Not playing with a full DEC: why development and evaluation committee methods for appraising new drugs may be inadequate

Nick Freemantle and James Mason

BMJ 1999;318;1480-1482

Updated information and services can be found at:
http://bmj.com/cgi/content/full/318/7196/1480

These include:

References
This article cites 6 articles, 5 of which can be accessed free at:
http://bmj.com/cgi/content/full/318/7196/1480#BIBL

4 online articles that cite this article can be accessed at:
http://bmj.com/cgi/content/full/318/7196/1480#otherarticles

Rapid responses
6 rapid responses have been posted to this article, which you can access for free at:
http://bmj.com/cgi/content/full/318/7196/1480#responses

You can respond to this article at:
http://bmj.com/cgi/eletter-submit/318/7196/1480

Email alerting service
Receive free email alerts when new articles cite this article - sign up in the box at the top left of the article

Topic collections
Articles on similar topics can be found in the following collections

Regulation (675 articles)
Resource allocation (including rationing) (286 articles)

Correction
A correction has been published for this article. The correction is available online at:
http://bmj.com/cgi/content/full/319/7203/185/a

Notes

To order reprints follow the "Request Permissions" link in the navigation box
To subscribe to BMJ go to:
http://resources.bmj.com/bmj/subscribers
between low risk of heart disease and high wine consumption had the same explanation. In the traditionally agricultural countries of southern Europe dietary fat has been low until relatively recently (which can explain the low mortality from heart disease); in addition, wine consumption is high and risk factors for heart disease vary little with occupation, so these are associated with low rates of heart disease.

Stampfer and Rimm claim an important role for alcohol, but we think it is minor. The protective effect of alcohol does not significantly increase beyond 1-2 units a day (see fig 1 in our article), so the higher average consumption per drinker in France confers no additional benefit. The data from the five largest cohort studies, recording in total 28 800 heart disease events (6500 in people who consumed ≥2 units of alcohol a day) establish this conclusively. Of the two studies cited by Stampfer and Rimm in support of a continuous association, one confirms the plateau (risk relative to lifelong abstainers was 0.7 in drinkers of 1-2 units per day and 0.8 in drinkers of ≥6 units a day) while the other recorded only 37 heart disease events in heavier alcohol drinkers. Such small studies are uninformative: the confidence intervals are so wide as to be consistent with either a continuous association or a plateau.

Stampfer and Rimm also cite the dietary interventions in the Lyon trial (more fruit and vegetables, less meat, substitution of unsaturated oils for butter and cream). They suggest that such dietary factors could help explain the French paradox, but our analysis of animal fat consumption between countries (adjusted for consumption of fruit and vegetables) took these into account. They speculate on other dietary factors not tested in the trial (cereal fibre, nuts, the glycaemic load of the diet), which are difficult to assess individually because of dietary confounding. These hypotheses, like Barker’s, assume that the paradox is unique to France. It is not; heart disease is low in relation to current risk factors in populations as diverse as Belgium, Japan, and Soweto, which share with France a recent increase in animal fat consumption but differ with respect to the other dietary factors.

After allowance for undercertification (and, in women, smoking), heart disease is 2.5 times more common in Britain than France. We believe that the time lag explanation is the major reason and that the alternative explanations offered in the commentaries are quantitatively unimportant.

Not playing with a full DEC: why development and evaluation committee methods for appraising new drugs may be inadequate

Nick Freemantle, James Mason

The consultation document A First Class Service: Quality in the New NHS heralded the introduction of the National Institute for Clinical Excellence (NICE). A key task of the institute is to provide rapid appraisal of new drugs in the period before licensing. New products may be accepted or refused NHS reimbursement, or they may be allocated “continuing research status.”

The review process described in A First Class Service and developed in a subsequent discussion paper mirrors that used for regionally funded development and evaluation committee (DEC) reports. Development and evaluation committee reports are produced by an independent arbitration committee comprising senior clinicians and others, who review the quality of available evidence, give explicit consideration to cost utility estimates, and make recommendations about interventions. Reports are published by NHS Research and Development (www.epi.bris.ac.uk/rd). Are the methods used by the development and evaluation committee up to the task?

Summary points

The National Institute for Clinical Excellence will appraise 30-50 drugs and technologies each year to inform decisions on whether these should be accepted or refused NHS reimbursement.

