ONLY IN RESEARCH WORKSHOP

23rd September 2010, NICE.

BRIEFING DOCUMENTS

Informing a decision framework for when NICE should recommend the use of health technologies only in the context of an appropriately designed programme of evidence development.

Claxton K[a,b]., Palmer S[a]., Longworth L[c]., Bojke L[a]., Griffin S[a]., Spackman E[a]., and Youn J[c].

a, Centre for Health Econopmics University fo York

b, Department of Economics and Related Studies, University of York

c, Health Economics Research Group, Brunel University

Contents:

1 Critical review of policies, practice and literature

1.1 Aims and objectives

There is a growing and diverse literature on OIR/AED and there is a need to critically review this literature to distil any common themes and core principles relevant to the NICE context. The main purpose of the review is to help to inform the development of a unifying conceptual framework within which these themes and principles can be located and understood and to enable consistent and clear terminology to be established. The specific aims of the review were:

i) To review alternative terminologies and taxonomies used to describe and classify approaches to OIR/AED and to establish their relevance to the NICE context.

ii) To identify any common themes and principles discussed in relation to OIR/AED.

1.2 Methods

The existing literature on OIR and AED is only partly represented in traditionally published papers and much is located in policy and discussion documents. The diversity in these sources was reflected in the range of search strategies employed, covering: i) traditional published literature; ii) grey literature; and iii) policy and discussion documents. In addition, relevant interest groups and policy websites were searched, reference lists of previous reviews[1] were checked and separate citation searches performed using key references and discussions with our advisory group. In reviewing the results of the systematic search and selecting relevant studies for inclusion, a relatively inclusive approach was adopted.

1.3 Literature search results

1.3.1 Identified references

A total of 55 references were subsequently included in the review. 39 of these were journal articles, 11 were policy documents (8 UK and 3 non-UK) and 5 were based on presentation slides or discussion documents.

1.3.2 Summary of the key issues for practice and policies for OIR and AED

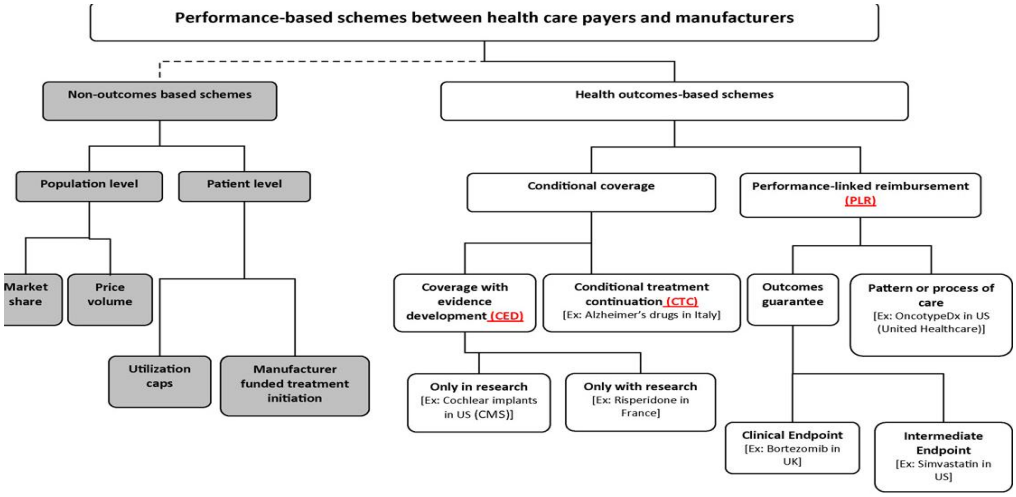
The following sections discuss the main findings in line with the key objectives of the review.

i) Review of terminology and taxonomies of OIR/AED

Multiple definitions of OIR and AED are reported in the literature, commonly provided within a broader consideration of conditional coverage or risk sharing schemes. Despite the variation in terminology that exists, a number of common aspects emerge. Most notably the use of OIR/AED is commonly defined as providing an alternative to a binary accept/reject decision for policy makers in situations where the technology does not appear to meet the standard criteria for reimbursement, predominantly because of

uncertainty surrounding the existing evidence base and where additional data collection could reduce this uncertainty [2]. The emphasis placed on uncertainty and the specific role that the collection/generation of additional evidence plays in reducing existing uncertainty is what distinguishes OIR/AED schemes from the broader range of conditional coverage or risk sharing approaches [3] (see Figure 1).

Figure 1.1 Taxonomy of performance-based schemes[3]



In considering how to appropriately define and categorise OIR/AED schemes, it is important to consider what particular terms mean in their various contexts. In the context of NICE, OIR is the term used when a recommendation is made to constrain the use of an intervention to those patients receiving it as part of a well designed programme of research. In contrast to OIR, the use of an AED recommendation, while still requiring a well designed programme research to be conducted as a condition of funding, does not necessarily limit coverage to those patients participating in the clinical study or registry. Hence, the distinction between OIR and AED is primarily the degree of coverage that each confers for reimbursement purposes. Importantly, both OIR and AED strategies are distinct from general recommendations for further research made as part of the appraisal process, where no formal link to generating evidence as a condition of coverage is made. In the UK, the use of an OIR recommendation is more common that the use of AED.

In the US, the term CED is often used as a catch all term for OIR/AED schemes. Medicare describes two forms of CED: coverage with appropriate determination (CAD) and coverage with study participation (CSP). The latter sub-type of CED would fit with OIR as it currently exists in the UK and CAD is synonymous with AED. In the taxonomy developed by Carlson et al [3] (see Figure 1), conditional coverage schemes are divided into CED and conditional treatment continuation (CTC) schemes. Within CED, two subtypes presented are OIR and 'only with research' (OWR), where OWR is similar to the term AED used in the UK.

Although there have been several previous attempts to develop taxonomies [1-4], none of these have been focussed specifically on OIR/AED schemes and typically form part of a broader categorisation of conditional coverage and risk sharing schemes. As a result, a more detailed consideration of OIR/AED schemes and the potential for further sub-types within these schemes has not been previously explored.

ii) General issues of OIR/AED

There are many issues that need to be resolved to enable the successful implementation of an OIR or AED scheme[5]. Central to this is the need to clarify the objectives of these schemes and the relevant criteria for their use. However, the critical review identified only limited information on the specific circumstances under which an OIR or AED scheme may be an appropriate policy option[6-8]. The lack of clear guidance has led to concerns expressed over ambiguity regarding their use[9, 10] and that OIR is currently being used as a 'polite no' by NICE[5]. Such concerns clearly highlight the importance of developing a clear set of principles for the use of OIR/AED by NICE.

In setting out a clear rationale for OIR/AED, NICE will also be able to work towards identifying which technologies may be suitable for such policies. Ideally these should be those with potential net benefit but also some degree of uncertainty[11]. It has also been argued that these schemes could also be used to 'fast track' particular treatments[9]. However, in addition to their role in new and emerging technologies[12], other commentators have also stressed their potential use for established interventions to inform recommendations for increased investment or for disinvestment.[13]

There are also numerous practical issues that need to be resolved for the successful use of such policies. The recent Lung Volume Reduction Surgery case study highlighted a number of challenges for OIR/AED, in particular significant opposition from the clinical community, significant level of funding required, the length of time required to complete data collection and limited access for patients in remote areas. As a result of the Multiple Sclerosis risk sharing scheme, the importance of inter-agency collaboration, achieving consensus on acceptable quality of evidence, external peer review, pre-defined clinical benefit and determining who pays for treatment was also apparent[14]. There also remain other important challenges, including the need to ensure that research is actually conducted and is fit for purpose, as well as ensuring the process is undertaken in a legal, ethical and acceptable manner[8]. Another important consideration is that these schemes need to be designed in order to develop appropriate incentives to produce evidence in a timely fashion and strategies need to be put into place to ensure that the research is actually carried out.

iii) Specific issues of OIR/AED

As well as the more general issues that need to be resolved to ensure the effective use of OIR/AED policy options, there are a number of specific issues that need to be addressed.

Evidence collection

Acquiring appropriate evidence following an OIR or AED policy is of paramount importance [3]. Without an appropriately designed and conducted study, it is likely that little will be achieved in terms of reducing the uncertainty that led to the use of such policies in the first instance. This raises a number of issues and potential challenges related to the design and funding of further research studies. Firstly, there is currently very little in the way of formalised arrangements following an OIR/ AED recommendation. A key issue identified in determining the success of these schemes is the development of working partnerships between stakeholders (clinical community, decision makers and manufacturers)[12]. Related to this is the issue of obtaining funding for OIR/AED studies and establishing who pays for the research. In relation to NICE, it has been recommended that the relevant study should either be planned or currently in progress, or alternatively that a new study could be established quickly[5]. Without secure funding the research may never be undertaken and thus the uncertainties leading to an OIR/AED still will remain.

The design of the OIR/AED study will ultimately determine its success[4] and some of the failures of existing schemes have been attributed to inappropriately designed studies[10]. Perhaps the most important consideration emerging from the literature is the issue of which type of study is most appropriate for an OIR/AED scheme[8]. OIR/AED research (unlike licensing research) is not confined to RCTs and depending on the source of uncertainties, other types of evidence may be sufficient[15]. The choice of study is ultimately context specific and related to the source of uncertainty; however it may also be influenced by factors such as cost and availability of suitable patients and collaborating clinical centres and potential ethical considerations. Clarification is also needed on how the evidence collected as a result of a OIR/AED policy will be used in an updated coverage decision[16] and also how much data is enough to inform subsequent decisions[17].

Investment and reversal costs

Investment and reversal costs have also been identified as relevant considerations in the existing literature. In particular NICE needs to determine whether a fully supportive decision (as opposed to OIR) would lead to significant irretrievable costs of implementation and if it would lead to termination of ongoing research or prevent future research[5]. There is also an ongoing challenge of disinvesting in technologies that have previously been approved[18]. Withdrawing coverage is logistically and politically difficult and it is considered more difficult to reverse a 'yes' than a 'no'[19]. Although no clear consensus

has emerged on how these costs could be factored into the decision making process[15] these could be based on formal options analysis.

Changing prices

Although discounting list prices can be thought of as an example of a risk sharing agreement[4], depending on how the OIR/AED system operates, it may also lead the manufacturer to reconsider the pricing of the technology. Allowing prices to change as part of an OIR/AED scheme also further extends the options available to decision makers. Evidence generated as part of an OIR/AED clinical study may also lead to a change in price if NICE believes that there is significant new evidence which will affect a drugs value. Something similar was observed with the MS Risk Sharing Scheme. Depending on the results observed, potential adjustments to the price of the drugs will be made at intervals to achieve an agreed cost per QALY of no more than £36,000[4]. The wider coverage associated with the MS risk sharing scheme meant that it was necessary to have an upfront agreement on price changes following provision of evidence. It is not clear, however, to what extent changing prices will reduce uncertainty regarding the coverage decision.

Ethical issues

The potential ethical issues arising from the use of OIR/AED schemes is another important theme emerging from the existing literature. For OIR, the issue of compulsory participation is often raised as a concern. Also because of practical arrangements under OIR, treatments may not be available in all areas, causing geographical inequalities[20]. If an RCT is commissioned following an OIR recommendation, this poses a greater issue in term of participation than a simple registry. These access issues in relation to an OIR policy linked a clinical trial, can be somewhat remedied by large scale, geographically diverse trial with broad inclusion criteria[21].

In addition, it has been argued that denying access to a treatment, demonstrated to be effective (however uncertain) is unethical. Patient advocacy groups may also be unwilling to accept this policy especially if the treatment is considered to be safe and efficacious[19]. These issues have important implications for both the design and successful conduct of research.

1.3.3 Summary

The critical review identified a number of important themes and principles outlined in relation to the use of OIR/AED schemes. However, much of the existing literature is relatively discursive and there is a need to provide a set of principles and to establish an analytic framework to help develop appropriate criteria for the use of OIR/AED schemes by NICE.

