for attention deficit hyperactivity disorder (ADHD) in children and adolescents (Review of Technology Appraisal 13)’, published March 2006. The analysis presented here is focused on the real-life application of NICE processes and was part of a more comprehensive study of the ADHD appraisal by this author. The present study had descriptive, explorative, and explanatory elements. First, the initial phase of the study consisted of defining a theoretical framework for analysis. This included a description of NICE technology appraisal processes, which fell in a period of substantial upgrade and definition of the so-called ‘reference case’ analysis by NICE (see below – Results, NICE technology appraisal process). During this phase, a thematic framework was defined, comprising use of the A4R concept as a process benchmark, an in-depth critique of the technology assessment report underlying the appraisal, as well as a review of the clinical and economic literature on attention-deficit/hyperactivity disorder in order to incorporate the complex interrelated issues involved in this technology appraisal (see Introduction).

The second phase of the study comprised data collection employing a number of closely related strategies: (1) from May 2004 to publication of guidance in March 2006, the NICE website (www.nice.org.uk) was visited at intervals of less than 1 month each and checked for newly posted information and documents (including meeting minutes and announcements) on (a) the technology appraisal process and related methods, (b) clinical guideline development, (c) deliberations of the NICE Citizens’ Council, and (d) ADHD; (2) scientific articles cited in these documents were obtained for analysis; (3) independent literature searches (using the PubMed and EBSCO databases as well as Google Scholar) were conducted for articles on ADHD diagnosis, treatment, compliance, cost, and cost-effectiveness; and were (4) complemented by a search for relevant abstracts presented at international meetings in the fields of clinical psychiatry, child and adolescent psychiatry, pediatrics, health economics, and pharmaco-economics. All searches for literature fully covered the technology assessment period (from June to December 2004, see below). After May 2005, no more systematic searches for scientific literature were conducted, and new papers were added to the database in an opportunistic manner only. Collected documents were indexed using categories including study type, product tested, and subject matter (e.g., treatment compliance) for further analysis and interpretation.

The present report is primarily concerned with the actual NICE appraisal process as observed in the case of ADHD treatments. All key steps of the ADHD appraisal process were identified and compared with NICE process descriptions. The A4R-derived criteria of publicity (transparency), reasonableness, and (revision and) appeal guide the discussion of observations derived from the case study, whereas the criterion of enforcement appeared less pertinent in this context.

Results

NICE technology appraisal process

NICE technology appraisals consist of three phases: (1) scoping; (2) assessment; and (3) appraisal. As an optional fourth phase, an appeal against a Final Appraisal Determination (FAD) by NICE may be filed by consultees and will be dealt with by an Appeal Panel. NICE has delineated this process in some detail in a series of related technical documents (Table 2).
Scoping

Topics for appraisal are suggested to NICE by relevant government Ministers (Department of Health [DoH] and Welsh Assembly Government [WAG]) – usually as part of a ‘wave’ of topics. NICE identifies experts and stakeholders as ‘consultees’ and ‘commentators’ and prepares a draft scope which is also provided to the Assessment Group – an independent academic group commissioned by the NHS Research and Development Health Technology Assessment Programme to assist in the appraisal. These groups also receive from NICE the draft remit (i.e., the initial brief given to NICE) and approximately 8 weeks thereafter, a scoping workshop is held by NICE. Components of the scoping procedure include a clear definition of the clinical problem (or disease) and the patient population, the technology (and its comparators) and their treatment setting, measures of health outcomes and costs, time horizon, and any special considerations appropriate to the topic. Following preparation of a final remit (produced by the DoH and the WAG) and a final scope (by NICE), a formal decision is made by Ministers to refer (or otherwise) the technology in question for appraisal by NICE. After referral NICE initiates the appraisal process, the timeline of which commences after NICE invites consultees and commentators to participate.