The appraisal process will mirror that of regional development and evaluation committees, but this may lead to poor decision making since the methods used are inappropriate.

The cost utility method of ranking treatments is based on strong assumptions and selective use of available evidence.

New drugs should be appraised in terms of physical outcomes that mean something to doctors and patients.
Evaluating new health technologies

Because of limited resources and medical cost inflation, new drugs can place a considerable strain on prescribing budgets. New and expensive drugs can benefit patients, but only if worthwhile advances can be differentiated from costly ways of achieving equivocal results. However, new drugs present particular difficulties; they have usually been evaluated in relatively few clinical trials in which measures of effectiveness are narrow—and there is often uncertainty about their value to patients. The risk of rare, unexpected, and untoward side effects is highlighted by the regularity with which licensed drugs are withdrawn from the market for safety reasons. This happened recently with the antipsychotic drug sertindole.

“Those conducting appraisals should have the complete findings of trials used for licensing new products”

Trial results are often published after the drugs have been introduced, and may not include enough data on all the outcomes. In some areas, reviewers can access substantial amounts of unpublished data, but the lack of success of the trials amnesty shows that this is not general. Retrieving unpublished data is a difficult but highly important task faced by the development and evaluation committee and will be a core task for the institute. In the appraisal process, it is proposed that pharmaceutical companies should submit trial results for consideration four months before licensing. But will all the data be required, or will a selection determined by the sponsoring company be sufficient?

Opportunity costs

The most important question is not whether a drug is expensive and can be afforded, but whether it is a worthwhile use of resources, given the competing claims on these resources. High quality evaluation of new drugs and technologies will enable clinicians and their advisors to prioritise emerging drugs so that they can either move quickly to incorporate true advances in practice or reject change where benefits remain uncertain. For example, angiotensin converting enzyme inhibitors have been shown to improve the length and quality of life of patients with heart failure, and they seem cost effective compared with other treatments. However, the overall profile of diagnosis, drug acquisition, and monitoring, and the use of other disease related resources, means that greater prescribing of angiotensin converting enzyme inhibitors requires additional funding. This is the rather crude but realistic sense in which a drug may be deemed cost effective.

Modelling the cost per QALY

The most striking feature of the appraisal methods used in development and evaluation committee reports is the explicit modelling of a cost per quality adjusted life year (QALY). The QALY tries to collapse the effects of a multidimensional phenomenon (aspects of quality of life) onto a single scale. This enables comparisons to be made across disease and treatment areas, and aids planning and resource allocation. However, ranking the value of treatments using cost per QALY is controversial, not least because of the strong assumptions that are often concealed. This is particularly likely when, as with new drugs, there is limited evidence.

The development and evaluation committee report considering donepezil hydrochloride for treating patients with Alzheimer’s disease allowed that the “range of cost-utility estimates is wide, ranging from approximately £21,000 to £200,000 per QALY,” based on data acknowledged to be “limited.” The pooled estimate of the effect on cognitive function of donepezil compared with placebo in the three trials available to the committee suggested a shift of about 2.8 points on the Alzheimer’s disease cognitive subscale (95% confidence interval 2.2 to 3.4). The development and evaluation committee equated this finding to a “three to six month delay in progression” when deriving a QALY score (figure); however, the drug trials obtained data directly on quality of life, and these showed no measurable improvement. It is simply not possible to square the arbitrary cost per QALY estimates or the conclusions in the committee’s report with the results provided by the trials.

Similarly, the development and evaluation committee report examining the place of olanzapine in treating schizophrenia based an estimated cost per QALY for one year on trials that did not exceed six weeks of properly randomised follow up. Although this report found “good evidence of excellent value for money,” claims of appreciable cost saving or a measurable gain in QALYs when olanzapine was compared with conventional treatment seem optimistic when all the available published reports on atypical antipsychotic drugs are considered. It is alarming how few of the assumptions made in the reports mentioned here could be validated from an evidence based approach.