Selected references

1. Stafinski, T., C.J. McCabe, and D. Menon, *Funding the unfundable: mechanisms for managing uncertainty in decisions on the introduction of new and innovative technologies into healthcare systems.* Pharmacoeconomics, 2010. **28**(2): p. 113-42.
2. McCabe, C.J., et al., *Access with evidence development schemes: a framework for description and evaluation.* Pharmacoeconomics, 2010. **28**(2): p. 143-52.
3. Carlson, J., et al., *Linking payment and health outcomes: A taxomomy and examination of performance-based reimbursement schemes between healthcare payers and manufacturers.* Health Policy, 2010.
4. Towse, A. and L.P. Garrison, Jr., *Can't get no satisfaction? Will pay for performance help?: toward an economic framework for understanding performance-based risk-sharing agreements for innovative medical products.* Pharmacoeconomics, 2010. **28**(2): p. 93-102.
5. NICE Citizens Council, *Only in research*. 2007, NICE: London.
6. Cooksey, D., *A review of UK health research funding*. 2006, HM Treasury: London.
7. House of Commons Health Committee, *National Institute for Health and Clinical Excellence. Report No: HC 27-1*. 2008, The Stationary Office: London.
8. NICE AHRQ, *Managing uncertainty in healthcare. Report of a meeting organised by NICE and AHRQ*. 2008.
9. Chalmers, I., *Addressing uncertainties about the effects of treatments offered to NHS patients: whose responsibility?* Journal of the Royal Society of Medicine, 2007. **100**(10): p. 440-441.
10. Slutsky, J. *CED for promising technologies: USA*. in *HTAi*. 2008.
11. Hutton, J., P. Trueman, and C. Henshall, *Coverage with evidence development: an examination of conceptual and policy issues.* International Journal of Technology Assessment in Health Care, 2007. **23**(4): p. 425-32.
12. Chalkidou, K., A. Hoy, and P. Littlejohns, *Making a decision to wait for more evidence: when the National Institute for Health and Clinical Excellence recommends a technology only in the context of research.* Journal of The Royal Society of Medicine, 2007. **100**(10): p. 453-60.
13. Tunis, S.R. and K. Chalkidou, *Coverage with evidence development: a very good beginning, but much to be done. Commentary to Hutton et al.* International Journal of Technology Assessment in Health Care, 2007. **23**(4): p. 432-5.
14. Ramsey, S.D. and S.D. Sullivan, *Evidence, economics, and emphysema: Medicare's long journey with lung volume reduction surgery.* Health Affairs, 2005. **24**(1): p. 55-66.
15. Chalkidou, K., et al., *Evidence-based decision making: when should we wait for more information?* Health Affairs, 2008. **27**(6): p. 1642-53.
16. *Medicare's coverage with evidence development: A policy-making tool in evolution.* Journal of Oncology Practice, 2007. **3**(6): p. 296-301.
17. Chalkidou, K., et al., *Evidence-informed evidence-making.* Journal of Health Services Research & Policy, 2008. **13**(3): p. 167-173.
18. Briggs, A., et al., *Access with evidence development in the UK: past experience, current initiatives and future potential.* Pharmacoeconomics, 2010. **28**(2): p. 163-70.
19. Trueman, P., D.L. Grainger, and K.E. Downs, *Coverage with Evidence Development: applications and issues.* International Journal of Technology Assessment in Health Care, 2010. **26**(1): p. 79-85.
20. Carino, T., et al., *Medicare's coverage of colorectal cancer drugs: a case study in evidence development and policy.* Health Affairs, 2006. **25**(5): p. 1231-9.
21. Miller, F.G. and S.D. Pearson, *Coverage with evidence development: ethical issues and policy implications.* Medical Care, 2008. **46**(7): p. 746-51.

2 Key principles and assessments

Since an important social objective is to improve health outcomes across the NHS, a technology can be regarded as valuable if its approval is expected to increase overall population health. The resources available to the NHS must be regarded as fixed (certainly by NICE), so it is not sufficient to establish whether a technology is more effective than the alternative interventions available because approving a more costly technology will displace other health care activities that would have otherwise generated improvements in health for other patients. Therefore, the health expected to be gained must be compared to the health expected to be forgone elsewhere as a consequence of additional NHS costs, i.e. the net health benefits offered by the technology or whether it is expected to be cost-effective. A social objective of health improvement and an ethical principle that all health impacts are of equal significance, whether they accrue to those who might benefit from the technology or other NHS patients, is an established starting point for the NICE appraisal process.

An assessment of expected cost-effectiveness or net health benefits relies on evidence about effectiveness, impact on long-term overall health and potential harms, as well as the costs which fall on the NHS budget with some assessment of what health is likely to be forgone as a consequence (the cost-effectiveness threshold). Such assessments are inevitably uncertain and without sufficient and good quality evidence, subsequent decisions about the use of technologies will also be uncertain, i.e., there will be a chance that the resources committed by the approval a new technology may be wasted if the expected net health benefits are not realised. Equally, rejecting a new technology will risk failing to provide access to a valuable intervention if the net health benefits prove to be greater than expected. Therefore, if the social objective is to improve overall health for *both* current and future patients then the need for and value of additional evidence is an important consideration when making decisions about the use of technologies.

This is even more critical once it is recognised that the approval of a technology for widespread use might reduce the prospects of conducting the type of research that would provide the evidence needed. In these circumstances there will be a trade-off between the expected net benefits for current patients from early access to a cost-effective technology and the health benefits for future patients from withholding approval until valuable research has been conducted. A key ethical question is whether the health impacts for future patients should be considered and regarded as of similar significance to impacts on current patients.

Since publically funded research also consumes valuable resources which could have been devoted to patient care or other more valuable research priorities there are a number of trade-offs which must be made. In making these trade-offs consideration also needs to be given to uncertain events in the near or distant future, which may change the value of the technology and the need for evidence. In addition,

implementing a decision to approve of a new technology is, in general, not a costless activity and may commit resources which cannot be recovered if the decision is subsequently revised in the future. For example, there may be costs associated with implementing guidance, training health care professionals, or other investment costs associated with equipment and facilities. The irrecoverable nature of these costs can have particular influence on a decision to approve a technology if it might be revised in the near future because research is likely to report or other events may occur (e.g., launch of new technologies or change in the prices of exiting technologies).

The primary purpose of this briefing document is to provide a non-technical summary of the conceptual framework developed in the draft report which identifies the key principles and assessments which are needed when considering both approval and research decisions. We do not presuppose how the assessments ought to be made since there are a range of different types of additional information, evidence and methods of analysis which might be useful. These are outlined in the 4th briefing document and will be more fully explored and evaluated through the subsequent case studies. Section 2.1 outlines the different types of assessment needed and how each sequence might lead to different categories of guidance. Section 2.2 examines how guidance might change if there are changes to the technology or other aspects of the policy environment. Section 2.3 highlights the social values and ethical principles associated with OIR and AED.

2.1 Key principles and assessments needed

The key principles and assessments fall into 4 broad areas: i) assessment of expected costs-effectiveness and the impact on population net benefit; ii) the need for evidence, iii) whether there are sources of uncertainty which cannot be resolved by research but only over time; and iv) whether there are significant investment or future reversal costs associated with approving the technology. Guidance will depend on the combined effect of all these assessments because they influence whether the benefits of research are likely to exceed the costs and whether net benefits of early approval are greater than withholding approval until research is conducted or other uncertainties resolve.

This can be complex since these different considerations interact, e.g., the effect of investment and reversal costs will depend on the need for additional research and also influence whether research is worthwhile. The sequence of assessments, decisions and resulting guidance can be represented by a flow chart or algorithm. Although a simplification of the necessary trade-offs it: i) helps to identify how different guidance might be arrived at; ii) indicates the order in which assessments might be made; iii) identifies how similar guidance might be arrived at through different combinations of considerations; and iv) identifies how guidance might change (e.g., following a reduction in price), and when it might be reviewed and decisions reconsidered. The complete algorithm is complex (reported in Appendices A, B and C for completeness), representing the sequences of assessments and associated decisions, each leading to a particular category and type of guidance. Four broad categories of guidance are

represented and include 'Approve', 'AED', 'OIR' and 'Reject'. Each of the categories are numbered to indicate the different types of apparently similar guidance. 'Delay' is not considered a particularly useful category since NICE always has the opportunity to revise it's guidance, i.e., a decision to 'Reject' can always be revised but it is only with hindsight that 'Reject' might appear to be delayed 'Approval'. The distinction made between assessment and decision reflects the NICE appraisal process; first critically evaluate the information, evidence and analysis (an assessment) which can then assist the judgements required in appraisal when formulating guidance (a decision) .

2.1.1 Technologies without significant investment and reversal costs

Some element of investment or future reversal cost is almost always present.. However, the significance of these types of costs depends on their scale relative to expected population net benefits associated with the technology (see section 2.1.2). In this section we consider the relatively simple sequence of assessments, decisions and guidance for those technologies which do not have significant investment and reversal costs.

i) Technologies expected to be cost-effective

The sequence of assessments and decisions, which ultimately leads to guidance, starts with cost-effectiveness and the expected impact on population net benefit (see Figure 2.1). This is an assessment of expected (on average) cost-effectiveness based on the balance of the evidence and analysis currently available. Any assessment may be very uncertain with the scale and consequences of uncertainty assessed subsequently in the need for additional evidence. The sequence of assessments and decisions illustrated in Figure 2.1 demonstrates that an assessment of cost-effectiveness is only a first step and does not itself, inevitably lead to particular guidance, e.g., a technology which might on balance be expected to be cost-effective might nevertheless receive OIR guidance if the evidence that is needed cannot be acquired with approval.

Need for evidence

Some initial assessment of the need for further evidence and a decision about whether further research might be potentially worthwhile is important because a 'No' at this point can avoid further and complex assessments, e.g., a technology offering substantial and well evidenced health benefits at modest additional cost is likely to exhibit little uncertainty about whether the expected population net benefits are positive. In these circumstances, further research may not even be potentially worthwhile (the opportunity costs exceed its potential value) so guidance to Approve [4] could be issued on the basis of existing evidence and at the current price of the technology. If additional evidence is needed and further research might be worthwhile, then further assessments and decisions are required before guidance can be issued. Critically, some assessment is required of the type of evidence that is needed and whether or not the type of research required to provide it can be conducted if approval is granted (see Figure 2.1).

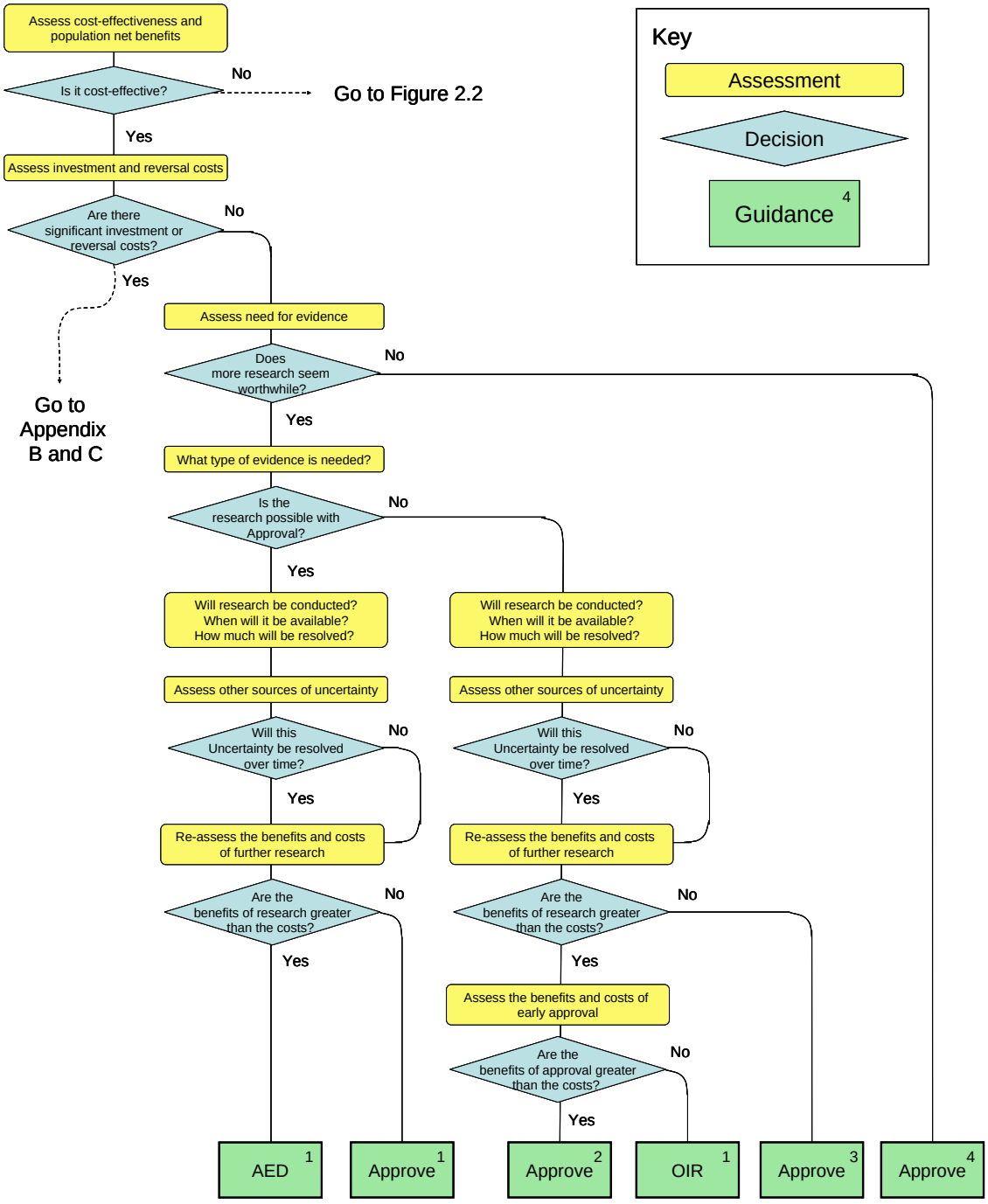
Research is possible with approval

If research is possible with approval, some further assessment of the long term benefits of research is required including: i) the likelihood that the type of research needed will be commissioned by research funders or conducted by manufacturers; ii) how long until such research will recruit and report and iii) how much of the uncertainty might be resolved by the type of research which is likely to be undertaken. An assessment of other sources of uncertainty which will only resolve over time is also needed (e.g., changes in prices or the launch of new technologies). These sources of uncertainty will influence the future benefits of research that could be undertaken as part of AED, e.g., even if the current benefits of research, which might be very likely to be undertaken are considerable, if the price of the technology is likely to fall significantly before or shortly after the research reports (or future innovation makes the current technology obsolete) then the future benefits, once the research reports, might be very limited. In these circumstances it might be better to approve (rather than AED) and reconsider whether and what type of research is needed at a later date once these uncertainties have resolved. The judgement of whether the long term benefits of research are likely to exceed its expected costs determines guidance, with AED [1] and Approve [1] dependent on 'Yes' and 'No' respectively.

