THE UNIVERSITY *of York*

CHE
Centre For Health Economics

Uncertainty, Evidence and Irrecoverable Costs: Informing Approval, Pricing and Research Decisions for Health Technologies

CHE Research Paper 69

Uncertainty, evidence and irrecoverable costs: Informing approval, pricing and research decisions for health technologies

[1,2]Karl Claxton
[1]Stephen Palmer
[3]Louise Longworth
[1]Laura Bojke
[1]Susan Griffin
[1]Claire McKenna
[1]Marta Soares
[1]Eldon Spackman
[3]Jihee Youn

[1]Centre for Health Economics, University of York, UK
[2]Department of Economics and Related Studies, University of York, UK
[3]Health Economics Research Group, Brunel University, UK

October 2011

Background to series

CHE Discussion Papers (DPs) began publication in 1983 as a means of making current research material more widely available to health economists and other potential users. So as to speed up the dissemination process, papers were originally published by CHE and distributed by post to a worldwide readership.

The CHE Research Paper series takes over that function and provides access to current research output via web-based publication, although hard copy will continue to be available (but subject to charge).

Acknowledgements

Disclaimer

Further copies

Centre for Health Economics
Alcuin College
University of York
York, UK
www.york.ac.uk/che

Contents

Abstract

The general issue of balancing the value of evidence about the performance of a technology and the value of access to a technology can be seen as central to a number of policy questions. Establishing the key principles of what assessments are needed, as well as how they should be made, will enable them to be addressed in an explicit and transparent manner. This report presents the key finding from MRC and NHIR funded research which aimed to: i) Establish the key principles of what assessments are needed to inform an only in research (OIR) or Approval with Research (AWR) recommendation. ii) Evaluate previous NICE guidance where OIR or AWR recommendations were made or considered. iii) Evaluate a range of alternative options to establish the criteria, additional information and/or analysis which could be made available to help the assessment needed to inform an OIR or AWR recommendation. iv) Provide a series of final recommendations, with the involvement of key stakeholders, establishing both the key principles and associated criteria that might guide OIR and AWR recommendations, identifying what, if any, additional information or analysis might be included in the Technology Appraisal process and how such recommendations might be more likely to be implemented through publicly funded and sponsored research. The key principles and the assessments and judgments required are discussed in Section 2. The sequence of assessment and judgment is represented as an algorithm, which can also be summarised as a simple set of explicit criteria or a seven point checklist of assessments. The application of the check list of assessment to a series of four case studies in Section 3 can inform considerations of whether such assessments can be made based on existing information and analysis in current NICE appraisal and in what circumstances could additional information and/or analysis be useful. In Section 4, some of the implications that this more explicit assessment of OIR and AWR might have for policy (e.g., NICE guidance and drug pricing), the process of appraisal (e.g., greater involvement of research commissioners) and methods of appraisal (e.g., should additional information, evidence and analysis be required) are drawn together.

Executive summary

The general issue of balancing the value of evidence about the performance of a technology and the value of access to a technology can be seen as central to a number of policy questions. This research was commissioned to inform when NICE should approve health technologies only in research (OIR) or with research (AWR). It has implications for policy (e.g., NICE guidance and drug pricing), the process of appraisal (e.g., greater involvement of research commissioners) and methods of appraisal (e.g., should additional information, evidence and analysis be required). However, establishing the key principles of what assessments are required and how they might be informed has much wider relevance beyond NICE and the UK NHS (e.g., informing the questions posed by coverage with evidence development initiatives).

Key principles and assessment needed
The key principles and assessments needed fall into four broad areas: i) expected cost-effectiveness and population net health effects (including benefits, harms and NHS costs); ii) the need for evidence and whether the type of research required can be conducted once a technology is approved for widespread use; iii) whether there are sources of uncertainty which cannot be resolved by research but only over time; and iv) whether there are significant (opportunity) costs which will be committed and cannot be recovered once the technology is approved.

Decisions (NICE 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 any benefits of early approval are greater than withholding approval until additional research is conducted or other sources of uncertainty are resolved. The sequence of assessment and judgments required is represented as an algorithm, which can be summarised as a simple seven point checklist.

Each sequence of assessment and decision, leads to different categories of guidance (e.g., Approve, AWR, OIR or Reject) for technologies with differing characteristics, indications and target populations. Different 'types' of apparently similar guidance can be identified. This illustrates how the same category of guidance might be arrived at in different ways, helping to identify the particular combination of considerations which might underpin decisions.

The principles suggest that restricting approval to OIR, or making it conditional on research through AWR, has wider application than is reflected in previous NICE guidance. For example, OIR may be appropriate when a technology is expected to be cost-effective. Even when research is possible with approval, OIR or even Reject maybe appropriate if there are significant irrecoverable costs. Therefore, the full range of categories of guidance 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 a necessary but not sufficient condition for approval and lack of cost-effectiveness is neither necessary nor sufficient for rejection.

Distinguishing principles (what assessment are needed) from methods of analysis (how they might be informed) allows potentially wide application of principles embodied in the algorithm and associated checklist, whist recognising that how the assessment might be made is likely to differ in different contexts.

Implications for value based pricing
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 a value based pricing scheme) will affect the key assessments, leading to different categories of guidance. The price at which a technology is just expected to be cost-effective is commonly regarded as its value based price. This describes the threshold price below which Approve rather than Reject would be appropriate if OIR or AWR are not available as policy options. However, if they are available there are often a number of relevant price thresholds. Once uncertainty, the need for evidence and the impact of irrecoverable costs are recognised, the threshold price that would lead to Approval rather than OIR will always be lower than a single value based price based on expected cost-effectiveness alone, i.e., disregarding uncertainty in costs and effects.

Even if price negotiation becomes possible alongside NICE appraisal, it will be important to retain OIR and AWR as available categories of guidance for two reasons: i) there is no guarantee that manufacturers will always agree to the lower price below which Approval rather than OIR or AWR would be appropriate; and ii) there may be many circumstances when no effective price reduction which would make Approval appropriate, e.g., Reject or OIR guidance may be appropriate even if the effective price of a technology was zero if there is substantial uncertainty about its effectiveness and/or potential for harms.

Incentives for evaluative research
An explicit assessment of OIR and AWR provides clear signals and an incentive to ensure the type of evidence, requiring research that cannot be conducted once approved for NHS use, is available and is sufficient at launch (e.g., relative effectiveness and subtle but important differences in side effect profiles). Therefore, a predictable OIR and AWR policy signals what type of evidence is likely to be most important at an early stage. It offers manufacturers a choice, to either: i) accept OIR Guidance at a higher price but restricted volume; ii) reduce the effective price to achieve Approval, or AWR where that is possible; or iii) conduct the evaluative research at an earlier stage so that additional evidence is available at launch.

How the NHS and manufacturers are likely to share the value of evidence might inform whether manufacturers should be expected to conduct the research specified in AWR or OIR guidance or contribute to the costs of publicly funded research which may ultimately benefit their product. The success of AWR when manufacturers are asked to conduct the research will depend on whether appropriate contractual arrangements can be established, i.e., those that can be monitored and enforced with credible penalties to ensure agreed research is conducted and in the way intended.   At present, NICE does not have a credible mechanism since removing approval of a technology simply because recommended research had not been conducted was not considered an ethical or credible threat.

The assessments required can be used to consider the value to the NHS of: i) being able to conduct research while a technology is approved (value of AWR); ii) making evidence that is needed by the NHS available at launch; and iii) being able to acquire evidence more quickly. This can inform investments in better data collection, registries or information systems that might make AWR possible. The value to the NHS of having access to the evidence needed at launch can inform a range of policies, such as early advice, public investment in transitional and evaluative research earlier in the development process or other incentives for research and development. Understanding the relationship between the time taken for research to report and the value of the evidence to future populations can help to inform: i) investments which might make research findings more quickly available; ii) the trade-off implicit in the choice of alternative research designs; and iii) those areas where if research is to be undertaken there must be confidence that it can report quickly.

The value of early evidence at launch and AWR can also be considered from the perspective of the manufacturer and inform whether they or the NHS might be expected to conduct the research needed. In principle, AWR and OIR research could be publicly funded rather than undertaken by manufacturers if the costs of research could be recovered directly from manufacturers or indirectly through other price discounts. Since the costs of public research are likely to be substantially lower than for manufacturers this might be mutually beneficial in many circumstances; providing appropriate support to innovation, while allowing wider access to the data generated and more transparency and accountability in the conduct of the research.

How should the assessment be undertaken?
The order of the assessments in the checklist relate to the sequence of decision nodes that fully describe the algorithm in Appendix A. This order of considerations means that all 7 assessments do not necessarily need to be made when an earlier judgement can lead directly to guidance. Therefore, one model for an efficient process of assessment would be to consider points 1-5 routinely. The Appraisal Committee would then be in a position to either rule out OIR or AWR and issue guidance in the usual way or indicate in the appraisal consultation document (ACD) that OIR or AWR was provisionally recommended subject to advice from a research advisory committee and subsequent analysis to support an assessment of points 6 and 7 of the checklist prior to FAD. This model would avoid unnecessary analysis and incorporate the judgments of the research community without necessarily delaying appraisal.

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 is required. Although the NICE appraisal process may be well suited to identifying the *need for evidence*, these other critical assessments (*the type of research and its priority*) are not necessarily ones for which NICE and its advisory committees, as currently constituted, have particular expertise. Informed judgements and better decisions might be possible through greater involvement of the research community. For example, a research advisory committee could be constituted which could consider provisional OIR or AWR guidance (at ACD), 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 publicly funded or undertaken by the manufacturer with appropriate contractual arrangements (which may require the involvement of DH at some stage).

What additional information and analysis might be required?
In the assessments, cost-effectiveness was presented as net health effects per patient treated and for the population of patients over time. This provides information in a way that is directly relevant to the assessments that need to be made using information generally already available during appraisal.

An early indication of potential importance of irrecoverable costs can be based on their scale relative to expected net health effects, the point at which any initial losses are expected to be compensated by later gains, whether treatment decisions are reversible and what opportunities to improve health might be forgone by a delay to initiating treatment.

The question of whether further research might be worthwhile requires some assessment of: i) how uncertain a decision based on expected cost-effectiveness might be; and ii) what the consequences, in terms of population NHE, are likely to be if an incorrect decision is made. This can be made in a series of steps each presenting what is already available within current methods of appraisal but in ways that can more directly inform the assessment required. How the consequences of uncertainty between as well as within scenarios can be presented and interpreted is also explored.

An assessment of the type of evidence needed requires judgements about: i) how important particular types of parameters (inputs to the economic model) are to estimates of cost and QALY; ii) what values these parameters would have to take to change a decision based on expected cost-effectiveness; iii) how likely is it that parameters might take such values and iv) what would be the consequences if they did, i.e., what might be gained in terms of population NHE if the uncertainty in the values of these parameters could be immediately resolved? The methods of analysis presented in Section 3 take these steps in turn; presenting what is available within current appraisal but in ways that more directly inform the assessment required. It is only when assessing the consequences of uncertainty associated with particular parameters that additional analysis is required to provide quantitative estimates.

The current appraisal process generally already provides the information and much of the analysis required to complete all the quantitative assessment reported in Section 3. However, the information required to assess whether other sources of uncertainty will resolve over time requires information that is not commonly sort as part of NICE appraisal. NICE many need to consider how access to this type of information can be provided or whether it should extract this type of information itself at an earlier stage of appraisal.

Any additional analysis to support a more explicit consideration of OIR and AWR would need to be included in manufacturers' submissions and be reviewed by the ERG. Although the additional analysis itself is limited (most is already required but sometimes presented in different forms), more explicit consideration of OIR and AWR and their link to price would make the critique of how uncertainty and its consequences has been characterised more important.   An assessment of whether the point estimate of cost–effectiveness is reasonable is inevitably a more limited task than also assessing whether the uncertainty surrounding that assessment is credible. Any additional burden on ERGs (and manufacturers) might be eased with clear guidance on the details of how analysis should be conducted and presented, what common assumptions are deemed reasonable and provision of additional information by the Institute as well as only considering points 6 and 7 on the checklist after ACD and following advice from a research advisory committee.

1.   Introduction and overview

NICE is increasingly making decisions about health technologies close to licence through the single technology assessment (STA) process. Inevitably these decisions are being made when the evidence base to support these technologies is least mature and when there may be substantial uncertainty surrounding their cost-effectiveness, including their effectiveness and potential for harms. In these circumstances further evidence may be particularly valuable as it would lead to better decisions which improve patient outcome and/or reduce resource costs. However, a decision to approve a technology will often have an impact on the prospects of acquiring further evidence to support its use. This is because the incentives on manufacturers to conduct research, once positive guidance has been issued, are limited. Also the clinical community is unlikely to regard further randomised controlled trials (RCTs) to be ethical once positive guidance provides access with a funding mandate. Therefore, the decision to approve a technology should account for both the potential benefits of access to a cost-effective technology, and the potential costs to future NHS patients in terms of the value of evidence that maybe forgone by early adoption. The general issue of balancing the value of evidence about the performance of a technology and the value of access to a technology can be seen as central to a number of policy questions. Establishing the key principles of what assessments are needed for 'only in research' (OIR) or 'approval with research' (AWR) recommendations, as well as how these assessments should be made, will enable them to be addressed in an explicit and transparent manner by the Institute.

The MRC and NIHR methodology programme recently funded the Universities of York and Brunel to undertake research to help inform when NICE should recommend the use of health technologies only in the context of an appropriately designed programme of evidence development. This report presents the key findings of this research, including i) the key principles of what assessments are needed to inform an OIR or AWR recommendation; ii) what additional information and/or analyses might help inform the assessments needed; and iii) the implications that this more explicit assessment of OIR and AWR might have for policy (e.g., NICE guidance and drug pricing), the process of appraisal (e.g., greater involvement of research commissioners) and methods of appraisal (e.g., should additional information, evidence and analysis be required) are drawn together.

This report is intended to be accessible to a wide audience; providing intuitive explanations of why certain assessments are important and illustrating how they might be informed with examples. The full HTA report[1]   provides substantial additional material, including: i) a critical review of policy, practice and literature in this area; ii) a review of NICE Technology Appraisal guidance; iii) a Technical Appendix which provides a more formal treatment of why, in principle, each type of assessment is important; iv) an Addendum which provides the full details of the analysis undertaken for each of the four case studies and v) a series of Technical Notes which deal with some conceptual and analytic details which are common to the case studies reported in the Addendum. At each stage this research was informed by a diverse and international Advisory Group and the feedback from participants at a series of two workshops involving a wide range of key stakeholders. Briefing documents and summaries of feedback from workshop participants, is available at:
http://www.york.ac.uk/che/research/teams/teehta/workshops/only-in-research-workshop/.

2.   What assessments are needed?

Since an important objective of the NHS is to improve health outcomes across the population it serves, 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 that a technology is more effective (the health benefits compensate for any potential harms) 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 [2]. Therefore, even if a technology is expected to be more effective, the health gained must be compared to the health expected to be forgone elsewhere as a consequence of additional NHS costs, i.e. a cost-effective technology will offer positive *net health effects [3-5]*. 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[6].

An assessment of expected cost-effectiveness or net health effects 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 together with some assessment of what health is likely to be forgone as a consequence (the cost-effectiveness threshold)[7]. Such assessments are inevitably uncertain and without sufficient and good quality evidence, subsequent decisions about the use of technologies will also be uncertain. There will be a chance that the resources committed by the approval of a new technology may be wasted if the expected positive net health effects are not realised. Equally, rejecting a new technology will risk failing to provide access to a valuable intervention if the net health effects 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[8-10]. 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[11]. In these circumstances there will be a trade-off between the net health effects 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[12].

Since publicly 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. 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 [13]. In addition, implementing a decision to approve a new technology is, in general, not a costless activity and may commit resources which cannot subsequently be recovered if guidance changes in the future [14-16]. For example, there may be costs associated with implementing guidance, training health care professionals, or other investment costs associated with equipment and facilities [17-18]. The irrecoverable nature of these costs can have a particular influence on a decision to approve a technology if new research is likely to report or other events may occur in the future (e.g. launch of new technologies or change in the prices of existing technologies).

The primary purpose of this Section is to provide a non-technical exposition, which identifies the key principles and assessments which are needed when considering both approval and research decisions. Section 2.1 outlines the key principles and 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 in the effective price of the technology or evidence. Section 2.3 highlights the social values and ethical principles associated with OIR and AWR.[1] Importantly, 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 alternatives are examined in Section 3 where they are more fully explored and evaluated through four case studies.

---

[1] Only in research (OIR) is defined as when a technology is only approved for NHS use (i.e., with the current funding directive) within the context of suitable research study. Approval with research (AWR) refers to approval while research is being conducted. In principle it is when approval is conditional on research being undertaken, which in the context of NICE guidance is where the research recommendation is part of the guidance in section 1. Whether such research will be undertaken and successfuly report is examined in later sections.

2.1   Key principles and assessments needed

The key principles and assessments fall into four broad areas:

i)   Expected cost-effectiveness and population net health effects (including benefits, harms and NHS costs).

ii)   The need for evidence and whether the type of research required can be conducted once a technology is approved for widespread use.

iii)   Whether there are sources of uncertainty which cannot be resolved by research but only over time.

iv)   Whether there are significant (opportunity) costs which will be committed and cannot be recovered once the technology is approved.

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 any benefits of early approval are greater than withholding approval until additional research is conducted or other sources of uncertainty are resolved.

This can be complex since these different considerations interact. For example, the effect of irrecoverable costs will depend on the need for additional research and will also influence whether research is worthwhile. The sequence of assessments, decisions and resulting guidance can be represented by a flow chart or algorithm.   Although such a representation is an inevitable simplification of the necessary trade-offs it helps to: i) identify how different guidance might be arrived at; ii) indicate the order in which assessments might be made; iii) identify how similar guidance might be arrived at through different combinations of considerations; and iv) identify 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 Appendix A, Parts I to III), representing the sequences of assessments and associated decisions, each leading to a particular category and type of guidance. However, the key decision points in the algorithm, reflecting the main assessments and judgments required during appraisal, can be represented as a simple 7 point check list (see Section 3.1).

Four broad categories of guidance are represented within the algorithm and include 'Approve', 'AWR', 'OIR' and 'Reject'. Each of the categories is further subdivided and numbered to indicate the different types of apparently similar guidance that could arise from different considerations. 'Delay' is not considered a particularly useful category since NICE always has the opportunity to revise its 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 (decisions) which are required in appraisal when formulating guidance.

2.1.1   Technologies without significant irrecoverable costs

Some element of cost which once committed by approval cannot be subsequently recovered is almost always present. However, the significance of these types of costs depends on their scale relative to expected population net health effects associated with the technology as well as the nature of subsequent events (see Sections 2.1.2 and 3.3.2) [19]. In this section we consider the relatively simple sequence of assessments and decisions which lead to guidance for those technologies that are not judged to have 'significant' irrecoverable costs associated with approval.

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 health effects (see Figure 2.1), i.e., where existing NICE appraisal currently ends. This is an assessment of expected cost-effectiveness, i.e., 'on average', based on the balance of the evidence and analysis currently available. It will include an assessment of effectiveness, potential for harms as well as NHS costs[6]. Any assessment may be uncertain with the scale and consequences of uncertainty assessed subsequently when considering the need for additional evidence. The sequence of assessments and decisions is illustrated in Figure 2.1. This demonstrates that an assessment of cost-effectiveness is only a first step and does not itself, inevitably lead to particular category of guidance. For example, a technology which might on
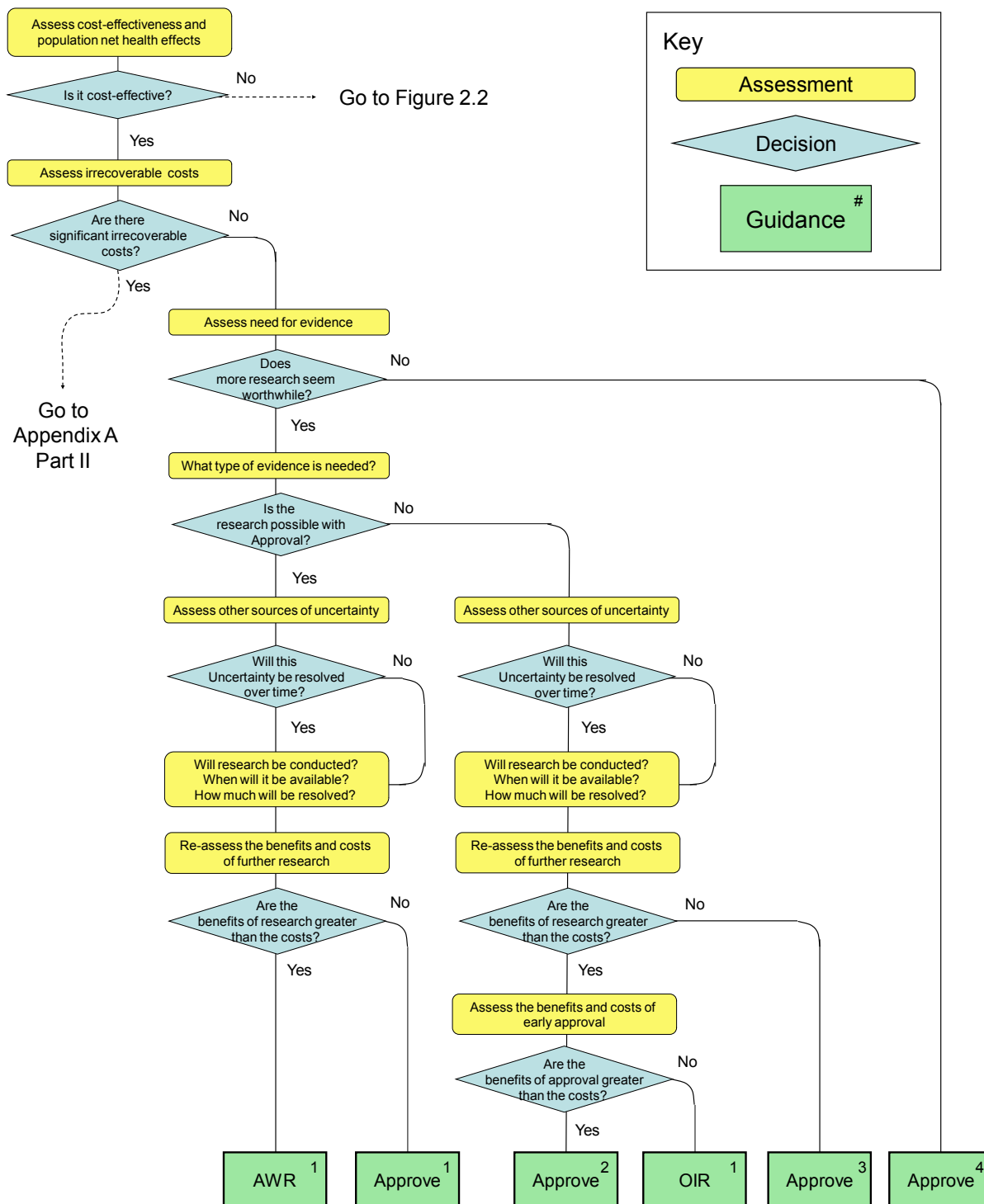
Figure 2.1   Technologies expected to be cost-effective

balance be expected to be cost-effective might nevertheless receive OIR guidance if the additional evidence that is needed cannot be acquired if the technology is approved.

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 health effects are positive. In these circumstances, further research may not even be potentially worthwhile (the opportunity costs of conducting this research exceed its potential value) so guidance

to approve could be issued on the basis of existing evidence and at the current price of the technology (e.g. Approve [4] in Figure 2.1). 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 is likely to be conducted if approval is granted [20].

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 successfully report; and iii) how much of the uncertainty might be resolved by the type of research which is likely to be undertaken[11]. 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[13]. These sources of uncertainty will influence the future benefits of research that could be undertaken as part of AWR. For example, 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 if 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 AWR) 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 AWR [1] and Approve [1] in Figure 2.1 dependent on 'Yes' and 'No' respectively.

Research is not possible with approval
The type of research needed may not be possible once a technology is approved for widespread NHS use, e.g. randomised clinical trials (RCTs) 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. If the benefits of research are judged to be less than the costs (i.e. research is not worthwhile anyway), the technology can be approved based on current evidence and prices (Approve [3] in Figure 2.1). However, judging that research is worthwhile at this point 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 then OIR guidance would be appropriate (OIR [1] in Figure 2.1). Alternatively, if the expected benefits of early access for current patients are judged to be greater than the opportunity costs for future patients, then approval would be appropriate (Approve [2] in Figure 2.1). 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 2.2.1).

ii)   Technologies not expected to be cost-effective
A technology which is not expected to be cost-effective will, on balance, impose negative population net health effects if it is approved. These negative net health effects can arise because the technology may not be effective, the potential for harm exceeds any benefits and/or the additional NHS costs are not justified by the magnitude of the expected health benefits offered. In these circumstances Approval can be ruled out, but which of the other categories of Guidance might be appropriate will depend on subsequent assessments and decisions (see Figure 2.2).

Need for evidence
Any assessment will be uncertain, so it remains possible that a technology which is not expected to be cost-effective on the balance of existing evidence might offer positive net health effects. 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 the technology can be rejected based on existing evidence and its current price (Reject [4] in Figure 2.2). Alternatively, if further research might be worthwhile then an additional assessment is

required 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 would be appropriate (OIR [2] in Figure 2.2). Alternatively, if the costs of research are likely to exceed the longer term expected benefits then the technology should be rejected at this point (Reject [1] in Figure 2.2).

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, greater understanding of learning and incremental improvements in the use of a technology, or identifying the particular types of patients that might benefit most[21]. 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] in Figure 2.2). 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 health effects 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 the technology should be rejected even though research, had it been possible without approval, would have been worthwhile (Reject [3] in Figure 2.2). Alternatively, if the net benefits of research more than offset the opportunity costs then AWR would be appropriate even though the technology is expected to be cost-ineffective (AWR [2] in Figure 2.2).

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

2.1.2 Technologies with significant irrecoverable costs

Irrecoverable 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 (per patient) 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 in the near or distant future [14, 16]. These types of cost are commonly thought of as capital expenditure on equipment or facilities which have a long life expectancy. They might also include the resources required to implement guidance, to train staff to use a new health technology or a period of 'learning' where health outcomes are lower. Although these costs are incurred 'up-front', they tend to be included in NICE assessments as if they are paid per patient treated over the life time of the equipment or facility. This common assumption will have no effect, so long as guidance is certain not to change during this period. However, if it is possible that initial approval might be withdrawn at some point, then, although future patients will no longer use the technology, these upfront costs cannot be recovered (see Section 2.3.2). Therefore, the possibility that Approve or AWR might be reconsidered after research reports, for example, and the impact this would have on expected costs needs to be considered, i.e. it may be better to withhold approval and avoid commitment of resources until the uncertainty is resolved.
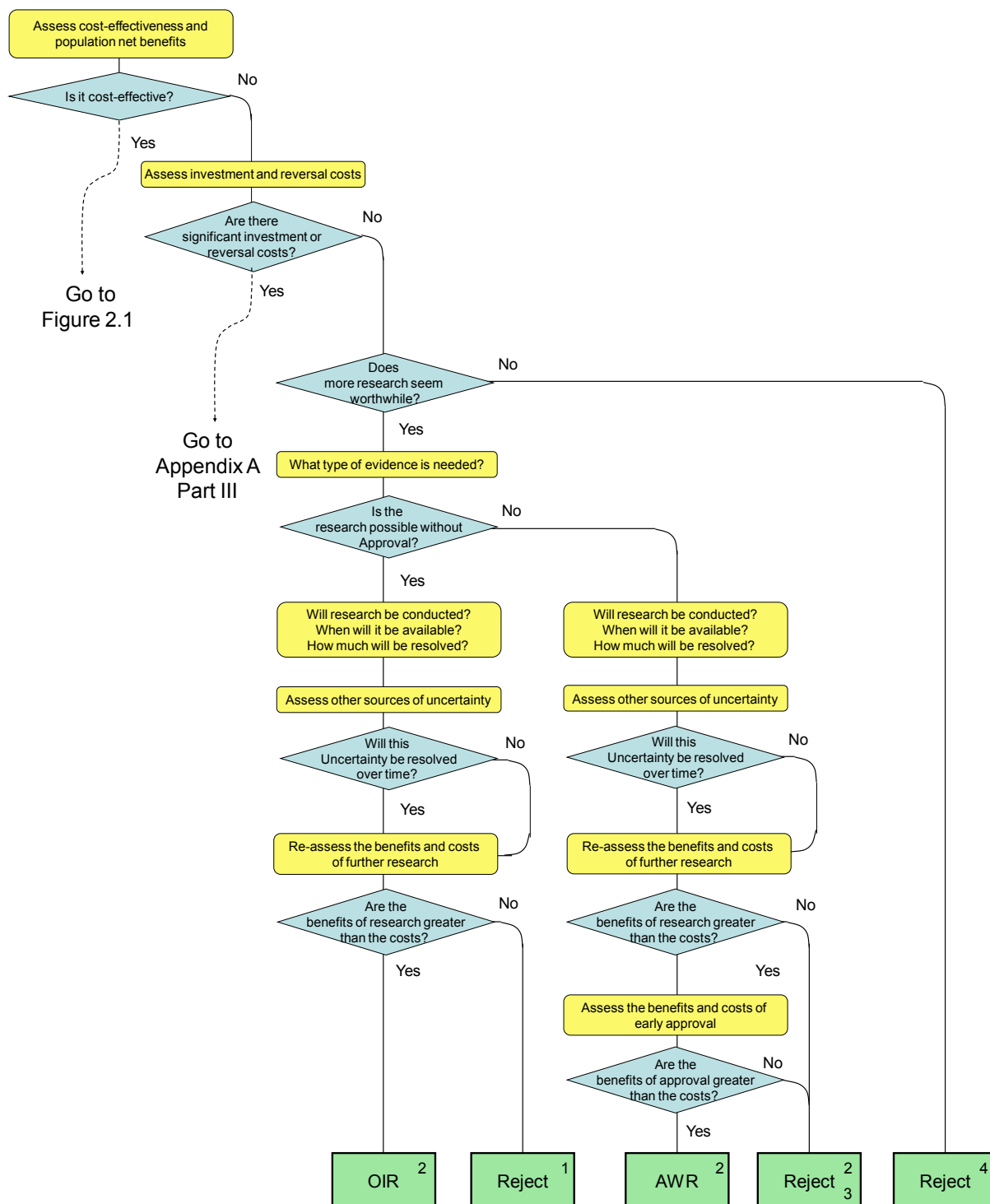
Figure 2.2   Technologies not expected to be cost-effective

However, irrecoverable costs may be much more common. Even in the absence of capital investment in equipment and facilities, most new technologies offer a 'risky investment profile' for each patient treated. Generally they impose initial per patient treatment costs which exceed the immediate health benefits (see Section 3.3.1). These irrecoverable treatment costs are only offset by cost savings and health benefits in the longer run, i.e. initially negative net health effects (losses) are only gradually compensated by later positive ones (gains). Therefore, a technology expected to be cost-effective may be expected to 'breakeven' (when accumulated 'gains' compensate earlier 'losses') after some considerable time. If guidance is likely to change it is possible that initial losses will not be compensated by later gains and the expected additional net health effects will not be realised[19]. This type of 'investment profile' becomes significant (has some influence on a decision to approval) if

the decision to treat a presenting patient can be delayed until uncertainty is resolved (e.g. research reports or other events occur) because the commitment of irrecoverable opportunity costs (negative net health effects) can be avoided (see Section 3.3.2). In these circumstances, OIR or reject avoids this commitment and preserves the option to approve the technology at a later date when its purchase by the NHS represents a 'less risky investment' [19].[2]

Although aspects of irrecoverable cost are almost always present, their potential significance also depends on their scale relative to expected population net health effects of the technology. Critically, their impact depends on the chance 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 more complex (see Parts II and III of the Algorithm in Appendix A), so here we focus on the key differences from section 2.1.1.