Assessment

The key activity in the assessment phase is the evaluation of the evidence relating to the technologies in question by the Assessment Group. The most recent update of the NICE methods guide\(^\text{10}\) is highly prescriptive of admissible evidence, of its analysis and of the presentation of findings. Specifically, a reference case has been defined
with the objective of achieving consistency across assessments, providing a detailed description of the methods considered most appropriate for the Appraisal Committee’s subsequent deliberations (see below – Appraisal). These include, *inter alia* and within the scope developed by NICE, the use of all health effects on individuals as outcome measure, to determine health benefits in terms of QALYs (using a standardized and validated generic instrument), to derive preferences for health state valuation from a representative sample of the public using a choice-based method (i.e., as opposed to a rating scale) for elicitation, the use of an annual discount rate of 3.5% for both costs and health effects, and finally the equity position that an additional QALY has the same weight regardless of the other characteristics of the individuals benefiting. Synthesizing evidence on outcomes should enable an unbiased estimate of clinical effectiveness. To achieve this, NICE expects a systematic review and meta-analysis, requiring that an assessment of the degree of and the reasons for heterogeneity be undertaken before any statistical pooling is carried out. The need is acknowledged to construct a decision analytical framework in order to estimate clinical and cost effectiveness relevant to the decision-making context in a clinical setting. Accordingly, modeling, ‘an unavoidable fact of life in economic evaluation’[23], is explicitly accepted and is likely to be required, among other situations, when trial populations are atypical, intermediate outcomes data from trials are used, relevant comparators have not been used in trials, or when long-term consequences extend beyond trial follow-up. NICE further expects parameter uncertainty to be presented using probabilistic sensitivity analysis (or, where appropriate, stochastic analysis of patient-level data). Moreover, patient subgroups should be identified and clinically justified, and uncertainty in subgroup results should be fully reflected. In addition, the Assessment Group is required to incorporate submissions from manufacturers and sponsors, alongside details of models used in these submissions. Such submissions are expected to meet the same criteria, and any electronic models need to be provided to NICE and the Assessment Group. Commercially sensitive data may be designated ‘commercial-in-confidence’ and will remain confidential, i.e., will not be published with the Assessment Report. A timeframe of 28 weeks is allowed by NICE for completion of the Assessment Report, although this is reduced to 14 weeks when the deadline for receipt of external submissions is taken into account (see Table 3). The Assessment Group may produce a *de novo* economic model, which will be protected by intellectual property rights. Although it may be provided to stakeholders upon their written request, it will be supplied as a read-only copy and must not be re-run with alternative assumptions or inputs[24].

**Appraisal**

The Appraisal Committees are standing advisory committees of NICE whose members are appointed for a 3-year term. Members are drawn from the NHS, patient and caregiver organizations, relevant academic disciplines, and the pharmaceutical and medical device industries. The appraisal stage of the process comprises four elements. One element is the consideration of the evidence in the Assessment Report (including confidential material) together with that submitted by other parties, the aim being to develop an Appraisal Consultation Document (ACD), with the participation of members of the independent academic Assessment Group. The preparation of and consultation on the ACD should respect specified benchmarks for incremental cost-effectiveness ratios, and take into account the longer-term interests of the NHS in encouraging innovation in technologies that will benefit patients. A further element of the appraisal process is the review by the Appraisal Committee of the ACD in the light of comments received during consultation. The ultimate element of the appraisal process is the preparation of the FAD. Subject to any appeal, the FAD will form the basis of the guidance by NICE on the use of the appraised technology. Ongoing activities including meeting proceedings (agenda, minutes) are published on the NICE website.

**Appeal**

Consultees are given 15 working days from receipt of the FAD to lodge an appeal which will be considered only if it falls within one or more of the following categories: (1) NICE has failed to act fairly and in accordance with its published procedures; (2) the FAD is perverse in the light of the evidence submitted, with ‘perverse’ meaning that the FAD is ‘obviously and unarguably wrong, in defiance of logic, or so absurd that no reasonable Appraisal Committee could have reached such conclusions’[25]; or (3) NICE has exceeded its powers. New evidence or simply disagreement with a FAD will ‘almost certainly’[26] not be accepted in this last stage of the appraisal process. Nor is it possible to reopen arguments and issues on which a determination by NICE has been reached. The phases of appraisal and, where applicable, that of appeal also follow defined timelines (see Table 2).