Research is not possible with approval

The type of research needed may not be possible with approval, e.g., if a randomised clinical trial (RCT) may not be possible once the technology is approved (due to ethical concerns, recruitment problems and limited incentives for manufacturers). In these circumstances the expected benefits of approval for current patients must be balanced against the benefits to future patients of withholding approval to allow the research to be conducted. Initially, the same assessment of the long term value of the type of research that might be conducted if approval is withheld is still required. Similarly, the impact of other sources of uncertainty on the longer term benefits of research is also needed. However, judging that the benefits of research exceed its costs is not sufficient for OIR guidance. In addition, an assessment of whether the benefits of early approval (expected population net benefits for current patients) are greater than the opportunity costs (the net benefit of the evidence likely to be forgone for future patients as a consequence of approval) is required. If the expected benefits of early approval are judged to be less than the opportunity costs (the expected net benefits of research forgone by approval) then OIR [1] guidance would be appropriate. Alternatively, if the expected benefits of early access for current patient are judged to be greater than the opportunity costs for future patients, then Approve [2] would be appropriate. If the benefits of research were judged to be less than the costs (research is not worthwhile anyway), the technology can be approved based on current evidence and prices (Approve [3]). All these assessments, including the benefits of early approval and the value of evidence will change if the effective price of the technology is reduced (see section 2.2).

Figure 2.1 Technologies expected to be cost-effective



ii) Technologies not expected to be cost-effective

A technology not expected to be cost-effective will, on balance, impose negative expected population net benefits if it is approved. Guidance, however, will ultimately depend on subsequent assessments and decisions (see Figure 2.2).

Need for evidence

Any assessment of net benefit will be uncertain, so it remains possible that it might in fact offer positive net benefits. Therefore, the scale and consequences of this uncertainty must be considered to make an initial assessment of the need for additional evidence and whether additional research might, in principle, be worthwhile. If it is not, then guidance to Reject [4] based on existing evidence and its current price would be appropriate. Alternatively, if further research might be worthwhile then an additional assessment of whether the type of evidence and research that is needed can be conducted without approval.

Research is possible without approval

Generally, most types of research (including RCTs) will be possible without approval. Further assessment of the longer term benefits of the type of research which is likely to be conducted and when it might report is required, including the impact of other sources of uncertainty which will resolve over time. If, following this re-assessment, the expected benefits of research are judged to exceed the associated costs then OIR [2] would be appropriate. Alternatively, if the costs of research are likely to exceed the longer term expected benefits then Reject [1] would be appropriate.

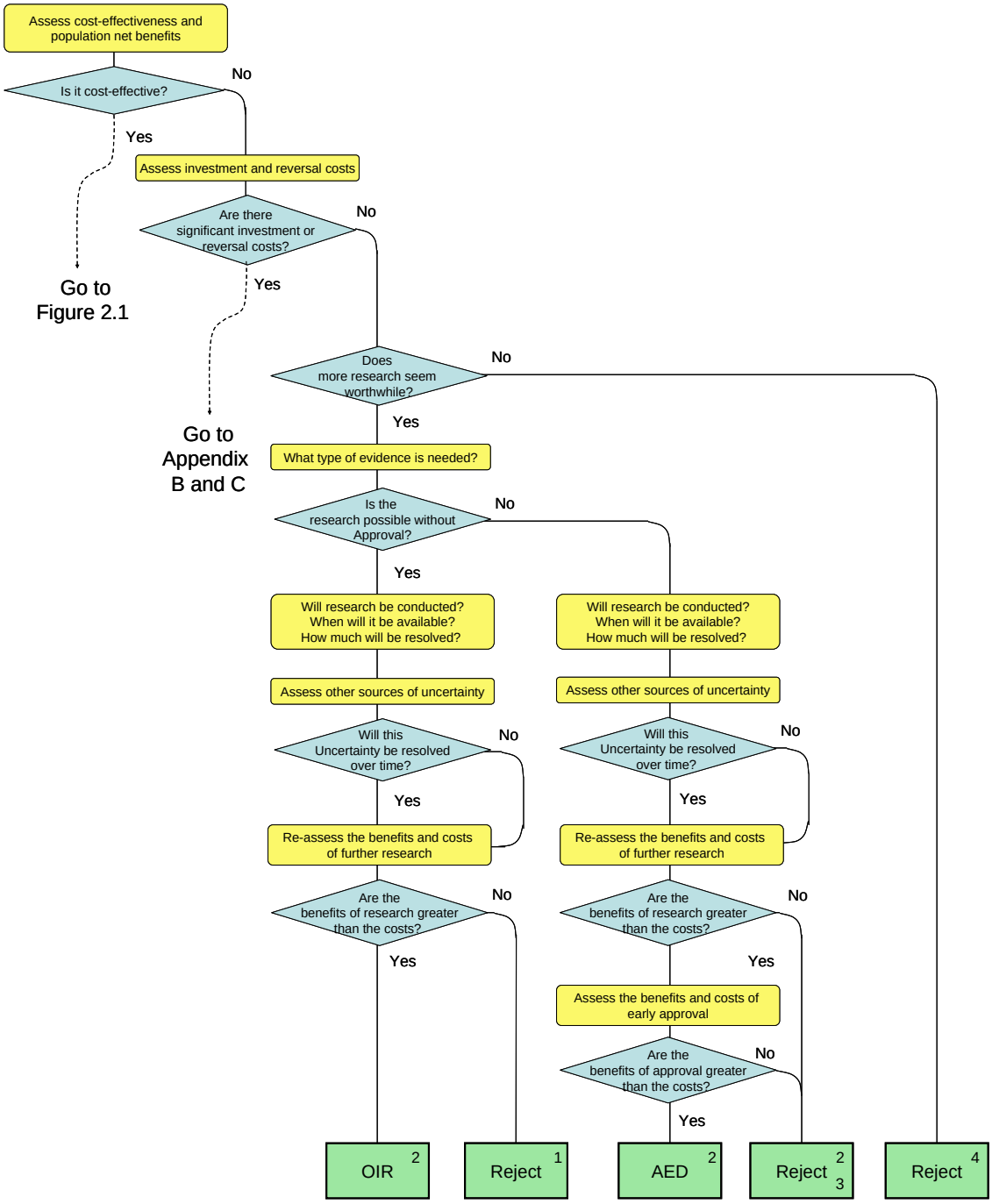
Research is not possible without approval

In some circumstances it is possible that certain types of evidence might only be acquired, or be more easily acquired (more quickly and at lower cost), once a technology is in widespread use, e.g., linking surrogates (specific to the technology) to longer term health outcomes, longer term and/or rare adverse events, or greater understanding of learning and incremental improvements in the use of a technology. In this less common situation, where the type of research needed is not possible (or is significantly more costly) without approval, the same assessment of the longer term benefits of research is required. If further research is judged not to be worthwhile following this re-assessment, the technology can be rejected (Reject [2]). Alternatively, if research is judged worthwhile an additional assessment of whether the benefits of approval exceed the costs is required. In this case, approval, which would make the research possible, will impose opportunity costs (negative expected population net benefits of widespread use of a cost-ineffective technology). The key question is whether the net benefits of the research exceed these opportunity costs. If they don't, then Reject [3] would be appropriate even though research, had it been possible without approval, would have been worthwhile. Alternatively, if the net benefits of research more than offset the opportunity costs then AED [2] would be appropriate.

Therefore, AED guidance for technologies not expected to be cost-effective is certainly possible but only appropriate in certain circumstances, where: i) the type of research needed is not possible without approval; ii) the long term benefits of the research are likely to exceed the expected costs and iii) the additional net benefits of such research exceeds the opportunity costs of approving a cost-ineffective technology. More commonly, research might be possible but more costly without approval. In these

circumstances, AED guidance could only be considered if the additional costs of research without approval exceed the opportunity costs of approving a cost-ineffective technology.

Figure 2.2 Technologies not expected to be cost-effective



2.1.2 Technologies with significant investment and reversal costs

Investment costs are those which once committed cannot be recovered should guidance be revised at a later date. In most NICE appraisals these are included in the expected cost of a technology. However, rarely is their potential additional impact explored when future events, such as research reporting or

other sources of uncertainty resolving, might mean that guidance will be revised. Investment costs are commonly thought of as expenditure on equipment or facilities which have a long life expectancy. However, there are other opportunity costs which are also irrecoverable, including: i) resources required to implement guidance; ii) training staff to use a new health technology; and iii) any period of learning where health outcomes maybe lower. However, irrecoverable investments costs may be more much common. Many technologies impose substantial initial NHS costs which are offset by cost savings and health benefits in the longer run (i.e., initially the per patient net benefit is negative), so if guidance is likely to change these initial 'losses' will not be compensated by later 'gains'. There are also costs associated with revising guidance at a later date which have a similar effect. Future reversal costs include: i) the resources required to ensure that revised guidance is implemented; and ii) the opportunity costs of any delay in fully implementing revised guidance, i.e., continued use of a cost-ineffective technology when withdrawing initial approval is difficult to fully implement.

Although aspects of investment or reversal cost are almost always present their potential significance depends on their scale relative to expected population net benefits offered by the technology. Critically, their impact depends on the risk that guidance will be revised in the near or distant future due to new evidence becoming available or changes in prices and technologies. The full algorithm becomes much more complex (Appendix B and C), so here we focus on the key differences from section 2.1.1.

i) Technologies expected to be cost-effective

If research is possible with approval and it is expected to be worthwhile, AED does not necessarily follow as previously (see 2.1.1) because the impact of investment and reversal cost must also be considered, e.g., OIR may be more appropriate than AED, even though the research would be possible with approval, because OIR avoids the commitment of irrecoverable costs until the results of research are known. This is especially so when there are also other sources of uncertainty which might resolve while the research is being conducted.

If research is not possible with approval and it is expected to be worthwhile then OIR will be appropriate if the opportunity costs of early approval are judged to exceed the benefits. These opportunity costs will now include the impact of investment and reversal costs when guidance might be revised as well as the value of evidence that will be forgone by early approval. Therefore, investment and reversal costs will tend to make OIR rather than approval more likely and especially if there are other sources of uncertainty which might resolve while the research is being conducted.

Whether or not research is possible with approval, if it is not judged worthwhile the technology should only be approved if there are no other sources of uncertainty. If there are, then an assessment of the benefits and costs of early approval is needed which takes account of investment costs and the risk that guidance might be revised in the future. Therefore, reject rather than approval is possible, even though

a technology is expected to be cost-effective, because the decision to commit the investment costs can be reconsidered once the other sources of uncertainty have resolved.

ii) Technologies not expected to be cost-effective

The presence of investment and reversal costs for technologies not expected to be cost-effective does not change the categories of guidance, or how they might be arrived at. However, it does mean that reject is more likely to be appropriate than AED (if research is not possible without approval). This is because a decision to reject, although it may be revised to approve, generally does not commit irrecoverable investment costs. There may be resources associated with making sure subsequent approval is properly implemented. But such costs are properly regarded as an investment cost associated with approval (rather than a reversal cost of reject). There may be circumstances when implementing guidance to reject a technology also requires resources if it has already diffused into clinical practice. If these are significant they should be taken into account in the same way as other investment costs, tending to make AED more likely to be appropriate.

2.1.3 Different types of guidance

Each sequence of assessment and decision, leads to different categories and 'types' of guidance for technologies with differing characteristics, indications and target populations. The different 'types' of guidance illustrates how similar guidance might be arrived at in different ways, helping to identify the particular combination of considerations which might underpin guidance, contributing to the transparency of the appraisal process. The possible categories and types of guidance are summarised in Table 2.1 where the numbers in the body of the table refer to the numbered guidance in Figures 2.1 and 2.2 and Appendices A, B and C.