(i)   Technologies expected to be cost-effective
The presence of irrecoverable costs associated with a technology that is expected to be cost-effective will only influence guidance and be regarded as 'significant' if there are future events (research reporting or other sources of uncertainty resolving) which might change guidance. For example, if research is possible with approval and is expected to be worthwhile, AWR does not necessarily follow as previously (e.g., see AWR [1] in Figure 2.1) because the impact of irrecoverable cost must also be considered. Now OIR may be more appropriate than AWR (e.g., the choice between OIR [4] or AWR [4] in Part II), 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 in so far as they increase the chance that guidance will be revised (e.g., OIR [3] or AWR [3] in Part II).

If research is not possible with approval, but is expected to be worthwhile, then OIR will be appropriate if the opportunity costs of early approval are judged to exceed the benefits (e.g. OIR [6] rather than Approve [9] in Part II). These opportunity costs will now also include the impact of irrecoverable costs when guidance might be changed as well as the value of evidence that will be forgone by early approval. Therefore, irrecoverable costs will tend to make OIR rather than approval more likely, particularly when there are other sources of uncertainty which might resolve while the research is being conducted (e.g., OIR [5] rather than Approve [7] in Part II).

If research is not judged worthwhile, approval does not necessarily follow as previously (e.g., Approve[1,3,4] in Part I). Now the technology should only be approved if there are no other sources of uncertainty. If there are other sources of uncertainty, then an assessment of the benefits and costs of early approval is needed which takes account of irrecoverable costs and the risk that guidance might change 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 irrecoverable costs can be reconsidered once the other sources of uncertainty have resolved (e.g., Reject [5, 6] in Part II).

(ii)   Technologies not expected to be cost-effective
The presence of irrecoverable 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 AWR when research is not possible without approval (see AWR [5] in Appendix A, Part III). This is because a decision to reject, although it may be revised to approve, generally does not commit irrecoverable costs. Although there may be resources associated with making sure subsequent approval is properly implemented, these costs are properly considered as an irrecoverable 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 irrecoverable costs, tending to make AWR more likely to be appropriate.

---

[2] When considering OIR rather than Approve in these circumstances full account must be taken of the irreversible effects of withholding access to the technology for the population of patients that could have benefited from it, e.g., some may not survive to benefit from the results of the research or disease may have progressed so that the expected health benefits are lower.

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 the full algorithm in Appendix A.

The categories of guidance available to NICE have wider application than is reflected in previous guidance [1]. 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 irrecoverable costs. AWR can only be considered when research is possible with approval but Reject remains a possibility even for a cost-effective technology if there are irrecoverable costs. Therefore, the full range of categories of guidance (OIR and Reject as well as AWR and Approve) ought to be considered for technologies, which on the balance of existing evidence and current prices, are expected to be cost-effective.

Table 2.1a   Different types of guidance (technologies expected to be cost-effective)

Research	No significant irrecoverable costs					Significant irrecoverable costs				
	Not needed	Possible with approval		Not possible with approval		Not needed	Possible with approval		Not possible with approval	
		Benefits > costs	Benefits < costs	Benefits > costs	Benefits < costs		Benefits > costs	Benefits < costs	Benefits > costs	Benefits < costs
Approve (12)	4		1	2	3	11, 12		5,6	7, 9	8, 10
AWR (3)		1					3,4			
OIR (5)				1			3,4		5,6	
Reject (3)						7		5		6

Table 2.1b   Different types of guidance (technologies not expected to be cost-effective)

Research	No significant irrecoverable costs					Significant irrecoverable costs				
	Not needed	Possible without approval		Not possible without approval		Not needed	Possible without approval		Not possible without approval	
		Benefits > costs	Benefits < costs	Benefits > costs	Benefits < costs		Benefits > costs	Benefits < costs	Benefits > costs	Benefits < costs
Approve (0)										
AWR (2)				2					5	
OIR (2)		2					7			
Reject (8)	4		1	2	3	11		8	9	10

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 AWR may be 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. Importantly, which category of guidance will be appropriate only partly depends on an assessment of expected cost-effectiveness and hence this assessment 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 irrecoverable costs.

2.2   Changes in prices and evidence

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 AWR guidance.

2.2.1   Changes in effective prices

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 a future value based pricing scheme) will affect key assessments and decisions, leading to different 'paths' through the algorithm, consequently changing the category of guidance that would be appropriate [22-23]. For example, provisional OIR guidance for a technology, which is expected to be cost-effective, 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 (e.g., from OIR [1] to Approve [2] in Figure 2.1).[3] Similarly, AWR might be revised to Approve if the benefits of early approval now exceed the value of additional evidence (e.g., from AWR [1] to Approve [2] in Figure 2.1). [4]

Equally, provisional guidance to reject a technology which is not expected to be cost-effective, might be revised to OIR if the reduction in price was not sufficient to make it cost-effective, but made the costs associated with a reject decision more uncertain and hence made the value of research worthwhile (e.g., from Reject [1] to OIR [2], in Figure 2.2).[5] If the reduction in price was greater and was sufficient to make the technology cost-effective, then guidance might be revised to AWR, if research remains worthwhile and possible with approval (e.g., from Reject [1] or OIR [2] in Figure 2.2 to AWR [1] in Figure 3.1). Clearly, with an even greater reduction in price, it is possible that provisional guidance to reject could be altered to early approval (e.g., Approve [1] in Figure 2.1). Even if research is not possible with approval a sufficient reduction in price could also lead to early approval (e.g., from Reject [1] or OIR [2] in Figure 2.2 to Approve [2, 3, 4] in Figure 2.1).[6]

Therefore, consideration of the effect of price changes on OIR and AWR is needed when assessing the potential impact of patient access schemes and more direct price negotiation through value based pricing [5, 24-26].   It should be noted that, all other things equal, the presence of significant irrecoverable costs will require greater reductions in effective price to achieve the same revision to a more permissive category of guidance.

Threshold prices and VBP
The price at which the technology would just be expected to be cost-effective is commonly regarded as the value based price for the technology, i.e., the maximum price the NHS can afford to pay without imposing negative health effects [5,27].   This single price describes the threshold for Approve/Reject decisions and would be the relevant threshold price or VBP where: i) OIR or AWR guidance is not available to the decision maker or there is no uncertainty in cost-effectiveness; or ii) the research, if needed, can be conducted with approval; and iii) there are no irrecoverable costs. In all other circumstances there are a number of other threshold (value based) prices. The number and value of these thresholds depends on the characteristics of the technology (the path through the algorithm), however, the threshold prices for Approval will always be lower than the single Approve/Reject price based on expected cost-effectiveness.

---

[3] If the primary source of uncertainty is whether the technology is effective (i.e., whether there any health benefits compared to it's comparators and, if so, do they compensate for any potential harms) then reductions in price will have a more limited impact on uncertainty and the need for evidence compared to where there might be confidence of improved overall effectiveness but there is uncertainty about whether then magnitude of benefit is sufficient to justify the NHS costs. In all cases a reduction in price will increase the expected benefits of early approval for current patients so it will change the final assessment (where it is relevant) of whether the benefits of approval are greater than the opportunity costs (see point 7 in the check list in Section 3.1).
[4] See footnote 3.
[5] Any reduction in price will make a cost-ineffective technology less so (the net health effects, even if remaining negative will be greater, so a decision to reject will be more uncertain. However, there are limits to the effects of price reductions since even at a zero price the technology might not be cost-effective and/or further research still may be required, because there is no confidence that it is effective (harms may not be compensated by benefits) and/or it imposes other non acquisition costs on the NHS
[6] The price reduction required for these different types of approval will generally be greater if research is not possible with approval. However, these also differ. Approve [2] would require the greatest reduction in price and Approve [4] would require the lowest. However, any price reduction (price greater than zero) may not make approval appropriate in these circumstances.

For example, for a technology (without significant irrecoverable cost) where research could be conducted without approval but not with it, there are two threshold prices: i) the threshold which would move guidance from Reject to OIR; and ii) from OIR to Approve. The latter will always be lower than the price which would move the same technology from Reject to Approve if OIR was excluded from consideration. If a technology also imposes significant irrecoverable costs then there may be more threshold prices. For example, when research can be conducted with or without approval there are three thresholds: i) Reject to OIR; ii) OIR to AWR; and iii) AWR to Approve. Again the latter will be lower than the Approve/Reject threshold for the same technology if AWR was excluded from consideration. All other things equal the presence of irrecoverable costs will tend to reduce the threshold price for Approval.

Even in circumstances where price negotiation becomes possible alongside NICE appraisal, it will be important to retain the OIR and AWR as available categories of guidance for two reasons. Firstly, there is no guarantee that manufacturers will always agree to the lower price threshold which would lead to Approval rather than OIR or AWR. Secondly, and possibly more importantly there may be many circumstances when there is no effective price reduction which would make Approval appropriate.[7] For example, Reject or OIR Guidance may still be appropriate even if the effective price of a technology was zero if there is substantial uncertainty about its effectiveness and/or potential for harms.

2.2.2   Incentives for evaluative research

These threshold prices represent the maximum effective price at launch to achieve a particular category of guidance when the results of any subsequent research, which might be undertaken, are not yet known. This is different to the type of flexible pricing agreements, described in the current PPRS, where price is revised once the research reports and the results are known; increasing prices if the evidence suggests that benefits were originally underestimated, or reducing it if they were overestimated [28]. This means that manufacturers retain an incentive to conduct further evaluative research if they believe that there are additional benefits which could not be evidenced at launch. Publicly funded evaluative research, however, will still be required where these incentives are insufficient and especially in those cases where the original evidence is likely to have overestimated the benefits or underestimated the potential for harm. However, it should be noted that linking effective prices to the results of publicly funded research means that the NHS will only benefit (realise the value of evidence) if the results lead to a lower price or more restrictive Guidance because the technology is found not be cost-effective (thus avoiding the losses associated with negative net health effects). Manufactures will, however, be able to appropriate the value of evidence when it suggests that net health effects were originally underestimated through higher prices within a flexible or value based pricing scheme. Even under current arrangements this value can be appropriated when the technology is reappraised by NICE (e.g., any PAS could be withdrawn or less restrictive positive Guidance issued). Consideration of how the NHS and manufacturers are likely to share the value of evidence might inform whether manufacturers should be expected to conduct the research specified in AWR or OIR guidance, as long as incentive consistent contractual arrangements can be set in place, i.e., those that can be monitored and enforced with credible penalties to ensure any agreed research is conducted in the way intended. Alternatively, manufactures might be expected to make some contribution to the costs of publicly funded research which may ultimately benefit their product (see Section 3.6).

It is important that policy provides (or at least does not undermine) appropriate incentives for manufacturers to conduct the type of research needed to support NICE guidance at launch. The use of OIR and AWR Guidance, as described in the Algorithm provides clear signals and incentives. For example, the threshold price for Reject/OIR will be higher than for OIR/Approve but guidance restricted to OIR offers very limited NHS volumes to manufacturers. This provides a strong incentive to ensure the type of evidence, which would require research that cannot be conducted once approved for NHS use, is available and is sufficient at launch (e.g., relative effectiveness and subtle but important differences in side effect profiles). Therefore, a predictable OIR and AWR policy signals what type of evidence is likely to be most important at an early stage.

The use of OIR and AWR, as described in the Algorithm, offers manufacturers a choice, to either: i) accept OIR Guidance at a higher price; ii) reduce the effective price to achieve Approval, where that is

---

[7] See footnotes 1, 3 and 5.

possible; or iii) conduct the evaluative research at an earlier stage so that cost-effectiveness is not uncertain at launch. Other things being equal, those new technologies which are supported at NICE Appraisal by more, better quality and relevant evidence will be more likely to be approved (rather than OIR or AWR) and at higher prices than those that are not, because additional evidence is less likely to be needed. Therefore, greater consideration of OIR and AWR will tend to reward those manufacturers who have invested in good quality and relevant evidence, with earlier approval of their technology. In addition, the effect of price on OIR and AWR recommendations suggests that those technologies supported by better evidence will tend to get approval at higher effective prices, providing an incentive for manufacturers to invest in the type of evidence needed earlier in the development process.

2.2.3 Assessing the prospects of research

When considering OIR or AWR 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 may be 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, as currently constituted, have particular expertise, not least because they reflect the decisions of those responsible for research design, prioritisation and commissioning [29-30]. Without sufficient coordination between these communities there is a danger that OIR or AWR 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. Since publicly 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 a commercial interest. Therefore, a decision of whether OIR or AWR research should be undertaken by the manufacturer or through publicly funded research is one that NICE cannot properly take alone.

Although some judgement about how the research community might respond to OIR or AWR recommendations when NICE is formulating guidance is clearly possible, more informed judgements and better decisions might be possible through greater involvement of the research community. For example, a research advisory committee could be constituted which could consider provisional OIR or AWR 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 publicly funded or undertaken by the manufacturer with appropriate contractual arrangements. 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 AWR guidance are, in fact, research decisions which fall outside its remit, it would seem sensible to draw 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[8]

The question considered here is whether OIR and AWR recommendations are consistent with the values and principles that currently underpin standard NICE practices[6]. It is not in the remit of this report to evaluate those underpinning values and principles themselves. In particular, it is assumed here that the health budget is necessarily limited; also that, generally speaking – and whilst also taking into account issues of need and equity as discussed in the NICE values statements[31] – scarce health care resources ought to be broadly allocated so as to maximise health outcomes of the population as a whole; and hence that treatments that benefit one group of patients will be funded at an opportunity cost to other patients. Given these assumptions, the emphasis in this discussion is on new ethical challenges created by OIR and AWR decisions[32], as distinct from issues shared with standard NICE recommendations.

---

[8] We would like to thank Steve Holland and Tony Hope, as ethics advisors to the project, and Iain Chalmers, as a member of the Advisory Group, for their overall contribution, but especially to this section of the report. We would like to acknowledge that they have substantially revised the original version of this discussion of ethical principles presented in the briefing documents to the first workshop (http://www.york.ac.uk/che/research/teams/teehta/workshops/only-in-research-workshop/).

2.3.1   Issues common to OIR and AWR recommendations

Generally speaking, the benefit of attaching research conditions to NICE recommendations is an improved evidence base for resource allocation decisions in the future. The beneficiaries of the research are members of future populations who will profit from better informed allocation decisions. But achieving this benefit can impose significant opportunity costs on current patients. This is true for both some OIR decisions and some AWR decisions.

Two issues need to be carefully considered. The first is what is meant by present and future populations. The present population comprises people whose interests are directly affected by a NICE recommendation (for example, they receive an innovative treatment approved by NICE, or benefit from resources made available because NICE rejects an innovative treatment). Future populations comprise people whose interests are indirectly affected by decisions, in particular by the subsequent research results that improve the evidence base for future NICE judgements. It is important to note that sometimes individuals in the present population may benefit from the research condition because they will also be members of the future population (e.g., those with chronic conditions). This will not be true of all, so the issue of balancing the interests of some individuals in the present population against some individuals in future populations remains.

The second issue is, under what conditions the present population is disadvantaged by the research condition compared with the alternative recommendation that NICE might make. This will depend both on what the alternative recommendation would be, and on the level of current evidence about cost-effectiveness for the intervention. In judging whether it is right to make a research conditional recommendation there are four key issues to consider: i) what is the likely effect on the current population of a research conditional recommendation compared with whichever would be the alternative recommendation; ii) what is the likely benefit to the future population from the research; iii) what proportion of individuals in the present population is also likely to be in the future population; and iv) how should we weigh up the disadvantages to individuals in the present population in relation to the advantages to individuals in the future population (some of whom will be in the present population)? The first three of these considerations are essentially empirical issues and NICE will have to assess, possibly informed by the information and analysis presented in Section 3. The fourth issue is an ethical one. Do the interests of members of future populations count? If so, how are they to be weighed against the interests of members of the present population? Which set of interests should prevail?

One way of addressing this question is to consider how radical a departure from current NICE values, principles and practices it would be to accord weight to the interests of future populations. Arguably, doing so would not be much of a departure at all. For one thing, taking the interests of future populations into account is consistent with fundamental NICE assumptions about how to make allocation decisions. Specifically, NICE takes the identifiability of patients who benefit from an intervention to be irrelevant. So, the fact that beneficiaries of putting a research condition on approval are unidentifiable because they will only exist in the future does not seem to add anything to considerations NICE currently recognises and weighs. Similarly, some consideration of future populations is already implicit in standard NICE judgements. It would seem, then, that decisions that take into account the interests of future populations are consistent with NICE's values, principles and practices.

2.3.2   Issues specific to OIR recommendations

To tease out the ethical issues raised by OIR it will be useful to construct an illustrative case. Suppose NICE appraises a new treatment for which there is strong evidence of effectiveness – that is, the innovation is known to be clinically superior to existing treatments – but there is considerable uncertainty over its cost-effectiveness. Whether the new treatment would prove to be cost-effective depends not only on how expensive it is, but also on how much health benefit it would produce. NICE might consider an OIR judgement in these circumstances in order to establish more exactly the size of benefit which, in turn, is deemed necessary to establish cost-effectiveness. For example, NICE could approve the new treatment only in the context of an RCT comprising two trial arms, the innovative treatment arm and the standard treatment arm. Crucially, on this decision, patients outside the trial, and participants randomly allocated to the standard arm of the RCT, would be denied what is almost certainly the better treatment for their condition. This scenario creates a number of important ethical issues.

i)   Equipoise

A criterion established in research ethics for the legitimacy of carrying out an RCT is that there is substantial uncertainty as to which of the treatments being compared – that is, an innovative treatment and a standard treatment – is the more effective. This is sometimes known as the principle of equipoise. The principle is meant to capture the intuition that no one – patient or participant – should knowingly be offered less than the best treatment for their condition. OIR decisions may be made when there is such substantial uncertainty. In such cases, the principle is respected and ethical review of the relevant research poses only issues already considered as standard by researchers and research ethics committees.   However, in the situation envisaged here, NICE is considering recommending OIR when the intervention in question is clearly superior to alternatives, but the degree of its superiority remains uncertain, and so its cost-effectiveness is uncertain. Evidently, in this scenario researchers are not in equipoise about the relative effectiveness of the two interventions. The substantial uncertainty relates to whether the more effective, but more expensive, treatment produces sufficient extra benefit compared with alternatives for it to be recommended by NICE; but it does not relate to whether it is more effective. Is it permissible to flout the principle of equipoise concerning effectiveness and give an OIR decision in such circumstances?

An intuitive response is that patients are harmed by an OIR decision that denies them the best known treatment for their condition in the interests of research. Theoretical support for this intuition is provided by the well known principle of maleficence: above all, do no harm. But harm based objections to OIR are inconclusive for two reasons. First, the concept of harm is contested. The three main accounts define someone's being harmed by contrast with, respectively, (i) their state before the harm was perpetrated, (ii) the state they could have been in, and (iii) a minimum or baseline standard of wellbeing. When clinicians and patients are not sufficiently uncertain about the effectiveness of the treatments being compared, patients denied a better treatment by OIR are harmed according to the definition of harm based on (ii), because they are put in a worse state than was possible. But the patients are not harmed on the basis of the construal of harm based on (i) and (iii) because they will receive the standard NHS treatment (that is, the same treatment as patients who do not take part in the research, or as all NHS patients would have received had NICE rejected the treatment rather than approved it 'only in research'). So, harm based objections to OIR are as inconclusive as the current debate on the definition of harm. Second, harm based objections to OIR are question-begging, in the following way. Suppose there were a consensus on the nature of harm and, further, on the fact that OIR without equipoise harms some patients. To conclude that this makes OIR impermissible is to assume that the harm in question outweighs the benefits of research. But this is precisely what is in dispute, namely, the relative values of the benefits of early approval and of further research. Given NICE values, principles and practices, it is perfectly feasible to conclude that the harm perpetrated by OIR is justified by the benefits to future patients of a better evidence base for allocation decisions.

There are more fruitful lines of thought about OIR without equipoise. First, the principle of equipoise itself is under considerable strain from pressures that have nothing to do with OIR and AWR. For example, it has been argued that it is permissible to trial less than the best known treatment for HIV in developing countries unable to afford the most effective interventions[33]. It is generally accepted that, as a result of such pressures, the principle of equipoise needs to be refashioned by addressing the question; 'about what must researchers be uncertain?' The traditional requirement is that researchers must be in equipoise about the best known treatment; but, for bodies such as NICE the answer maybe that, the really salient uncertainty is not about effectiveness per se, but about the extent of effectiveness, and hence about the cost-effectiveness. Since, in the scenario envisaged here, the OIR recommendation is made precisely because of uncertainty about the extent of effectiveness, the refashioned principle of equipoise is respected. This argument is part of a larger research ethics question about what is required in terms of equipoise, so will not be addressed here. In the present context, there is a more important practical consideration to emphasise. An RCT required by OIR would have to be reviewed by a Research Ethics Committee (REC); RECs are used to requiring traditional equipoise (that is, substantial uncertainty about which is the better treatment). So, if RECs are to approve research of the kind considered here, where researchers are not in equipoise as traditionally understood, they will have to be informed of, and agree with, the rationale for conducting these distinctive studies.

Another way of looking at this issue is as follows. If NICE does not advise OIR then it must advise either to approve or not to approve. If it advises not approve, because, although the treatment is clinically the more effective it is judged not to be sufficiently effective to be cost-effective, then no one

receives this treatment on the NHS. Compared with that situation, an OIR decision would benefit some patients, and harm none. If NICE were to approve the treatment then all patients for whom the treatment is relevant would benefit, and, by comparison, an OIR decision would harm some patients (on definition (ii) above). But other patients might be unfairly harmed by the decision to approve because if, in fact, the treatment is not sufficiently effective to be cost-effective, the opportunity costs of providing the treatment outweigh the benefits. And because the relevant research is not being carried out, it will remain unknown that this is the case. In any event it would be problematic for a REC to refuse to sanction the type of research we are considering once NICE had made an OIR decision because that condemns both present and future populations to receive only the inferior treatment, whether or not the more effective treatment is sufficiently effective.

ii)   Coercion

Another research ethics principle is that competent patients have the right to consent to participate in, and withdraw from, a research project. This is akin to the competent patient's right to consent to treatment, both rights being underpinned by the principle of respect for individual autonomy. Conversely, it is impermissible to coerce competent patients to participate in research. In OIR, the patient can have the more effective intervention on the NHS only if she agrees to be a research participant. Does this coerce patients to participate in the trial?

Any research study involving care clinically superior to that available on the NHS provides an incentive to enrol in that study. Whether such an incentive constitutes coercion depends on whether the patient is being presented with a threat, as opposed to an offer to participate. Importantly, in the sort of OIR decisions under discussion, patients who do not receive the new treatment – that is, patients not enrolled on the trial, and participants not allocated to the new treatment arm – will receive standard NHS care. Arguably, then, the trial provides an offer, namely, the chance of receiving better than standard treatment, as opposed to any threat, and so does not constitute coercion. To clarify, suppose, by contrast, that patients will be refused access to normal NHS care unless they agree to participate in the research; this would present a threat as opposed to an offer to participate, and thereby constitute coercion. In sum, providing that standard research ethics requirements are met – principally, that prospective participants are properly informed about, and give valid consent to participate in, the trial; and that if they do not participate their access to standard NHS care will not be affected – the research required by OIR is not coercive and respects the principle of individual autonomy because patients retain the right to choose whether or not to accept the chance of better than standard treatment offered by the trial.

iii)   Equity

A further ethical worry is that OIR decisions result in inequity because participants in one arm of the trial would receive better treatment than both those in the other arm of, and those not enrolled on, the trial. Here it is important to distinguish two versions of this inequity charge. The first is that it is always wrong to allocate health resources in ways that will lead to an unequal distribution of health benefits. This version of the worry is bound to founder in this context because NICE's values and principles, which are taken as granted here, entail that limited health resources will and should be allocated so as to maximise benefit to the whole population, even at the expense of subgroups within it. So, the fact that OIR results in an unequal distribution of health benefits can be justified on NICE's assumptions if the research will provide sufficiently valuable evidence.

A second version of this inequity worry is more involved. Under OIR, some patients will receive the intervention (paid for on the NHS) and others will not, simply because the former happen to be in a position to participate in the research. Many of the factors that determine access to a formal treatment comparison – such as geographical location, socio-economic status, and patient characteristics – should not be considered relevant to whether patients have access to the treatments being studied. This is a significant consideration but one that can be overridden. Specifically, in deciding whether or not an OIR recommendation is ethically acceptable, a judgement would need to be made as to whether the benefits of such a recommendation outweigh the lack of equity (although, in fulfilling an OIR recommendation, it would always be important to minimise such inequities).

2.3.3   Issues specific to AWR recommendations

AWR raises a problem about consent, which is related to the discussion of coercion, above; AWR also raises a further distinctive issue centring on the likelihood of the research taking place.

i)   Consent

Two established principles of medical ethics are that competent patients have a right to confidentiality (including a right to decide whether to disclose their personal medical information) and a right to informed consent to participate in research (including a right to decline to participate in, or to decide to withdraw from, a study with impunity). It might be argued that some AWR decisions transgress these rights because the required research may involve collecting data on long-term outcomes and adverse events on patient registries (or some similar system of epidemiological data collection) without the explicit consent of patients. This argument is unsound. In many cases a patient's decision to have a treatment that was approved with research can be taken to imply consent to the relevant data collection. Furthermore, although it comprises personal medical information, the data collected will be anonymous, ameliorating concerns about breaching confidentiality.   Finally, the data collection required by an AWR decision is equivalent to current large-scale epidemiological research studies that are considered ethically permissible. So, concerning consent, and on grounds of consistency, AWR should be permitted as long as the research is conducted to the ethical standards normally required of data collection of this sort.

ii)   Incentivising research

A further ethical issue raised specifically by AWR recommendations involves the mechanisms used to give 'teeth' to the research requirement: how will NICE ensure that the relevant research is carried out? One option has a subtle ethical dimension. NICE might threaten that, if the research is not satisfactorily completed (for example, by the relevant manufacturing company), the intervention would cease to be made available on the NHS. Although this would provide an incentive for the manufacturer to carry out the research, it raises the following problem. At time $T(1)$ the AWR decision is made; that is, the intervention is funded by the NHS. Suppose that, at time $T(2)$ – the time when NICE reconsiders the decision to fund – the research has not been carried out (or has failed to provide any further relevant information). At time $T(2)$ NICE will be making a decision on exactly the same information and evidence as at $T(1)$. In this case, it would seem that NICE should make exactly the same decision, namely, to provide the intervention on the NHS. But if NICE decides to reject the intervention (on the grounds, for example, that the manufacturer had failed to carry out the relevant research) then patients could claim unfairness. The unfairness is that they were provided with the treatment on the NHS between $T(1)$ and $T(2)$ but not after $T(2)$ even though the evidence is exactly the same in both situations.

This is a serious problem for AWR decisions but it is not essentially an ethical matter. Rather, it is another point at which the ethical discussion segues into practical considerations familiar from other sections of this report. Specifically, various problems will result from the research condition put on an AWR decision not being met, and not just the ethical quandary outlined here. This situation should be avoided by doing everything reasonable to ensure that the relevant research will be conducted and reported (e.g. contractual arrangements with penalties that would be deemed ethical and therefore credible). Without credible mechanisms to ensure research conditions are met, AWR with research undertaken by manufacturers may not be feasible. If the research is not a sufficient priority for public funding then OIR may be the only credible alternative to Approve or Reject, i.e., the ethical issue described here has a direct influence on when OIR rather than AWR might be considered.