Undertaking more valid appraisals

In the context of the National Institute for Clinical Excellence, a number of fundamental changes need to be made to development and evaluation committees’ methods. Those conducting appraisals should have the complete findings of trials used for licensing new

---

**Trials of donepezil hydrochloride**

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Mean difference</th>
<th>95% confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rogers; 5 mg</td>
<td>+</td>
<td>(0.1 to 0.9)</td>
</tr>
<tr>
<td>A301; 5 mg (unpublished)</td>
<td>-0.5</td>
<td>(-0.9 to -0.1)</td>
</tr>
<tr>
<td>A301; 10 mg (unpublished)</td>
<td>0.5</td>
<td>(0.1 to 0.9)</td>
</tr>
<tr>
<td>Pooled overall</td>
<td>0</td>
<td>(-0.1 to 0.1)</td>
</tr>
</tbody>
</table>

Donepezil hydrochloride trials: pooled standardised weighted mean difference (95% confidence interval) for a quality of life scale, random effects model.
products. It is reasonable for the NHS to expect such information from the pharmaceutical industry, unreasonable to expect clinicians to prescribe without such evidence, and important that the potential value of new treatments receives timely and independent scrutiny. This will avoid the sort of situation that occurred with the report on donepezil hydrochloride, where evidence of quality of life measured in trials and used for licensing purposes was not apparently available to the appraisers.

Producing systematic profiles
The objective of appraisal should be to enable clinicians and patients to reach informed decisions about treatments. A cost per QALY framework which requires substantial assumptions is the wrong approach. Instead, the institute’s appraisals should produce a systematic profile of the attributes of treatment (effectiveness, tolerability, safety, health service delivery issues, resource implications, and cost), and comment on the absence or inadequacy of critical information. The profile should include best estimates of effect over a range of relevant outcomes. With such a profile, it should be possible to establish the “ball park” value for money of new treatments without reference to the questionable assumptions required to model an estimate of cost per QALY. Additionally, this approach will ensure that research priorities for those drugs which are categorised as “continuing research status” are identified correctly to further our understanding rather than build a better model.

“The government’s stated aims describe a shift from clinical freedom towards care driven by protocol”

More extensive review
Rather than examining a single drug, considerable information could be gained by considering other related products, particularly those with a similar pharmacological action. The development and evaluation committee report on donepezil hydrochloride could usefully have considered trials of tacrine, which has a similar pharmacological action but is associated with greater side effects. The olanzapine report should have considered the broader range of atypical antipsychotic drugs. Some of these have been licensed for use for several years, and properly randomised longer term data are available. Although this principle implies a more extensive review, it would strengthen the validity of recommendations.

“Horizon scanning” is inefficient
A First Class Service suggests that the horizon scanning approach, by which new and emerging agents are identified early and plans made for their assessment, should be extended. This may be an inefficient use of resources that could be better deployed in the rapid appraisal of drugs. Problems with horizon scanning are exemplified by experience with drugs such as lubeluzole, whose further development was halted because of poor results of phase III trials. Before lubeluzole was withdrawn, “new drugs for acute stroke” had been identified as a priority for the NHS Health Technology Assessment Programme in the expectation that lubeluzole would be licensed. An alternative approach could be for the National Institute for Clinical Excellence to require structured outline pharmacoeconomic impact profiles for all new products at a specified time (for example, one year) before the intended date of marketing to the NHS, so that major challenges to prescribing budgets can be anticipated.

Including the right people
Finally, rapid appraisal of new technologies conducted by the institute should include representation from relevant NHS health professionals. For example, if a new or existing drug will mostly be prescribed in primary care, it is important to include general practitioners in the interpretation of the research evidence.

Conclusion
Guidance provided by the National Institute for Clinical Excellence is one aspect of a range of schemes aiming to promote clinical governance and to make NHS organisations accountable for the quality of services provided. The belief underpinning clinical governance is that acceptable and unacceptable variations in health care are separable. However, audit to ensure that the right patient gets the right treatment at the right time is unrealistic, and simple performance measures are likely to be misleading. Evidence establishing the worth of many treatments currently provided by the NHS is inadequate. The government’s stated aims suggest a shift from clinical freedom towards care driven by didactic protocol; it should give careful consideration to the dual issues of measurability and enforceability implicit in these aims. If health professionals are to respond appropriately to the concept of clinical governance, it is essential that methods of assessment are sound and transparent, reflecting both the available evidence and the uncertainties.

Competing interest: None declared.


(Accepted 10 February 1999)