The categories guidance available to NICE have wider application than is reflected in previous guidance (see briefing document 3). For example, there are 5 different types of OIR which may be appropriate when a technology is expected to be cost-effective. Indeed, OIR maybe appropriate even when research is possible with approval if there are significant investment and reversal costs. AED can only be considered when research is possible with approval but Reject remains a possibility even for a cost-effective technology if there are investment and reversal costs. Therefore, the full range of categories of guidance (OIR and Reject as well as AED and Approve) ought to be considered for technologies, which on the balance of existing evidence and current prices, are expected to be cost-effective.

It is only approval that can be ruled out if a technology is not expected to be cost-effective, i.e., cost-effectiveness is necessary but not sufficient for approval but lack of cost-effectiveness is neither necessary nor sufficient for rejection. Although likely to be uncommon, there are circumstances when AED maybe appropriate even when a technology is not expected to be cost effective. More commonly the choice of appropriate guidance will be either Reject or OIR. Which category of guidance will be

appropriate only partly depends on an assessment of expected cost-effectiveness and should only be regarded as an initial step in formulating guidance. Guidance will depend on a number of other key assessments which include: i) the need for evidence; ii) whether the type of research required is possible with approval; iii) the expected longer term benefits and costs of the type of research likely to be conducted; iv) the impact of other sources of uncertainty which will resolve over time; and v) the significance of any investment and reversal costs.

Table 2.1 Different types of guidance

Guidance	Technology is expected to be cost effective									
	No significant investment and reversal costs					Significant investment and reversal costs				
	Research					Research				
	Not needed	Possible with approval		Not possible with approval		Not needed	Possible with approval		Not possible with approval	
		Benefit > cost	Benefit < cost	Benefit > cost	Benefit < cost		Benefit > cost	Benefit < cost	Benefit > cost	Benefit < cost
Approve (12)	4		1	2	3	11, 12		5, 6	7, 9	8, 10
AED (3)		1					3, 4			
OIR (5)				1			3, 4		5, 6	
Reject (3)						7		5		6

Guidance	Technology is not expected to be cost effective									
	No significant investment and reversal costs					Significant investment and reversal costs				
	Research					Research				
	Not needed	Possible without approval		Not possible without approval		Not needed	Possible without approval		Not possible without approval	
		Benefit > cost	Benefit < cost	Benefit > cost	Benefit < cost		Benefit > cost	Benefit < cost	Benefit > cost	Benefit < cost
Approve (0)										
AED (2)				2					5	
OIR (2)		2					7			
Reject (8)	4		1	3	2	11		8	10	9

2.2 Changes to the technology and the policy environment

The type of guidance that might be appropriate will be influenced by changes in the effective price of the technology, the type of evidence available to support its use and whether further research is likely to be undertaken, either by manufacturers or research commissioners, as a result of OIR or AED guidance.

i) Effective price reduction

Any change in the effective price of the technology, either through patient access schemes (which offer some form of discount that reduces NHS costs), or direct price changes (possibly negotiated though value based pricing schemes) will affect key assessments and decisions, leading to different 'paths' through the algorithm, changing the guidance that would be appropriate. For example, provisional OIR guidance for a technology which is expected to be cost-effective (e.g., OIR[1] in Figure 2.1), might be revised to Approve with a sufficient price reduction because the benefits of early approval will be greater and uncertainty about its cost-effectiveness and therefore the value of additional evidence will tend to

be lower (see below). Similarly, AED[1] might be revised to Approve[1] in Figure 2.1. Equally, provisional guidance to reject a technology which is not expected to be cost effective (e.g., Reject [1] in Figure 2.2), might be revised to OIR [2], if the reduction in price was not sufficient to make it cost-effective. If the reduction in price was sufficient (and depending on the magnitude of this reduction), guidance might be revised to AED [1], if the research was possible with approval, or even Approve [1 or 4].
Therefore, consideration of the effect of price changes on OIR and AED is needed when assessing the potential impact of patient access schemes and more direct price negotiation through value based pricing. Flexible pricing agreements, where price is revised depending on the results of research, mean that the value of additional evidence is captured by the manufacturer (the net health benefits to the NHS will be zero whatever the results of the research). Such schemes remove the benefit to the NHS of OIR and AED at least in the medium term (until patent expiry), which suggests that manufacturers should bear the costs of research within these types of schemes.

It should be noted that the impact of a reduction in price depends on whether the primary source of uncertainty is the effectiveness of the technology or the scale of improvements in effectiveness (i.e., whether the improved effectiveness is sufficient to justify the costs). If it is primarily the former then changes in price will not affect this uncertainty and will have limited impact on the need for evidence. More commonly both sources of uncertainty are present, so changes in effective price will influence the need for evidence. However, there will be a limit to how much a reduction in price can reduce uncertainty and the need for additional evidence.

ii) Additional evidence

New technologies which are supported at NICE Appraisal by more and better quality evidence will be more likely to be approved (rather than OIR or AED) because additional evidence is less likely to be needed. Therefore, greater consideration of OIR and AED will tend to reward those manufactures who have invested in good quality and relevant evidence, with earlier approval of their technology. In addition the effect of price on OIR and AED suggests that those technologies supported by better evidence will tend get approval at higher effective prices, providing an incentive for manufacturers to invest in the type of evidence needed earlier in the development process.

iii) Assessing the prospects of research

When considering OIR or AED guidance there must be some assessment of: i) the type of research needed to address the key uncertainties, ii) whether this will be regarded as ethical and can be undertaken while the technology is approved for use, iii) whether it is likely to be a priority for public funding and be commissioned; and iv) when it is likely to report.

Although the NICE appraisal process maybe well suited to identifying the need for evidence when assessing cost-effectiveness, these other critical assessments are not necessarily ones for which NICE

and its advisory committees have particular expertise, not least because they reflect the decisions of those responsible for research design, prioritisation and commissioning. Without sufficient coordination between these communities there is a danger that OIR or AED could be issued when either the type of research required would not be regarded as ethical or feasible, or not of sufficient priority compared to other competing research needs to be commissioned. Since publically funded research is also budget constrained, it is perfectly possible that research which might be valuable from a wider NHS perspective might nevertheless not be a priority if other more valuable research might be displaced. This might be a particular concern if there is a possibility that the research could be undertaken by the manufacturer rather than displacing other research without proprietary interest. Therefore, a decision of whether OIR or AED research should be undertaken by the manufacturer or through publically funded research is one that NICE cannot properly take alone.

Although some judgement about how the research community might respond to OIR or AED recommendations when formulating guidance is clearly possible, more informed judgements and better decisions might be possible though greater involvement of the research community. For example, a research advisory committee could be constituted which could consider provisional OIR or AED guidance, making recommendations about the type of research needed, its ethics, feasibility and likely priority during the consultation period before final appraisal and guidance. It might also make recommendations about whether research should be publically funded or undertaken by the manufacturer. There are of course many different ways in which greater coordination might be achieved. However, since some of the assessments that NICE must make in formulating OIR or AED guidance are, in fact, research decisions which fall outside its remit, it would seem sensible to daw on the expertise of those involved in, and responsible for, these types of research decisions to help make these assessments.

2.3 Social value judgements and ethical principles

An OIR decision will benefit future patients but in some circumstances impose opportunity costs on current patients by withholding early approval of a technology which on balance is expected to be cost-effective. Equally, an AED decision will benefit future patients but at the expense of other current NHS patients if it is not expected to be cost-effective. Therefore, OIR and AED decisions impact on different populations of patients: i) the current patient population who could benefit from the technology, a subset of whom might enrol in research; ii) the future patient population who will benefit from the results of research and iii) other unidentified NHS patients (current or future) who will forgo health if a more costly technology is approved. The ethical implications are explored by examining whether the different types of OIR and AED decisions described in 2.1 conflict with 4 ethical principles: i) known and unknown lives; ii) current and future patients; iii) do no harm and iv) mere means.

2.3.1 Known and unknown lives

An established starting point for NICE appraisal is that improving health is an important objective of the NHS and that all health impacts whether they accrue to identifiable individuals who might benefit from the technology being appraised or other unidentified NHS patient who might forgo health are of equal significance. This principle does conflict with common emotional reactions favouring known individuals. However, such sentiment may not provide a sound ethical, or coherent, basis for social decisions, because who happens to be known in any particular instance is a matter of perspective, time and ignorance. That is, those unknown to the decision maker will be know to others and, with enough information or simply with sufficient time, those currently unidentified could become known. Therefore, decisions made on behalf of everyone served by the NHS should not distinguish between those who happen to be identified and unidentified at the time a particular decision is made.

2.3.2 Current and future patients

A similar ethical question is whether health impacts for future patients should be regarded as of equal importance and given equal weight (subject to discounting) as current patients. NICE and other NHS decisions implicitly accept that future health benefits might come at the expense of current patients, e.g., a new technology maybe approved if the future health benefits offset the opportunity costs imposed on current (albeit unknown) patients due to high initial costs. Investment in facilities, public health, prevention and clinical research are examples where the benefits accrue to future patient populations.

These two ethical principles were taken as the starting point in developing the key assessments in 2.1. The first is clearly embodied in the assessment of health forgone as a consequence of additional NHS costs and the second in the assessment of whether the benefits of early approval exceed the opportunity costs (see Figures 2.1 and 2.2). Therefore, unless other ethical principles are deemed relevant, the implications ought to be acceptable unless an argument can be made to disregard or down weight benefits to future patients or favour known over unknown patients. Assuming these two principles are acceptable there may, however, be conflict with other established ethical principles which might also be applied.

2.3.3 Do no harm

A principle of 'do no harm' could define harm as compared to: i) current state; ii) some minimum standard or ii) what patients could have had (counterfactual harm). Some type of harm maybe unavoidable but must be justified by benefits or avoidance of harm to others. For technologies expected to be cost-effective, AED would not harm current patients using any of these definitions. An OIR decision would not harm the current population compared to their current state or a minimum standard because they will continue to receive current NHS treatment. However, there will be counterfactual harm (the technology could have been approved). The question is whether this harm

can be justified. In section 2.1 OIR would only be appropriate if the expected net benefits of research to future patients exceed the opportunity cost (counterfactual harm) to current patients. This justification is consistent with i) and ii) above. Also the alternative to OIR (Approve) will not avoid harm either, because approval of this technology will harm other NHS patients compared to their current state (justified by the benefits to current and future patients). For technologies not expected to be cost-effective, an OIR decision would not harm current patients, However, an AED decision would harm other NHS patients (compared to current state) and this would only be justified by the net benefits to future patient when the research reports.

In general, the ethical implications of OIR and AED decisions appear uncontroversial. In many circumstances OIR or AED does not impose harm (however defined) on the current population and if there is harm, the key principles ensure it must be justified. Some form of harm to some population is often unavoidable and where OIR or AED is appropriate, the alternatives (Approve or Reject) impose more significant and less justified harm.

2.3.4 Mere means

A key ethical principle, based on the nature of the action rather than its consequences, is to avoid using individuals as a mere means to an end, i.e., without their consent. This principle of informed consent is central to the ethics of clinical research. Here we focus on OIR or AED decisions made by NICE, assuming that any research following OIR or AED guidance will only be conducted if deemed ethical by the clinical and research communities.

Patients are unlikely to give informed consent to participate in an RCT when they already have access to a new technology which is expected to be more effective. For this reason the type of research needed might not be possible once approval is granted, so withholding approval to allow research (OIR) might be appropriate. The question arises of whether the population of current patients are being used as a 'mere means' because access to the new technology is being withheld simply to make research possible (encourage informed consent) to the benefit future patients. Equally, AED when a technology is not expected to be cost-effective might be considered as using other NHS patents who will forgo health care, as a mere means of benefiting future patients. There are a number of responses to this problem:

i) Absurd implications

The implications of the principle suggest it is not very useful, e.g., the alternative to OIR - approval of the new technology - would deprive other (albeit unknown) patients of health care, i.e., other NHS patients would be used as a 'mere means'. The consistent application of this ethical principle implies that only technologies which saved NHS costs in each period could ever be ethically approved.

ii) Consent

The question of mere means is one of consent. Although individuals do not consent to particular NICE guidance they do consent to use the NHS, accepting that their individual interests will not always be met because decisions within collectively funded health care must balance the interests of current, future and other patient populations. Informed consent would still be required for many types of research to be regarded as ethical in the context of OIR.

iii) Double effect

The only intention of OIR guidance is to improve net health benefits for future patients by improving the evidence base for future clinical practice. If a technology is expected to be cost effective this will have the unintended but foreseeable consequence that current patients will not benefit from the technology until the research reports. However, should circumstances change (e.g., if research becomes possible with approval) the technology would be approved for current patients.

iv) Equipoise

The problem of mere means in clinical research (some patients are not allocated to the treatment expected to be 'best') is overcome by invoking the notion of equipoise; that any difference in the effectiveness of the technologies is 'sufficiently uncertain' that it is 'not possible' to judge which is better, i.e., patients are not being used as mere means because it is 'not known' which is more effective. The notion of equipoise, if applied to NICE decisions, would define 'better' as the overall impact on health (i.e., expected cost-effectiveness). Whether expected cost-effectivness is 'sufficiently uncertain' is really a question of whether further evidence is needed and research is worthwhile. Therefore, if OIR or AED was deemed appropriate (using the principles in 2.1) NICE could also be described as in equipoise. Therefore, whether or not the new technology will improve or reduce overall health is unknown so withholding approval is not using current patients as mere means.