3. Informing the assessments

How the sequence of assessments required might be made is examined by take existing methods of NICE appraisal as an accepted starting point and focus instead on what additional information and analysis might feasibly be included in appraisal and how it might be interpreted to inform the judgements required. We also consider whether this type of additional information and analysis might be routinely required within appraisal or only conducted when OIR or AWR appear to be particularly relevant, e.g., more sophisticated additional analysis might only be required if it is established that further research might in principle be worthwhile.

3.1 A checklist of assessment

The sequence of assessment and judgement required can be summarised as a simple checklist that could be considered by the TAR team/ERG and Appraisal Committee as well as manufacturers during appraisal. There are two checklists: one for technologies expected to be cost-effective (Table 3.1a) and one for those not expected to be cost-effective based on the balance of existing evidence and current effective prices (Table 3.1b). The only difference between the checklists is at point 4, where, for technologies expected to be cost-effective, the judgment is whether the research is possible *with* approval whereas a judgment of whether research is possible *without* approval is required if the technology is not expected to be cost-effective.

Each of the seven points on the checklist relate to the sequence of decision nodes that fully describe the algorithm in Appendix A. Therefore, each sequence of Yes or No judgements defines a single pathway leading to a particular type and category of guidance (the type and category of guidance implied by each combination is described in Table B1, Appendix B). However, all 7 assessments do not necessarily need to be undertaken because sometimes earlier decisions will lead directly to guidance. For example a 'No' at point 3 always leads directly to either Approve or Reject and hence further assessment is unnecessary. Similarly, a 'No' at point 6 also leads directly to Approve or Reject if there are no significant irrecoverable costs associated with the technology (See Table B1, Appendix B).

Table 3.1a   Checklist for OIR and AWR (technologies expected to be cost-effective)

Point	Assessment	Judgement Yes	No
1	Is it cost-effective?	Yes	
2	Are there significant irrecoverable costs?		
3	Does more research seem worthwhile?		
4	Is the research possible with approval?		
5	Will other sources of uncertainty resolve over time?		
6	Are the benefits of research greater than the costs?		
7	Are the benefits of approval greater than the costs?		

Table 3.1b   Checklist for OIR and AWR (technologies not expected to be cost-effective)

Point	Assessment	Judgement Yes	No
1	Is it cost-effective?		No
2	Are there significant irrecoverable costs?		
3	Does more research seem worthwhile?		
4	Is the research possible without approval?		
5	Will other sources of uncertainty resolve over time?		
6	Are the benefits of research greater than the costs?		
7	Are the benefits of approval greater than the costs?		

3.2   Introduction to case studies

The objective of developing a series of case studies was to: i) demonstrate how the key principles and assessments might inform the development of guidance through application of the checklist and ii) establish whether existing methods of appraisal are sufficient, or whether (and when) additional information and analysis might be useful.

3.2.1   Selection of case studies

Case studies were selected to ensure that the full range of possible analysis was feasible within the time and resource constraints of this research project, while exploring situations where OIR or AWR might be particularly relevant and challenging. Therefore, de novo or substantial re-analysis of original assessments is not possible. 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 candidate case studies which met the following feasibility criteria were considered: i) the economic analysis was regarded as a suitable basis for developing guidance; ii) an analysis of uncertainty in expected cost-effectiveness (PSA as specified in the NICE reference case) was conducted and iii) ready access was available to the electronic versions of the models which informed guidance.

There are three groups of potential case studies where the key principles and assessment described above might have influenced guidance: i) where OIR or AWR was included in the FAD; ii) where OIR or AWR was considered during appraisal (e.g., included in ACD or section 6 of TA Guidance); and iii) where OIR or AWR was not obviously considered at any stage. As well as examples of AWR 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. For technologies expected to be cost-effective these include: i) OIR rather than Approve when research *is not possible* with approval; and ii) OIR or even Reject rather than AWR or Approve even if research *is possible* with approval because there are significant irrecoverable costs.

To fully explore the implications of these principles and assessments it is useful to select case studies which reflect the range of possible and interesting characteristics. For example, i) technologies which are and are not expected to be cost-effective; ii) with and without irrecoverable costs; iii) where other sources of uncertainty are and are not present; iv) where the research needed is and is not possible with approval; v) consideration of non pharmaceutical interventions and vi) those appraised under the MTA and STA process. Four studies will not be enough to demonstrate the full range of possible combinations of interesting characteristics or illustrate all of the potential impacts of interest.

Therefore, in selecting case studies there was a need to balance feasibility and coverage of those characteristics of greatest interest.

3.2.2 Background to the case studies

The following four case studies were selected. A range of additional information was sought, and further analysis conducted, to inform the sequence of assessment and judgements required when completing the OIR/AWR checklist in Tables 3.1a and 3.1b.

i)   Enhanced External Counterpulsation for chronic stable angina (EECP)

The NIHR-HTA programme identified EECP as an important topic and commissioned a short report to examine the clinical effectiveness and cost-effectiveness of EECP as an adjunct to standard therapy in patients with chronic stable angina. Although the topic was not ultimately considered by NICE it was commissioned in the same way and with the same resources as other assessment reports which inform NICE guidance. The assessment followed the NICE reference case and is consistent with the type of analysis which would have been required in an MTA appraisal. Like other MTA TARs it was published in full as a HTA monograph [19, 34].

EECP is a non-invasive procedure (adjunct to standard therapy) used to provide symptomatic relief from stable angina. The analysis compares EECP to standard therapy alone. RCT evidence suggests an improvement in HRQL with EECP at 12 months. To characterise the uncertainty associated with possible longer durations of treatment effect, formal elicitation of expert clinical judgement was undertaken. This provided an estimate of the probability, with uncertainty, of continuing to respond to treatment with EECP in subsequent years.

The possible pathways through the algorithm that EECP illustrate are reported in Figure B1 in Appendix B. In this case study the new technology is expected to be cost-effective but with potentially significant irrecoverable costs. These irrecoverable costs include both: i) long lived costs associated with the purchase of equipment; and ii) large initial per patient treatment costs, combined with a chronic condition where a decision not to treat a particular patient with EECP can be changed at a later date (decisions are not irreversible) when research reports or other events occur. Consequently these irrecoverable costs might influence the category of guidance, e.g., OIR rather than Approve. EECP also provides an opportunity to explore the impact of research design (length of follow-up) on guidance and to examine the potential role of elicitation rather than extreme scenarios to characterise uncertainty.

ii)   Clopidogrel for the management of patients with non-ST-segment elevation acute
coronary syndromes (CLOP)

The use of CLOP (for up to 12 months) in combination with low dose aspirin was recommended by NICE following an MTA appraisal for patients with non-ST-segment-elevation acute coronary syndrome (NSTE-ACS) presenting with a moderate to high risk of ischemic events (TA80 in 2004 and updated in 2010 in CG94).   In TA80 the Appraisal Committee considered 12 month or lifetime treatment with CLOP, but recommended research to inform optimal treatment duration. The original Assessment Report had included an analysis of shorter treatment durations (<12 months) and the NIHR-HTA programme subsequently commissioned additional re-analysis based on this original work to inform this research recommendation in 2009. This case study is based on the re-analysis of TA80 undertaken in 2009 which included standard therapy compared to 4 alternative treatment durations of clopidogrel of 1, 3, 6, and 12 months. Importantly, while the case study is based on the later re-analysis of TA80, the analysis considered here has been undertaken from the standpoint of the original TA80 appraisal asking, what assessments might have been made at that time when standard therapy was low dose aspirin?

The research recommendation was made in Section 6 of TA80, therefore, CLOP is not an example of AWR at FAD, but where AWR was considered during appraisal. The possible pathways through the algorithm that the CLOP case study illustrates are reported in Figure B2 in Appendix B, where the new technology is expected to be cost-effective and with no significant irrecoverable costs. The CLOP case study also illustrates a number of other important characteristics, including: (i) the impact that other sources of uncertainty (price change following patent expiry) can have on the value of further research; (ii) the interpretation of analyses where there are multiple alternatives; and (ii) the use of scenarios to represent alternative but credible assumptions.

iii)   Omalizumab for the treatment of severe persistent allergic asthma in children aged 6 to 11 years (OMAL)

The use of OMAL for the treatment of severe persistent allergic asthma in children aged 6 to 11 years was not recommended by NICE following an STA appraisal (TA201 in 2010).   The analysis compared OMAL as an add-on to standard care versus standard care alone. The primary analysis was based on a pre-specified severe asthma population within an international, multicentre, placebo-controlled RCT.   However, a high-risk subgroup within this population (recent hospitalisation for an asthma exacerbation) was also identified post-hoc.

Omalizumab was not found to be cost-effective in either the severe or severe/high risk populations. However, an RCT was recommended comparing OMAL to oral corticosteroids (OCS) in children to establish reduction in OCS use. This was made in Section 6 of TA201; therefore, OMAL is not an example of OIR at FAD, but where OIR was considered during appraisal. The possible pathways through the algorithm that OMAL illustrate are reported in Figure B3 in Appendix B, where the new technology is not expected to be cost-effective and with no significant irrecoverable costs. OMAL also illustrates assessment in small patient populations (rare disease) for the assessment and how sub group analysis can be considered.

iv)   Etanercept, infliximab and adalimumab for patients with active and progressive Psoriatic arthritis (PsA)

Following an MTA appraisal (TA199 in 2010), the use of biologic treatment with etanercept, infliximab and adalimumab was recommended by NICE for patients with active and progressive PsA and who have an inadequate response to standard treatment, including two conventional disease-modifying antirheumatic drugs (DMARDs). However, the guidance also recommended that treatment should start with the least expensive biologic, taking account of dose, route of administration and price. This guidance updated an earlier MTA appraisal in 2006 (TA104), which had recommended etanercept and restricted guidance on the use of infliximab to only those patients shown to be either intolerant or contraindicated to etanercept.[9]   The analysis in this case study is from the standpoint of TA199, using the updated model which included new evidence and adalimumab as an additional comparator. At this point NICE guidance recommended etanercept, so the first question posed in the checklist can be interpreted as, are the other technologies available (infliximab, adalimumab or palliative care) expected to be cost-effective compared to etanercept?

In Section 6 of TA199 the importance of data on long term outcomes and adverse events from patient registries was highlighted. Therefore, PsA is not an example of AWR at FAD, but where AWR was considered during appraisal. The possible pathways through the algorithm that PsA illustrate are reported in Figure B4 in Appendix B. In this case study the alternatives to etanercept are not expected to be cost-effective. However, etanercept as well as infliximab and adalimumab have potentially significant irrecoverable costs because of the high initial per patient treatment costs, combined with a chronic condition where treatment decisions are not irreversible. PsA, like EECP, also provides an opportunity to examine the potential role of elicitation in the appraisal process.

3.3   Is it cost-effective and what are the risks?

The judgements made at points 1 and 2 of the checklist are critical because, although neither leads directly to a particular category of guidance, they do determine the subsequent path that might be taken, sometimes avoiding further and potentially complex assessments. For example, the absence of significant irrecoverable costs means that only 4 out of the 12 possible pathways require all 7 assessments to be made (see Table B1, Appendix B).

---

[9] TA104 included an AWR recommendation in the ACD but this was removed in the FAD. The recommended research was to enter patients into the BSR register, on the grounds of the possibility of severe side effects and little information on the use of these agents beyond the duration of RCTs

3.3.1   Point 1 - Is it expected to be cost-effective?

The sequence of assessments starts with cost-effectiveness and the expected impact on population net health effects. i.e., at the following point in the algorithm:
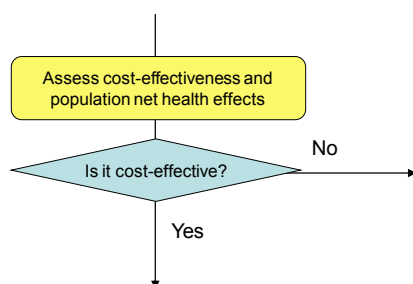
This requires an assessment of expected cost-effectiveness based on the balance of the evidence and analysis currently available.   Methods to estimate expected cost-effectiveness are well established within the NICE appraisal process and are extensively described in the Guide to Methods of Technology Appraisal[6]. [10]   Commonly, expected cost-effectiveness is summarised and presented using incremental cost-effectiveness ratios (ICERs). Equivalently, but more usefully in this context, cost-effectiveness can be expressed in terms of expected net health effects (NHE), which can be expressed per patient treated or for a population of patients. This is especially important when later assessments require a comparison of benefits to current or future patient populations and when assessing the significance of irrecoverable costs (see Section 3.3.2). All the information required to express expected cost-effectiveness in these ways is already available during appraisal.

i)   Cost-effectiveness at the patient level

Estimates of the expected NHS costs and QALYs for each patient treated over an appropriate time horizon - the 'patient time horizon' [11] - can be summarised as an ICER, which must be compared to a cost-effectiveness threshold to judge cost-effectiveness. Equivalently, this can be expressed as the per patient NHE of each intervention, i.e., the difference between any health gained and health forgone elsewhere. [12]

Technologies expected to be cost-effective

The results for EECP are summarised in Table 3.2a. There are only two alternatives (EECP and standard care, so only one ICER. EECP is just expected to be cost-effective at a threshold of £20,000 per QALY.[13]   Consequently the NHE of EECP are greater than standard care but the difference per patient treated (the incremental NHE) is small.

---

[10] In each case-study the estimates of expected costs and QALYs reported and used throughout are the mean costs and QALYs derived from probabilistic analysis using Monte Carlo simulation. The costs and QALYs from a deterministic analysis will be incorrect unless the model is multi-linear with independent parameters.

[11] This is the time horizon over which costs and benefits are likely to differ for an individual patient (commonly termed the model time horizon). In some circumstances (e.g., where there is a mortality effect) this will be the lifetime of the patient. Expected costs and QALYs each period are the expectations (means) from the results of probabilistic analysis. All future costs and QALY (per patient or population) are discounted at 3.5% throughout.

[12] The expected per patient net health effects for each intervention (i) is the difference between the expected health (QALYs) with the intervention (h_i) and the health likely to be 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 expected net health effects of each intervention ($NHE_i = h_i - c_i/k$) can be expressed using the same information required to present the more familiar ICERs. It can also be expressed in terms of the NHS resources required to generate the NHE_i ($k.h_i - c_i$) The intervention which is expected to be cost-effective is the one with the highest expected net health effects. This is entirely equivalent to drawing conclusions about cost-effectiveness based on ICERs but has many advantages once an assessment of uncertainty and its consequences is required. It is also needed when considering the impact of irrecoverable costs and is especially important when decisions require a trade-off to be made between benefits to current or future patients.

[13] All analysis has been conducted at the upper and lower bound for the range NICE has adopted for the threshold. However, unless otherwise stated results in the text relate to a threshold of £20,000 per QALY.

Table 3.2a   Expected cost-effectiveness of EECP per patient treated

				Cost-effectiveness threshold at:			
				£20,000 per QALY		£30,000 per QALY	
Treatment	Costs	QALYs	ICER	NHE, QALY (£)	Incr NHE, QALY (£)	NHE, QALY (£)	Incr NHE, QALY (£)
EECP	£4,744	7.6045	£19,391	7.3673 (147,346)	0.0074 (£149)	7.4464 (£223,391)	0.0865 (£2,595)
Std	-	7.3598	-	7.3598 (147,197)	-	7.3598 (£220,795)	-

It is also important to consider how NHE accumulate over time or the 'investment profile' per patient treated with EECP. Figure 3.1a illustrates the cumulative incremental NHE over the patient time horizon. The initial per patient costs of EECP are high and are far in excess of the immediate health benefits in the initial period of treatment. These negative NHE are gradually offset by positive NHE in later periods. In this case, it is only after 14 years that the initial losses are compensated by later gains. i.e., EECP is not expected to 'breakeven' until 14 years from initial treatment. It is only beyond 30 years that the modest incremental NHE reported in Table 3.2a are eventually achieved.[14]
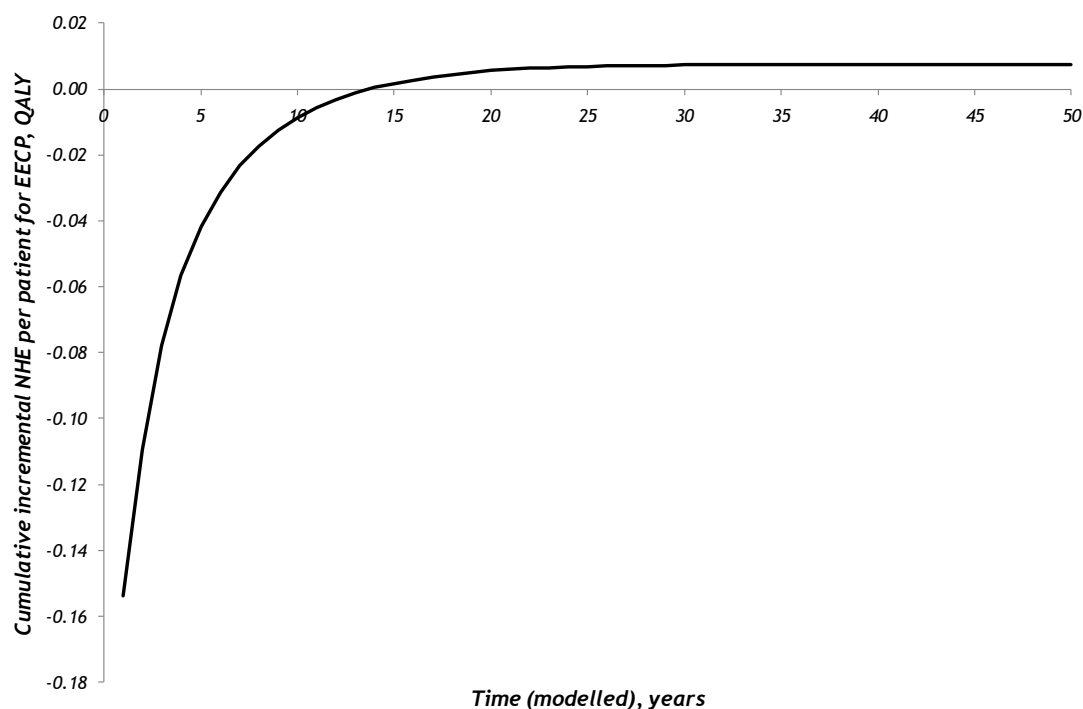
Figure 3.1a   Cumulative incremental NHE of EECP over the patient time horizon

Multiple alternatives
Similar analysis can be conducted when there are more than two alternatives.   For example, in CLOP four treatment durations as well as current NHS treatment (aspirin alone) were considered at the time of TA80. The results in Table 3.2b indicate that 12 month treatment with CLOP is expected to be cost-effective, although the difference in NHE between 12 and 6 months treatment duration is small.

---

[14] The time at which initially negative NHE are expected to be offset by cumulating positive NHE, or the 'breakeven' point, is only an indicator or proxy for the scale of irrecoverable opportunity costs, e.g., the scale of initial loses as well as the breakeven point also matters. The presence of even very large irrecoverable opportunity costs does not necessarily mean they are significant and will influence guidance. That will depend on whether treatment decisions are reversible and the impact of withholding treatment for patients who might receive it (see Section 2.2.2 and 3.3.2), as well as whether uncertainty is likely to resolve (research reporting or other sources of uncertainty resolving.

Table 3.2b   Expected cost-effectiveness of CLOP per patient treated

Treatment	Costs	QALYs	ICER (£/QALY)	£20,000 per QALY NHE, QALY (£)	£30,000 per QALY NHE, QALY (£)
clop12	£20,127	8.122	18,663	7.115 (142,307)	7.451 (223,525)
clop6	£19,860	8.107	10,477	7.114 (142,288)	7.445 (223,362)
clop3	£19,712	8.093	9,396	7.108 (142,154)	7.436 (223,087)
clop1	£19,598	8.081	4,961	7.101 (142,025)	7.428 (222,837)
NHS	£19,502	8.062	-	7.087 (141,734)	7.412 (222,353)

The 'investment profile' of CLOP, per patient treated, is illustrated in Figure 3.1b. The per patient costs of CLOP are in excess of the health benefits during the period of treatment. These negative NHE are eventually offset by positive NHE in later periods. In this case, it is only after 5 years that 12 months of treatment with CLOP 'breaks even' against current NHS care and it is not until 21 years that it is better than a shorter treatment duration of 6 months. Notice that shorter treatment durations with CLOP offer a much less 'risky profile', e.g., the 'breakeven' point for one month of treatment is 2 years against current NHS care.
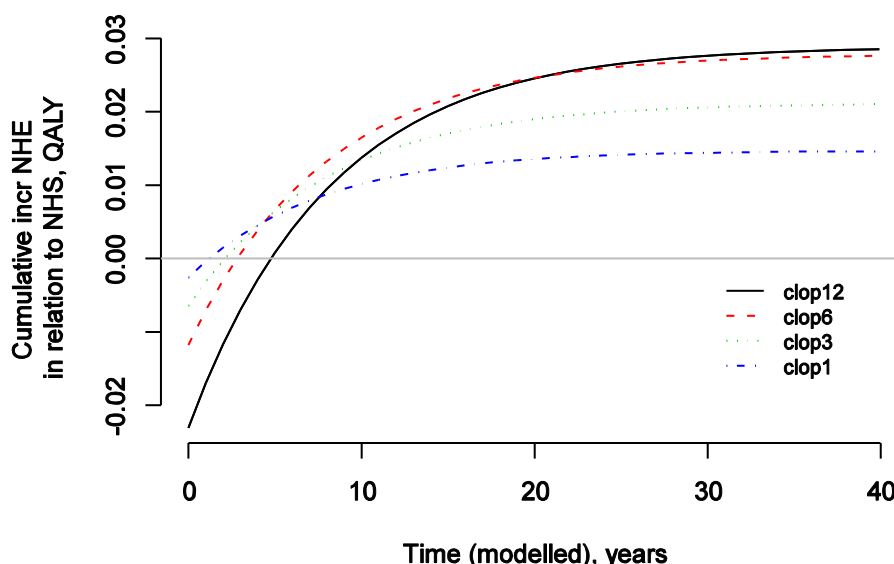
Figure 3.1b   Cumulative incremental NHE of CLOP over the patient time horizon

Technologies not expected to be cost-effective
The ICER for omalizumab in Table 3.2c is greater than the threshold so it is not expected to be cost-effective compared to standard care alone. Consequently the incremental NHE of OMAL is negative.

Table 3.2c   Expected cost-effectiveness of OMAL per patient treated

Treatment	Costs	QALY	ICER (£/QALY)	£20,000 per QALY NHE, QALY (£)	£20,000 per QALY Incr NHE, QALY (£)	£30,000 per QALY NHE, QALY (£)	£30,000 per QALY Incr NHE, QALY (£)
Omal + Std	£94,992	16.64	93,844	11.8861 (237,721)	-2.1908 (-43,815)	13.4693 (404,078)	-1.2627 (-37,882)
Std	£39,310	16.04	-	14.0768 (281,536)	-	14.7320 (441,960)	-

The per patient 'investment profile' for OMAL is illustrated in Figure 3.1c and shows that it is always expected to offer negative NHE compared to standard care over the entire patient time horizon, i.e., the high costs of treatment are never compensated by future health gains. In this example, the initial treatment costs with OMAL continue for 10 years (10 years is assumed to represent the duration a patient would continue to receive treatment with OMAL) with health effects predominately while on

treatment. Therefore, OMAL is not so much a 'risky purchase' but one that is simply not cost-effective at its current price.
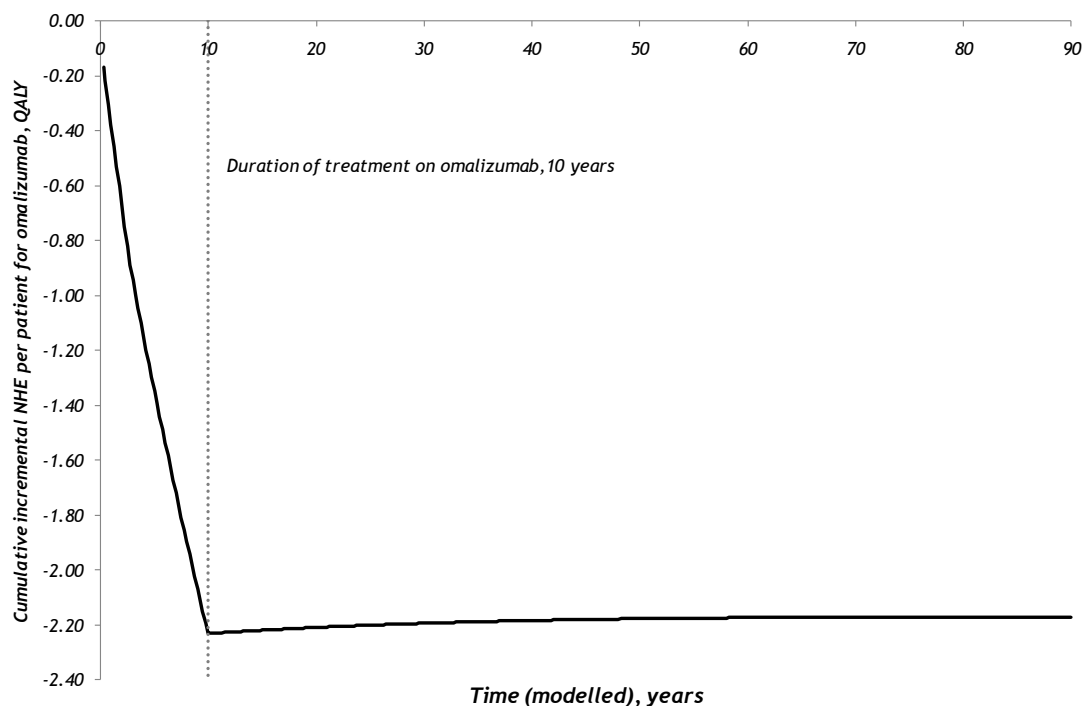
Figure 3.1c   Cumulative incremental NHE of OMAL over the patient time horizon

Multiple alternatives
PsA offers an example where the alternatives to the treatment already recommended by NICE (etanercept at the time of TA199) are not expected to be cost-effective, i.e., the results in Table 3.2d indicate that etanercept is expected to be cost-effective. Notice that although adalimumab is less effective than etanercept it is also cheaper. However, the resources savings it offers do not compensate for the reduction in health benefits.

Table 3.2d   Expected cost-effectiveness in PsA per patient treated

				Cost-effectiveness threshold at:	
				£20,000 per QALY	**£30,000 per QALY**
Treatment	**Costs**	**QALYs**	**ICER (£/QALY)**	**NHE, QALY (£)**	**NHE, QALY (£)**
1: Infliximab	£90,343	7.269	60,965	2.752 (5504)	4.258 (8516)
2: Etanercept	£78,150	7.069	17,733	3.161 (6322)	4.464 (8928)
3: Adalimumab	£72,972	6.777	14,622	3.129 (6258)	4.345 (8690)
4: Palliative care	£51,800	5.329	-	2.739 (5478)	3.602 (7204)

Consequently the 'investment profile' of the alternatives to etanercept, illustrated in Figure 3.1d, differs in appearance. However, all the biologic treatments for PsA have high initial costs which are only gradually compensated by later health benefits. All three ultimately offer positive NHE compared to palliative care but only 'breakeven' at 17, 17.5 and 34.5 years for adalimumab, etanercept and infliximab respectively. Adalimumab offers a slightly less risky profile than etanercept, so it is only at 21.25 years that etanercept is expected to offer the highest NHE.
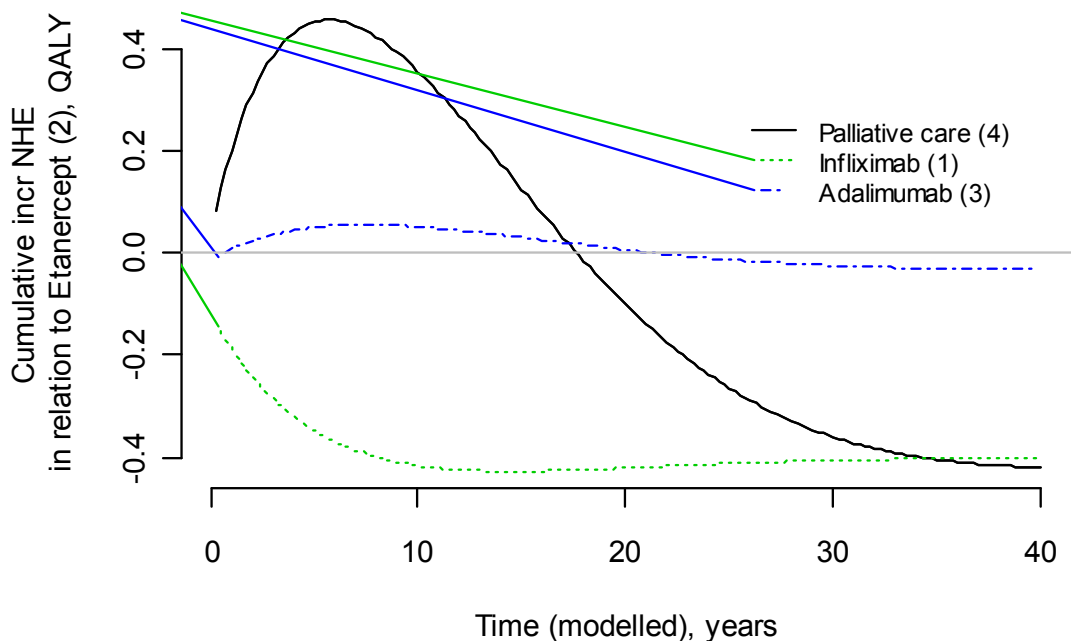
Figure 3.1d   Cumulative incremental NHE in PsA over the patient time horizon

ii)   Cost-effectiveness at the population level

Per patient NHEs can also be expressed for the population of current and future patients. This requires information about prevalence and future incidence of the target population (already required in appraisal). It also requires a judgement about the time horizon over which the technology will be used. This 'technology time horizon' ought to reflect the period over which the technology is likely to be part of clinical practice and generate the expected NHEs.[15] An estimate of the scale of the total population NHEs and how they cumulate over time is important for subsequent assessments, including: i) where the NHE for current patient populations must be compared with the benefits to future patients; and ii) where the treatment decision can be changed so the irrecoverable opportunity costs of initially negative NHE become significant, i.e., might influence the category of guidance.