**Clinical guidelines**

A separate role of NICE, not to be confused with technology appraisals, is the issue of clinical guidelines that provide recommendations for the treatment and care of people by health care professionals. Clinical
guidelines are usually broader in scope than technology assessments, one consequence being that gaps in the available scientific evidence are addressed by expert opinion. Clinical guidelines are developed by a Guideline Development Group comprising health professionals and patient/caregiver representatives. Guideline Development Groups are set up by one of currently seven National Collaborating Centers, which have been established by NICE to harness the expertise of the Royal medical colleges, professional bodies and patient/caregiver organizations. Accordingly, in contrast to technology appraisals, the clinical guideline development process is predominantly administered by clinical experts, rather than economists.

### NICE appraisal of treatments for ADHD

#### Appraisal in 2000

The first appraisal of methylphenidate for ‘hyperactivity’ (HKD according to ICD-10) was conducted by NICE in 2000. In October 2000, NICE issued guidance recommending the use of methylphenidate as part of a comprehensive treatment program for ‘severe ADHD’, which had been considered roughly equivalent to HKD. (HKD according to ICD-10 criteria actually corresponds best to the ‘impaired combined subtype’ of ADHD according to DSM-IV criteria). The evidence basis for this appraisal was a technology review commissioned by NICE that

### Table 3. NICE technology appraisal process (timelines)

<table>
<thead>
<tr>
<th>Phase</th>
<th>Milestone</th>
<th>Key activities</th>
<th>Timing</th>
<th>Schedule</th>
<th>Actual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scoping</td>
<td>Remit from DoH to NICE</td>
<td>Scope published</td>
<td>2003 (Jul)</td>
<td>2003 (Aug)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Final scope published</td>
<td>NA</td>
<td>NA</td>
<td>2003 (Oct)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[ADHD appraisal temporarily halted]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assessment</td>
<td>Official start of appraisal</td>
<td>Consultees/commentators invited to participate</td>
<td>Week 0</td>
<td>2004 (Jun, 10)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Final protocol available (from assessment group)</td>
<td>Week 3</td>
<td>2004 (Jun, 30)</td>
<td>2004 (Jun, 22)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consultees’ meeting</td>
<td>Week 14</td>
<td>2004 (Sep, 17)</td>
<td>2004 (Sep, 17)</td>
</tr>
<tr>
<td></td>
<td>AR available</td>
<td>AR received by NICE</td>
<td>Week 28</td>
<td>2004 (Dec, 10)</td>
<td>2004 (Dec, 9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Consultees invited to comment on AR</td>
<td>Week 30</td>
<td>2004 (Dec, 24)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Comments on AR from consultees received</td>
<td>Week 34</td>
<td>2005 (Jan, 21)</td>
<td>2005 (Feb)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Evaluation report sent to appraisal committee</td>
<td>Week 36</td>
<td>2005 (Feb, 4)</td>
<td>2005 (Feb)</td>
</tr>
<tr>
<td>Appraisal</td>
<td>First meeting of appraisal committee</td>
<td></td>
<td>Week 37</td>
<td>2005 (Feb, 11)</td>
<td>2005 (Feb, 15)</td>
</tr>
<tr>
<td>ACD available</td>
<td>ACD produced and distributed</td>
<td>Week 40</td>
<td>2005 (Mar, 4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ACD published on website</td>
<td>Week 41</td>
<td>2005 (Mar, 11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Second meeting of appraisal committee</td>
<td></td>
<td>Week 45</td>
<td>2005 (Apr, 8)</td>
<td>2005 (Apr, 21)</td>
</tr>
<tr>
<td>FAD available</td>
<td>FAD produced and distributed</td>
<td>Week 51</td>
<td>2005 (May, 20)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FAD published on website</td>
<td>Week 52</td>
<td>2005 (May, 27)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appeal</td>
<td>Appeal announcement published on website</td>
<td></td>
<td></td>
<td></td>
<td>2005 (Jul, 7)</td>
</tr>
<tr>
<td></td>
<td>Appeal details published on website</td>
<td></td>
<td></td>
<td></td>
<td>2005 (Jul, 7)</td>
</tr>
<tr>
<td></td>
<td>Appeal panel meeting</td>
<td>Week 52</td>
<td>2005 (Aug, 25)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guidance</td>
<td>Completion of appraisal</td>
<td>Expected date of issue, after appeal hearing</td>
<td>2005 (Aug)</td>
<td></td>
<td>2006 (Mar, 22)</td>
</tr>
</tbody>
</table>