The notion of equipoise used in clinical research defines 'better' in terms of effectiveness rather than the overall impact on the NHS. Therefore, even when OIR or AED is appropriate and regarded as ethical, research may nevertheless not be considered ethical if equipoise (in effectiveness) cannot be established and if consent would not be given by informed patients. Therefore, some consideration of whether particular types of research are likely to be considered ethical by clinical and research community is needed, including, whether guidance is likely to influence this ethical judgement. An important question is whether clinical and research communities will regard research as ethical in the context of OIR where informed consent would be given but evidence suggests that the technology, which is not currently approved and available for widespread NHS use, is more effective, i.e., equipoise cannot be established.

If the principle that all health impacts whether to identifiable or unidentified patients or current and future patients should be regarded as of equal importance and given equal weight (subject to discounting) is accepted, then the ethical implications of OIR and AED guidance are in general uncontroversial. An OIR decision when a technology is expected to be cost-effective does, however, pose the question of whether current patients are being used as a mere means to benefit future patients (similarly, an AED decision when the technology is not cost-effective poses the question of whether other NHS patients are being used as a mere means to benefit future patients). However, closer examination of this principle of action suggests there are a number of reasons why this concern might be set aside.

2.4 Questions for group discussion

2.4.1 Do the key principles and assessments cover the range of considerations which might affect guidance (OIR, AED, Approve and Reject)?

i) Do they provide useful guidance on when OIR and AED might be considered and are there other important considerations that should be included as key principles?

ii) Is the distinction between assessing the need for evidence and whether research is likely to be worthwhile and likely to be commissioned useful? What implications might this have for appraisal?

2.4.2 Do the possible changes to the technology or the policy environment that might influence OIR and AED (described in section 2.2) include those that are most relevant?

2.4.3 Are the social values and ethical principles associated with OIR and AED acceptable?

i) Is the principle that all health impacts whether to identifiable or unidentified patients or current and future patients be regarded as of equal importance acceptable?

ii) How and who should assess whether the research needed as part of OIR or AED is ethical and will this depend on whether the technology is approved for widespread NHS use?

3 A review of NICE Technology Appraisal guidance

NICE issues Technology Appraisal (TA) guidance on the use of new and existing health technologies in the NHS. The recommendations are formulated by independent Appraisal Committees after a review of evidence and other submissions from interested parties on the technology of interest. NICE provides its Committees with general guidance on the HTA methodologies and social value judgements it considers to be most appropriate for the formulation of NICE guidance (NICE 2008a, NICE 2008b). These documents include guidance on the assessment of cost-effectiveness and other considerations that should be taken into account when formulating recommendations. The document describing the social value judgements that its committees should consider states that "NICE's advisory bodies may sometimes recommend that an intervention is used only within a research programme. They should consider whether the intervention is reasonably likely to benefit patients and the public, how easily the research can be set up or whether it is already planned or in progress, how likely the research is to provide further evidence, and whether the research is good value for money." (NICE 2008b)

We have reviewed all published NICE guidance and draft guidance (where publicly available) to identify which appraisals have recommended the collection of further evidence in the guidance. The documents have been reviewed to establish what led to the formulation of the research recommendations.

3.1 Definitions and methods used in the review

NICE's guidance documents are published in a standardised format. The guidance to the NHS is presented in Section 1 and the rest of the document provides an overview of the evidence, an explanation of how the evidence was interpreted by the Committee (Section 4.3) and additional information to assist the implementation of the guidance. Its guidance documents ('Final Appraisal Determinations') are made publicly available and can be appealed by specific stakeholders before becoming final guidance to the NHS. In 2002, the NICE process was amended to also publish draft guidance ('Appraisal Consultation Documents') for public consultation. It is these documents that formulate the basis for this review.

The following definitions are used in the review:

- Only in research (OIR) – a recommendation which states that the technology should not be used routinely and which recommends that further research is conducted in the guidance section (Section 1 of the FAD or ACD).
- Approval with evidence development (AED) – a recommendation which states that the technology should be used routinely for the population or a subgroup, and which recommends further research is conducted in the guidance section.

Recommendations that refer to ongoing or planned research in the guidance section of the documents are included as OIR/AED recommendations. Draft recommendations that request further clarification or analysis from the sponsor of the technology (sometimes referred to as 'minded no' recommendations in the Single Technology Appraisal process) are not included as OIR/AED recommendations as they usually require the reanalysis of existing data rather than additional data collection. Early draft guidance documents were not made publicly available and these draft documents are therefore not included in this review (appraisals 1-43, except 32). Documents that have been publicly released but subsequently withdrawn and removed from the NICE website are included in the review (for example, guidance that has been replaced by a subsequent review), and have been obtained directly from NICE where appropriate. The documents containing OIR/AED recommendations were cross-checked with a categorisation of appraisals conducted by NICE to check for potential omissions. Data from each document were extracted using a template to record information for each of the appraisals that included OIR and/or AED recommendations. The data were analysed to identify common characteristics of appraisals that included OIR and/or AED recommendations. Where recommendations changed between draft and final guidance, the explanations for the change were reviewed and assessed against the key principles.

3.2 How has NICE used OIR and AED recommendations to date?

Of the 184 appraisals conducted up to November 2009, forty included OIR/AED recommendations in the draft and/or final guidance. This included 29 FADs and 31 ACDs. Multiple ACDs (and in some cases FADs) were released for some appraisals; the 31 ACDs related to 25 appraisals. Table 3.1 shows the frequency of OIR and AED recommendations in the documents. A list of all the appraisals including OIR and AED recommendations is provided in the appendix.

Table 3.1 Frequency of OIR/AED recommendations in NICE guidance

	ACDs	FADs
OIR	26	25
AED	5	4
Total	31	29

Most appraisals that included research recommendations within the guidance phrased the requirements for the technology to be used 'only in research' rather than as 'approval with evidence development'. However, there were changes in the inclusion of OIR/AED recommendations between ACD and FAD for some appraisals. Eleven appraisals included OIR/AED recommendations in the draft guidance, but not in the final FAD. Three appraisals included OIR/AED recommendations in the final guidance, but not in the draft guidance (ACDs were unavailable for a further 12 appraisals).

Most pieces of NICE guidance included several recommendations. These related to multiple technologies, multiple indications or different settings for the use of the technology. Over half of the OIR/AED recommendations specified the need for further research in specified subgroups of patients (52% of OIR/AED recommendations in final guidance documents). In approximately a quarter of cases, the OIR/AED recommendations targeted a subset of the technologies included in the appraisal.

Table 3.2 shows the frequency of OIR/AED recommendations by year of issue. Sixteen percent of all guidance included an OIR/AED recommendation. No final guidance included OIR/AED recommendations in 2007. The proportions of published guidance including OIR/AED recommendations is lower during the last few years compared to earlier appraisals, however they are based on relatively small numbers.

Table 3.2 Number of OIR/AED recommendations by year of publication

Publication year	**ACDs with OIR/AED**	**FADs with OIR/AED**	**% OIR/AED of all final guidance (Total number of TA guidance published)**
2000	N/A	6	35% (17)
2001	N/A	2	14% (14)
2002	6	6	26% (23)
2003	3	4	21% (19)
2004	2	1	8% (13)
2005	7	3	43% (7)
2006	6	4	21% (19)
2007	3	0	0% (21)
2008	4	2	6% (32)
2009	0	1	5% (19)
Total	31	29	16% (184)

Differences in the frequency of OIR/AED recommendations were observed between the two NICE appraisal processes. Of appraisals issued through the MTA process, OIR or AED recommendations were included in the final guidance of 28 appraisals and in the draft guidance of 23 appraisals. These FADs account for 19% of all final guidance issued within the MTA process. Fewer appraisals conducted through the STA process included OIR/AED recommendations: 2 ACDs and 1 FAD. This accounts for just 2% of all final guidance issued through the STA process up to the time this review was conducted.

In absolute terms, OIR/AED recommendations were more common for cancer treatments (n= 10; 34% of all FAD OIR/AED recommendations) and musculoskeletal conditions (n=7; 24% of all FADs). However NICE has appraised a large number of treatments for cancer; 28% of all published appraisals over the review period. Only 7% of all NICE TA guidance has related to musculoskeletal conditions and so it appears that a disproportionate amount of these have included OIR/AED recommendations compared to other appraisals for other conditions. The majority of appraisals with OIR/AED recommendations related to appraisals of drugs (n=16). However, as a proportion of the total number of final guidance issued, a greater proportion of guidance for procedures (47%) and devices (27%) included OIR/AED recommendations compared to drug appraisals (11%).

3.3 What considerations contributed to the OIR and AED recommendations?

3.3.1 Type of evidence requested

The accounts of the NICE Committee's considerations were reviewed to establish the rationale for the OIR/AED recommendations. Table 3.3 shows the stated reasons for issuing the OIR/AED recommendations. In some cases, the documents cited more than one reason for the OIR/AED recommendation. Of the five appraisals that did not explain the rationale for the OIR/AED recommendation, four were issued prior to a section on the Committee's considerations being routinely included in the documents. The OIR in the other appraisal related to three specific subgroups of patients: two were not referred to in the Committee's considerations at all and it was stated that there was 'no clinical or modelling evidence, or expert opinion' to support the use of the technology in the third subgroup (TA75: hepatitis C).

A need for further evidence on the relative effectiveness of the intervention in the overall population or the OIR/AED subgroup was the most commonly cited reason for issuing the OIR/AED. A need for longer-term data was also frequently cited. Uncertainty in the cost-effectiveness estimates was also a common consideration; however in all cases this was coupled with a need for further clinical evidence. Concern about the budget impact of introducing the technology, investment and reversal costs and the potential impact on ongoing research did not lead to the OIR/AED recommendation in any of the appraisals.

Table 3.3 Types of reasons for including research recommendations within the guidance

Reason for requesting further research	Number of ACDs	Number of FADs
None stated	1	5*
Clinical effectiveness		
Need for more evidence on relative effectiveness	19	16
Need for data on clinical effectiveness in the target OIR population	15	9
Need for longer-term data	13	7
Need for information on adverse effects	6	4
Need for data on natural history/ progression of disease	2	0
Need further evidence to support mechanism of treatment action	4	3
Cost-effectiveness		
Uncertainty in cost effectiveness estimates	13	6
Need for cost effectiveness data with an appropriate comparator	2	2
Need for more data on quality of life impact	6	3
Need for more data on costs	1	1
Other uncertainties		
Budget impact	0	0
Investment and reversal costs	0	0
Potential impact on ongoing research	0	0

Note that there may be multiple ACDs for each appraisal and that there may be more than one stated rationale for requiring further research.
* This includes some appraisals published before the Committee's considerations were routinely reported in a section of the guidance documents.

3.3.2 Assessment of cost-effectiveness

The NICE Methods Guide states that all appraisals should include an assessment of cost-effectiveness as a standard part of the NICE appraisal process (NICE, 2008a). Most of the guidance documents reported several different estimates of incremental cost-effectiveness based on analyses submitted by different stakeholders, relating to different uses of the technology or based on different sets of assumptions or evidence. The formal assessment of cost-effectiveness was not always conducted or reported in the ACD or FAD for the use of the technology specified in the OIR/AED recommendation. Table 3.4 below shows the ICERs (incremental cost per QALY gained) for the overall population and for the specific OIR/AED indication where this differs. As the ICER considered to be most plausible by the

Committee was not stated in all circumstances, the basecase estimate from the Evidence Review Group (ERG) or Assessment Group (AG) is also reported for information.