For example, there is a large prevalent population eligible for EECP relative to future incident populations in this chronic condition. The total population NHE, assuming the technology will be used to treat prevalent and incident patients over 10 years, are reported in Table 3.3a. Expected cost-effectiveness is unchanged (ICER is the same as Table 3.2a) but the incremental NHE although small per patient, is more significant at a population level.

Table 3.3a   Expected cost-effectiveness of EECP for the population

				Cost-effectiveness threshold at:			
				£20,000 per QALY		£30,000 per QALY	
Treatment	Costs	QALY	ICER	NHE	Incr NHE,	NHE	Incr NHE,
	(£m)		(£/QALY)	QALY (£m)	QALY (£m)	QALY (£m)	QALY (£m)
EECP	896	1,435,787	£19,391	1,391,001 (27,820)	1,405 (28)	1,405,930 (42,177)	16,334 (490)
Std	-	1,389,596	-	1,389,596 (27,792)		1,389,596 (41,688)	

The 'investment profile' for EECP when used to treat patients over 10 years is illustrated in Figure 2.2. At a population level it is not until 17 years (rather than 14 years at a patient level) that initial losses are compensated by later gains and EECP 'breaks even'. In other words, EECP appears a more risky investment when evaluated at a population rather than individual level. This is because, although each patient treated with EECP is expected to offer the same profile of NHEs shown in Figure 3.1a,

---

[15] The time horizon for the technology might be longer or shorter than the patient time horizon. Technology time horizons might be based on historical evidence of the obsolescence of health technologies but any estimate will be a proxy for a complex and uncertain process of future changes in new technologies, prices and evidence. Therefore the impacts of different technology time horizons have been explored in each case study.

the negative NHE associated with patients' incident and treated in year 10 won't be offset by later gains until year 24. The population level 'investment profile' would exhibit greater risk (breakeven later) if the prevalent population was smaller relative to the incident population and/or the technology time horizon was longer. For example the 'breakeven' point extends to 23 years when the technology time horizon is increased to 20 years.
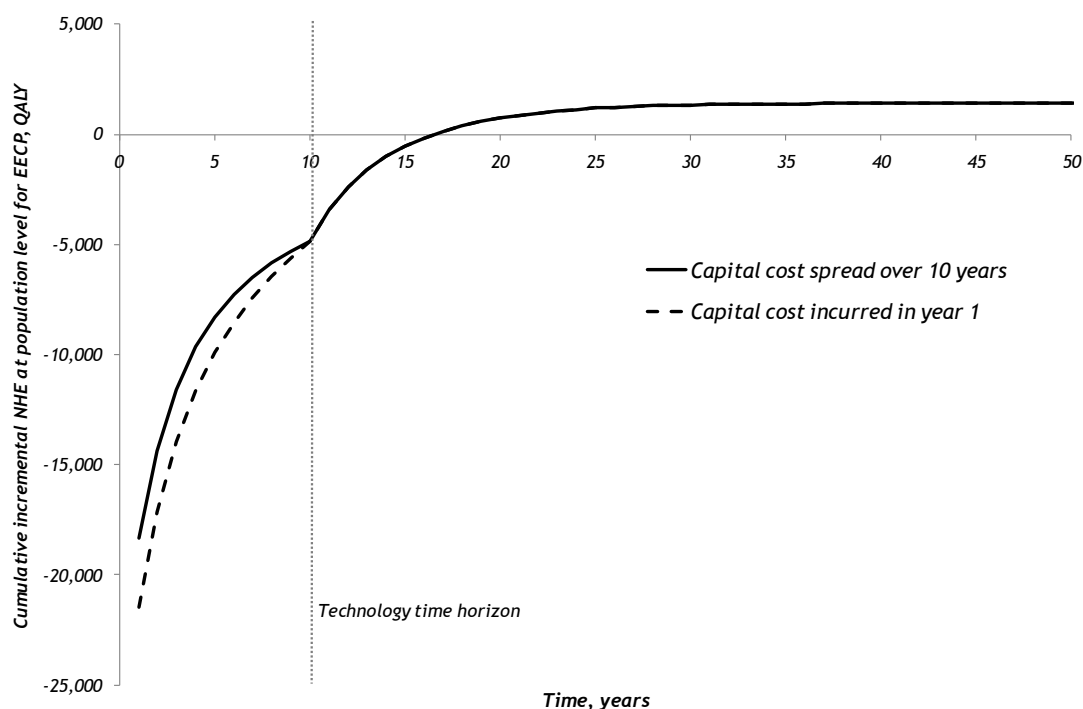
Figure 3.2 **Cumulative incremental NHE of EECP for the population**

The effect on the other case studies of assessment at the population level is similar to EECP. It simply increases the magnitude of differences in per patient NHE (to a greater extent for longer technology time horizons), but leaves expected cost-effectiveness unchanged. However, the 'investment profiles' at a population level also differ, exhibiting greater 'risk' indicated by later 'breakeven' points for the same reasons as EECP. For example, the 'breakeven' points for CLOP when evaluated at a population level are reported in Table 3.3b. At a technology time horizon of 10 years it is only at 11 years, rather than 5 years for a single patient, that 12 months of CLOP treatment 'breaks even' against current NHS care and not until 27 years (rather than 21 years) that it is better than a shorter treatment duration of 6 months. Even the shorter durations of treatment offer a 'risky profile', e.g., the 'breakeven' point for one month of treatment is 4 years (rather than 2 years).

Table 3.3b   Expected cost-effectiveness of CLOP for the population

Technology time horizon	Treatment	Incr NHE, QALYs (£m)	'Breakeven points' (years)		
			12 months vs 6 months	12 months vs NHS	1 month vs NHS
5 years	1: clop12	269 (5.4)			
	2: clop6	1,881 (37.6)			
	3: clop3	1,804 (36.1)	24	8	4
	4: clop1	4,073 (81.5)			
	5: NHS	-			
10 years	1: clop12	495 (9.9)			
	2: clop6	3,465 (69.3)			
	3: clop3	3,324 (66.5)	27	11	4
	4: clop1	7,502 (150)			
	5: NHS	-			
15 years	1: clop12	686 (13.7)			
	2: clop6	4,799 (96)			
	3: clop3	4,603 (92.1)	30	12	4
	4: clop1	10,389 (207.8)			
	5: NHS	-			
20 years	1: clop12	846 (16.9)			
	2: clop6	5,921 (118.4)			
	3: clop3	5,680 (113.6)	33	12	4
	4: clop1	12,820 (256.4)			
	5: NHS	-			

3.3.2   Point 2 - Are there significant irrecoverable costs?

The second point on the checklist requires: i) an assessment of whether there are irrecoverable costs and ii) a judgement of their potential significance, i.e., at the following point in the algorithm
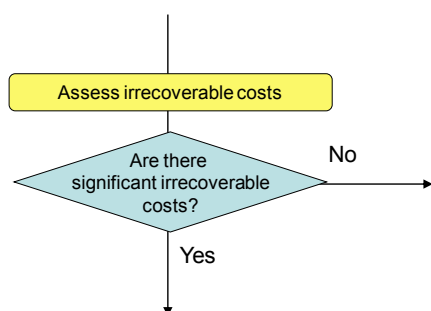
Irrecoverable costs are those which once committed cannot be recovered if guidance is changed at a later date. Irrecoverable costs are most commonly thought of as 'up-front' or capital costs of new facilities or equipment with long life expectancy (they might also include any practitioner training and the costs of implementation efforts). In NICE appraisal these types of cost are first annuitized [16] and then allocated pro-rata to the number of patients likely to be treated during the lifetime of the equipment. That is, capital costs are treated *as if* they are paid per patient treated over the life time of the equipment.   If guidance remains unchanged throughout this period (i.e., research does not report or other sources of uncertainty resolve) then this common assumption has no influence. However, should guidance change (initial approval is withdrawn) before the end of the lifetime of the equipment then, although future patents will no longer use the technology, the cost of the equipment allocated to them cannot be recovered. The possibility that initial guidance might change and its impact on expected costs needs to be considered before costs are made irrecoverable through approval or AWR. The impact of irrecoverable costs will tend to be greater if they represent a greater proportion of the total costs, if guidance is more likely to change and to change in the near rather than distant future.

---

[16] The annual payments required each year over the life time of the equipment (discounted at 3.5%) which would be equivalent to the capital cost at the start of the 1st year.

EECP is the only case study in which these types of cost are present to any great extent because treatment requires capital investment in the EECP machines themselves. The expected per patient and population costs reported in Tables 3.1a and 3.2b allocated this capital cost in the usual way (i.e., annuitized over the 10 year life time of the machines and allocated to the number of patients treated each year). The irrecoverable costs are reported separately in Table 3.4 and represent 19% of the total. However, this will have no influence on expected cost-effectiveness so long as guidance does not change during the lifetime of the equipment.

Table 3.4   Capital costs associated with EECP

					Cost-effectiveness threshold	
					£20,000 per QALY	
Treatment	**Capital costs**	**Non capital costs**	**QALY**	**ICER** (£/QALY)	**NHE** QALY (£m)	**Incr NHE** QALY (£m)
EECP	£170,304,591	£725,408,798	1,435,787	19,391	1,391,001 (27,820)	1,405 (28)
Std	-	-	1,389,596	-	1,389,596 (27,792)	-

'Investment profile' of NHE
Even in the absence of capital costs of equipment and facilities, NHE accumulate over time both at a patient and population level. With the possible exception of OMAL [17] the analysis in Section 3.3.1 indicates a common pattern of initially negative NHE that are only gradually offset by positive NHE in later periods. Therefore, approval or AWR commonly commits opportunity costs of negative NHE which are irrecoverable.

i)   Are they likely to be significant?
Whether or not irrecoverable costs are significant, i.e., might influence guidance, depends critically on whether guidance is likely to change and whether that is more likely in the near or distant future. That will depend on whether research is likely to be undertaken and when it is likely to report, as well as other events that might occur, e.g., a change in price following patent expiry. These are assessed later, at points 5 and 6 in the checklist. However, the *potential* significance of any irrecoverable costs can be assessed at this point. For example, capital costs can be judged based on the proportion of total population cost which are irrecoverable and their scale relative to the additional population NHE offered (e.g., see Table 3.4).

Judging the potential significance of the investment profiles of NHE is more nuanced. It depends whether treatment decisions for individual patients are irreversible, which in part depends on the nature of the disease. For example, in an acute condition the decision to treat a particular presenting patient with a technology cannot be reconsidered at a later date – it is irreversible. Although it is possible that the later benefits are not realised, it is also possible that they will realise more (the profiles of NHE in Figure 3.1a to 3.1d are the average over these possibilities). Similarly the possibility that guidance might change in the future (e.g., research suggests that the longer term benefits will not offset initial losses), will not influence the irreversible decision to treat a presenting patient with a technology that is expected to be cost-effective prior to the research reporting.

Implication for the case studies
CLOP is a treatment for acute coronary syndromes and, although decisions about treatment and its duration are not irreversible in the very short run, over the time scales more likely for research being conducted (and reporting) or other events occurring, which would change guidance, they can be regarded as such. Therefore, although the investment profile of CLOP (at a patient and more so at a population level) exhibits irrecoverable opportunity costs these should not be judged significant in the sense that they have little potential to influence guidance. There are also no significant irrecoverable costs associated with OMAL but for different reasons; although treatment decisions are reversible in this chronic condition, any irrecoverable opportunity costs appear very limited (see Figure 3.1c).
Both EECP and the biologics in PsA are for chronic conditions where the decision to treat a particular patient can be changed at some later date (decisions are not irreversible). Therefore, the type of

---

[17] The profile of NHE at a patient level did not exhibit significant irrecoverable opportunity costs. Assessment at a population level and for longer technology time horizons simply increases the magnitude of the expected negative NHE. Therefore, there are no irrecoverable costs in this case study.

'investment profile' of NHE at a patient and population level is significant because, instead of committing irrecoverable costs by deciding to use technologies expected to be cost-effective now, the decision and commitment of costs can be made later, after research reports, other events occur and/or guidance changes. Of course, proper account must be taken of the impact of withholding initiation of treatment on expected health benefits and costs (see Section 3.6), e.g., some patients who might have been treated may not survive to benefit from the results of the research or disease may have irreversibly progressed so that the expected health benefits are lower[19].

EECP is the only case where both types of irrecoverable costs are potentially significant. Figure 3.2 illustrates the impact of accounting for the actual timing of expenditure on EECP machines rather than treating it as if it was paid when each patient was treated, i.e., where expenditure is treated like a consumable cost by spreading the capital cost over 10 years.[18] If approval of EECP might be withdrawn before 10 years, the potential losses in NHE will be greater than initially indicated in Figure 3.2 because the equipment costs allocated to treating future patients cannot be recovered. The earlier such a change might occur the greater the additional loss. The impact of these possibilities should be considered at point 7 of the checklist before guidance to approve or AWR commits both types of irrecoverable costs.

Pricing and irrecoverable costs
The significance of irrecoverable treatment costs should also consider the scale of initially negative NHE as well as the duration of such losses (how long until the use of the technology 'breaks even' for an individual patient and for the population of patients who are likely to be treated if it is approved). Health technologies with patent protection are more likely to be priced close to the point at which the expected incremental NHE are close to zero, i.e., where the ICER is close to or equal to the threshold. A value based pricing scheme would formalise these existing incentives. The use of a technology which is only just expected to be cost-effective will not 'break even' until close to the end of the patient time horizon and much longer for the population of patients likely to benefit from its use (up to the technology time horizon *plus* the patient time horizon - less if patent expiry and cheaper generics enter before the technology time horizon). The scale of initial losses will also tend to be greater. Therefore, those technologies already priced close to the threshold, and all new technologies considered in a value based pricing scheme, will tend to increase the scale of irrecoverable costs committed by approval, making OIR or Reject more likely even when a technology is just expected to be cost-effective at point 1 of the checklist.[19]

3.4   Is further research required?

The judgements made at points 3 and 4 of the checklist are critical because if more research is not judged to be worthwhile no further assessments are required, unless there are significant irrecoverable costs (see table D1 in Appendix D). If research is worthwhile, then what type of evidence is needed and whether the research required to generate it can be conducted while the technology is approved will determine whether AWR or OIR are possibilities.

3.4.1   Point 3 – Does more research seem worthwhile?
The third point on the checklist requires an assessment of the potential benefits of conducting further research, i.e., at the following point in the algorithm:

---
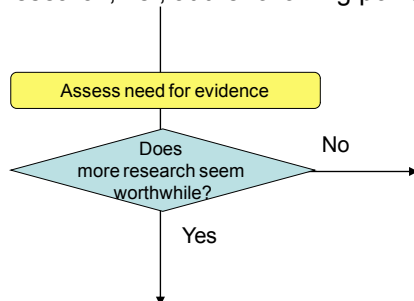
[18] In Figure 3.2 the technology time horizon happens to coincide with the life time of the equipment but it need not

[19] It is also important to consider the risk profile of the health technologies and activities likely to be displaced. Insofar as additional NHS costs do not just displace new technologies the net effect will still tend to increase 'risk' (see the technical note in the Addendum to the HTA report).

This requires judgements about: i) how uncertain a decision to approve or reject a technology might be based on the estimates of expected cost-effectiveness; and ii) whether the scale of the likely consequences of this uncertainty might justify further research. Some assessment of the potential consequences of uncertainty is important because it indicates the scale of the population NHE that could be gained if the uncertainty surrounding this decision could be immediately resolved, i.e., it represents an expected upper bound on the benefits of more research.[20] If the potential benefits of further research are unlikely to justify the costs, then a judgement that more research does not seem worthwhile will lead directly to guidance in the following circumstances (extracted from Table B1 in Appendix B):

Assessment	1	2	3	4	5	6	7	Guidance
6	Yes	No	No	-	-	-	-	Approve 4
12	No	No	No	-	-	-	-	Reject 4
35	No	Yes	No	-	-	-	-	Reject 11

(Pathway number is indicated to the left spanning the three data rows.)

i)   Assessing the consequences of uncertainty

Some assessment is required of: i) how uncertain a decision based on expected cost-effectiveness might be; and ii) what the consequences, in terms of population NHE, are likely to be if an incorrect decision is made.

EECP is expected to be cost-effective compared to standard care (see Tables 3.2a and 3.5a) but the estimates of cost and QALYs are uncertain so there is a chance that a decision to approve EECP based on existing evidence will be incorrect, i.e., standard care might offer greater NHE. Some assessment of the likely consequences of approving EECP when standard care might be better could be based on the difference in expected NHE, i.e., the expected incremental population NHE reported in Tables 3.3a and 3.5a). This is illustrated in Figure 3.3a where a judgement about the probability that a decision based on expected cost-effectiveness is correct translates into expected consequences based on the expected incremental population NHE. For example, if the decision was judged to be 100% certain then there are no consequences and so there would be nothing to be gained by more research. However, as the probability that the decision is correct becomes less certain, the expected consequences (and hence potential value of more research) increase.

This judgment, of how uncertain a decision might be, can be informed by the probabilistic analysis (PSA) already used to estimate costs and QALY and is required as part of the NICE reference case[35-36]. The probability that EECP is cost-effective is 0.428 (see Table 3.5a),[21] which would translate into approximately 800 QALYs (see Figure 3.3a) over the technology time horizon,[22] based on the expected or average difference between NHE. However, the difference in NHE when EECP is not the correct decision is not necessarily the average. In fact, it is very unlikely to be the average and such estimates may substantially under or overestimate the expected consequences of uncertainty.[23]

---

[20] In mathematics and economics this is referred to as expected opportunity loss. In decision theory and its applications including economic evaluation it is referred to as the expected value of perfect information (EVPI). It is also directly related to option value in financial economics.
[21] The alternative which is expected to be cost-effective may not have the highest probability of being cost-effective if, as in this case, the distribution of NHE are skewed, i.e., when the NHE of EECP are greater than std they are much greater but when std offers higher NHE they are only a little higher than EECP.
[22] The time horizon over which evidence generated by research about a technology might be valuable may be longer (or shorter) than the period over which the technology is used. Therefore there is a distinction between the technology time horizon and the time horizon for the benefits of research. To simplify the exposition in this summary of the case studies they are assumed to be equal but other credible assumptions are explored more fully in the addendum to the main report.
[23] If an assessment of expected consequences based on mean NHE was always an underestimate this would be a useful, simple assessment of a lower bound to the potential benefits of research. However, such estimates can also overestimate expected consequences, e.g., in the analysis of EECP at a threshold of £30,000. Unfortunately such circumstances cannot be specified in advance without conducting a proper analysis of the expected consequences anyway (see Technical Notes to the Addendum to the HTA Report).

Table 3.5a   Expected consequences of uncertainty for EECP

Treatment	ICER (£/QALY)	Cost-effectiveness threshold at:					
		£20,000 per QALY			£30,000 per QALY		
		Incr NHE QALY (£m)	Probability cost-effective	Expected consequences, QALY (£m)	Incr NHE, QALY (£m)	Probability cost-effective	Expected consequences, QALY (£)
EECP	19,391	1,405 (28.1)	0.428	9,287 (185.7)	1,405,930 (490)	0.7	2,774 (83.2)
Std		-	0.572		-	0.3	

The same probabilistic analysis can be used to record the difference between the NHEs of EECP and standard care and the frequency of such errors. This distribution of consequences is illustrated in Figure 3.4a. Commonly there are no consequences, because EECP is the correct decision (42.8%). However, when EECP offers lower NHE than standard care the consequence of error may be relatively small, e.g., 9% are less than 5,000 QALYs. However, they may be very large, although less likely, e.g., there is a small chance (5.7%) that they are greater than 30,000 QALYs. The average over this distribution provides the expected consequences of uncertainty, which in this case is 9,287 QALYs.[24]
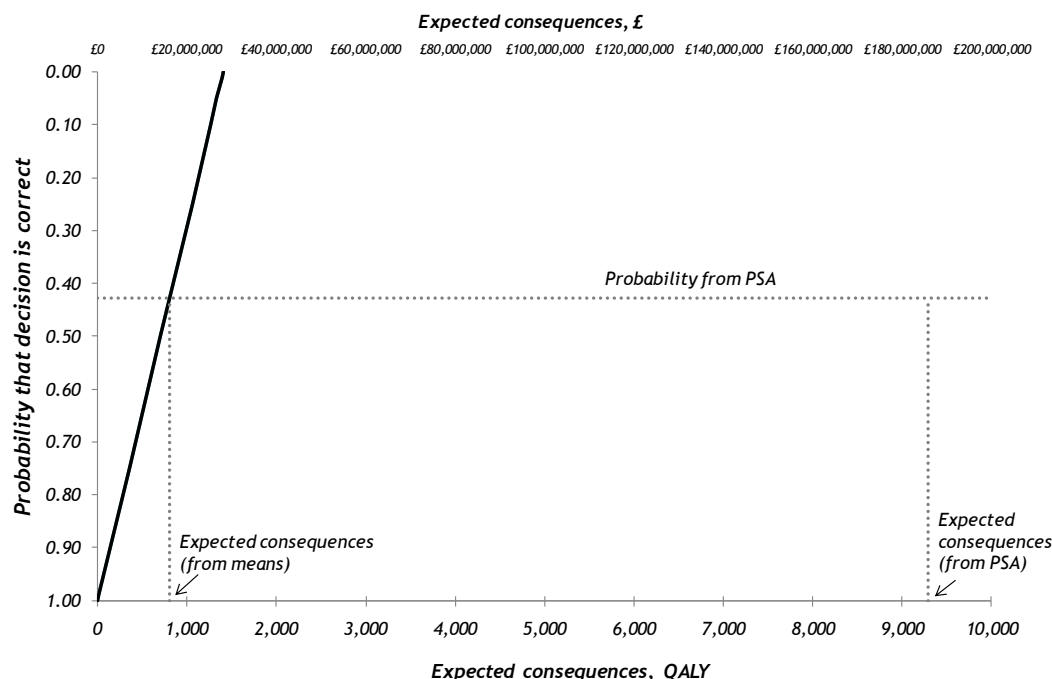
Figure 3.3a   Probability that EECP is cost-effective and the consequences of uncertainty

These expected consequences can be interpreted as an estimate of the population NHE over the technology time horizon that could be gained if the uncertainty surrounding this decision could be resolved immediately, i.e., it indicates an expected upper bound on the benefits of more research[35, 37].[25] The consequences can also be expressed as the equivalent NHS resources required to generate the same population NHE (£185.7m in Table 3.5a). They will increase with the size of the patient population and the technology time horizon. In the case of EECP the consequences fall with the cost-effectiveness threshold because a decision to approve EECP will be less uncertain (see Table 3.5a). A judgment at this point that more research might be worthwhile seems reasonable, since the upper bound on its potential benefits exceed the likely costs.

---

[24] This is substantially greater than the estimate of 800 QALYs based on mean incremental population NHEs, demonstrating that such simple estimate may be misleading - see Figure 3.3a.
[25] It should be noted that these estimates of QALYs that might be gained are for the population over the time horizon for the benefit of research (in this case equal to the technology time horizon – see foot note 22) if all sources of uncertainty could be immediately resolved. It includes both improvements in health outcomes for this population but also NHS resource saving that could be made and used to generate QALYs elsewhere.
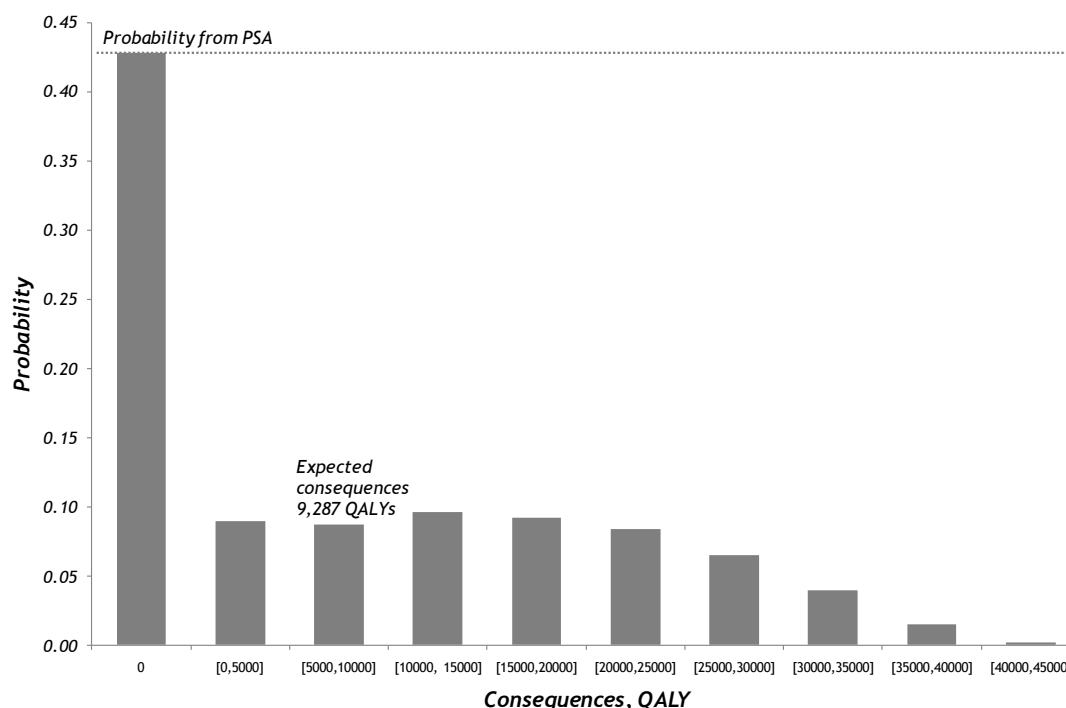
Figure 3.4a   Distribution of the consequences of uncertainty for EECP

Multiple alternatives
Similar analysis can be conducted when there are more than two alternatives but greater difficulties are encountered unless the results of PSA are used to assess both uncertainty and its consequences. For example, in the CLOP case study, 12 month treatment duration with CLOP is expected to be cost-effective but this is also uncertain. A judgement is required about the chance that 12 months of treatment is incorrect and if so which of the other four alternatives are likely to offer higher NHE, and how much higher.   In other words, for decisions involving multiple alternatives, a judgement is required on the level of uncertainty surrounding the decision, how this uncertainty is distributed across the various alternatives and what the consequences are likely to be. The results of PSA can inform this judgement. The probabilities that each of the 5 alternatives is cost-effective are reported in Table 3.5b. This indicates that 12 months treatment is uncertain (probability that it is incorrect is 0.476). However, much of this probability of error is allocated to 6 months treatment with CLOP (0.18) where the difference in NHEs is likely to be relatively modest.