ACD = appraisal consultation document; ADHD = attention-deficit/hyperactivity disorder; AR = assessment report; DoH = UK Department of Health; FAD = final appraisal determination; NA = not applicable; NICE = UK National Institute for Health and Clinical Excellence.
drew heavily on two previously published systematic reviews, one by the US Agency for Healthcare Research and Quality (AHRQ) and a second by the Canadian Coordinating Office for Health Technology Assessment (CCOHTA). In addition, submissions were made by the manufacturers of the two immediate-release methylphenidate (MPH-IR) products available at that time (see Table 1). One UK cost-utility analysis was also available based on a Wessex Development and Evaluation Committee (DEC) report at the time of appraisal, which indicated that for methylphenidate (MPH-IR) in the treatment of hyperactivity (HKD, ICD-10) the cost per QALY estimate was £7400 to £9200 at 1997 prices for a 12 months time horizon. The NICE technology review carried an expiry date of July 2003, and the review date for the guidance was scheduled for August 2003.

Scope

In mid-2003, NICE published the scope for the imminent review, which was expanded to cover the full range of drug treatments for ADHD in children and adolescents; specifically methylphenidate (including new formulations), atomoxetine (a non-stimulant drug for ADHD that had been licensed in the US since 2002 but which was still in development in the UK at the time of the 2003 scoping), and dexamphetamine (an older stimulant drug). The scope specified the Department of Health remit to NICE as follows:

Comparators should include placebo and usual care. Outcomes should include the incidence and severity of core symptoms, problem behaviors, educational performance, measures of depression and/or anxiety, measures of conduct/oppositional-disorder-related outcomes, adverse events, and quality of life. A recommendation was also included that consideration should be given to the impact of co-morbid disorders, quality of life of other family members, and the optimal duration of treatment, where the evidence permits.

In October 2003, the appraisal process was temporarily paused to synchronize the appraisal timelines with the anticipated licensing of one of the technologies in this appraisal, and it was resumed in May 2004 (see Table 3).

Assessment

In June 2004, the final protocol for the technology assessment was provided by the Assessment Group, reflecting the scope delineated above. This was published on the NICE website in October 2004. Treatment outcomes to be included were specified confirming the scope, including incidence and severity of core symptoms, measures of depression and/or anxiety, adverse effects, and quality of life. If evidence allows, consideration should be given to the impact of comorbid disorders. The assessment protocol stated explicitly that ‘studies that have used parent and teacher ratings of hyperactivity’ would be assessed in the first instance; ‘in addition, physician ratings of clinical global impression’ would be examined. The deadline for industry submissions was September 17, 2004, and the final Assessment Report was scheduled for December 9, 2004. It also stated in detail the search strategy for evidence, which would include the following sources and study designs: conference proceedings, gray literature, randomized controlled clinical trials (of at least 3 weeks duration), full economic evaluations that compare at least two options and consider both costs and consequences, including cost-effectiveness, cost-minimization, cost-utility and cost-benefit analysis; it further explicated that ‘full paper manuscripts of any titles/abstracts that may be relevant’ would be obtained where possible.