Table 3.4 Cost-effectiveness of technologies with OIR/AED recommendations

Incremental cost per QALY	OIR/AED indication (n)*		Total population (n)	
	Committee's estimate	AG/ERG's estimate	Committee's estimate	AG/ERG's estimate
Not reported	30 (68%)	22 (50%)	23 (52%)	10 (23%)
Dominates	0	1 (2%)	0	1 (2%)
ICER <£20,000	0	2 (5%)	1 (2%)	4 (9%)
ICER £20-30,000	4 (9%)	1 (2%)	5 (11%)	3 (7%)
ICER >£30,000	9 (20%)	15 (34%)	12 (27%)	22 (50%)
Dominated	0	2 (5%)	0	1 (2%)
Other	1 (2%)	1 (2%)	3 (7%)	3 (7%)
Total*	44 (100%)	44 (100%)	44 (100%)	44 (100%)

Key: Dominates = the technology is more effective and less costly than the alternative. Dominated = the technology is less effective and more costly than the alternative. Other = ICER was not framed in terms of a cost per QALY or the basecase was presented as a range that could not be classified into the categories.
*Total from 40 appraisals with OIR/AED recommendations (29 FADs and 11 ACDs). Two appraisals reported ICERs separately for two technologies with OIR/AED recommendations and one reported ICERs separately for three technologies.

Most documents did not cite the incremental cost-effectiveness ratio considered by the Appraisal Committee to be the most realistic. The ERG/AG estimates were more frequently reported and it is likely that they were available in supporting documents where not directly referred to in the ACD or FAD. In some cases ICERs were reported but were based on analyses that did not use the QALY as the outcome measure. For example, TA5 on the use of liquid based cytology reported ICERs of £1100 and £2500 per life year gained depending upon the length of the screening interval. Where ICERs were not directly reported, there was usually an indication of whether the technology was considered to be cost-effective. For example TA 65 stated that "*The clinical and cost effectiveness of rituximab in patients with localised disease has not been established*" and TA44 stated that "*Appraisal Committee believed that metal on metal hip resurfacing arthroplasty was likely to be of similar cost effectiveness to conventional total hip replacements in people who were expected to outlive the device".*

Table 3.5 shows the frequency of technologies considered to be cost-effective when used in the context of the OIR/AED recommendation; the type of recommendation is also presented. In most cases (79% of FADs with OIR/AED) the technology was not cost-effective and an OIR recommendation was issued. The technology was considered likely to be cost-effective in three of the four cases where an AED

recommendation as issued. There were examples of OIR being used where the technology was probably cost-effective based on the accepted analyses. Both of these appraisals (TA5 on liquid based cytology and TA51 on computerised cognitive behavioural therapy) requested that pilot implementation programmes be undertaken prior to routine introduction of the technologies in the NHS. One appraisal included an AED recommendation where the ICER was higher than the usual threshold range: in this case the ICERs were £27,000 to 35,000 and close to the upper end of the cost-effectiveness threshold range (TA36 – etanercept and infliximab).

Table 3.5 Type of recommendation and conclusion regarding cost-effectiveness (in FADs only)

	OIR	AED	Total
Considered cost-effective	2	3	5
Not considered cost-effective	23	1	24
Total	25	4	29

3.3.3 Investment and reversal costs

Investment and reversal costs were not quantified in the guidance documents and concern about these costs were not cited as a rationale for any of the OIR/AED recommendations. However, TA51 on computerised cognitive behavioural therapy (CCBT) did suggest concerns regarding the levels of training required for the implementation of a recommendation to routinely introduce CCBT in the NHS. "*Further information is required about the extent of training needed and circumstances under which different staff could provide support for users of CCBT" (TA51).*

3.3.4 Possibility of conducting research with and without approval

No appraisals cited a concern for the impact of recommendations on ongoing trials as a rationale for the OIR/AED guidance. In addition, the possibility of conducting research was not explicitly noted in most appraisals.

Evidence on relative effectiveness may be more difficult to collect if a technology is in routine use as patients may be less willing to be included in a randomisation procedure that could allocate them to previous ('old') standard care. Table 3.6 reports the frequency of appraisals where a need for more data on relative effectiveness was cited as a consideration for the OIR/AED recommendation. Some of these appraisals cited other considerations in addition to relative effectiveness (see Table 3.3).

Table 3.6: Type of recommendation and stated need for evidence on relative effectiveness (FADs only)

	OIR	AED	Total
Data on relative effectiveness required	19	1	20
Data on relative effectiveness not required	6	3	9
Total	25	4	29

A need for data on relative effectiveness was cited in 69% of FADs. Only one of these FADs included an AED recommendation: TA113 (inhaled insulin) noted a gap in the evidence on clinical effectiveness for the highly selective subgroup of patients targeted in the recommendations, but also stated that data would be most appropriately collected through a registry study. Of the nine appraisals which did not require further evidence on relative effectiveness, six included OIR recommendations. Five of these did not report the Committee considerations leading to these recommendations. One appraisal referred to clinical studies "*designed to generate robust and relevant outcome data, including data on optimum treatment regimens, long-term outcomes, quality of life and costs*" but did not specifically mention relative effectiveness (TA68: photo-dynamic therapy - age-related macular degeneration).

Although explicit reference to the Committee's consideration of the likelihood of research being conducted was rare, there were some instances where this was recorded. For example, the ACD for TA129 (Bortezomib for the treatment of relapsed multiple myeloma) stated that bortezomib monotherapy was not recommended except for use in well-designed clinical studies; however this research recommendation was removed in the FAD (which was amended following appeal). The change from an 'OIR' to a 'reject' decision was due to anticipated difficulties in the research being conducted.

3.3.5 The impact of price on OIR/AED recommendations

NICE considers the list price of technologies (e.g. as reported in the British National Formulary for drugs) and possible changes in price over time are not usually taken into account. However, NICE has recently established a formal process for the consideration of 'patient access schemes' (NICE, Sept 2009). These schemes may involve a manufacturer formally offering a reduction in price of the technology to the NHS or may offer other schemes that reduce the overall cost of the technology to the NHS (for example, by providing some courses of treatment at no cost). Whilst this formal process is new, 'access' or 'risk sharing' schemes have previously been adopted, for example the DH risk sharing scheme for beta interferon (DH circular HSC 2002/04).

Of the appraisals that included changes in research recommendations between draft and final guidance stage of production, one included a patient access scheme. As noted previously, the ACD for TA129 (Bortezomib for the treatment of relapsed multiple myeloma) stated that bortezomib monotherapy was not recommended except for use in well-designed clinical studies in the first ACD and that the Committee was not persuaded of its cost-effectiveness. Although this research recommendation was withdrawn in the first FAD issued (i.e. a 'reject' decision was made), the final guidance approved the technology following a reconsideration of its cost-effectiveness of its provision in accordance with a patient access scheme in which the manufacturer reimbursed the cost of the drug for patients whose disease inadequately responds. In another appraisal, the OIR recommendation was revised to an approval after the Committee revised their estimates of cost-effectiveness based on discounted prices of the technology along with further information on quality of life improvements (TA166: cochlear implants).

3.3.6 Other considerations

Considerations around whether uncertainties in the evidence base would resolve over time were not explicitly mentioned as reasons for issuing OIR or AED recommendations. In addition, the relative costs and benefits of conducting research were not reported as considerations of the Committee when formulating its research recommendations.

3.4 How successful have the recommendations been?

Each piece of NICE guidance is considered for update at a specified length of time after publication. In order to examine the impact of OIR/AED recommendations on evidence collection and the possibility of evidence generation with approvals, appraisals with OIR/AED recommendations were examined for differences in the data available between appraisals and their reviews. Among the OIR/AED recommendations in FADs, ten were reviewed, including two that were incorporated into clinical guidelines (CG). Table 3.7 provides details of the appraisals, whether additional evidence was provided and the change to the OIR/AED recommendation (new evidence for other recommendations included within the guidance is not noted in the table).

In the majority of reviewed appraisals (n=7), new evidence informing the OIR/AED recommendation was available for the review. In four of these reviews, the OIR or AED restriction was removed and the technology was recommended routinely. In two cases, the additional evidence was considered insufficient to warrant a change in the OIR recommendation. In the remaining appraisal the OIR was revised so that some technologies within the class were recommended routinely whereas OIR recommendations were issued for others (TA51: CCBT).

In three cases no new evidence was provided on the OIR/AED indication. For the review of TA6, no new RCT data were available for the OIR recommendation, which was made more restrictive in the review guidance. New evidence on clinical effectiveness was not available for the review of TA33 but further information on adverse effects was provided. This was considered inadequate and no change was made to the OIR recommendation. The OIR recommendation was removed from the review of TA37 despite a lack of new evidence presented. The documents state that the reasons for this were a reduction in demand for the drug in this setting (it had since become licensed and NICE approved for treatment of an earlier stage of disease) and concerns about the feasibility of future data collection.

Table 3.7 Details of appraisals that underwent review

Original	Review	Additional evidence provided for the OIR/AED indication?	Summary of change to OIR/AED guidance
TA 5	TA 69	New evidence available . Pilot implementation programmes were requested in the OIR. A Scottish implementation study and other evidence became available.	**OIR removed** Technology recommended.
TA 6	TA 30	No additional evidence presented.	**OIR amended** TA30 includes an OIR recommendation for a more restricted indication.
TA 16	TA 89	Updated RCT data and new non-RCT evidence.	**OIR unchanged** **(**Some amendments to types of evidence required)
TA 17	TA 105	New evidence (RCTs) available	**OIR removed** Technology recommended.
TA 30	CG81	New evidence (RCT & registry data) available	**OIR removed** CG81 did not include the OIR indication in the scope of the guideline.
TA 33	TA 93	No new RCTs, but updated adverse effect data	**OIR unchanged**
TA 36	TA 130 (only in ACD)	New RCT and registry data available	**AED removed.** Technology recommended. A new OIR for another use of the drugs in ACD, but this was removed in the FAD (no guidance provided for this use of the drugs)
TA 37	TA 137	No new evidence presented	**OIR removed** Technology recommended.
TA 51	TA 97	New evidence (RCT & non-RCT) available	**OIR amended** Original OIR was for CCBT as a class. Amended OIR was for specific CCBT packages.
TA 72	CG 79	New evidence (RCTs) available	**OIR unchanged**

3.5 Conclusions

NICE is able to recommend that technologies are used in a research context as part of its remit and it has issued OIR/AED recommendations in 16% of its published guidance. These recommendations have

most frequently taken the form of OIR recommendations for technologies considered to be cost-ineffective based on the evidence available at the time of the appraisal. However, a range of other recommendations have been made including OIR for (likely) cost-effective technologies and AED for both cost-effective and cost-ineffective technologies. The most common reason cited for OIR/AED recommendations was the need for further evidence on relative effectiveness. Potential investment and reversal costs have not explicitly led to OIR/AED recommendations. OIR/AED recommendations have been rarely used in appraisals conducted through the STA process. It is unclear why this should be given that technologies appraised through this process are usually newer and have a more limited evidence base than technologies appraised through the MTA process. Changes in the evidence base of reviewed appraisals show some limited success in the implementation of OIR/AED recommendations. However, using the reviews as an indication of success is somewhat limited as a lack of new evidence could have led to the postponement of planned reviews.

References

NICE (2008a). Guide to the Methods of Technology Appraisal. National Institute for Health and Clinical Excellence. London.

NICE (2008b). *Social Value Judgements: Principles for the Development of NICE Guidance.* National Institute for Health and Clinical Excellence. London.

Department of Health. Health Service Circular 2002/004 - Cost effective provision of disease modifying therapies for people with Multiple Sclerosis. London; 2002

3.6 Questions for group discussion

3.6.1 Is the review a fair summary of how NICE has, to date, considered the use of recommendations including evidence collection?

i) Are there other examples of appraisals where OIR/AED should have been considered? Why were OIR/AED recommendations not used?

ii) Are there other considerations that NICE currently takes into account when formulating OIR/AED recommendations and which are not reflected in the review?

3.6.2 Are there any procedural or organisational constraints that limit the ability of NICE to use OIR/AED recommendations?

i) What could be done to overcome these constraints?

ii) Are the constraints different in the STA and MTA processes?

iii) Should the increased use of patient access schemes affect the way in which OIR/AED recommendations are used by NICE? If so, how?

3.6.3 To what extent are OIR and AED recommendations evident in other NICE programmes (e.g. clinical guidelines, public health)?

i) To what extent are the considerations used in these other programmes in using OIR/AED recommendations different to those from the appraisals programme?