Table 3.5b   Expected consequences of uncertainty for CLOP

Treatment	ICER (£/QALY)	Incr NHE * QALY (£m)	Probability cost-effective	Expected consequences QALY (£m)	Incr NHE * QALY (£m)	Probability cost-effective	Expected consequences QALY (£m)
		Cost-effectiveness threshold at:					
		£20,000 per QALY			**£30,000 per QALY**		
1: clop12	18,663	495 (9.9m)	0.524		2,798 (56.0m)	0.677	
2: clop6	10,477	3,465 (69.3m)	0.180		4,736 (94.7m)	0.092	
3: clop3	9,396	3,324 (66.5m)	0.018	5,194 (103.9)	4,305 (86.1m)	0.009	3,657 (109.7)
4: clop1	4,961	7,502 (150.0m)	0.075		8,327 (166.5m)	0.052	
5: NHS	-	-	0.202		-	0.170	

* The mean additional population NHE of moving from the least to most effective alternative, i.e., the incremental NHE of 12 month compared to NHS is the sum of these increments (14,786 QALY or £295.7m at £20,000 per QALY)

The distribution of consequences is illustrated in Figure 3.4b.   Most commonly (52.4%) there are no consequences, because 12 months duration of treatment with CLOP is the correct decision. When it is not, there is a greater chance of relatively small consequences (30% are less than 10,000 QALYs)

which occur predominantly when 6 months treatment duration offers the highest NHE. But there is a small chance of larger consequences (less than 5% chance that they are greater than 30,000 QALYs) when standard NHS treatment offers the highest NHE, i.e., there remains important uncertainty about the cost-effectiveness of CLOP itself, not just its duration. The expected consequence of uncertainty (5,194 QALYs) is simply the average over this distribution. Again this can be interpreted as an estimate of the population NHE that could be gained, over the time horizon of this technology, if the uncertainty about treatment and its duration could be immediately resolved. Therefore, like EECP, a judgement at this point that more research might be worthwhile seems reasonable, since the potential benefits exceed the likely costs.
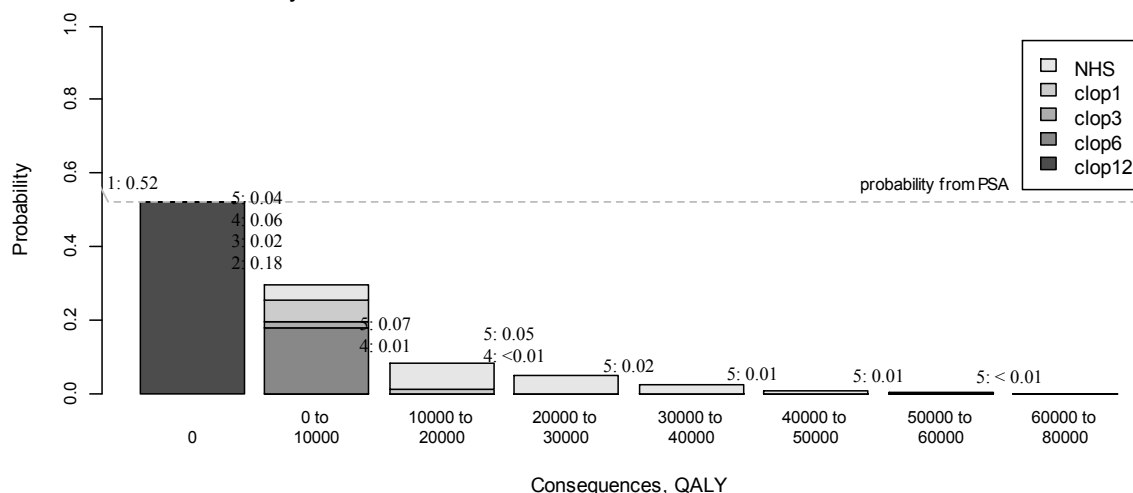
Figure 3.4b   Distribution of the consequences of uncertainty for CLOP

PsA provides a similar picture to CLOP, where approval of the alternative which is expected to be cost-effective (etanercept) is uncertain (probability that approval is incorrect is 0.557), but in this case most of this probability of error is associated with palliative care (probability of 0.4 that it is cost-effective). Again, there is a greater chance of relatively small consequences (19% are less than 28,000 QALYs), most of which occur when adalimumab has the highest NHE, but a smaller chance of very large consequences (4.7% chance that they are greater than 138,000 QALYs), which occur when palliative care offers the greatest NHE. The expected consequences of uncertainty and the upper bound on the population NHE that might be gained by immediately resolving uncertainty (35,342 QALYs or £707m over the technology time horizon) supports a judgement that more research maybe worthwhile.

ii)   Analysis of subgrou*ps*

OMAL was not expected to be cost-effective based on existing evidence. The ICER in Table 3.2c was substantially greater than the threshold and a decision to reject this technology does not appear uncertain. This judgement is supported by the results of PSA (the probability that OMAL is cost-effective is zero in Table 3.5c). Therefore, a decision to reject (Reject [4] in the algorithm) is not uncertain; there are no consequences of uncertainty and nothing to be gained by more research. However, it is possible to consider a high risk subgroup within this population. Subgroups, once credibly defined, need to be considered in the same way; starting at point 1 on the checklist, i.e., entering at the top of the algorithm. Although the ICER for this high risk subgroup is somewhat lower, it is still significantly higher than the threshold. The results of PSA suggest that even at a threshold of £30,000 the probability that OMAL is cost-effective is very small and the upper bound on the gains from more research are very limited (10.61 QALYs). Therefore, even after an analysis of subgroups OMAL is not expected to be cost-effective and more research does not seem worthwhile. OMAL can be rejected at this point and no further assessment is required.

Table 3.5c   **Expected consequences of uncertainty for OMAL**

Severe population		Cost-effectiveness threshold at:					
		£20,000 per QALY			£30,000 per QALY		
Treatment	ICER (£/QALY)	Incr NHE QALY (£m)	Probability cost-effective	Expected consequences QALY (£)	Incr NHE QALY (£m)	Probability cost-effective	Expected consequences QALY (£)
Omal + Std	93,844	-5,789 (-116)	0.0	0	-3,337 (-100)	0.0	0.0
Std		-	1.0		-	1.0	
High risk subgroup		Cost-effectiveness threshold at:					
		£20,000 per QALY			£30,000 per QALY		
Treatment	ICER (£/QALY)	Incr NHE QALY (£m)	Probability cost-effective	Expected consequences, QALY (£)	Incr NHE QALY (£m)	Probability cost-effective	Expected consequences, QALY (£m)
Omal + Std	69,463	-3,851 (-77)	0.0	0	-2,048 (-61)	0.013	10.61 (0.32)
Std		-	1.0		-	0.987	

iii)   **Alternative scenarios**

There are often alternative views about the quality and relevance of evidence as well as other assumptions that might be made when estimating expected costs and QALYs. These are commonly presented as separate scenarios, with estimates of costs and QALY presented for each. Much of the deliberation by the Appraisal Committee often surrounds the scientific value judgments required to judge the credibility of the alternative assumptions represented by such scenarios. The type of probabilistic analysis reported represents the uncertainty within each scenario and will be sufficient to indicate the potential benefits of research when only one scenario is regarded as credible. However, when more than one scenario might be credible and carry some 'weight', there will be uncertainty *between* as well as *within* scenarios. The 'weighting' of scenarios can be made explicit by assigning probabilities to represent how credible each is believed to be. The weighted average of costs and QALY across scenarios can easily be calculated. It is also tempting to take a simple weighted average of the expected consequences of uncertainty across these scenarios as well. However, a simple weighted average may under or overestimate the combined consequences of uncertainty within and between scenarios [1, 38]. The correct estimate requires the probabilities (weights) to be applied directly to the simulated output from PSA rather than to the mean values. Although this doesn't require additional simulation and is quick and easy to implement, it does require that either the probabilities are made explicit in advance or for estimates to be presented for a range of probabilities that might represent the judgement of the Appraisal Committee following deliberation.

For example, the CLOP analysis presented above assumes a constant relative treatment effect for different durations of treatment (scenario A). An alternative assumption (scenario B) was that the relative treatment effect also differed by duration based on the data reported in the SIGN guidelines. This alternative assumption made longer durations less cost-effective and reduced the expected consequences of uncertainty from 5,195 to 3,969 QALYs. Although scenario A was regarded as more credible by the AC, scenario B might nevertheless carry some weight or have some probability associated with it. In this case the simple weighted average of expected consequences (linear combination of mean estimates) is very similar to the correct estimate based on weighting the output of PSA in Figure 3.5a. This also shows how these estimates can be presented for a range of probabilities.
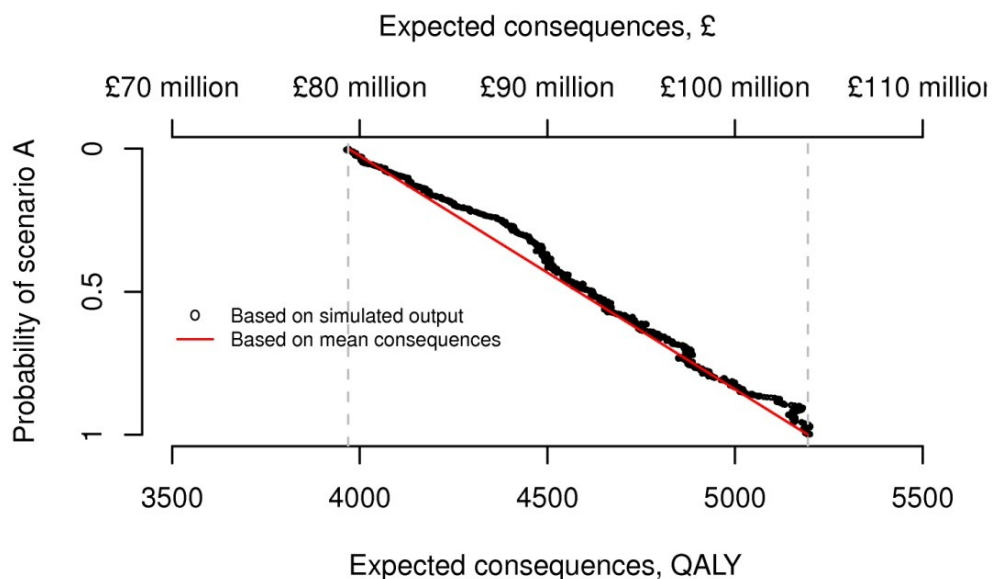
Figure 3.5a **Expected consequences of uncertainty with alternative scenarios (CLOP)**

An alternative assumption of a common class effect across the three biologics was considered in the PsA case study (scenario B), but was judged less credible than the analysis which allowed differential effects (scenario A).   The alternative scenario made etanercept less likely to be cost-effective and increased the expected consequences of uncertainty from 34,930 to 38,521 QALYs (see Figure 3.5b). In this case a simple weighted average of expected consequences based on the probability assigned to each scenario is, in general, lower than the correct estimate of expected consequences based on the output from PSA.
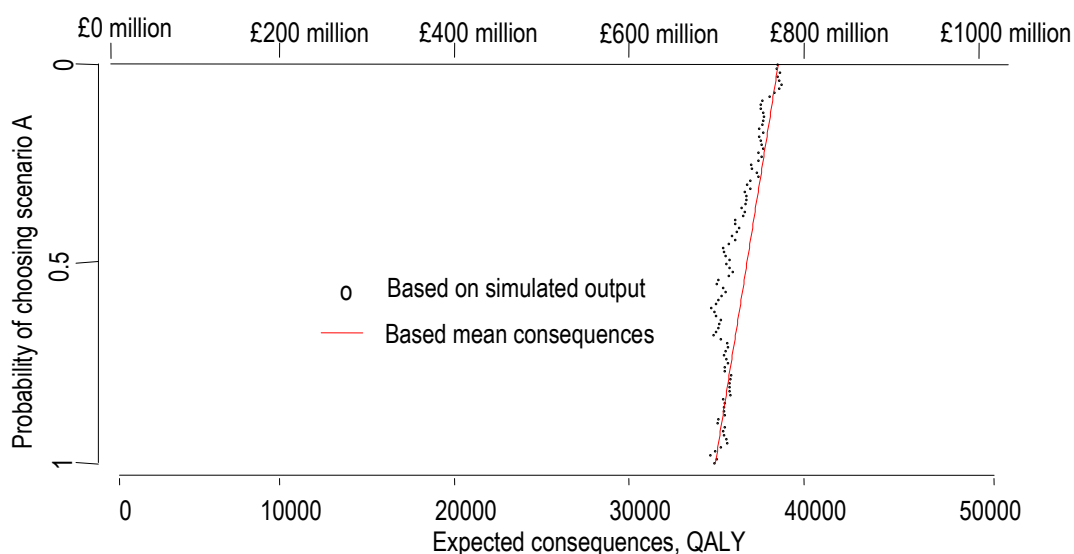
Figure 3.5b **Expected consequences of uncertainty with alternative scenarios (PsA)**

Elicitation
The single RCT of EECP showed evidence of improvements in quality of life at 12 months; however, the degree to which these are sustained in the long run is uncertain. Rather than make alternative assumptions and present extreme scenarios, formal elicitation of the judgement of clinical experts about the likelihood of QALY gains in subsequent years was undertaken.[26] The uncertainty in these

---

[26] Five experts with experience and knowledge of EECP in the UK, independently completed an Excel based exercise. The uncertainty associated with any judgement is critical, so a frequency chart format, where experts place 20 crosses on a frequency chart to represent a distribution was adopted. The results from each expert were linearly pooled, with equal weight, providing the probability of continuing to respond to treatment in each subsequent year. The uncertainty associated with these pooled estimates was characterised by fitting Beta distributions to pooled responses.

elicited values is included in the estimates of the expected consequences of uncertainty reported in Table 2.5a, which might otherwise have been represented by alternative scenarios. For example: no QALY benefits beyond 12 months could be assumed for scenario A; benefits sustained for a patient's life time for scenario B; and sustained for 4 years for scenario C. The results of elicitation implied probabilities of: 0.243; 0.353; and 0.404 associated with each of these scenarios respectively.   A simple weighted average of the expected consequences within each scenario using these probabilities (1,442 QALYs) significantly underestimates both the estimate of expected consequences based on the all the information from elicitation (9,287 QAYs) and the estimate based on weighting scenarios using the simulated output rather than the mean estimates (13,081 QALYs). This illustrates that: i) a simple weighted average of expected consequences may be misleading; and ii) that elicitation may provide a richer characterisation of uncertainty as well the probabilities associated with alternative assumptions (see Technical Notes in the Addendum to the HTA Report[1].

3.4.2   Point 4 - Is research possible with approval?

The fourth point on the check list requires an assessment of what type of evidence is needed and a judgement of whether the research required to generate it can be conducted while the technology is approved, i.e., at the following point in the algorithm:
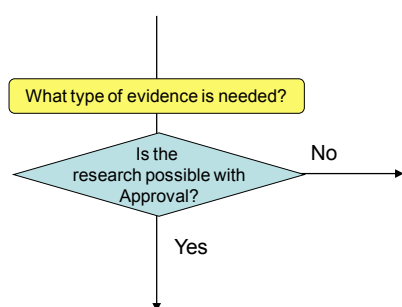
Although the decision at this point does not lead directly to guidance, it does determine whether AWR or OIR are possibilities. This judgement will depend, in part, on whether the type of evidence that is needed will require experimental research design. For example, more precise estimates of relative treatment effect are likely to require an RCT if the dangers of selection bias are to be avoided. However, further RCTs for this particular indication and patient group are unlikely to be possible once a technology is approved for widespread NHS use.

This requires judgements about: i) how important particular types of parameters (inputs to the economic model) are to estimates of cost and QALY; ii) what values these parameters would have to take to change a decision based on expected cost-effectiveness; iii) how likely is it that parameters might take such values and iv) what would be the consequences if they did, i.e., what might be gained in terms of population NHE if the uncertainty in the values of these parameters could be immediately resolved?

i)   Assessing the importance of parameters

The type of economic model used to estimate expected cost-effectiveness in NICE appraisal specifies the relationship between the inputs (the parameters) and outputs (costs and QALYs). A simple summary of the direction and strength of these relationships can be provided by calculating elasticities for each, i.e., the proportionate change in the NHE of each alternative, or differences in NHE, due to a one percent change in the value of the parameter, e.g., those parameters with high elasticities (especially with respect to differences in NHE) might be regarded as more 'important'. These elasticities are presented for CLOP case study in Table 3.6a.   They give some indication of i) relative importance for certain comparisons (e.g., RR_death seems particularly important for all comparisons); ii) identifies those that are of no or very limited importance (e.g., parameters 1-6 in the comparison of 12 and 6 months treatment duration); and iii) the direction of the relationship (e.g., the elasticity for C_Well is negative indicating that if the costs of NHS care in the Well state are greater, then 12 month treatment will be less cost-effective compared to 6 months or current NHS care).

Table 3.6a   Elasticities associated with parameters (CLOP)

		Parameter	Elasticity over the NHE (QALY) of					Elasticity over the INHE (QALY) of		
			clop12	clop6	clop3	clop1	NHS	clop12 vs. NHS	clop12 vs. clop6	clop12 vs. all
Natural history	1	P_die_0.1	-0.208	-0.207	-0.207	-0.207	-0.222	0.014	-	0.003
	2	P_NFMI_0.1	-0.012	-0.012	-0.011	-0.011	-0.015	0.004	-	-
	3	P_die_1.3	-0.137	-0.137	-0.137	-0.147	-0.145	0.008	-	0.004
	4	P_NFMI_1.3	-0.002	-0.002	-0.002	-0.002	-0.002	0.001	-	-
	5	P_die_3.6	-0.146	-0.146	-0.157	-0.157	-0.154	0.008	-	0.007
	6	P_NFMI_3.6	-0.005	-0.005	-0.007	-0.007	-0.007	0.002	-	0.001
	7	P_die_6.12	-0.148	-0.159	-0.158	-0.157	-0.155	0.007	0.011	0.01
	8	P_NFMI_6.12	-0.005	-0.007	-0.007	-0.007	-0.007	0.002	0.002	0.002
	9	TP_AC	-0.121	-0.12	-0.12	-0.12	-0.118	-0.003	-0.001	-0.002
	10	TP_AD	-3.637	-3.622	-3.604	-3.594	-3.541	-0.096	-0.016	-0.047
	11	TP_CD	-0.233	-0.235	-0.239	-0.24	-0.253	0.02	0.002	0.009
	12	TP_BD	-0.586	-0.593	-0.602	-0.605	-0.641	0.055	0.007	0.024
Utilities	13	U_Well	0.746	0.745	0.743	0.742	0.737	0.009	0.001	0.004
	14	U_Well1	6.09	6.064	6.034	6.017	5.929	0.16	0.026	0.079
	15	U_NFMI	0.133	0.134	0.136	0.136	0.144	-0.011	-0.001	-0.005
	16	U_POSTMI	1.138	1.15	1.165	1.171	1.236	-0.099	-0.012	-0.043
RE	17	RR_death	-0.639	-0.491	-0.344	-0.207		-0.641	-0.15	-0.38
	18	RR_NFMI	-0.024	-0.018	-0.013	-0.011		-0.025	-0.006	-0.014
Costs	19	C_Well	-0.74	-0.737	-0.733	-0.731	-0.72	-0.019	-0.003	-0.009
	20	C_MI_LT	-0.051	-0.052	-0.053	-0.053	-0.056	0.004	0.001	0.002
	21	C_PostMI	-0.142	-0.143	-0.145	-0.146	-0.154	0.012	0.002	0.005
	22	TC_Well_Dead	-0.027	-0.027	-0.027	-0.027	-0.027	-	-	-
	23	C_t1	-0.045	-	-	-	-	-0.045	-0.045	-0.045
	24	C_t2	-	-0.033	-	-	-	-	0.033	0.008
	25	C_t3	-	-	-0.026	-	-	-	-	0.007
	26	C_t4	-	-	-	-0.022	-	-	-	0.005
	27	C_t5	-	-	-	-	-0.016	0.016	-	0.004

Although these measures of importance are more instructive than a series of arbitrary one way sensitivity analysis, they do not directly help the assessment of what values parameters must take to change decisions and how likely such values might be.   A simple summary of the values particular parameters must take to make each of the alternatives cost-effective can also be provided. These 'threshold values' for parameters are presented for CLOP case study in Table 3.6b.   This provides additional information to the elasticities in Table 3.6a, e.g., there are only 6 parameters which could possibly take values that would lead to current NHS care (without CLOP) generating higher NHE than 12 months of treatment with CLOP. However, although instructive, such 'threshold values' do not indicate how likely it is that threshold will be crossed or the combined effect of groups of related parameters.

Table 3.6b   Thresholds associated with parameters (CLOP)

		Parameter	Mean value	Clop12	Clop6	Clop3	Clop1	NHS
Natural history	1	P_die_0.1	0.032	0 to 0.10	0.11 to 0.54	0.54 to 0.63	0.63 to 1	-
	2	P_NFMI_0.1	0.04	0 to 0.14	0.14 to 0.71	0.71 to 0.82	0.82 to 1	-
	3	P_die_1.3	0.022	0 to 0.10	0.10 to 0.55	0.55 to 1	-	-
	4	P_NFMI_1.3	0.004	0 to 0.10	0.10 to 0.7	0.7 to 1	-	-
	5	P_die_3.6	0.023	0.01 to 0.10	0.10 to 1	0 to 0.01	-	-
	6	P_NFMI_3.6	0.011	0 to 0.11	0.11 to 1	-	-	-
	7	P_die_6.12	0.024	0.02 to 1	0 to 0.02	-	-	-
	8	P_NFMI_6.12	0.009	0.005 to 1	0 to 0.005	-	-	-
	9	TP_AC	0.018	0 to 0.06	0.06 to 1	-	-	-
	10	TP_AD	0.072	0 to 0.08	0.08 to 0.10	-	-	0.10 to 1
	11	TP_CD	0.188	0.12 to 1	0 to 0.12	-	-	-
	12	TP_BD	0.07	0.06 to 1	0.04 to 0.06	-	-	0 to 0.04
Utilities	13	U_Well	0.798	0.29 to 1	0 to 0.29	-	-	-
	14	U_Well1	0.93	0.90 to 1	0.74 to 0.90	-	-	0 to 0.74
	15	U_NFMI	0.801	0 to 1	-	-	-	-
	16	U_POSTMI	0.931	0 to 1	-	-	-	-
RE	17	RR_death	0.931	0 to 0.93	0.94 to 0.97	0.97 to 0.98	0.98 to 0.99	1.00 to max[*]
	18	RR_NFMI	0.71	0 to 0.82	0.83 to 1.55	1.56 to 1.83	-	1.84 to max[*]
Costs	19	C_Well	2061.5	0 to 2690	2690 to 5611	-	-	5611 to max[*]
	20	C_MI_LT	6050	0 to max[*]	-	-	-	-
	21	C_PostMI	2309.7	870 to max[*]	0 to 870	-	-	-
	22	TC_Well_Dead	871.5	0 to 20474	20474 to max[*]	-	-	-
	23	C_t1	895.1	0 to 910	910 to max[*]	-	-	-
	24	C_t2	651.6	630 to max[*]	0 to 630	-	-	-
	25	C_t3	524.2	370 to max[*]	-	0 to 370	-	-
	26	C_t4	434.8	150 to max[*]	-	-	0 to 150	-
	27	C_t5	329.8	0 to max	-	-	-	-

ii)   Assessment of uncertainty

The judgement about how likely it is that parameters might take values which will change the technology expected to be cost-effective can be informed by the results of probabilistic analysis. This is because the distributions assigned to parameters in PSA ought to reflect the amount and quality of exiting evidence and describe how uncertain the parameter estimates are. The probability that each parameter might take values which would lead to each of the alternatives being cost-effective are reported for the CLOP case study in Table 3.6c. This, essentially, decomposes the overall probabilities reported in Table 3.5b into the contribution that each parameter makes.[27] Interestingly, it indicates that it is uncertainty in the estimates of relative effect (RR_Death) that contributes most to the probability of error associated with 12 months of treatment. It is the only parameter which (alone) might take values that could make any of the other alternatives cost-effective. It is also worth noting that there is a very small chance that cost in the 'well state' (C_Well) might be sufficiently high that

---

[27] The probability of error associated with 12 month of treatment reported in table 3.5b will, in general, not equal the sum of probabilities of error across the parameters, because the overall probability from PSA takes account of the joint effect of uncertainty in all parameters simultaneously. Even if parameters are independent they will be related to differences in NHE in different ways (indicated by the sign and magnitude of the elasticities – see Table 3.6a), so sometimes the effect of uncertainty in one may, to some extent, 'substitute' or 'complement' the effect of uncertainty in others.

standard NHS care would be cost-effective, i.e., if the NHS costs associated with the 'well state' are higher than any cost savings associated with moving more patients more quickly to the well state will tend to be lower.

Table 3.6c Probabilities associated with parameter values (CLOP)

		Parameter	Clop12	Clop6	Clop3	Clop1	NHS
Natural history	1	P_die_0.1	1	-	-	-	-
	2	P_NFMI_0.1	1	-	-	-	-
	3	P_die_1.3	1	-	-	-	-
	4	P_NFMI_1.3	1	-	-	-	-
	5	P_die_3.6	1	-	-	-	-
	6	P_NFMI_3.6	1	-	-	-	-
	7	P_die_6.12	0.65	0.35	-	-	-
	8	P_NFMI_6.12	0.91	0.09	-	-	-
	9	TP_AC	1	-	-	-	-
	10	TP_AD	0.83	0.17	-	-	-
	11	TP_CD	1	-	-	-	-
	12	TP_BD	0.85	0.15	-	-	-
Utilities	13	U_Well	1	-	-	-	-
	14	U_Well1	0.94	0.06	-	-	-
	15	U_NFMI	1	-	-	-	-
	16	U_POSTMI	1	-	-	-	-
RE	17	RR_death	0.55	0.18	0.01	0.10	0.16
	18	RR_NFMI	0.97	0.03	-	-	-
Costs	19	C_Well	0.78	0.19	-	-	0.03
	20	C_MI_LT	1	-	-	-	-
	21	C_PostMI	0.89	0.11	-	-	-
	22	TC_Well_Dead	1	-	-	-	-
	23	C_t1	0.95	0.05	-	-	-
	24	C_t2	0.99	0.01	-	-	-
	25	C_t3	1	-	-	-	-
	26	C_t4	1	-	-	-	-
	27	C_t5	1	-	-	-	-

iii) What type of evidence is needed?

Although an understanding of uncertainty and importance of parameters separately is helpful, an assessment of the likely consequences of this uncertainty, and therefore what might be potentially gained, in terms of population NHE, if uncertainty could be immediately resolved, is required. This assessment can directly inform the judgement of what evidence is needed and whether the type of research required to generate it will be possible with approval. In a similar way to Section 3.4.1, the results of PSA can inform this judgement since estimates of the expected consequences of uncertainty associated with each parameter combines both uncertainty in its potential values and their importance in terms of changing decisions and differences in NHE. The expected consequences of uncertainty associated with each parameter in CLOP are reported in Table 3.6d.   This decomposes the overall expected consequences reported in Table 3.5b into the contribution that each parameter makes and which other alternatives might offer higher NHE than 12 month treatment.[28] It confirms that it is uncertainty in the estimates of relative effect (RR_Death) that contributes most and where there is potentially the most to be gained by resolving this uncertainty through additional research (4,433 QALYs or £88.7m). Since more precise estimates of relative effects are likely to require a RCT, a judgement that the type of research need will not be possible if 12 month treatment duration is approved may be reasonable. However, the potential benefits of resolving the uncertainty associated with other groups of parameters, e.g., costs (547 QALYs or £10.9m) and the natural history (369 QALYs or £7.4m), might mean that other types of cheaper, non experimental research could be worthwhile as well or might be conducted prior to commissioning potentialy expensive experimental research which may take some time to complete and report[39].[29]

---

[28] For similar reason to foot note 27 the overall expected consequences of uncertainty reported in table 3.5b (5,194 QALYs) will not, in general, equal the sum of the expected consequences for each of the parameters separately (5,432 QALYs).

[29] The sequence in which research might be conducted can also be considered. This is discussed at greater length and more formally in the technical appendix to the main report, i.e., OIR for treatment effects followed by AWR for natural history if that research continues to be necessary once research recommended in OIR reports. It is feasible to withdraw approval of a

Table 3.6d   Consequences of uncertainty associated with parameter values (CLOP)

		Expected consequences (QALYs)					
		Decomposed by treatment choice					
	Parameter	clop12	clop6	clop3	clop1	NHS	Overall
Natural history*	1 P_die_0.1	0	-	-	-	-	0
	2 P_NFMI_0.1	0	-	-	-	-	0
	3 P_die_1.3	0	-	-	-	-	0
	4 P_NFMI_1.3	0	-	-	-	-	0
	5 P_die_3.6	0	-	-	-	-	0
	6 P_NFMI_3.6	0	-	-	-	-	0
	7 P_die_6.12	0	250	-	-	-	250
	8 P_NFMI_6.12	0	9	-	-	-	9
	9 TP_AC	0	-	-	-	-	0
	10 TP_AD	0	47	-	-	-	47
	11 TP_CD	0	-	-	-	-	0
	12 TP_BD	0	35	-	-	-	35
Utilities*	13 U_Well	0	-	-	-	-	0
	14 U_Well1	0	10	-	-	-	10
	15 U_NFMI	0	-	-	-	-	0
	16 U_POSTMI	0	-	-	-	-	0
RE	17 RR_death	0	284	16	518	3614	4433
	18 RR_NFMI	0	3	-	-	-	3
Costs*	19 C_Well	0	153	-	-	321	474
	20 C_MI_LT	0	-	-	-	-	0
	21 C_PostMI	0	8	-	-	-	8
	22 TC_Well_Dead	0	-	-	-	-	0
	23 C_t1	0	8	-	-	-	8
	24 C_t2	0	0	-	-	-	0
	25 C_t3	0	-	-	-	-	0
	26 C_t4	0	-	-	-	-	0
	27 C_t5	0	-	-	-	-	0

* Expected consequences for groups of parameters are: natural history 369 QALY (7.4m); RE 4,504 QALYs (£90.1m); 15 QALYs (£0.3m) and costs 547 QALYs (£10.9m). These are not equal to the sum of expected consequences for component parameters for the reasons explained in footnotes 23 and 24.

EECP provides a similar pattern of results, with the most significant consequences of uncertainty associated with parameters related to relative treatment effect; suggesting that the research needed might not be possible following approval of EECP. Interestingly, although the probability of sustaining the QALY benefits of EECP in the long run is very uncertain, the greater part of potential value is in more precise estimates of QALY gains in the first 12 months (2,709 QALYs or £54m and 8,511 QALYs or £170m respectively).

In PsA, on the other hand, the greater potential value is associated with uncertainty in natural history of HAQ progression (8,697 QALY or £17.4m) rather than relative treatment effect (1,201 QALYs or £2.4m). Although this might suggest that AWR, which recommended research on HAQ progression is possible and worthwhile, the combined potential benefits of resolving uncertainty associated with natural history (both in HAQ and PsARC) and treatment effect together is much greater than the 'sum of its parts'.[30] This suggests that both types of research could be conducted while etanercept continues to be approved but infliximab and adalimumab are not, i.e., a possible OIR rather than reject for infliximab and adalimumab but AWR for etancercept.