On December 9, 2004, the Assessment Report was completed by a group of 10 authors, one of whom was a clinical expert. It was subsequently published by NICE on March 9, 2005, together with the Appraisal Consultation Document. The Assessment Report (AR), comprising 605 pages with 13 appendices, included a systematic review of the evidence and a statistical data synthesis using advanced mixed treatment comparison (MTC) techniques, a review of the submissions by manufacturers, and an economic evaluation model developed de novo by the Assessment Group. The main conclusions of the Assessment Report were that ‘(i) drug therapy seems to be superior to no drug therapy; (ii) no significant differences between the various drugs in terms of efficacy or side effects were found – mainly due to lack of evidence; (iii) the additional benefits from behaviour therapy (in combination with drug therapy) are uncertain’ and ‘Given the lack of evidence for any differences in effectiveness between the drugs, the [economic] model tends to be driven by drug cost, which differ considerably’. More specifically, it was stated that ‘for a decision taken now, with current available data, the results of the economic model clearly identify an optimal treatment strategy’ and that ‘this analysis showed that a treatment strategy of 1st line dexamphetamine, followed by second line methylphenidate immediate-release for treatment failures, followed by third line atomoxetine for repeat treatment failures was optimal’. This is noteworthy, for dexamphetamine in the United Kingdom is licensed as an ‘adjunct in the management of refractory hyperkinetic states’ only, and that no reference was made to methylphenidate modified-release preparations (MPH-MR08, MPH-MR12; see Table 1). The economic model itself was based on six
randomized clinical studies reporting Clinical Global Impression/Improvement (CGI-I) subscores after a treatment duration of 3–8 weeks, one of which was ‘commercial-in-confidence’. An unspecified number of studies excluded from the effectiveness review were nevertheless included in the cost-effectiveness analysis. Data from these studies were mathematically synthesized despite design heterogeneity. For secondary extensions of the model, further studies were included using different clinical effectiveness measures.

**Appraisal**

NICE convened the first Appraisal Committee meeting on February 15, 2005, and published an Appraisal Consultation Document (ACD) on its website on March 9, 2005. The preliminary recommendation was in favor of all three compounds – methylphenidate, atomoxetine, and dexamphetamine – as therapeutic options within their licensed indications. It stated that the decision about which product to use should be based on the presence of comorbid conditions (for example, tic disorders, Tourette syndrome, epilepsy), the different adverse effect profiles of each drug, specifically identified issues regarding compliance, for example possible problems created by the need to administer a mid-day treatment dose at school, the risk potential for drug-diversion (where the medication is forwarded on to others for non-prescription uses) and/or misuse, and the individual preferences of the child/adolescent and/or their parent/guardian. Compared to no treatment, estimated costs per QALY were below £7000 for all therapeutic options evaluated. The meeting minutes provide little information about details other than mentioning that topics of the discussion included, among others, ‘issues such as variations in measures of efficacy across trials’, ‘the availability of long-term studies’, and ‘the issue of single daily dose regimens versus multiple-dose regimens.’

Following a second Appraisal Committee meeting on April 21, 2005, the FAD was issued by NICE on its website on June 06, 2005 and these recommendations were essentially upheld. The Appraisal Committee considered evidence from the Assessment Report as well as submissions and comments from manufacturers/sponsors, professional/specialist and patient/caregiver groups, and commentator organizations on the draft scope, Assessment Report and ACD. A substantial number of (generally minor) adjustments of the ACD were incorporated into the FAD. Few of these changes are noteworthy here. First, a sentence was added to the FAD stating ‘the evidence from short-term randomized placebo-controlled trials suggests that methylphenidate is an effective treatment to reduce core symptoms of ADHD in children who continue to take the medication.’ This narrow focus upon short-term placebo-controlled designs is potentially misleading since there is also compelling evidence from two long-term studies demonstrating significant benefits from methylphenidate over two years.

Second, a statement in the FAD that ‘most studies did not indicate statistically significant differences in terms of effectiveness when comparing the immediate-release and modified-release formulations with each other’ can raise concern about the distinction between efficacy and effectiveness. This issue has perpetuated from the Assessment Report and will be addressed in more detail in its analysis.

In summary, the Appraisal Committee found that on the basis of the evidence reviewed, it was ‘not possible to distinguish between the different [treatment] strategies on the grounds of cost-effectiveness’ [FAD, p. 13], and ‘accepted the importance of having a range of drug treatment options’ [FAD, p. 17]. They ‘concluded that all three drugs are cost-effective relative to no drug treatment’ [FAD, p. 18]. The Committee also noted ‘that there were a number of important factors to be taken into account when selecting a treatment for an individual … with ADHD [including] … consideration of concordance and compliance issues, particularly with respect to the timing of doses, … previous adverse effects, comorbidities, and the preferences of patients and carers’ [FAD, p. 18]. In effect, this NICE guidance ultimately provides patients, their parents/guardians, and their physicians with a very high degree of discretion regarding the choice treatment.