Appendix: Appraisals including OIR or AED recommendations in draft or final guidance

TA	Title	Date	ACD/ FAD
2	Prostheses for primary total hip replacement	April 2000	FAD only*
5	Cervical cancer - liquid based cytology	May 2000	FAD only*
6	Taxanes for Breast cancer	June 2000	FAD only*
8	Hearing aid technology	July 2000	FAD only*
16	Knee joints (defective) - autologous cartilage transplantation	Dec 2000	FAD only*
17	Colorectal cancer - laparoscopic surgery	Dec 2000	FAD only*
23	Recurrent malignant glioma (Brain cancer) - temozolomide	April 2001	FAD only*
30	Taxanes for Breast cancer	Sept 2001	FAD only*
33	Colorectal cancer (advanced) - irinotecan, oxaliplatin & raltitrexed	March 2002	FAD only*
35	Arthritis (juvenile idiopathic) - etanercept	March 2002	FAD only*
36	Rheumatoid arthritis - Etanercept and infliximab	March 2002	FAD only*
37	Lymphoma (follicular non-Hodgkin's) - rituximab	March 2002	FAD only*
44	Hip disease - metal on metal hip resurfacing	June 2002	ACD & FAD
50	Leukaemia (chronic myeloid) - imatinib	May 2002	ACD only
51	Depression and anxiety - computerised cognitive behaviour therapy (CCBT)	Oct 2002	ACD & FAD
60	Diabetes - patient education models	Nov 2002	ACD only
65	Aggressive Non-Hodgkin's lymphoma - rituximab	Sept 2003	ACD & FAD
68	Macular degeneration (age-related) - photodynamic therapy	Sept 2003	ACD & FAD
70	Leukaemia (chronic myeloid) - imatinib	Oct 2003	FAD only
72	Rheumatoid arthritis - anakinra	Nov 2003	ACD & FAD
75	Hepatitis C - pegylated interferons, ribavirin and alfa interferon	Jan 2004	ACD & FAD
86	Gastro-intestinal stromal tumours (GIST) - imatinib	May 2004	ACD only
89	Cartilage injury - autologous chondrocyte implantation (ACI)	May 2005	ACD & FAD
92	Tooth decay - HealOzone	July 2005	FAD only
93	Colorectal cancer (advanced) - irinotecan, oxaliplatin and raltitrexed	Aug 2005	ACD & FAD
97	Depression and anxiety - computerised cognitive behavioural therapy (CCBT)	Feb 2006	ACD & FAD
99	Immunosuppressive therapy for renal transplantation in children and adolescents	April 2006	FAD only
104	Psoriatic arthritis - etanercept and infliximab	June 2005	ACD only
111	Alzheimer's disease - donepezil, galantamine, rivastigmine and memantine	Nov 2006 (Update: 2009)	ACD & FAD
113	Diabetes (type 1 and 2) - inhaled insulin	Dec 2006	ACD & FAD
121	Glioma - carmustine implants and temozolomide	Dec 2005	ACD only
129	Multiple myeloma - bortezomib	July 2006	ACD only
130	Rheumatoid arthritis - adalimumab, etanercept and infliximab	Feb 2006	ACD only
135	Mesothelioma - pemetrexed disodium	March 2006	ACD only
142	Anaemia - erythropoietin (alpha and beta) and darbepoetin	July 2005	ACD only
143	Ankylosing spondylitis - adalimumab, etanercept and infliximab	July 2007	ACD only
159	Pain (chronic neuropathic or ischaemic) - spinal cord stimulation	Oct 2008	ACD & FAD
163	Ulcerative colitis (acute exacerbations) - infliximab	Dec 2008	ACD & FAD
166	Hearing impairment - cochlear implants	Dec 2007	ACD only
167	Abdominal aortic aneurysm - endovascular stent-grafts	Feb 2009	ACD & FAD

* ACDs were not publicly available for these appraisals

4 Informing assessments and decisions

The key principles and assessments which are needed when considering OIR or AED guidance were outlined in briefing document 2. The possible sequences of assessment and decision which lead to a particular categories and types of guidance were represented as an algorithm in section 2.1 (see Figure 2.1 and 2.2 and Appendices A, B and C). How these assessments ought to be informed and decisions made, and whether existing methods of appraisal are sufficient, or whether addition information, evidence and analysis might be useful was not addressed. In this document we outline additional information and evidence which might be useful and a range of methods of analysis which could be used to inform each of the assessment and decisions within the algorithm. We take existing methods of NICE appraisal as an excepted starting point and focus instead on what additional information and analysis might feasibly be included in appraisal and how it might be usefully interpreted to inform the decisions described in the algorithm. We also consider whether this type of additional information and analysis might be routinely required within appraisal or only conducted when OIR or AED appear to be particularly relevant, e.g., more sophisticated additional analysis might only be required if it is established that additional evidence is required and further research might in principle be worthwhile.

Section 4.1 provides a brief, non-technical summary of the type of information and range of additional analysis which will be conducted for each type of assessment. This provides a broad analysis plan for the case studies which will explore the utility and feasibility of different ways to inform the key assessment within the existing NICE appraisal process. Section 4.2, discusses the criteria which will be used to identify suitable case studies, ensuring that the full range of analysis is feasible within the time and resource constraints of this research, while exploring situations where OIR or AED are particularly relevant and challenging to NICE.

4.1 Methods of assessment

4.1.1 Expected cost effectiveness

Methods to estimate expected cost-effectiveness are well established within the NICE appraisal process and are extensively described in the Guide to Methods of Technology Appraisal. Commonly, expected cost-effectiveness is summarised and presented as incremental cost-effectiveness ratios. Equivalently, but more usefully in this context, cost-effectiveness can be expressed in terms of expected net health benefit, which can be presented per patient, per time period and for a population. All the information required to express expected cost-effectiveness in these ways should already be available in existing appraisals or can be extracted from the existing economic analysis.

i) Per patient net benefits

The expected per patient net health benefit for each intervention (i) under consideration is simply the difference between the expected health (usually expressed as QALYs) with the intervention (h_i) and health forgone elsewhere are a consequence of the costs of the intervention (c_i), which requires an estimate of the cost-effectiveness threshold (k). Therefore, the per patient expect net benefit for each intervention ($NB_i = h_i - c_i / k$) can be expressed using the same information required to present the more familiar ICERs. The intervention which is expected to be cost-effective is the one with the highest expected net benefit. This is entirely equivalent to drawing conclusions about cost-effectiveness based on ICERS but has many advantages once an assessment of uncertainty and it consequences is required. It is also needed when considering the impact of investment costs and is especially important when decisions require a trade-off to be made between benefits to current or future patients.

ii) Per period net benefits

The expected net health benefit for a patient or cohort of patients can be expressed each period over the time horizon of the analysis used to estimate cost-effectiveness. This can already be made available from the type of economic analysis used in appraisal but is not commonly reported. However, understanding how net benefit changes over time may be an important aspect of investment costs when other events (price changes or research reporting) means that guidance might be revised (see 4.1.2). Estimates of the population expected net health benefit in each period is also required when assessing the impact of investment and reversal costs, the longer term benefits of research, and any trade-off between net benefits of access to a technology for current patients or benefits to future patient populations from research findings. This requires information about prevalence and incidence of the target population as well as some judgement about whether incidence is likely to change over time.

iii) Population net benefits

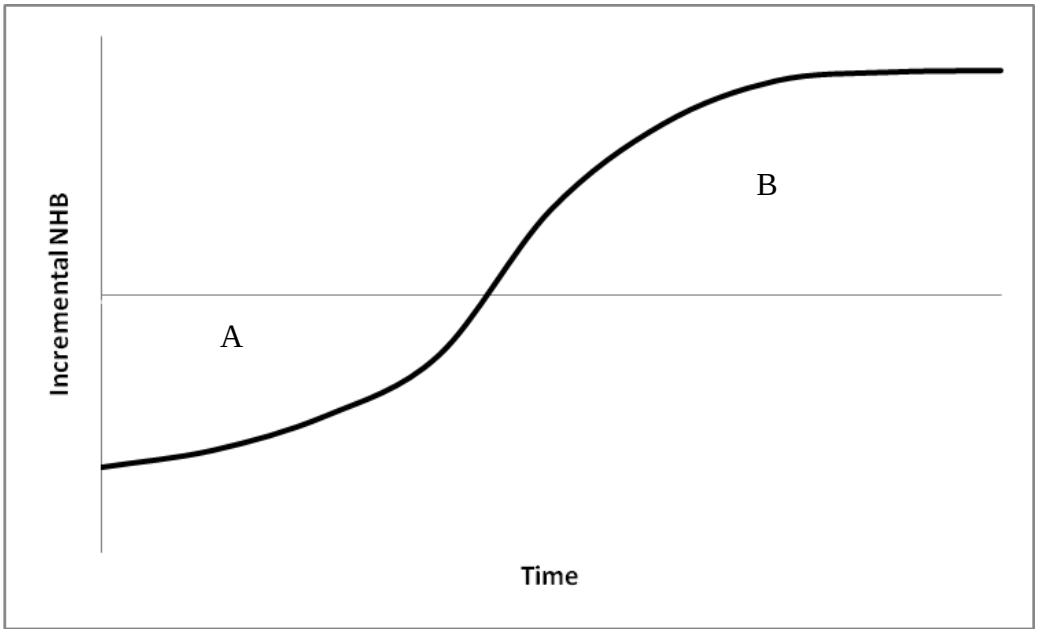
The total population net health benefit is the sum over population net benefits each period (appropriately discounted). Some judgement is required about the time horizon over which the estimated net benefits are likely to accrue (i.e., where the technology likely to be part of clinical practice or at least considered a relevant comparator), especially when considering investment and reversal costs. A range of time horizons can be based on: i) the review date for guidance; ii) historical evidence of the obsolescence of health technologies and iii) unbounded. Each has different arguments to commend it depending on the type of technology and context, so the impact of different time horizons will be explored. Similarly a range of time horizons for the benefits from research will also be explored. However, these two time horizons are not necessarily the same since evidence about a technology can continue to be valuable after the technology itself is obsolete.

4.1.2 Investment and reversal costs

Investment costs are those which once committed cannot be recovered if guidance is revised at a later date. Investment costs are generally considered to be those with a long life expectancy such as new facilities or equipment necessary to use the new technology. These are commonly annuitised and included in the expected costs of the technology. However, the impact of their irrecoverable nature is not often explored. Less often included in appraisal are the costs of implementation efforts, which might include development and dissemination of new guidelines, practitioner training and the opportunity costs of any delay in fully implementing new guidance.

However, irrecoverable costs also include situations where initially negative net heath benefits of a new technology are offset by later positive net benefits. This is illustrated in Figure 4.1 where the initial losses (area A) are more than offset by the later gains (area B) - the technology is expected to be cost effective overall. However, if guidance is revised (e.g., due to research revealing that the technology is not as effective as expected) then initial losses will have been incurred but they will not be compensated by later gains. Overall losses will tend to be greater if guidance is more likely to change and in the more immediate future.

Figure 4.1 Irrecoverable costs



There are many circumstances where initial losses are only offset by later gains. For example, when practitioners must learn how to best use a new technology by experience the initial learning costs imposed on patients (who will not get the maximum benefit from the new technology) are compensated by later gains once the technology can be used to its full potential. Similarly, when it is not possible for practitioners to identify those patients likely to respond to treatment without first treating them, the initial

population net health benefit will be lower as non-responders impose a cost without any additional benefit. As non-responders are identified and treatment is withdrawn overall treatment costs fall and net benefit increases. In many circumstances the initial per patient costs of a technology can be very high (e.g., oncology drug costs) and far in excess of the immediate health benefits in the initial period of treatment. However, these losses tend to be offset by future health benefits and sometimes reductions in future NHS costs. Therefore, Figure 4.1 represents a common pattern for technologies with mortality affects, which (although not commonly reported or its implications explored) can be extracted from existing economic analysis.

The range of possible investment and reversal costs which might be present will be explored in each of the case studies. Some types of irrecoverable costs (e.g., timing of costs and benefits and equipment costs) should be available or can be generated by the existing economic analysis. Others will require searching for additional information. Whether any investment and reversal costs are significant will depend on their scale relative to the expected net benefits offered by the technology. However, it will also depend on how likely guidance is to be revised in the near future. Therefore, we will present thresholds representing how likely and quickly guidance would need to be revised for the investment costs have a significant impact on decisions.

4.1.3 Need for evidence

Whether additional evidence is needed and what impact it might have on improving net health benefits depends on how uncertain a decision to approve or reject a technology might be and what the consequences of making the 'wrong' decision, e.g., there will be a chance that the resources committed by approval of a technology maybe wasted if the expected net health benefits are not realised. Equally, rejecting a new technology will risk failing to provide access to a valuable intervention if the net health benefits prove to be greater than expected. Therefore, evidence is needed and is valuable because by reducing uncertainty (the probability of decision error) better decisions which improve net health benefit are possible.

Although explicit and robust (but sometimes computationally demanding) analytic methods to estimate the value of evidence have been increasingly applied in the evaluation of health technologies, including within the NICE appraisal process, there are a range of ways in which this assessment might be informed. We intend to apply a range of approaches within each case study to explore how they perform in directly informing this assessment and what, if any, additional demands would be placed on NICE appraisal. Possible approaches to assessment include those based on i) simple criteria to identify those cases where additional evidence is very unlikely to be needed and further analysis is unnecessary ii) one way sensitivity and scenario analysis; iii) the results of probabilistic sensitivity analysis (PSA) alone; iv) value of information analysis based only on PSA; and v) value of information analysis which

also incorporates uncertainty between alternative scenarios representing alternative but plausible assumption (i.e., structural uncertainty). Although each approach represents a different level of sophistication and exploitation of available evidence, two elements are common: i) some assessment of how uncertain a guidance decision is likely to be and ii) some assessment of the impact on net health benefit if a decision based on current evidence turns out to be 'wrong'. We intend to highlight the strengths and weaknesses of each approach.