Implications of between scenario uncertainty
In part iii) of Section 3.4.1 the contribution alternative scenarios might make to the overall expected consequences of uncertainty and therefore the potential gains from further evidence was considered and discussed. In situations where more than one scenario might be regarded as credible, there will

---

technology that is expected to be cost-effective to allow research to be conducted then in this case AWR followed by OIR if necessary is likely to offer greater expected population NHE.
[30] These values represents an upper bound on what might be gained be resolving each alone. The potential value of resolving these uncertainties together is much greater for the reasons given in footnote 28 and 27.

be uncertainty *between* as well as *within* each of the scenarios. It was demonstrated in Section 3.4.1 that an assessment of the combined consequences of both sources of uncertainty requires 'weights' (probabilities) to be assigned to represent their credibility, which can then be applied directly to the simulated output from PSA. However, the same analysis can also be used to identify the expected consequences of uncertainty associated with the alternative scenarios themselves, i.e., what might be gained if evidence could immediately distinguish which scenario was 'true'[1]. This can help to inform the assessment of what type of evidence might be needed and whether the research required to generate it is likely to be possible once a technology is approved for widespread NHS use.

For example in the CLOP analysis, scenarios A and B (treatment effect was constant or differed by treatment duration respectively) were associated with expected consequences of uncertainty of 5,195 and 3,969 QALYs respectively. If both scenarios where regarded as equally likely, the overall expected consequences of uncertainty (combining consequences within and between scenarios) would be 4,667QALYs. However, the expected consequences of uncertainty associated with the two alternative scenarios themselves and what might be potentially gained if the uncertainty between them could be immediately resolved is relatively modest at 85 QALYs, i.e., most of what might be gained from further evidence is associated with the parameters in Table 3.6d rather than the alternative scenarios. This suggests that more evidence about overall relative effect on mortality is more important than resolving uncertainty about whether such an effect differs by treatment duration.

In the EECP case study, formal elicitation of the judgement of clinical experts about whether observed QALY gains at 12 months are likely to be sustained in subsequent years was undertaken. Since the uncertainty in these elicited values was incorporated into the analysis in the same way as other parameters, the use of alternative scenarios was not necessary. However, scenarios were used to illustrate the type of analysis, which, without elicitation, might otherwise have been required. The scenarios included: A - no QALY benefits beyond 12 months; B - benefits sustained for a patient's life time; and C - sustained for 4 years. The results of elicitation implied probabilities of: 0.243; 0.353; and 0.404 associated with each of these scenarios respectively. Based on these 'weights' for each scenario the overall expected consequences of uncertainty (combining the consequences within and between scenarios) would be 14,146 QALYs. In this case the expected consequences of uncertainty between the scenarios (13,202 QALYs) are much greater than what might be potentially gained from resolving the uncertainty within each scenario (1,765 QALYs). Therefore, unlike CLOP, most of what might be gained from further evidence about EECP (in the absence of formal elicitation) would be evidence that could help distinguish between the scenarios rather than the parameters associated with each.

3.5   Do the benefits of research exceed the costs?

The judgements made at points 5 and 6 of the check list are critical because if the benefits of research are not judged to exceed the costs then no further assessment are required (unless there are significant irrecoverable costs, see Table B in Appendix B). If they are and research can be conducted with approval, then AWR would be appropriate. However, other sources of uncertainty need to be assessed first, as they will influence the potential benefits of research and, even when research is not conducted, they will also influence the appropriate category of guidance when there are significant irrecoverable costs.

3.5.1   Point 5 – Will other sources of uncertainty resolve over time?

The fifth point on the check list requires an assessment of whether changes are likely to occur in the future which will influence the cost-effectiveness of the alternative technologies and the potential benefits of research, i.e., at the following point in the algorithm:
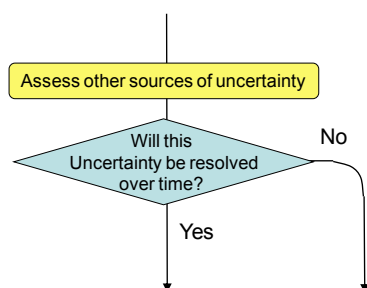
The judgement made at this point will influence the potential benefits of research and therefore subsequent decisions which lead directly to a particular category of guidance (see point 6 in Section 3.5.2 below). Even when research was not considered worthwhile (at point 3) the presence of other sources of uncertainty will determine whether significant irrecoverable costs are likely to influence the category of guidance. In some circumstances it can lead directly to guidance, i.e., if there are no other sources of uncertainty even significant irrecoverable costs will have no influence and a technology which is expected to be cost-effective can be approved:

	Assessment	1	2	3	4	5	6	7	Guidance
Pathway number	29	Yes	Yes	No	-	No	-	-	Approve 12

This assessment requires information about: i) changes in prices of the technology and its comparators; ii) the emergence of new technologies which might make existing ones obsolete or change their cost-effectiveness; and iii) other relevant research reporting. A number of potential sources of information and evidence were examined to inform this assessment for each case study[1]. However, many potentially useful sources were either proprietary or public access was restricted, making it surprisingly difficult to inform these assessments with publicly available information. When information and estimates were available they were often not complete or directly relevant to a UK context.

i)   Changes in the price (the technology and its comparators)
Changes in prices not only influence expected cost-effectiveness but also uncertainty and the potential benefits of research to future patients, e.g., if the price of a technology expected to be cost-effective is likely to fall significantly just before research reports the potential benefits will not be realised because approval of the technology will be less uncertain and there may be much less or little to gain from the results of the research. This assessment requires information about when major changes in prices are likely and some evidence about the likely extent of the change. A major event in the life cycle of a pharmaceutical technology is the date at which the patent expires and cheaper generic versions of the brand become available. Although the date of patent expiry is, of course known, it is surprisingly difficult to obtain the relevant date for particular products in the UK from publicly available sources. Evidence of the extent to which the price of generic versions are below the original brand price are also difficult to obtain and are likely to differ by health care system, type of technology, indication and time since patent expiry. Therefore, the estimate, reported by the Office of Fair Trading, that on average generic prices tend to be 25% of the original price was used in the subsequent analysis.

At the time of TA80 the patent for CLOP was expected to expire 7 years later and subsequent analysis assumes that at that time equivalent generic prices will be 25% of the original price of CLOP at the time of TA80.[31]   Although it was possible for the PsA case study to find patent expiry dates for Etanercept (Enbrel), Infliximab (Remicade) and Adalimumab (Humira) in the US (2012, 2014 and 2017 respectively), they were not available for the UK on the National Patent Database (IPO). It is even more difficult to locate patent information relevant to devices, such as EECP, since a device may only have a CE mark, which, unlike a patent, does not offer protection but can be renewed every

---

[31] This assumes that either prescribing will switch from the brand to equivalent generic (brands tend to maintain, or even increase premium prices in some health care systems, after patent expiry) or that any new branded technologies will be appraised and/or priced using generic versions of the old brand as a comparator.

10 years. Any patent is likely to relate to some aspect of the device rather than the device itself. Although, prices may change over time they can also be relatively stable but with incremental innovation of the original device. Again this is likely to differ by health care system, technology and indication. For these reasons future changes in prices are only quantitatively explored in subsequent analysis in CLOP in Section 3.5.2. There is a need to consider how access to the type of information required during NICE appraisal can be provided and how estimates of likely changes in prices relevant to the UK can be made readily available, if these assessments are to be routinely made.

ii)   Entry of new technologies

The entry of a new technology may make the existing technology that is expected to be cost-effective obsolete (no longer the most cost-effective alternative). Even when it does not, it will tend to change the relative cost-effectiveness of the alternatives, influencing how uncertain a decision to approve the original technology will be for future patients and the potential gains from research. A number of potential sources of information were examined to identify new technologies relevant to the indications which were likely to become available. These included a variety of sources related to NICE topic selection, information about licence applications, clinical research in phase I, II and III as well as evidence of the probability that earlier phase research leads to entry (probability of successful licence) and the likely time of entry (time to launch from initiating phase I, II and III research). Again this information and evidence is fragmented, and in some cases restricted, e.g., NHS Horizon Scanning. Nevertheless, the information that was available indicated that one new technology relevant to CLOP and one relevant to PsA might have been expected to enter. Information about these technologies was limited so scenarios are used to explore the implication for CLOP and PsA in Section 3.5.2.

iii)   Other research reporting (the technology and its comparators)

Research which is already underway, commissioned or likely to be undertaken whether in the UK or elsewhere, is relevant for two reasons. Firstly, if it is research based in the UK then guidance might impact on recruitment and the successful completion of this research (see Section 3.6). Secondly, when this research reports there is a chance that it will change the estimates of cost-effectiveness and resolve some of the current uncertainties. In other words, there is little to be gained by recommending OIR or AWR if the uncertainty is likely to be resolved in the near future when other research reports. A number of potential sources of information were examined to identify clinical research underway at the time of the relevant appraisal, including: national and international trial registries; as well as other databases which report NHS funded research and not just clinical trials (e.g., NRR and UKCRN). Despite an assiduous search no records relevant to the case studies were identified. This may suggest that no other research was ongoing or expected for these comparators in these indications, or it may indicate that currently available sources are fragmented, incomplete and/or difficult to access.

3.5.2   Point 6 – Are the benefits of research greater than the costs?

The sixth point on the checklist requires a re-assessment of the potential benefits of conducting further research which were initially considered at point 3 (see Section 3.4.1), and a judgment of whether the benefits of research are likely to exceed the costs i.e., at the following point in the algorithm:
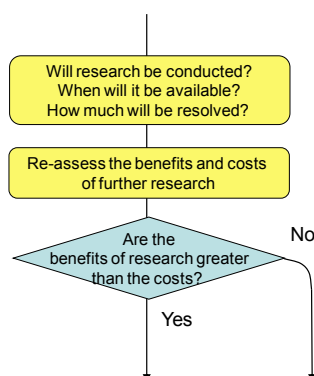
A judgment about whether the potential benefits of research identified in Section 3.4 will be realised requires an assessment of: i) whether the type of research that is required is likely to be conducted; ii)

if conducted, when the results are likely to be available; iii) how much uncertainty is likely to be resolved and iv) the likely impact of the other sources of uncertainty identified in Section 3.5.1 on the longer term benefits of research.

The decision at this point may not necessarily lead directly guidance, e.g., where the benefits of research exceed the costs but research is not possible with approval or there are significant irrecoverable costs. Which category of guidance will ultimately be appropriate will depend on whether the benefits of approval are judged to exceed the costs, i.e., point 7 of the checklist in Section 3.6. However, in many other circumstances the decision at this point will lead directly to a particular category of guidance. These circumstances or pathways through the algorithm are detailed below (extracted from Table B1 in Appendix B):

Assessment	1	2	3	4	5	6	7	Guidance
1	Yes	No	Yes	Yes	Yes/No	Yes	-	AWR 1
2	Yes	No	Yes	Yes	Yes/No	No	-	Approve 1
5	Yes	No	Yes	No	Yes/No	No	-	Approve3
7	No	No	Yes	Yes	Yes/No	Yes	-	OIR 2
8	No	No	Yes	Yes	Yes/No	No	-	Reject 1
11	No	No	Yes	No	Yes/No	No	-	Reject 3
19	Yes	Yes	Yes	Yes	No	No	-	Approve 6
26	Yes	Yes	Yes	No	No	No	-	Approve 10
30	No	Yes	Yes	Yes	Yes/No	Yes	-	OIR 7
31	No	Yes	Yes	Yes	Yes/No	No	-	Reject 8
34	No	Yes	Yes	No	Yes/No	No	-	Reject 10

(Pathway number)

The expected consequences of uncertainty reported in Section 3.4.1 represented the NHE that could be gained over the lifetime of the technology if the uncertainty surrounding the decision based on expected cost-effectiveness could be immediately and completely resolved. This represents an upper bound on the potential benefits of research for a number of reasons: i) research, although recommended, might not be commissioned and/or recruit and report; ii) any research will take some time to complete before results are available; and iii) not all of the uncertainty is likely to be resolved. In addition, future events (identified in Section 3.5.1) might change the NHE expected to be gained by future patient populations. Finally, the expected benefits of research once properly re-assessed must be compared to the likely costs.

i)   Will the research be conducted?
Even if research is recommended in OIR or AWR, it might not be undertaken by manufacturers or commissioned by research funders. Even if undertaken or commissioned, there is no guarantee that research will be able to recruit or it may not complete for other reasons. The expected consequences of uncertainty for CLOP and EECP reported in Section 3.4.1 are illustrated for a range of probabilities that research will be successfully undertaken in Figures 3.6a and 3.6b respectively. This indicates that the potential gains depend on a judgment of whether the research recommended as part of OIR or AWR will be successfully completed. They also illustrate that the cost of research (in this case considered to be either £1.5m or £10m[32]) can be compared directly to the potential benefits by either: i) expressing the potential gains in population NHE as the equivalent NHS resources, i.e., the resources that would be required to generate the same NHE; or ii) expressing the cost of research in terms of the QALYs that could be gained elsewhere in the NHS by using the same resources to provide access to health care.

---

[32] Based on the range of costs of trials commissioned by NCCHTA – personal communication.
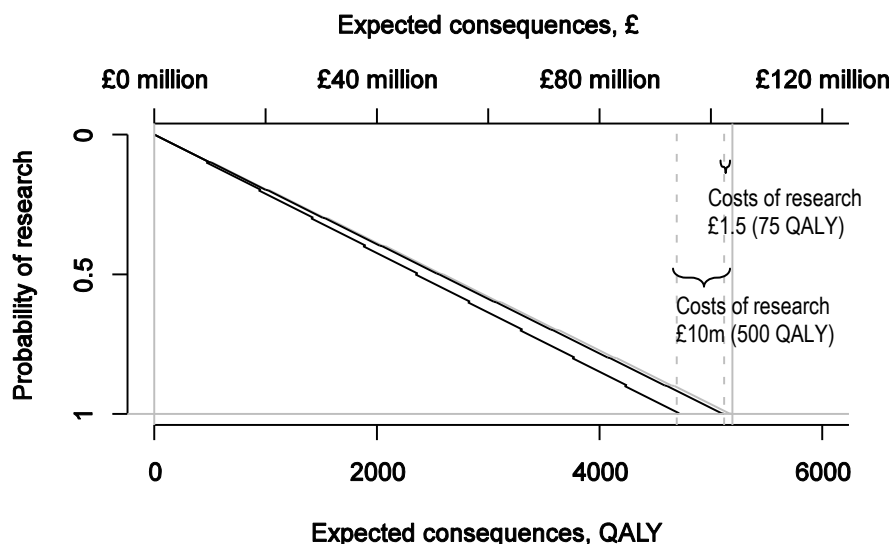
Figure 3.6a   Expected potential benefits of research (CLOP)

ii)   When will it be available?

Research, even if commissioned and successfully completed, will take time to complete and report. Therefore, any assessment of the potential benefits should account for the fact that patient populations will not benefit from the results of research until they are available. Whether those patients who are prevalent while research is underway will be able to benefit from the results will depend on whether treatment decisions for presenting patients are irreversible or not (see Section 3.3.2). If treatment decisions are irreversible, e.g., in CLOP it is only those patients' incident after the research reports that will realise any of the potential benefits. In contrast, treatment decisions in EECP are not irreversible (it is a chronic condition), so although patients prevalent while research is undertaken will not benefit immediately, those that survive can benefit from the results once the research is completed. How long research might take to report will depend in part on the design (follow-up, sample size and endpoints), recruitment rates and size of the eligible patient population, as well as how efficient the organisation and data collection might be. The potential value of research in CLOP and EECP over a range of possible time horizons is reported in Figures 3.7a and 3.7b respectively. In both cases the potential value of further research declines with the time to research reporting. This relationship gives some indication of the value of improving the timeliness of research though, for example, investment in research infrastructure or adopting a research design, which, although offering less potential benefits, can be conducted more quickly.
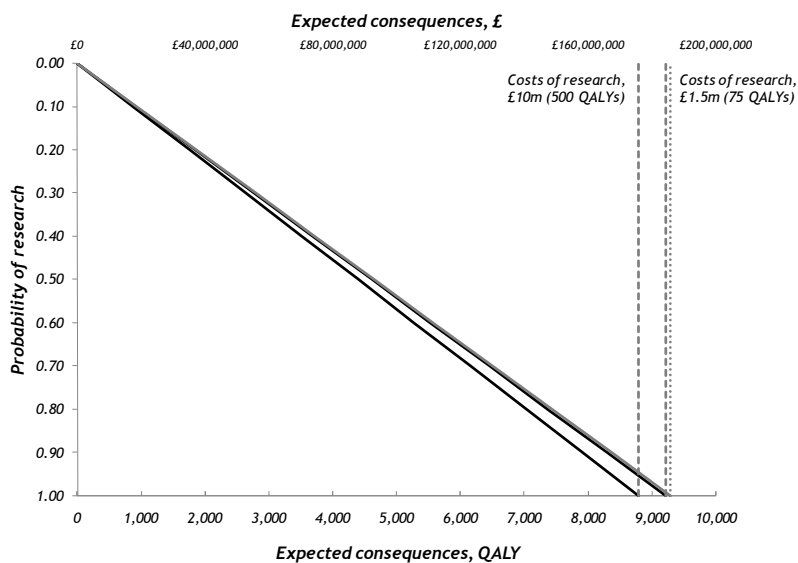

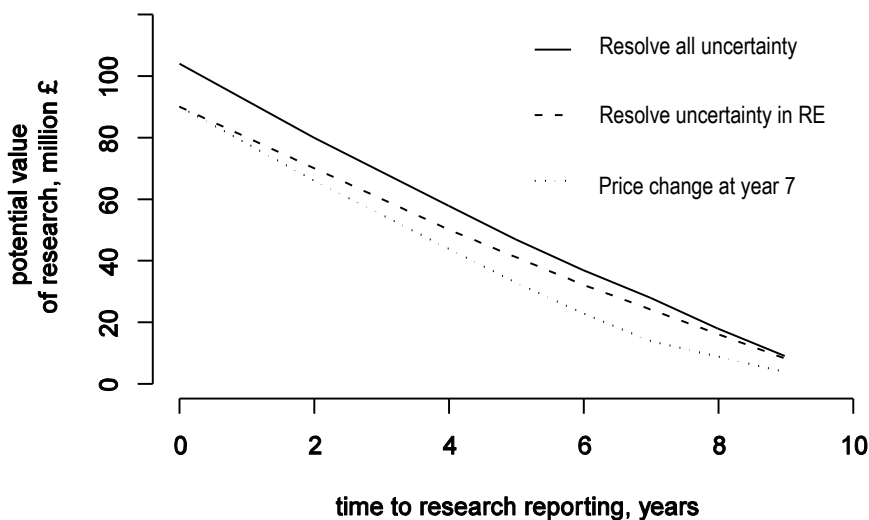**Figure 3.6b   Expected potential benefits of research (EECP)**

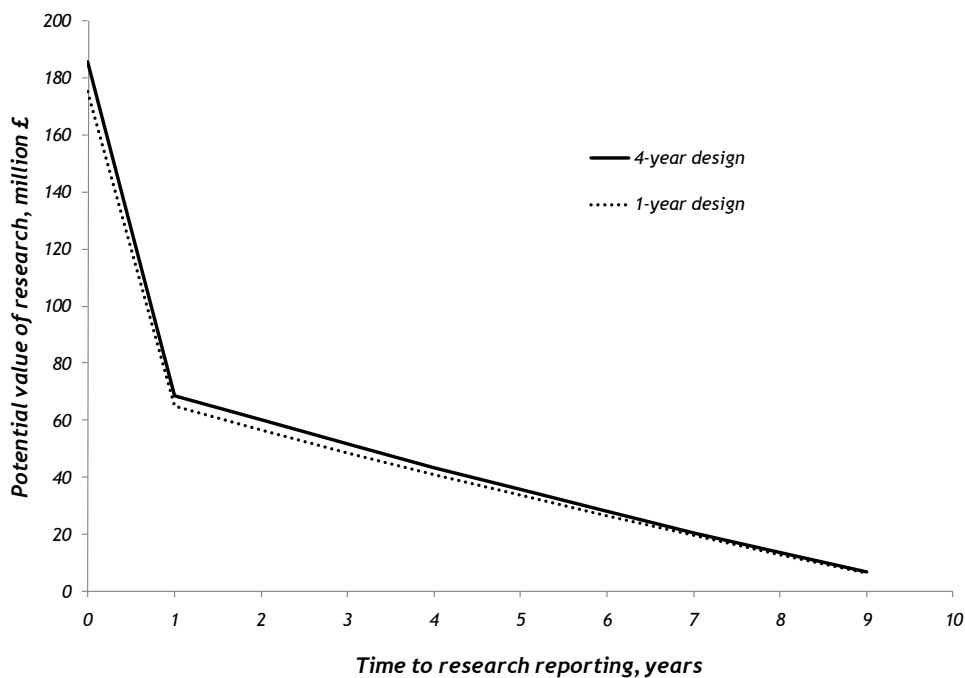Figure 3.7a   Potential value of research and time to report (CLOP)

Figure 3.7b   Potential benefits of research and time to report (EECP)

iii)   How much will be resolved?

Most research will not inform all the parameters that determine expected cost and QALYs but usually a subset of them. Therefore, the potential benefits of research that might be conducted will not be the total expected costs of uncertainty surrounding expected cost-effectiveness, but some part of it. In Section 3.4.3 the potential benefits of different types of evidence was assessed. In CLOP it was additional evidence about relative treatment effects that were most valuable and therefore experimental research may be required to provide a more precise estimate of RR_Death.   The potential value of research, which only resolved uncertainty about this relative treatment effect over a range of times to report, is also represented in Figure 3.7a (denoted by the legend 'resolve uncertainty in RE'). Although the potential value of research is lower at every time point, unless

research is likely to take more than 8 years, the potential value is still likely to exceed the costs. In EECP there was most benefit to be gained by resolving the uncertainty in the improvement in quality of life at 12 months, which in common with CLOP, is likely to require experimental design. Figure 3.6b represents the potential benefits of alternative trial designs with either one or four years of follow-up (1, 2, 3 and 4 year follow-up designs were evaluated)[1]. Although longer follow-up offers greater potential benefits they are relatively small compared to the loss of potential value if longer follow-up delays the time until research findings are available, i.e., a 4 year design will require a minimum of 4 years to complete. Again, as long as research reports before 8 years, the potential benefits are likely to exceed the costs

iv)   What is the impact of other source of uncertainty?

In Section 3.5.1 the information that was publicly available identified that the patent for CLOP was due to expire 7 years after the appraisal. Based on the OFT estimate that generic prices tend to be 25% of the original brand price, this other source of uncertainty can be integrated quantitatively when estimating the potential value of research over the life time of the technology. In this case a significant fall in price in year 7 will substantially reduce the uncertainty surrounding 12 months of treatment with CLOP. Therefore, after year 7 there is less to be gained from resolving uncertainty, so the potential and value of research findings for patients' incident after year 7 are thereby reduced. The effect of a price change on research which could potentially resolve uncertainty in cost natural history, and utilities as well as relative effect is also illustrated in Figure 3.7a. The potential value of the research is lower whenever the research reports, because it includes the value to future as well as current patient populations. Nevertheless, even if research didn't report until 7 years the potential value is likely to exceed the costs (see Figure 3.7a). The expected price reduction reduces the potential value of research at each time point for both scenarios, e.g., for scenario B from 174,519 QALYs when research is immediately available (see Figure 3.8a) to 165,701 with a price change at year 7.[33]

There was some evidence of possible entry of a new technology (comparator) in the indication described in the CLOP case study. However, there was limited information on its characteristics. Therefore two alternative but somewhat extreme scenarios are illustrated in Figure 3.8a. In scenario A the new technology enters at years 5 and makes CLOP entirely obsolete, i.e., not cost-effective and not uncertain (equivalent to a shorter technology time horizon of 5 years). At this point there is no value in the evidence generated by research about CLOP.[34] In these circumstances the potential value of research is only likely to exceed its costs if it reports quickly. In scenario B the new technology has similar NHE to 12 months of treatment with CLOP[35] and the uncertainty surrounding its expected cost-effectiveness is also similar. Now research about CLOP has more potential value in the future since it will also help resolve some of the uncertainty in the choice between CLOP and the new technology for patients that become incident after that time.   Although there was no evidence of new technologies emerging in EECP, the same scenarios are explored as the development and launch of new devices are more difficult to identify in advance. The impact on the potential value of research is illustrated in Figure 3.8b and demonstrates similar qualitative effects as CLOP. In scenario A (EECP becomes obsolete) the potential benefits of further research about EECP are only likely to exceed the costs if the research reports quickly. Nevertheless, even in this extreme scenario the benefits of research with only 1 year follow-up are likely to exceed the costs so long as it reports before 4 years.

The potential value of research presented in these figures, even after accounting for the type of evidence, follow-up and time until research reports, should still be regarded as an upper bound to the value that is likely to be realised by actual research for two reasons: i) even well designed research with large sample sizes will not fully resolve the uncertainty in the value that a parameter might take, especially in specific target populations and in a particular (future) context; and ii) insofar as implementation of NICE guidance is not 'perfect' and all clinical practice might not immediately respond to the results of research, the full benefits will only be realised over time or with additional implementation efforts. For these reasons a judgment of whether benefits of research are likely to

---

[33] These much higher values of immediate research than the 4,495 QALYs or £89.9m in Figure 3.7a without the entry of a new technology but with a similar price change.

[34] There may continue to be value if evidence about CLOP remains an important link in mixed or indirect treatment comparisons required to evaluate the new technology.

[35] This is likely to be an increasingly common scenario if value based pricing effectively makes all branded technologies equally cost-effective.

exceed the costs might be made conservatively, requiring evidence that, even in pessimistic scenarios, the research would still be worthwhile.
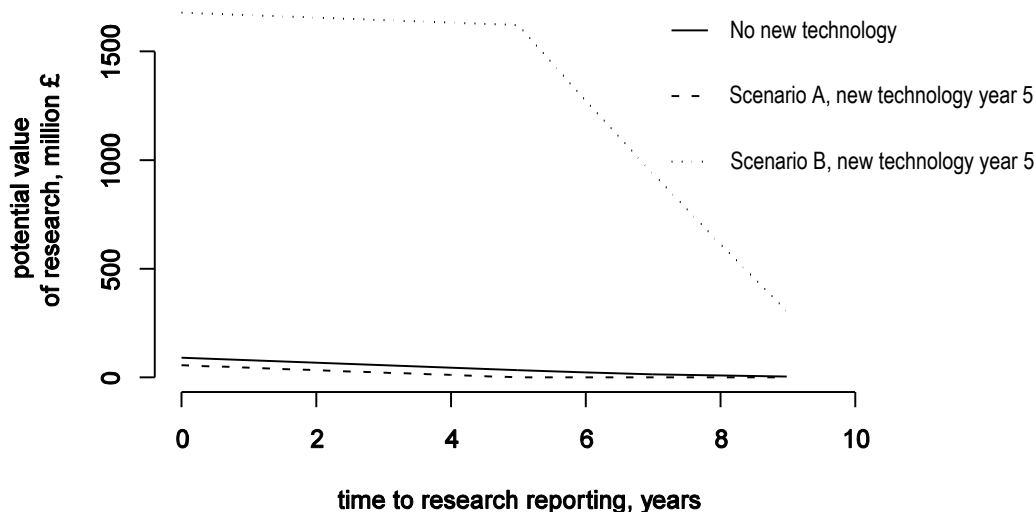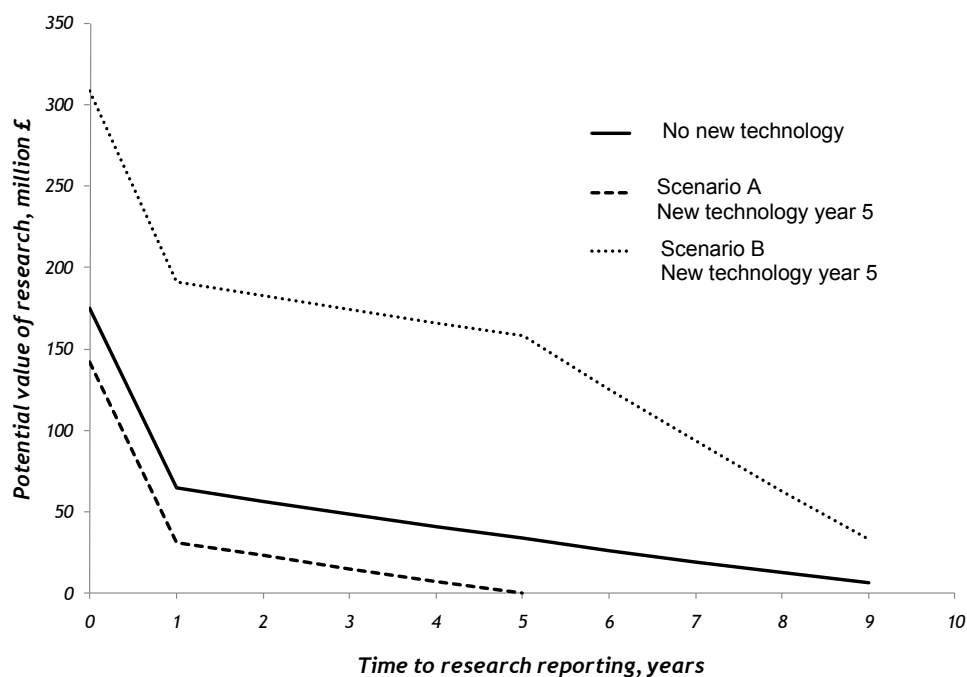

Figure 3.8a   Potential value of research and other sources of uncertainty (CLOP)


* These potential values are based on a 1 year follow-up design

Figure 3.8b   Potential value of research and other sources of uncertainty (EECP)*

3.6   Point 7 – Are the benefits of approval greater than the costs

The seventh and final point on the checklist requires an assessment and comparison of the benefits and costs of early approval. The costs of approval, are not financial ones, but opportunity costs, and will include the potential value of any research that may be forgone as a consequence, e.g., if the research needed cannot be conducted once the technology is approved for use. It will also include any costs that are irrecoverably committed by approval. As well as the capital costs of equipment and facilities (or training and learning), they will also include the irrecoverable opportunity costs of initially negative NHE (if treatment decisions are not irreversible - see Section 3.3.2). A judgment of whether the benefits of approval and early access for current patients are likely to exceed the opportunity costs for future patients is required i.e., at the following point in the algorithm:
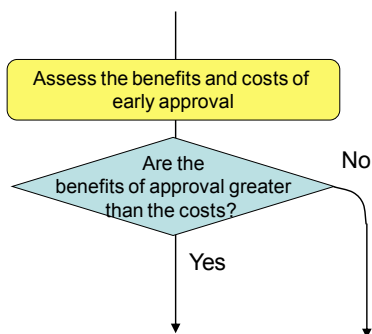
The decision at this point always leads directly to guidance; allocating all remaining possible pathways to a particular type and category of guidance. These remaining (20) pathways through the algorithm are detailed below (extracted from Table B1 in Appendix B):

Assessment	1	2	3	4	5	6	7	Guidance
3	Yes	No	Yes	No	Yes/No	Yes	Yes	Approve 2
4	Yes	No	Yes	No	Yes/No	Yes	No	OIR 1
9	No	No	Yes	No	Yes/No	Yes	Yes	AWR 2
10	No	No	Yes	No	Yes/No	Yes	No	Reject 2
13	Yes	Yes	Yes	Yes	Yes	Yes	Yes	AWR 3
14	Yes	Yes	Yes	Yes	Yes	Yes	No	OIR 3
15	Yes	Yes	Yes	Yes	Yes	No	Yes	Approve 5
16	Yes	Yes	Yes	Yes	Yes	No	No	Reject 5
17	Yes	Yes	Yes	Yes	No	Yes	Yes	AWR 4
18	Yes	Yes	Yes	Yes	No	Yes	No	OIR 4
20	Yes	Yes	Yes	No	Yes	Yes	Yes	Approve 7
21	Yes	Yes	Yes	No	Yes	Yes	No	OIR 5
22	Yes	Yes	Yes	No	Yes	No	Yes	Approve 8
23	Yes	Yes	Yes	No	Yes	No	No	Reject 6
24	Yes	Yes	Yes	No	No	Yes	Yes	Approve 9
25	Yes	Yes	Yes	No	No	Yes	No	OIR 6
27	Yes	Yes	No	n/a	Yes	n/a	Yes	Approve 11
28	Yes	Yes	No	n/a	Yes	n/a	No	Reject 7
32	No	Yes	Yes	No	Yes/No	Yes	Yes	AWR 5
33	No	Yes	Yes	No	Yes/No	Yes	No	Reject 9

3.6.1   Technologies without significant irrecoverable costs

Only four of the 20 possible pathways illustrated above are associated with technologies without significant irrecoverable costs. In these four pathways, research was either: i) not considered possible *with* approval for those expected to be cost-effective (i.e., Approve[2] or OIR[1]); or ii) research was not possible *without* approval for those not expected to be cost-effective (i.e., AWR[2] or Reject [2]). CLOP provides an example of the former. It is research that would provide more precise estimates of the relative effect of CLOP and of shorter treatment durations, which is potentially valuable (see Section 3.4.2). As a consequence, the type of experimental design that is likely to be needed is unlikely to be possible if 12 months of treatment with CLOP is already approved for widespread NHS use. Although

treatment with CLOP does commit initially negative NHE that are irrecoverable, these should not be regarded as significant since the treatment decision for a presenting patient is irreversible in relevant time frames (see Section 3.3.2).   Therefore, AWR may not be possible, so the benefits of early access to 12 months of treatment with CLOP (Approval) must be compared to the potential value of OIR.