**Appeal**

One consultee (the manufacturer of a modified-release methylphenidate product [MPH-MR12], Table 1) lodged an appeal against the FAD and NICE Guidance on the technologies under consideration. A public hearing was convened at NICE on August 25, 2005. Although the Appeal Panel acknowledged statistically significant differences between the effects of MPH-MR12 and atomoxetine discussed in the FAD, it upheld the view of the Appraisal Committee that it ‘had to make an overall judgment about the clinical superiority of MPH (either as IR or MR formulations) compared to atomoxetine on the totality of the available evidence’ (Decision of the Panel, p. 5). On this basis, including the observation of only extremely small QALY differences (extending only to the third decimal place) calculated by the Assessment Group, the Appeal Panel rejected the appellant’s claim that MPH-MR12 was more effective and less expensive than atomoxetine. The outcome of the appeal process was published by NICE on December 08, 2005, dismissing the
appeal while at the same time recognizing a failure of the Assessment Group to conform to the agreed assessment protocol. After the final decision of the Appeal Panel, NICE postponed the issue of guidance in order to be able to incorporate anticipated advice on the use of atomoxetine resulting from an ongoing review of its health risks and benefits by the Medicines and Healthcare products Regulatory Agency (MHRA). Final guidance was published by NICE March 22, 2006 and reflected the FAD, obviously without deviation.

Clinical guidelines

In parallel to the appraisal process, on June 16, 2004 the Department of Health (DoH) and the WAG requested that NICE develop clinical guideline on the ‘management of attention deficit hyperactivity disorder’

As part of its ‘Tenth Wave’ working program, NICE cited the remit on its website on August 11, 2004, as ‘to prepare a guideline for the NHS in England and Wales on the effectiveness of methylphenidate and other pharmacological and psychological interventions in combination or separately for the treatment of ADHD’ and that ‘the guideline should apply to the treatment of children, young people and adults where evidence of treatment effectiveness is available’

This process will be led by the National Collaborating Centre for Mental Health and is broader in scope than the technology appraisal, intended to cover ‘the full range of care routinely made available by the NHS, notably including psychological interventions as well as treatment of adults (Draft Scope, published January 31, 2006, and subsequently confirmed by the Final Scope of August 8, 2006).

Discussion

In light of the above, the NICE technology appraisal process may be compared with the ‘Accountability for Reasonableness’ conditions specified by Daniels and Sabin.

NICE process and publicity

Without doubt the first condition, ‘publicity’, was met to a great extent (see Table 4). Key documents were continuously posted on the NICE website, enabling tracking the progress and providing stakeholders with well-defined opportunities to participate throughout all phases of the process. A timetable was also published and continuously updated on the NICE website, creating a high level of predictability for stakeholders wishing to submit information. NICE demonstrated flexibility in adjusting timelines in response to changes in the environment, notably related to uncertainty surrounding the tolerability of atomoxetine (see above, and Table 3). Reasons for any changes were provided by NICE. At the same time, except for sensible adaptation to a changing environment, NICE consistently kept published deadlines. This appears especially remarkable given the substantial complexity of the clinical problem under consideration (see Introduction). Thus, there remain relatively few concerns related to transparency, which relate to rather uninformative Appraisal Committee meeting minutes, to the treatment of ‘commercial-in-confidence’ data, and to the economic models developed by Assessment Groups.