In many cases the analysis will turn on a judgement about how plausible particular scenarios are likely to be. This is essentially a judgment the Appraisal Committee comes to when formulating guidance. It is unlikely that we will be able to fully reflect this often implicit judgement based on Assessment Reports and Guidance documents. Instead we will explore a number of 'what if' scenarios in each case, e.g., 'evidence would be needed and OIR would be appropriate if scenario A or B were considered equally plausible'.

4.1.4 Is research possible with or without approval?

A key question is whether the research that is needed can be conducted while the technology is approved for widespread NHS use. The type of research that might be needed depends critically on the source of uncertainty and type of additional evidence that would be of most value. For example, if the key uncertainty is the size of the relative effect of the technology compared to a comparator then additional evidence is likely to require a randomised clinical trial (RCT) if selection bias is to be avoided. Since an RCT is unlikely to be regarded as ethical and patients are unlikely to agree to participate once approval is granted, the research needed is unlikely to be possible with approval. Other sources of uncertainty, such as underlying risks or quality of life associated with certain events or states might only require observational studies which could be conducted while the technology is approved.

Therefore, if additional evidence is considered to be potentially worthwhile, some assessment of the most important sources of uncertainty is required. Although value of information analysis provides explicit and robust methods to establish the value of additional evidence about particular groups of parameters representing different sources of uncertainty, such analysis can be computationally expensive in some circumstances. Therefore, we intend to apply a range of approaches within each case study to explore how they perform. These will be based on: i) thresholds and one way sensitivity analysis; ii) alternative scenarios representing different assumptions used to extrapolate from current evidence; iii) measures of association and contribution to variance which can be gathered directly from PSA; iv) analysis of the value of information associated with groups of parameters based on the PSA and, where possible, supplemented with an analysis of structural uncertainty. We anticipate some computational challenges so approximations to make more demanding analysis tractable will; be explored.

Identifying which sources of uncertainty are more important, and which parameters need to be more precisely estimated, is not sufficient. Some further consideration needs to be given to the types of research design that could reasonably inform them, including the use of selection models as well as more traditional epidemiological designs. We intend to highlight the judgements that are required in deciding what type of studies are needed and whether particular designs might be regarded as ethical with or without approval.

4.1.5 Other sources of uncertainty

An assessment of whether there are other sources of uncertainty which will only resolve over time is important because they will influence the future benefits of research that might be undertaken as part of OIR or AED and, if investment or reversal costs are also present, might make OIR more appropriate than AED even if the research could be conducted while the technology is approved (see 2.1.2). Key changes that commonly occur over time that influence both the cost-effectiveness of the technology and the future relevance and value of evidence generated by research include: changes in prices; the entry of new technologies and other research reporting.

i) Changes in prices

Future changes in the price of the technology under appraisal or changes in the price of comparator technologies are likely to occur and may be important, especially if they are imminent. A key assessment is when particular prices are likely to change and by how much. Most significant price changes occur on patient expiry and the entry of generics. How much cheaper future generic version of current brands are likely to be is uncertain but can be informed by historical evidence. Therefore, we will seek information on the likely time of patent expiry and use evidence of past price changes to assess their impact on cost-effectiveness and the future benefits of research.

ii) Entry of new technologies

The future launch of new technologies which may make the technology being appraised obsolete might mean that evidence about its performance from research commissioned as part of OIR or AED will only be relevant and valuable in the short run. Although future innovation in particular clinical areas maybe difficult to predict in the longer run, some developments can be anticipated. For example, we will seek information about potential competitor technologies identified in Topic Selection and/or scheduled for future appraisal. We will also seek information about applications for licensing in the same area (whether new technologies or extensions of the licensed indication for existing ones). We will also consider whether a shorter time horizon for the benefits of research should be used to reflect areas of historically rapid development and early obsolescence.

iii) Other research reporting

There are a number of reasons why presenting as complete a picture as possible of all research, whether underway, commissioned or planned, is important: i) it may be that the key evidence needed is likely to be provided by other research so OIR or AED is unnecessary; ii) guidance might change once other research reports, making evidence currently regarded as valuable, unnecessary, or incurring irrecoverable costs should initial guidance to approval or AED be revised; iii) early approval might undermine recruitment into valuable ongoing research or undermine the case to commission planned research that would resolve some of the key uncertainties. Therefore, in each case we will seek information from available trial registers and elsewhere to provide as complete a picture as possible of relevant research which is underway or is likely to be undertaken.

4.1.6 Re-assessing the benefits and costs of research

Whether or not research is possible with approval, some further assessment of its long term benefits is required. This will include assessment of: i) the likelihood that the type of research needed will be commissioned by research funders or conducted by manufacturers; ii) how long until such research will recruit and report; iii) how much of the uncertainty might be resolved by the type of research which is likely to be undertaken and iv) what impact the other sources of uncertainty (described in 4.1.5) will have on the future benefits of the research. With the possible exception of iv) these assessments are primarily judgements for which those directly involved in research design, prioritisation and commissioning might be best placed to make.

However, aspects of each case study will help to inform some of these assessments. For example , there may be a lack of incentive for manufacturers to conduct research and limited recourse if promised research is not undertaken or fails to properly report. In these circumstances it is unlikely that research will be successfully conducted unless publically funded. The benefits of research will (other things equal) tend to be reduced if it takes longer to report. The time until research is likely to report is uncertain but determined by a number of factors including rates of recruitment in the relevant clinical area and population, and the sample size and the length of follow-up required to address key uncertainties. In each case we intend to highlight the type of judgments needed (where possible, informed by evidence) and illustrate the implications of different but reasonable views that could be taken. Where ever possible this will include explicit and quantitative analysis of the impact on the expected benefits of research. The longer term expected benefits of research must be compared to the likely cost of conducting it.

However, even if the benefits are judged to exceed the costs the research might not necessarily be a priority and be commissioned since research funding is also budget constrained and other more valuable research might be displaced. Again this requires judgments which might best be made by

those involved in research prioritisation and commissioning. Therefore we will explore a wide range of reasonable views that could be taken.

4.1.7 Assessing the benefits and costs of early approval

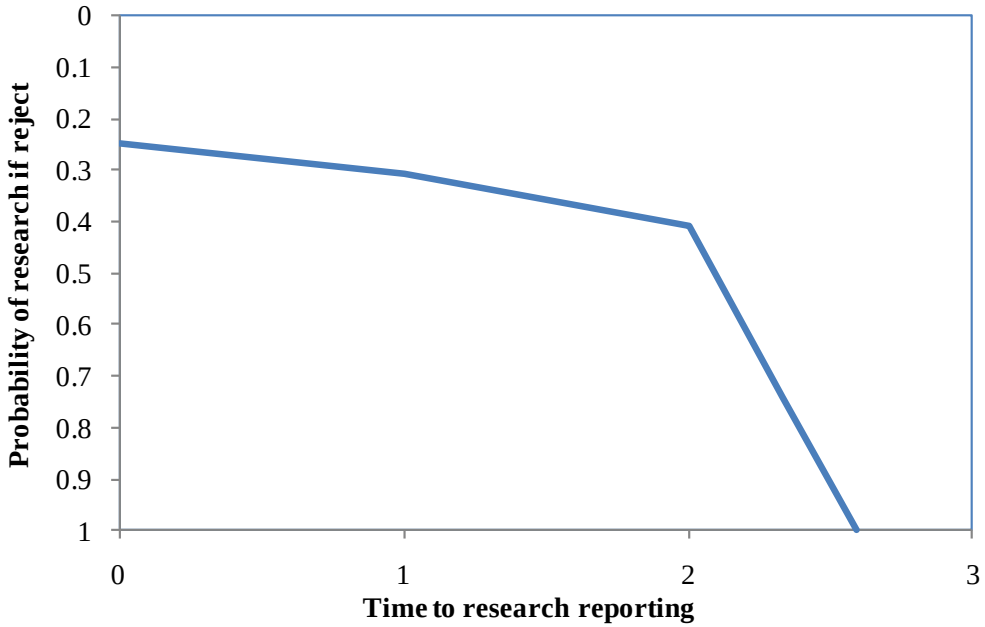
The assessment of whether technologies which on balance are expected to be cost-effective, should be approved depends on whether or not research will be possible once it is in widespread NHS use. If research is possible then the choice between approval or AED is relatively straight forward because it can be based on assessments that have already been made. No additional assessments are needed because AED would be appropriate if further research is worthwhile (see 4.1.6) but if not the technology should be approved (see Figure 2.1).

However, if the type of research needed is not possible with approval then some additional assessment is needed because judging that the benefits of research exceed its costs is not sufficient for OIR guidance. An assessment of whether the additional net benefits which are expected to accrue to future patients following research findings exceed the net benefits of early approval to current patients is required.

Existing economic analysis, supplemented with the additional types of information and methods of analysis described above can provide quantitative estimates of both, i.e., the expected net health benefits of approving the technology for the current patient population and the potential net health benefits for future patients if research that is judged worthwhile is conducted. However, whether the net benefits expected to accrue to future patients exceed the net benefits of early approval for current patients will depend on whether the research is commissioned and how long it will take to report. If research is less likely to be undertaken the expected net benefits for future patients will be lower. If it takes longer to report a larger current patient population (those prevalent while the research is undertaken) will receive lower net health benefits and for longer.

Judgements are required about the probability that research will be commissioned and when it is likely to report. These critical assessments reflect the decisions of those responsible for research design, prioritisation and commissioning, so are not necessarily ones for which NICE and its advisory committees have particular expertise or can easily be informed by other sources of evidence. Therefore, we intend to present a form of threshold analysis by estimating a 'boundary' for approval based on analysis already described above. This is illustrated in Figure 4.2 where early approval of a technology expected to be cost-effective is only appropriate at points north west of the boundary (where research is unlikely to be commissioned if approval is withheld and even if undertaken will take some time to report). To the south east of the boundary, where research is likely to be commissioned if approval is withheld and will report quickly, OIR is increasingly more likely to be appropriate.

Figure 4.2 A boundary for approval



4.2 Selection of case studies

The objective of developing case studies is to: i) demonstrate how the key principles and assessment might inform the development of guidance and ii) establish whether existing methods of appraisal are sufficient, or whether (and when) addition information and analysis might be useful. Suitable case studies will be selected to ensure that the full range of analysis is feasible within the time and resource constraints of this research, while exploring situations where OIR or AED are particularly relevant and challenging to NICE.

4.2.1 Feasibility

The resource and time constrains make de novo or substantial re-analysis of original assessments impossible. Nor would it be necessary or informative, since one of the objectives is to explore what *additional* information and analysis might be required. For this reason potential case studies which meet the following feasibility criteria will be considered: i) the economic analysis was regarded as a suitable basis for developing guidance; ii) an analysis of uncertainty in expected cost-effectiveness (PSA) was conducted and iii) there will be ready access to electronic versions of the versions of models which informed guidance.

4.2.2 Relevant and challenging

i) Policy interest

The are three groups of potential case studies where the key principles and assessment described above might influenced guidance: i) where OIR or AED was included in the FAD; ii) was considered

(e.g., included in ACD) but not included in the FAD; and iii) was not obviously considered at any stage. As well as examples of AED for technologies expected to be cost-effective and OIR for those not, there are also a number of particularly interesting ways in which guidance might be influenced by these additional considerations. These include: i) OIR rather than approval when a technology is expected to be cost-effective; ii) OIR rather than AED when there are investment costs and the technology is expected to be cost-effective and iii) AED rather than OIR when the technology is not expected to be cost-effective.

ii) Characteristics of the technology

To fully explore the implications of these principles and assessment it will be useful to select case studies which reflect a range possible and interesting characteristics, i.e., examples which are and are not expected to be cost-effective; ii) with and without investment and reversal costs; iii) where other sources of uncertainty are and are not present and iv) where the research needed is and is not possible with approval. Four studies will not be able to demonstrate the full range of possible combinations of interesting characteristics or illustrate all potential impacts of interest. Nor will it necessarily be clear, before additional analysis is complete, whether a case study will provide a good example of a particular impact on guidance. Therefore, in selecting case studies there will be a need to balance feasibility and those characteristics of greatest interest. To this end interviews with the Chairs and Vice Chairs of the Technology Appraisal Committees are planned to discuss proposed case studies.

4.3 Questions for group discussion

4.3.1 Have all the key assessments that need to be made been included in the range of proposed analysis

4.3.1 Have the range of proposed ways to inform these assessment included all potentially useful information and additional analysis which could feasibly be conducted within the resource and time constraints of NICE appraisal

4.3.2 Are the criteria for case study selection appropriate and do they identify a suitable range of interesting and policy relevant examples.