OIR is more likely to offer greater expected NHE than Approve if the research can be conducted quickly and report sooner, since fewer patients forgo their access to CLOP and more can have treatment choice informed by the research findings. This is illustrated in Figure 3.9 which reports the difference between Approve and OIR in population NHE over a range of times for when the research recommended in OIR might report. This takes account of both the expected changes in price at year 7 and research costs of £10m. It shows that OIR will only be appropriate if the research reports within 3 years of appraisal (T* = 3) because beyond this time the NHE forgone by withholding access to CLOP will exceed the potential gains to future patients.
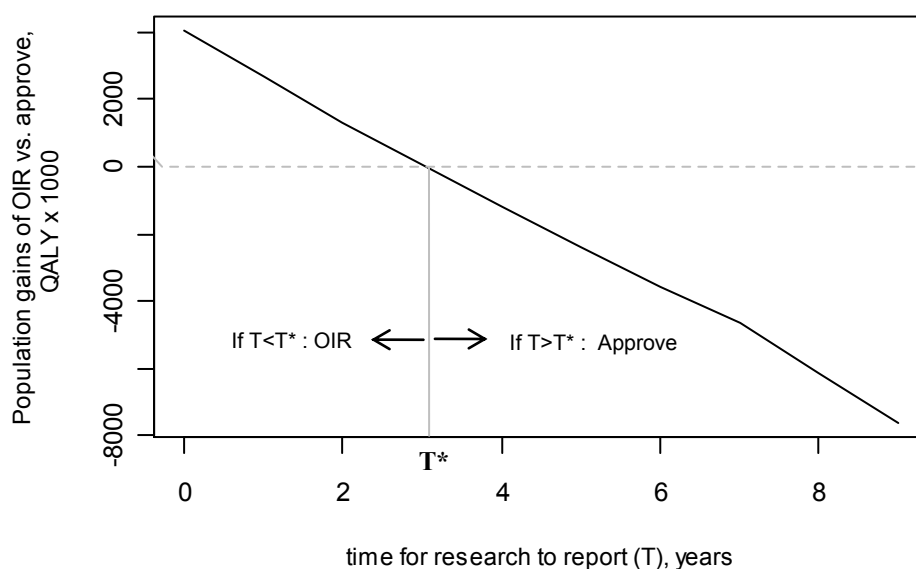

Figure 3.9   Population NHE of Approve and OIR for time to research reporting (CLOP)

The trade-off between NHE for current and future patients which lies behind Figure 3.9 is illustrated in Figure 3.10 using undiscounted values for ease of exposition. It illustrates the (per period) population NHE of approval and OIR, if the research recommended as part of OIR reports at year 3. At this point, the initial losses of NHE, caused by restricting access to CLOP (area A), start to be offset by the potential gains from the research findings (area B). The price change at year 7 increases the NHE of approval (i.e., CLOP is more cost-effective) but on balance reduces the NHE of OIR. In other words, since CLOP is more cost-effective and offers greater NHE the evidence generated by the research is less valuable because the choice of treatment and duration is less uncertain (see sections 3.5.1 and 3.5.2). With research reporting at 3 years the initial losses of OIR (area A) are just offset by the later gains (area B), so T* = 3.   If research reported earlier than 3 years (area A > area B) and OIR would be appropriate but if later than 3 years (area A > area B) and Approve would be more appropriate.

However, there is no guarantee that the research recommended as part of OIR guidance will be conducted by manufacturers or commissioned by research funders. Even if it is, it is not certain that it will be successfully completed (see discussion in Section 3.5.2).   Therefore, the probability that research will report at a particular time also needs to be considered. The implications of considering whether the recommended research will be conducted and when it might report are illustrated in Figure 3.11, which presents a boundary for when OIR might be appropriate or when approval should be granted. For example, if research is certain to report but will take 4 years, or when it will only take 1 year but with only a 50% chance of reporting, then OIR would not be appropriate and 12 month treatment of CLOP should be approved, i.e., points that fall to the north east of this boundary.   Points to the south west of the boundary indicate that OIR might be appropriate.
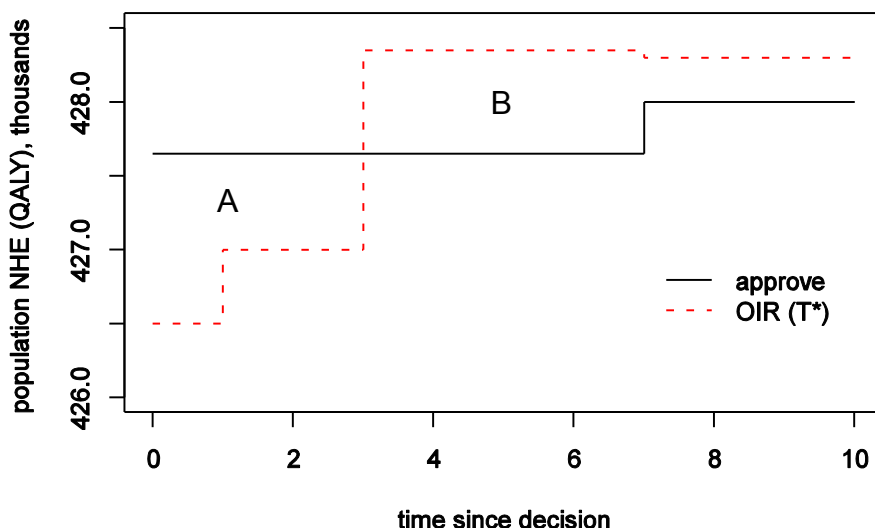
Figure 3.10   Population NHE of Approve and OIR at T* (CLOP)
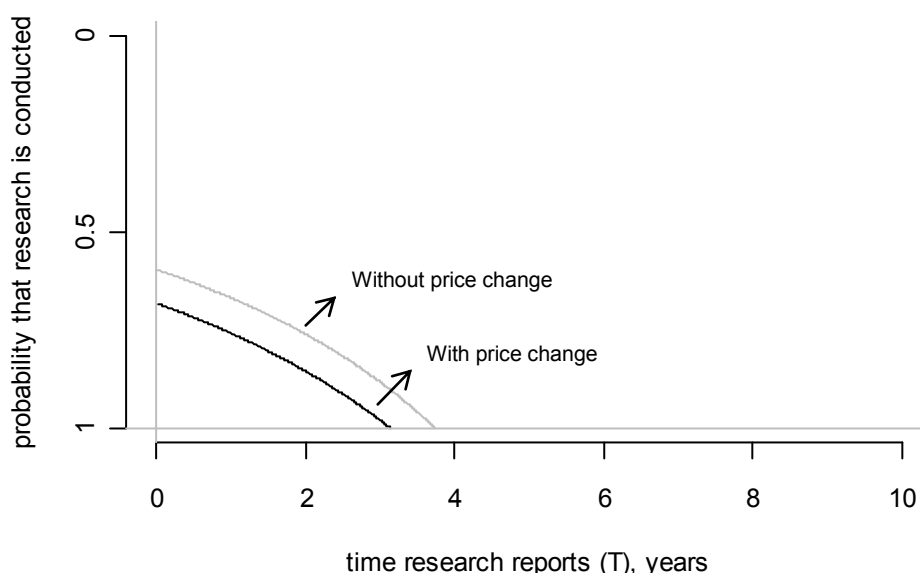

Figure 3.11   An OIR or Approve boundary (CLOP)

However, the estimates of the potential value of additional evidence on which these boundaries are based are still likely to overestimate the value that will be realised by research so they represent a necessary condition for OIR. Therefore, OIR guidance should require a conservative judgment that the point is almost certain to be below the boundary, rather than on balance close to it. For the same reason, points anywhere above the boundary represent a sufficient condition for approval. The boundary when the change in price is included is to the south west; reflecting the lower potential value of research, and OIR guidance, once CLOP becomes more cost-effective.   In this case it seems unlikely that the type of research required could report quickly enough and with sufficient confidence that OIR would be appropriate. Therefore, these assessments would support a judgment that the benefits of approval are likely to exceed the opportunity costs, and Approve[2] (pathway 3 for CLOP in Table B1, Appendix B) would seem more appropriate.

The assessments that have been undertaken for CLOP can be brought together to consider: i) what would be the value of being able to conduct research while CLOP is approved (value of AWR); and ii) what would be the value of making evidence that is needed by the NHS available at launch. These

questions can be informed by the results (already presented elsewhere) but also reported together in Table 3.7a.

Table 3.7a   **Population NHE over the technology time horizon for different policies (CLOP)**

	Approve	OIR	AWR*	Reject	Value of AWR	Uncertainty resolved at launch	Value of evidence at launch
	Expressed in QALY						
T<T* (T=2)	3,680,187	3,681,480	3,682,995	3,671,660	1,515	3,684,181	2,701
T>T* (T=7)	3,680,187	3,675,487	3,680,362	3,671,660	175	3,684,181	3,994
	NHE expressed in £m						
T<T* (T=2)	73,604	73,630	73,660	73,433	30	73,684	54
T>T* (T=7)	73,604	73,510	73,607	73,433	4	73,684	80

* This is the expected population NHE if AWR was to become a possibility

The difference in population NHE between AWR (if it had been possible) and the next best feasible policy (OIR when T<T* and Approve when T>T*) is £30m and £4m respectively and represents the value to the NHS of being able to conduct research while CLOP is approved for use, e.g., informing whether investment in better data collection, registries or information systems that might make this possible.[36]   The difference in population NHE if all uncertainty had been resolved prior to appraisal (at launch) and the next best available policy (OIR when T<T* and Approve when T>T*) is £54m and £80m respectively and represents the value to the NHS of having access to the evidence needed at launch. This can inform policies which might make better and more relevant evidence available.

It is also possible to consider the commercial as well as NHS value in each of the cells of this table. The value of early evidence at launch can then be considered from the perspective of the manufacturer (the expected revenue streams), taking account of prices (see Sections 2.2.1 and 2.2.2) and expected volumes over the remaining patent life and technology time horizon. Together with estimates of the costs of conducting research by manufacturers or through public funding, this assessment might inform when manufacturers might be expected to conduct the research needed (high commercial value that exceeds the cost to manufacturers) and when the NHS might be expected to undertake it (low commercial value but high potential value to the NHS that exceeds the costs to the NHS). In many circumstances both the commercial and NHS values will exceed their respective costs. In these circumstances the question of who should conduct, or pay for, the research might be informed by which sector has the comparative advantage, i.e., which has the highest 'relative efficiency' in generating social value? Of course, the value to the NHS and to manufacturers will depend; to a large extent, on what type of flexible pricing arrangements and value based pricing scheme might be in place (see discussion in Section 2.2.1 and 2.2.2). The question will also turn on how any agreements can be made and incentive consistent contracts written and enforced.

3.6.2   *Technologies with significant irrecoverable costs*

Most of the possible pathways illustrated above are associated with technologies with significant irrecoverable costs (16 out of the remaining 20).   This is because even when research is possible with approval (or even when not needed), the impact of committing irrecoverable cost through AWR (or approval) must be considered, so OIR (or reject) remains a possibility. EECP provides an example of this; where research that would provide more precise estimates of the effect of treatment on quality of life accounts for all the potential value (see Section 3.4.2). EECP does commit both capital costs associated with long lived equipment, as well as initially negative per patient NHE. Unlike CLOP these irrecoverable opportunity costs at a patient level are significant because treatment choice for a presenting patient is not irreversible over relevant time frames (see Section 3.3.2). As a consequence, even if research is possible with approval it is not clear that AWR would be appropriate, because OIR avoids the commitment of irrecoverable costs until research findings are available and a more informed decision can be made.

i)   **Research is possible with approval**
Even when research is possible with approval OIR offers greater expected NHE than AWR as long as research reports before 9 years in Figure 3.12a. This is because the consequences (losses of population NHE) of committing both aspects of irrecoverable costs through AWR are greater than the

---

[36] Even with such investment AWR might not be possible if there is insufficient variation in treatment assignment and no robust way of controlling for unobserved characteristics through selection models, e.g., use of instrumental variables.

NHE forgone by restricting access to EECP through OIR. The costs of research have not been included because they are incurred with both AWR and OIR guidance.[37]

As previously for CLOP, there is no guarantee that the research recommended as part of OIR or AWR guidance will be conducted and research report. A boundary for when OIR rather than AWR might be appropriate is illustrated in Figure 3.13a for four research designs with differing follow-up. A one year follow-up will generate evidence with the lowest potential value (so the boundary is to the south west) but it is likely to report sooner. Therefore, OIR might be appropriate even if the probability that the research will be conducted and report is relatively low. In this case it seems likely that the type of research required could report quickly enough and with sufficient confidence that OIR would be appropriate even though the research could be conducted while EECP is approved. Therefore, these assessments would support a judgment that the benefits of approval (through AWR) are unlikely to exceed the opportunity costs (the NHE of OIR), so OIR[4] (pathway 18 in Table B1, Appendix B) rather than AWR[4] (pathway 17) would seem more appropriate.
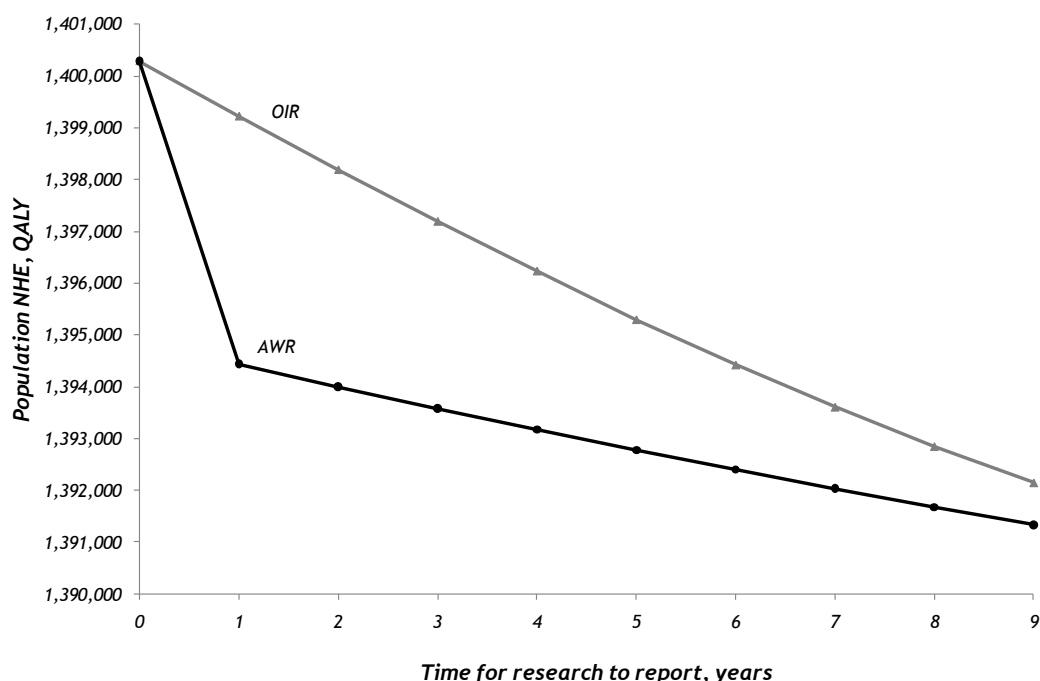

Figure 3.12a   Population NHE of AWR and OIR for time to research reporting (EECP)

ii)   Research is not possible with approval

For the general reasons discussed in Section 3 and those specific to EECP discussed in Section 3.4.2, the type of experimental research required to robustly estimate the effect of EECP on quality of life is unlikely to be possible once it is approved and in widespread use. Now approval (now through Approve rather than AWR) not only commits the type of irrecoverable costs discussed above it also means that the potential value of evidence to future patients must also be forgone. This is reflected in Figure 3.12b where the difference between OIR and approve are always greater than between OIR and AWR in Figure 3.12a. It suggests that as long as the cost of the research exceeds the difference between OIR and approve, when it is expected to report, OIR rather than approve would be appropriate. This is also reflected in the boundaries for OIR and Approve reported in Figure 3.13b. These boundaries are always to the north east of the OIR/AWR boundaries reported in Figure 3.13a, again reflecting the fact the approval not only commits irrecoverable costs but also forgoes the potential value evidence that might have been generated through an OIR recommendation. These assessments would support a judgment that the benefits of approval are unlikely to exceed the opportunity costs (the NHE of OIR), so OIR[6] (pathway 25 in Table B1, Appendix B) rather than Approve[9] (pathway 24) would be more appropriate.

---

[37] Any difference in costs of research under AWR or OIR guidance can easily be integrated into these assessments.
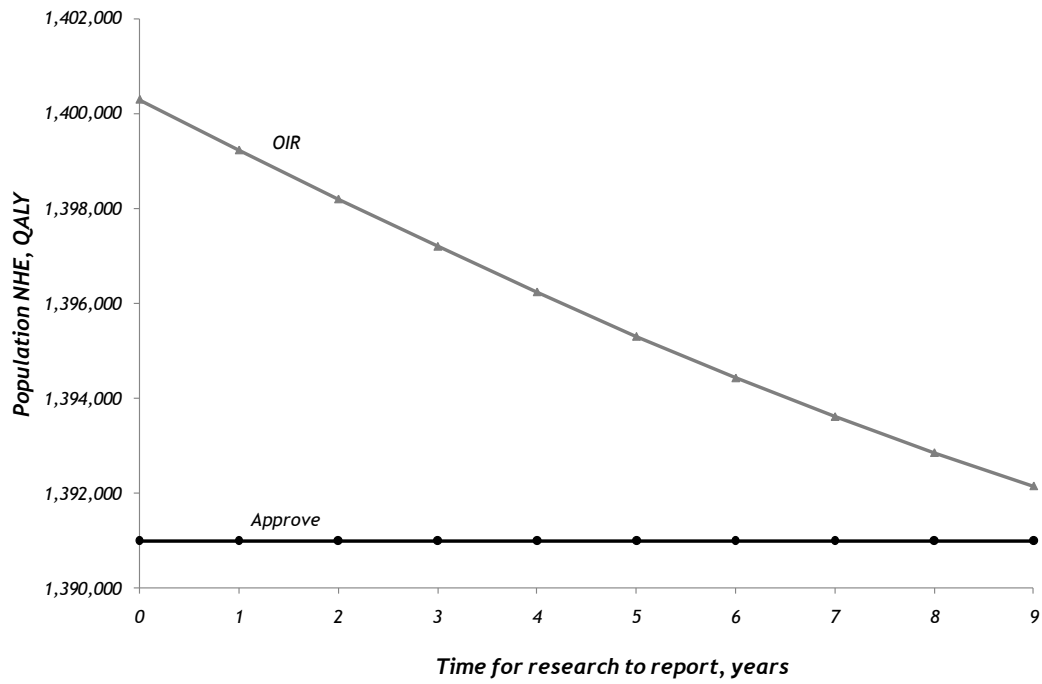
Figure 3.12b   Population NHE of Approve and OIR for time to research reporting (EECP)
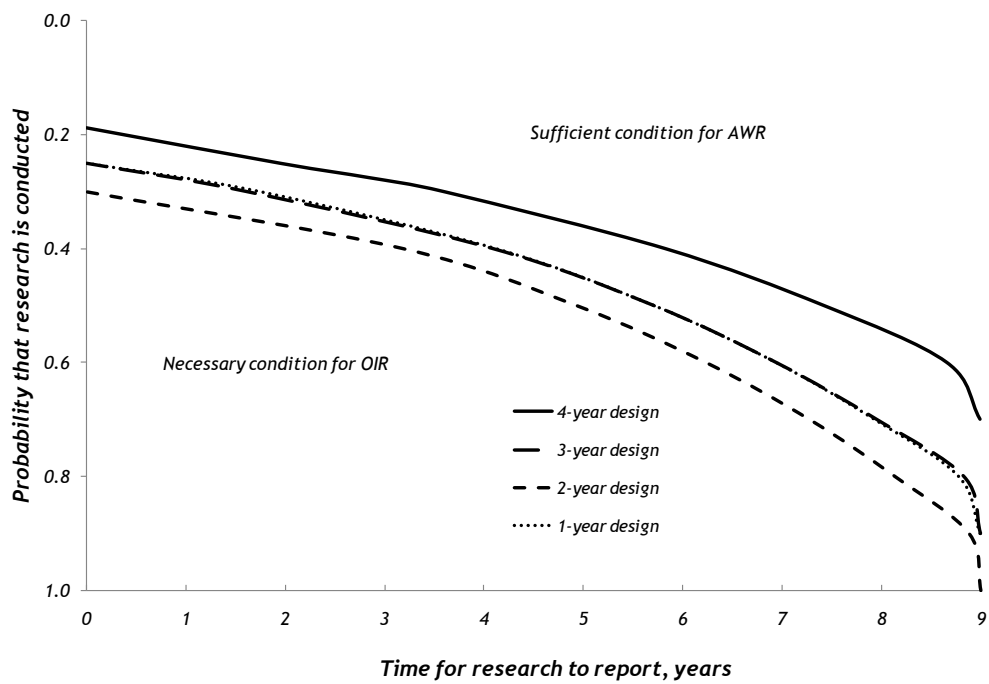

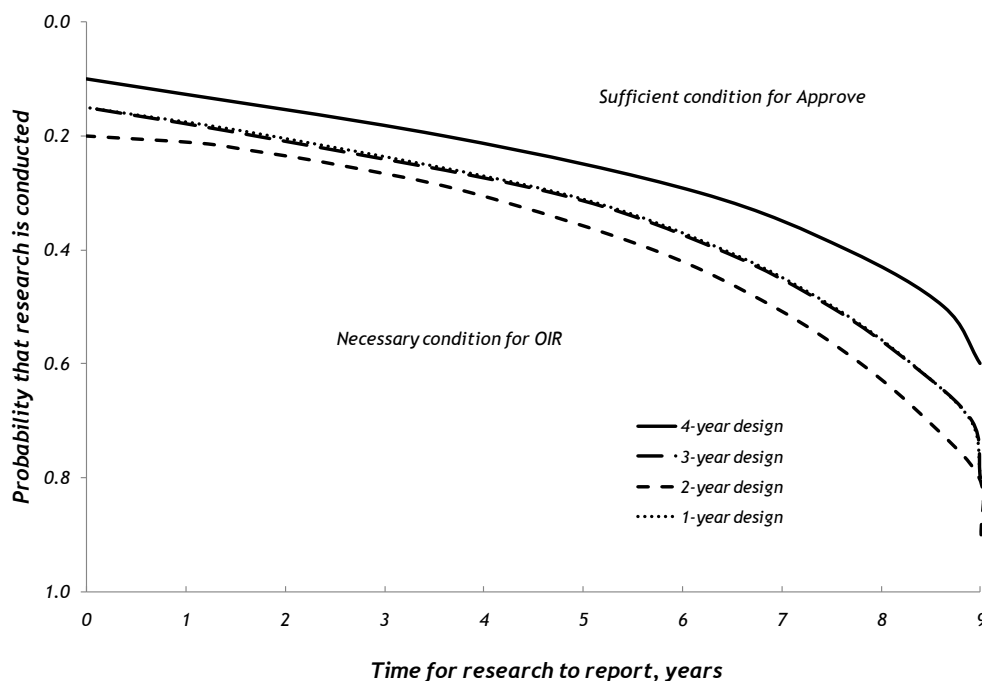**Figure 3.13a   An OIR or AWR boundary (EECP)**

Figure 3.13b   An OIR or Approve boundary (EECP)

As with CLOP, the assessments that have been undertaken for EECP are bought together in Table 3.7b and can help inform the same policy questions: i) what would be the value of being able to conduct research while EECP is approved and ii) what would be the value of making the evidence that is needed by the NHS available at launch. In this case, due to the irrecoverable costs associated with EECP, there is no value to the NHS of being able to conduct research while EECP is approved for use. In fact these figures are negative; indicating that even if AWR was possible it would not be appropriate. However, like CLOP, there is value to the NHS of having the evidence needed prior to appraisal. The value, expressed in the equivalent NHS resources, depends on how long it would otherwise have taken for an OIR recommendation to deliver the same evidence e.g., £62m if 3 years and £134m if 7 years. As with CLOP these assessments can also inform policies which might make better and more relevant evidence available and the question of how and who might contribute most to providing the evidence needed at the right time.

Table 3.7b   Population NHE over the technology time horizon for different policies (EECP)

	Approve	OIR	AWR	Reject	Value of AWR	Uncertainty resolved at launch	Value of evidence at launch
				Expressed in QALY			
T=3	1,391,001	**1,397,192**	1,393,578	1,389,596	-3,614	1,400,288	3,096
T=7	1,391,001	**1,393,608**	1,392,030	1,389,596	-1,578	1,400,288	6,680
				Expressed in £m			
T=3	27,820	**27,944**	27,872	27,792	-72	28,006	62
T=7	27,820	**27,872**	27,841	27,792	-32	28,006	134

4. Implications for policy, process and methods

A more explicit assessment of OIR and AWR has a number of implications for policy (e.g., NICE guidance and drug pricing), the process of appraisal (e.g., greater involvement of research commissioners) and methods of appraisal (e.g., should additional information, evidence and analysis be required).

Key principles and assessment needed
The key principles and assessments needed fall into four broad areas: i) expected cost-effectiveness and population net health effects (including benefits, harms and NHS costs); ii) the need for evidence and whether the type of research required can be conducted once a technology is approved for widespread use; iii) whether there are sources of uncertainty which cannot be resolved by research but only over time; and iv) whether there are significant (opportunity) costs which will be committed and cannot be recovered once the technology is approved.