In the meantime, NICE have addressed the debate about the use of confidential information in technology assessments, and NICE have reached an agreement with the pharmaceutical industry defining

Table 4. Transparency of NICE technology appraisal process (the ‘publicity’ condition of ‘accountability for reasonableness’)

<table>
<thead>
<tr>
<th>Key features of transparency</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overall process</strong></td>
<td></td>
</tr>
<tr>
<td>Well-defined structure, detailed timelines, continuously updated, predictable opportunities for stakeholders to provide input</td>
<td>Selection of topics for appraisal (sometimes)</td>
</tr>
<tr>
<td><strong>Assessment</strong></td>
<td></td>
</tr>
<tr>
<td>Assessment protocol published; assessment report published</td>
<td>‘Commercial-in-confidence’ data withheld. Economic model not released (‘intellectual property’)</td>
</tr>
<tr>
<td><strong>Appraisal</strong></td>
<td></td>
</tr>
<tr>
<td>Appraisal committee; meeting agendas published; meeting minutes published; ACD, FAD published</td>
<td>Uninformative appraisal committee meeting minutes</td>
</tr>
<tr>
<td><strong>Appeal</strong></td>
<td></td>
</tr>
<tr>
<td>Appeal panel holding public hearings; detailed meeting minutes (Note that NICE definition of ‘appeal’ differs from that of A4R)</td>
<td>Conditions for appeal more restrictive than A4R recommendations; this appears unlikely to be fully compensated for by opportunities for stakeholder participation</td>
</tr>
</tbody>
</table>

A4R = accountability for reasonableness; ACD = appraisal consultation document; FAD = final appraisal determination
the circumstances under which non-publication of data would be acceptable.[20] This agreement between NICE and the Association of the British Pharmaceutical Industry (ABPI) was made on October 27, 2004, i.e., after the September 17, 2004 deadline set for submissions by consultees for the ADHD appraisal reviewed here (Table 3). The restriction persists, however, that economic models remain confidential,[86] to the effect of insulating a key element of the Assessment Groups’ work from public scrutiny. While transfer of intellectual property rights to academic Assessment Groups for work commissioned by NICE may be primarily a concern relevant to British taxpayers, there remain at least two further issues with the current practice. First, it might soon be tested whether the practice is in compliance with the Freedom of Information act.[111] Second, secrecy prevents public academic debate about the relative merits of modeling approaches and may impede advances of methodological standards in the science of cost-effectiveness, to the potential detriment of its various stakeholders.

Relevance

The relevance of the NICE ADHD technology appraisal is less clear. First, its scope is narrow compared to clinical guidelines in development, necessarily reducing its relevance in this respect. Second, QALYs are understood to represent an intrinsically problematic instrument to measure clinical outcomes in children.[53,54] Third, the QALY aggregation rule implicitly underlying cost per QALY rankings (i.e., ‘league tables’), even though relaxed by NICE,[120,122,123] carries with it some morally controversial (if not unacceptable) assumptions.[80,124–126] which have been found to be empirically flawed.[14] Further, it should be taken into account that emphasis on due process does not necessarily provide an indication of the exact content of this process.[83] Next, there is a real possibility that the requirement to be evidence-based[81] might have been missed during assessment, since the economic model was built exclusively on a selection of short-term clinical data.[78,85] Although the Appraisal Committee moderated the clear conclusions brought forward by the Assessment Group, Health Technology Assessments by other agencies, such as the Scottish Medicines Consortium or the Australian Pharmaceutical Benefits Advisory Committee (PBAC), have reached different results.[78,85,127,129] This discrepancy clearly warrants further inquiry.

Appeal

Finally, the NICE provisions for appeal appear markedly more restrictive than those proposed for A4R by Daniels and Sabin. Appeals are limited to specific grounds and do not allow to reopen debate[86,87], which differs from A4R recommendations.[10–12] On the other hand NICE offers, beyond A4R requirements, ample opportunities for stakeholder participation during the process. In practice, however, these opportunities may be hampered by tight timelines – consultees and commentators are given 4 weeks to submit comments on the ACD and 15 working days to lodge an appeal against the FAD – in combination with limited transparency of commercial-in-confidence information and economic models. This difficulty may be exacerbated if the Technology Assessment Report (a document comprising 605 pages in the present case) is made publicly available simultaneously with the ACD (Table 3). This notwithstanding, NICE’s appeal system may have improved consistency and, to date, has prevented appellants from proceeding to legal challenge.[20]

Enforcement

Fulfillment of the fourth A4R condition, ‘enforcement’, is beyond the scope of the present case study. One might argue that enforcement might entail an effective quality assurance system for technology assessments, and there is no indication that such a system exists at NICE. Relevance of this observation will have to be determined by the subsequent review of the assessment itself[7]. In a broader sense, enforcement might be interpreted to include implementation of NICE guidance.[8] This is currently subject of debate and research in England[30] but beyond the scope of present study.

Conclusions

The present case study of the NICE appraisal of ADHD treatments indicates a high level of transparency of the process, albeit not perfect. For instance, stakeholder participation may be impeded by the use of commercial-in-confidence information and incomplete description of economic models. Nevertheless, the NICE approach was well-structured, highly predictable, and provided, throughout its phases, stakeholders with defined opportunities to participate. Specified timelines were met or modified in a sensible way, reflecting relevant changes in the environment. This is considered a remarkable record given the complexity of the clinical problem addressed. Issues remain, however, as to the relevance of the appraisal in light of A4R criteria.

Acknowledgements

Declaration of interest: There was no third-party or industry involvement in the present study, which
was funded by the Institute for Innovation and Valuation in Health Care (InnoVal-HC). InnoVal-HC is a not-for-profit organization accepting support under a policy of unrestricted educational grants only.

Potential competing interests: The Institute and/or its staff report having received public speaking and conference attendance as well as project support from payers', physicians', and pharmacists' associations, as well as from companies including E. Lilly and Johnson & Johnson.

The author wishes to thank Joachim E. Fischer (Heidelberg) and two anonymous reviewers for helpful comments on an earlier version of this paper.

Part of this material was presented as a poster at the 11th Annual International Meeting of the International Society for Pharmacoeconomics and Outcomes Research (ISPOR), Philadelphia, PA, May 20–24, 2006. An abstract ‘Learning from NICE Technology Assessments: A Case Study of its Recent Appraisal of ADHD Treatment Strategies’ was published in Value in Health 2006;9;A71-2. Key findings from the present case study were also presented at a seminar at Monash University, West Heidelberg, Vic, September 14, 2006. Some parts were also integrated in a review presentation on the economics of ADHD delivered at the child and adolescent psychiatry (‘IACAPAP’) World Congress in Melbourne, Vic, September 12, 2006.

References

5. Towse A. The efficient use of pharmaceuticals: does Europe have any lessons for a Medicare drug benefit? Health Aff 2003;22:42-5
12. Lipman T. NICE and evidence based medicine are not really compatible. BMJ 2001;322:489-90
44. Criado Álvarez JJ, Romo Barrientos C. Variabilidad y tendencias en el consumo de metilfenidato en España. Estimación de la prevalencia del trastorno por déficit de atención con hiperactividad. Revista de Neurologia 2003;37:806-10
49. Llanas ME, Crismon ML. Methylphenidate: increased abuse or appropriate use? J Am Pharm Assoc 1999;39:526-30
58. MTA Cooperative Group. A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. Arch Gen Psychiatry 1999;56:1073-86
68. Lam LT. Attention deficit disorder and hospitalization due to injury among older adolescents in New South Wales, Australia. J Atten Disord 2002;6:77-82
70. Wilens TE, Biederman J. Alcohol, drugs, and attention-deficit/hyperactivity disorder: a model for the study of addictions in youth. J Psychopharmacol 2006;20:580-8
NICE accountability for reasonableness Schlander 221

119. Mauskopf J, Drummond M. Publication of pharmacoeconomic data submitted to reimbursement or clinical guidelines agencies. Value Health 2004;7:515-6

120. National Institute of Clinical Excellence (NICE): agreement between the association of the British Pharmaceutical Industry (ABPI) and the National Institute for Clinical Excellence (NICE) on guidelines for the release of company data into the public domain during a health technology appraisal. London: NICE, October 27, 2004


130. Freemantle N. Commentary: is NICE delivering the goods? BMJ 2004;329:1003-4

CrossRef links are available in the online published version of this paper:
http://www.cmrojournal.com
Paper CMRO-3557_4, Accepted for publication: 24 October 2006
Published Online: 15 December 2006
doi:10.1185/030079906X159461