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 any benefits of early approval are greater than withholding approval until additional research is conducted or other sources of uncertainty are resolved. There was general consensus among workshop participants that the key principles represented by a sequence of assessment and judgments were the relevant ones, complete and in an appropriate order. In response to feedback from the first workshop this sequence of assessment and judgement was summarised as a simple seven point checklist (see Section 3.1) that could be considered by the TAR team/ERG and Appraisal Committee as well as manufacturers during appraisal.

Categories and type 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 apparently similar guidance illustrates how the same category of guidance (e.g., Approve, AWR, OIR or Reject) 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. Although improving transparency in communicating the considerations that underpin guidance, the application of the checklist alone is unlikely to be sufficient, especially in situations where OIR guidance was made for a technology that was, on balance, expected to be cost-effective. Evidence of how, not just what assessments and judgements were made, would be required.

The principles suggest that the categories of guidance available to NICE have wider application than is reflected in previous guidance. For example, there are 5 different types of OIR which may be appropriate when a technology is expected to be cost-effective (see Table 2.1a in Section 2.1.3). Indeed, OIR or even Reject maybe appropriate even when research is possible with approval if there are significant irrecoverable costs. Therefore, the full range of categories of guidance (OIR and Reject as well as AWR 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. Therefore, which category of guidance will be appropriate only partly depends on an assessment of expected cost-effectiveness and hence this assessment should only be regarded as an initial step in formulating guidance.

Other NICE programmes
The principles outlined in Section 2 and represented by the checklist do not presuppose how the assessments required might be informed and judgments made. Distinguishing principles from methods of analysis in this way means that the principles and the checklist, may be useful in other NICE programmes, whilst recognising that how the assessment might be made is likely to differ. It was recognised that some amendments might be required where cost-effectiveness is not the prime consideration. The complexity of multiple outcomes in Public Health, the greater scope and complexity of decision problems in Clinical Guidelines and relative paucity of evidence and speed of innovation in devices offer greater challenges for quantitative assessment but do not change the key principles and considerations.

Implications for patient access schemes and value based pricing

Any change in the effective price of the technology, either through patient access schemes or direct price changes will affect the key assessments, leading to different categories of guidance. Therefore, consideration of the effect of price changes on OIR and AWR is needed when assessing the potential impact of patient access schemes and more direct price negotiation through value based pricing.

The price at which the technology would just be expected to be cost-effective is commonly regarded as the value based price for the technology. It is the maximum price the NHS can afford to pay without imposing negative health effects. This describes the threshold price below which Approve rather than Reject would be appropriate when either: i) OIR or AWR guidance is not available to the decision maker; ii) there is no uncertainty in cost-effectiveness; or iii) the research, if needed, can be conducted with approval and there are no irrecoverable costs. In all other circumstances there are a number of other value based prices, each representing the threshold price below which guidance would change. Importantly, once uncertainty and the need for evidence as well as the impact of irrecoverable costs are recognised the threshold price that would lead to Approval will always be lower or equal to a single value based price based on expected cost-effectiveness alone, i.e., disregarding uncertainty in costs and effects.

For example, when research could be conducted without approval but not with it, there will be a price threshold above which Reject rather than OIR would be appropriate; and a lower price threshold below which Approve rather than OIR would be appropriate. This Approve/OIR threshold price will always be lower than a VBP based on expected cost-effectiveness. If a technology also imposes significant irrecoverable costs this threshold price will be lower still.

Health technologies with patent protection are more likely to be priced close to the point at which the ICER is close to or equal to the threshold, i.e., expected incremental NHE are close to zero. A value based pricing scheme would formalise these existing incentives so technologies will tend to impose greater irrecoverable opportunity costs. The use of a technology which is only just expected to be cost-effective will not 'breakeven' until close to the end of the patent time horizon and much longer for the population of patients likely to benefit from its use. Therefore, those technologies already priced close to the threshold, and all new technologies considered in a value based pricing scheme, will tend to increase the scale of irrecoverable costs committed by approval. If these irrecoverable costs are significant (because treatment decisions are not irreversible) then OIR or Reject maybe more likely even when a technology is just expected to be cost-effective.

Even in circumstances where price negotiation becomes possible alongside NICE appraisal, it will be important to retain the OIR and AWR as available categories of guidance for two reasons. Firstly, there is no guarantee that manufacturers will always agree to the lower price below which Approval rather than OIR or AWR would be appropriate. Secondly, and possibly more importantly, there may be many circumstances when there is no effective price reduction which would make Approval appropriate. For example, Reject or OIR guidance may still be appropriate even if the effective price of a technology was zero if there is substantial uncertainty about its effectiveness and/or potential for harms.

Incentives for evaluative research

Consideration of OIR and AWR recommendations provides a link between uncertainty, evidence and price which might appropriately align incentives for manufacturers conducting the type of evaluative research that would be most valuable for the NHS. It provides clear signals and an incentive to ensure the type of evidence, which would require research that cannot be conducted once approved for NHS use, is available and is sufficient at launch (e.g., relative effectiveness and subtle but important differences in side effect profiles). Therefore, a predictable OIR and AWR policy signals what type of evidence is likely to be most important at an early stage. It offers manufacturers a choice, to either: i) accept OIR Guidance at a higher price but restricted volume; ii) reduce the effective price to achieve Approval, or AWR where that is possible; or iii) conduct the evaluative research at an earlier stage so that additional evidence is available at launch.

Consideration of how the NHS and manufacturers are likely to share the value of evidence might inform whether manufacturers should be expected to conduct the research specified in AWR or OIR guidance, or whether manufactures might be expected to make some contribution to the costs of publicly funded research which may ultimately benefit their product. Two other issues need to be

considered. Firstly, the resource constraints on publicly funded research may mean that other research priorities (often without commercial interest) may be more valuable to the NHS. Secondly, the success of AWR recommendations when manufacturers are asked to conduct the research will depend on whether NICE and/or DH are able to establish incentive consistent contractual arrangements as part of an AWR recommendation, i.e., arrangements that can be monitored and enforced with credible penalties to ensure agreed research is conducted and in the way intended. It was widely recognised that, at present, NICE does not have a credible mechanism to ensure that the type of research recommended in AWR Guidance would actually be undertaken by manufacturers. Removing approval of a technology simply because recommended research had not been conducted was not considered an ethical or credible threat.

Although OIR provides a greater incentive to undertake research, it was recognised there may be circumstance where manufacturers would nonetheless choose not to undertake it, i.e., accepting an effective Reject. If the research is also not a sufficient priority to secure public funding, then Approval rather than OIR (an effective reject in these circumstances) would only be appropriate at the lower Approve/OIR price threshold. All AWR and OIR research could in principle be publicly funded rather than undertaken by manufacturers if the costs of research could be recovered directly from manufacturers or indirectly through other price discounts. Since the costs of public research are likely to be substantially lower than for manufacturers this might be mutually beneficial in many circumstances; providing appropriate support to innovation, while allowing wider access to the data generated and more transparency and accountability in the conduct of the research[40].

Value of AWR and evidence at launch
The assessments that need to be made (especially in sections 3.5.2 and 3.6) can be used to consider: i) what would be the value of being able to conduct research while a technology is approved (value of AWR); ii) what would be the value of making evidence that is needed by the NHS available at launch; and iii) what is the value of being able to acquire evidence more quickly. This can inform whether investment in better data collection, registries or information systems that might make AWR possible. Importantly this will differ by technology and will depend on the scale and significance of irrecoverable costs. The value to the NHS of having access to the evidence needed at launch can inform a range of policies, such as early advice, public investment in transitional and evaluative research earlier in the development process or other incentives for research and development. Understanding the relationship between the time taken for research to report and the value of the evidence to future populations can help to inform: i) investments which might make research findings more quickly available; ii) the trade-off implicit in the choice of alternative research designs (i.e., greater precision or timeliness); and iii) research prioritisation (identifying those areas where if research is to be undertaken there must be confidence that it can report quickly).

The value of early evidence at launch and AWR can also be considered from the perspective of the manufacturer; taking account of prices and expected volumes over the remaining patent life and technology time horizon. This might inform when manufacturers might be expected to conduct the research needed (high commercial value that exceeds the cost to manufacturers) and when the NHS might be expected to undertake it (low commercial value but high potential value to the NHS that exceeds the costs to the NHS). In many circumstances both the commercial and NHS values will exceed their respective costs. The question of who should conduct, or pay for, the research might be informed by which sector has the comparative advantage, i.e., which has the highest 'relative efficiency' in generating social value?

Social value judgements and ethical principles
Although OIR and AWR recommendations pose important ethical issues, the type of judgements required appear ethically permissible and consistent with the social values and principles that underpin existing NICE appraisal. It was recognised, however, that OIR guidance may be made when clinicians and researchers would not be in equipoise about the effectiveness (as opposed to the cost-effectiveness) of an intervention. Therefore, more work on how a useful notion of equipoise might be informed by the type of social decisions that bodies like NICE have to make may be valuable. Whether an AWR recommendation will be undertaken by manufacturers also has an ethical dimension. Threatening to withdraw approval if the research fails to report would be unfair on future patients because the evidence base would not have changed. This potential inequity strengthens the case to either: ensure enforceable, incentive consistent are in place; or acknowledge that AWR may not be possible for these reasons, and OIR is the only effective policy option.

How should the assessments be undertaken?

The order of the assessments in the checklist relate to the sequence of decision nodes that fully describe the algorithm in Appendix A. This order of considerations means that all 7 assessments do not necessarily need to be made when an earlier judgement can lead directly to guidance. The assessment of whether in principle further research might be worthwhile and what type of evidence might be required would need to be undertaken routinely. If research may be worthwhile, some indication of the type of evidence needed would also be useful for those making an assessment of the prospects of research (see below) and whether the type of research required to generate it would be possible with approval. Therefore, routine assessment up to point 4 of the checklist would seem appropriate before others with expertise and responsibility for research design and commissioning considered the prospects of the type of research needed.

One model for an efficient order of assessment would be to consider points 1-5 routinely (some early assessment of other sources of uncertainty at point 5 probably ought to be undertaken so that it can inform these deliberations). The AC would then be in a position to either rule out OIR or AWR and issue guidance in the usual way or indicate in the ACD that OIR or AWR was provisionally recommended subject to advice from a research advisory committee and subsequent analysis to support an assessment of points 6 and 7 of the checklist at FAD. This model would avoid unnecessary analysis and assessment and incorporate the judgments of the research community without necessarily adding to the time that an appraisal might take.

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 the report is required. Although the NICE appraisal process may be well suited to identifying the *need for evidence*, these other critical assessments (*the type of research and its priority*) are not necessarily ones for which NICE and its advisory committees, as currently constituted, 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 AWR 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 publicly 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 a commercial interest. Therefore, a decision of whether OIR or AWR research should be undertaken by the manufacturer or through publicly funded research is one that NICE cannot properly take alone.

Informed judgements and better decisions might be possible through greater involvement of the research community. For example, a research advisory committee could be constituted which could consider provisional OIR or AWR guidance (at ACD), 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 publicly funded or undertaken by the manufacturer with appropriate contractual arrangements (which may require the involvement of DH at some stage). Since some of the assessments which must be made in formulating OIR or AWR guidance are, in fact, research decisions which fall outside the NICE remit, it would seem sensible to draw on the expertise of those involved in, and responsible for, these types of research decisions to help make these assessments, i.e., translating the need for particular types of evidence into a need for particular type of research and its prospects.

Triggers for reappraisal also need to be linked to OIR and AWR guidance in two respects: i) to ensure that guidance is reconsidered once research findings are available and ii) to reconsider guidance if the type of research anticipated as part of AWR or OIR guidance is not undertaken and there is little prospect that it will be, e.g., where there is no agreement with a manufacturer or publicly funded research was not prioritised or commissioned to avoid unnecessary research.   In addition, other changes (prices, entry of new technologies and new evidence) which might mean that research is no longer required ought to be included as criteria of reappraisal.

What additional information and analysis might be required?

Cost-effectiveness was presented in terms of net health effects per patient treated and for the population of patients over time in Section 3. This provides information in a way that is directly relevant to the assessments that need to be made, especially at point 2 and point 7 for the checklist (see Section 3.3.1). All the information required to express expected cost-effectiveness in these ways are generally already available during appraisal and are entirely equivalent to the more familiar ICER.

Amending how irrecoverable capital costs are incorporated into the estimates of expect costs pose few technical difficulties. Considering the significance of the irrecoverable opportunity costs of initially negative net health effects of a technology is more nuanced (see Section 2.3.2) and their precise effect will depend on assessments made at points 5 and 6, which can be supported by additional analysis using the type of economic models developed during appraisal. However, some early indication of their potential importance can be based on their scale relative to expected net health effects, the point at which initial losses are expected to be compensated by later gains, whether treatment decisions are reversible and what opportunities to improve health might be forgone by a delay to initiating treatment.

The judgment at point 3 of the checklist requires some assessment of: i) how uncertain a decision based on expected cost-effectiveness might be; and ii) what the consequences, in terms of population NHE, are likely to be if an incorrect decision is made. The methods of analysis presented in Section 3.4.1 attempt to decompose this assessment into a series of steps each presenting what is already available within current methods of appraisal but in ways that can more directly inform the assessment required.

This assessment at point 4 of the checklist requires judgements about: i) how important particular types of parameters (inputs to the economic model) are to estimates of cost and QALY; ii) what values these parameters would have to take to change a decision based on expected cost-effectiveness; iii) how likely is it that parameters might take such values and iv) what would be the consequences if they did, i.e., what might be gained in terms of population NHE if the uncertainty in the values of these parameters could be immediately resolved? The methods of analysis presented in Section 3.4.2 decompose this assessment in to these series of steps; presenting in turn what is already available within current methods of appraisal but in ways that more directly inform the assessment required. It is only when assessing the consequences of uncertainty associated with particular parameters that additional analysis (using the results of existing probabilistic analysis) is required to provide quantitative estimates. There are circumstances where this additional computation is a significant burden, although increasingly, suitable simplifications and approximations are available.

It was recognised by participants that uncertainty in the parameters included in PSA generally do not represent all sources of uncertainty. Commonly there is also uncertainty between alternative assumptions judgements that might be made; often represented by alternative scenarios. How the consequences of uncertainty between as well as within scenarios can be presented was explored in Sections 3.4.1 and 3.4.2. Although this analysis doesn't require additional simulation, it does require that either the probabilities are made explicit in advance or estimates to be presented for a range of probabilities that might represent the judgement of the AC following deliberation. Rather than make alternative assumptions and present extreme scenarios, formal elicitation of the judgement of clinical experts about the unknown parameters for which assumption are required is possible. Such elicitation may provide a richer characterisation of uncertainty than weighting alternative scenarios but requires relevant experts to be identified in advance and for the AC to accept their judgements.

The current appraisal process generally already provides the information and much of the analysis required to complete all the quantitative assessment reported in Section 3. However, the information required to assess whether other sources of uncertainty will resolve over time (point 5 on the checklist) requires information that is not commonly sort as part of NICE appraisal. Many potentially useful sources were either proprietary or public access restricted, making it surprisingly difficult to inform these assessments with publicly available information. When information and estimates were available they were often not complete or directly relevant to a UK context. NICE many need to consider how access to this type of information can be provided or whether it should extract this type of information itself at an earlier stage of appraisal.

Since the STA process bases appraisal primarily on manufacturers submissions, any additional analysis would need to be included in the submission and be reviewed by the ERG. Although the additional analysis itself is limited (most is already required but sometimes presented in different forms), more explicit consideration of OIR and AWR and its link to price would make the critique of how uncertainty and its consequences has been characterised more important.   An assessment of whether the point estimate of cost–effectiveness is reasonable is inevitably a more limited task than also assessing whether the uncertainty surrounding that assessment is credible.   Any additional burden on ERGs (and manufacturers) might be eased with clear guidance on the details of how analysis should be conducted and presented, what common assumptions are deemed reasonable and provision of additional information by the Institute as well as only considering points 6 and 7 on the checklist after ACD and following advice from a research advisory committee.

References

1. Claxton K, Palmer S, Longworth L, Bojke L, Griffin S, McKenna C, et al. 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. *Health Technology Assessment*. Forthcoming.

2. Culyer A, McCabe C, Briggs A, Claxton K, Buxton M, Akehurst R, et al. Searching for a threshold, not setting one: the role of the National Institute for Health and Clinical Excellence. *Journal of Health Services Research & Policy*. 2007;12:56-8.

3. Phelps CE, Mushlin AI. On the (near) equivalence of cost-effectiveness and cost-benefit analyses. *Int J Technol Assess Health Care*. 1991;7(1):12-21.

4. Stinnett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. *Med Decis Making*. 1998 Apr-Jun;18(2 Suppl):S68-80.

5. Claxton K, Briggs A, Buxton M, Culyer A, McCabe C, Walker S, et al. Value based pricing for NHS drugs: an opportunity not to be missed? *BMJ*. 2008;336:251-4.

6. NICE. Guide to the methods of technology appraisal. London: NICE2008.

7. McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold: what it is and what that means. *Pharmacoeconomics*. 2008;26(9):733-44.

8. Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *Journal of Health Economics*. 1999;18(3):342-64.

9. Claxton K, Cohen J, Neumann P. When is evidence sufficient? *Health Affairs*. 2005;24(1):93-101.

10. Claxton K, Sculpher M, Drummond M. A rational framework for decision making by the National Institute for Clinical Excellence. *Lancet*. 2002;360:711-5.

11. Griffin SC, Claxton KP, Palmer SJ, Sculpher MJ. Dangerous omissions: the consequences of ignoring decision uncertainty. *Health Econ*. 2011 Feb;20(2):212-24.

12. Trueman P, Grainger DL, Downs KE. Coverage with evidence development: applications and issues. *Int J Technol Assess Health Care*. 2010 Jan;26(1):79-85.

13. Philips Z, Claxton K, Palmer S. The half-life of truth: what are appropriate time horizons for research decisions? *Med Decis Making*. 2008 May-Jun;28(3):287-99.

14. Eckermann S, Willan AR. The option value of delay in health technology assessment. *Medical Decision Making*. 2008 May-Jun;28(3):300-5.

15. Palmer S, Smith P. Incorporating option values into the economic evaluation of health care technologies. *Journal of Health Economics*. 2000;19:755-66.

16. Eckermann S, Willan A. Expected value of information and decision making in HTA. *Health Econ*. 2006;16(2):195-209.

17. Fenwick E, Claxton K, Sculpher M. The value of implementation and the value of information: combined and uneven development. *Medical Decision Making*. 2008 Jan-Feb;28(1):21-32.

18. Hoomans T, Fenwick EA, Palmer S, Claxton K. Value of information and value of implementation: application of an analytic framework to inform resource allocation decisions in metastatic hormone-refractory prostate cancer. *Value Health*. 2009 Mar-Apr;12(2):315-24.

19. McKenna C, Claxton K. Addressing adoption and research design decisions simultaneously: the role of value of sample information analysis. *Med Decis Making*. 2011 Mar 10.

20. Chalkidou K, Lord J, Fischer A, Littlejohns P. Evidence-based decision making: when should we wait for more information? *Health Affairs*. 2008 Nov-Dec;27(6):1642-53.

21. Basu A. Economics of individualization in comparative effectiveness research and a basis for a patient-centered health care. *J Health Econ*. 2011 May;30(3):549-59.

22. Hutton J, Trueman P, Henshall C. Coverage with evidence development: An examination of conceptual and policy issues. *International Journal of Technology Assessment in Health Care*. 2007;23(4):425-35.

23. Towse A, Garrison LP, 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):93-102.

24. Claxton K, Sculpher M, Carroll S. Value-based pricing for pharmaceuticals: its role, specification and prospects in a newly devolved NHS: Centre for Health Economics, University of York 2011.

25. Department of Health. A new value-based approach to the pricing of branded medicines: a consultation. London: Department of Health2010.

26. Department of Health. A new value-based approach to the pricing of branded medicines: Government response to consultation. London: Department of Health 2010.

27. Claxton K. OFT, VBP: QED? *Health Econ*. 2007;16:545-58.

28. Department of Health. The Pharmaceutical Price Regulation Scheme. London: Department of Health 2008.

29. Chalkidou K, Hoy A, Littlejohns P. 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;100:453-60.

30. Chalkidou K, Walley T, Culyer A, Littlejohns P, Hoy A. Evidence-informed evidence-making. *Journal of Health Services Research & Policy*. 2008;13(3):167-73.

31. NICE. Social Value Judgements Guideline. Principles for the development of NICE guidance. London: NICE2008.

32. NICE citizens council. "Only in research". London: NICE2007.

33. Levine RJ. The "best proven therapeutic method" standard in clinical trials in technologically developing countries. *IRB*. 1998 Jan-Feb;20(1):5-9.

34. McKenna C, McDaid C, Suekarran S, Hawkins N, Claxton K, Light K, et al. Enhanced external counterpulsation for the treatment of stable angina and heart failure: a systematic review and economic analysis. *Health Technol Assess*. 2009 Apr;13(24):iii-iv, ix-xi, 1-90.

35. Claxton K. Exploring uncertainty in cost-effectiveness results. *Pharmacoeconomics.* 2008;9(781-98).

36. Claxton K, Sculpher M, McCabe C, Briggs AH, Akehurst R, Buxton M, et al. Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. *Health Econ*. 2005;14(4):339-47.

37. Claxton K, Sculpher M. Using value of information analysis to prioritise health research: some lessons from recent UK experience. *Pharmacoeconomics*. 2006;24:1055-68.

38. Price MJ, Welton NJ, Briggs AH, Ades AE. Model averaging in the presence of structural uncertainty about treatment effects: influence on treatment decision and expected value of information. *Value Health*. 2011 Mar-Apr;14(2):205-18.

39. Griffin S, Welton N, Claxton K. Exploring the research decision space: the expected value of information for sequential research designs. *Med Decis Making*. 2010;*30*:155-62.

40. Jayadev A, Stiglitz J. Two ideas to increase innovation and reduce pharmaceutical costs and prices. *Health Aff (Millwood).* 2009 Jan-Feb;28(1):w165-8.
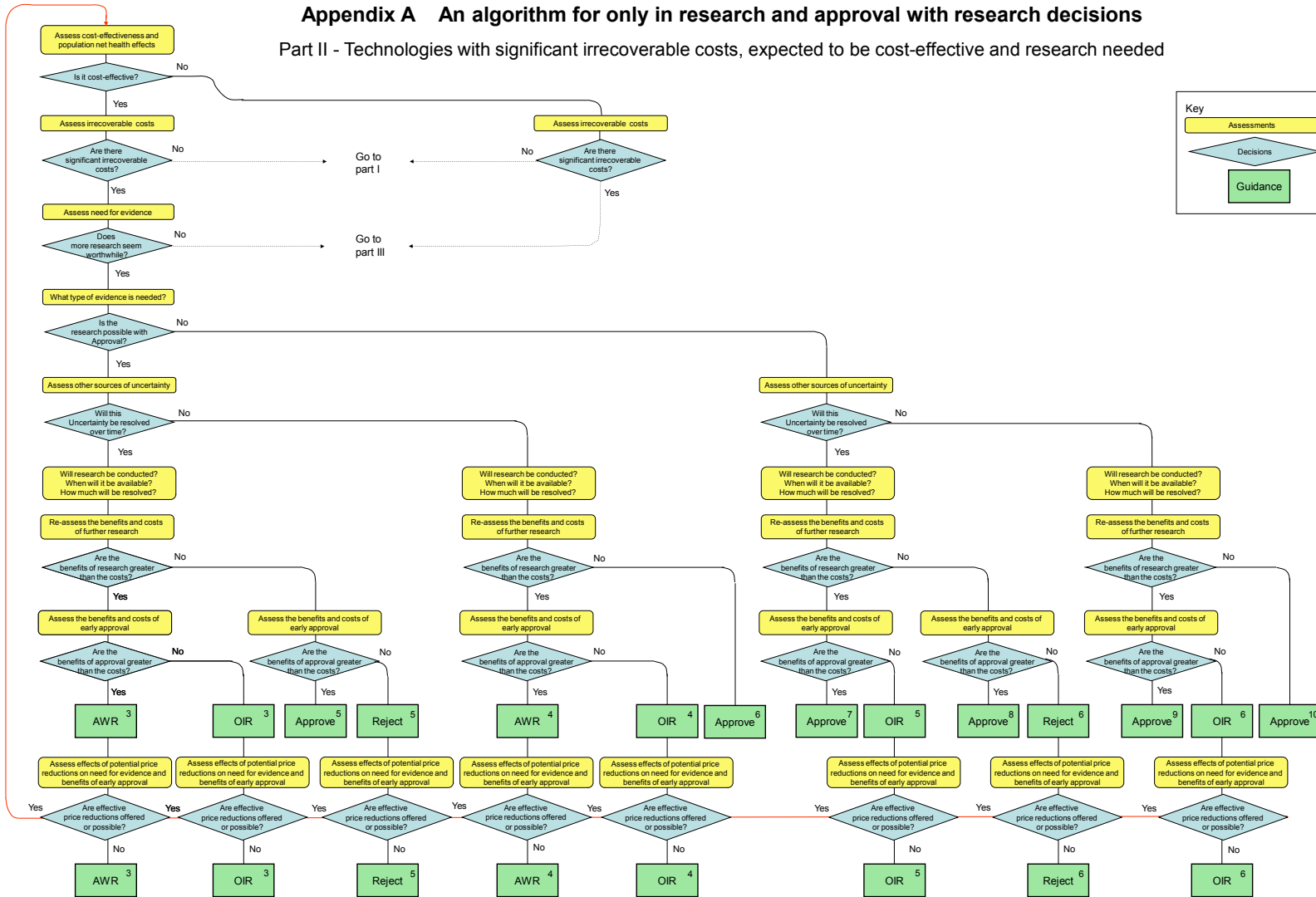
Appendix A

Appendix A   An algorithm for only in research and approval with research decisions

Part I – Technologies without significant irrecoverable costs
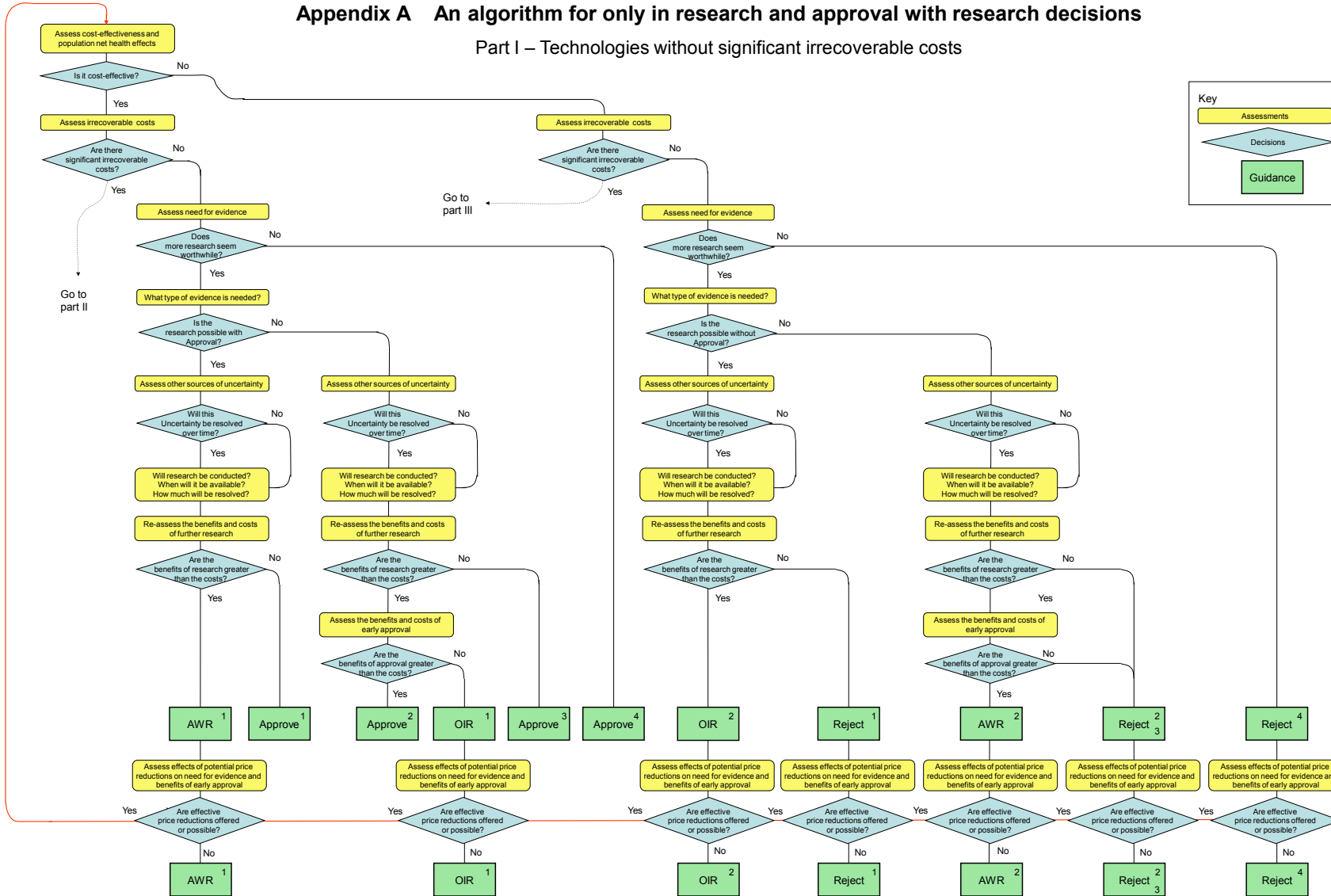
Appendix A   An algorithm for only in research and approval with research decisions

Part II - Technologies with significant irrecoverable costs, expected to be cost-effective and research needed

Key

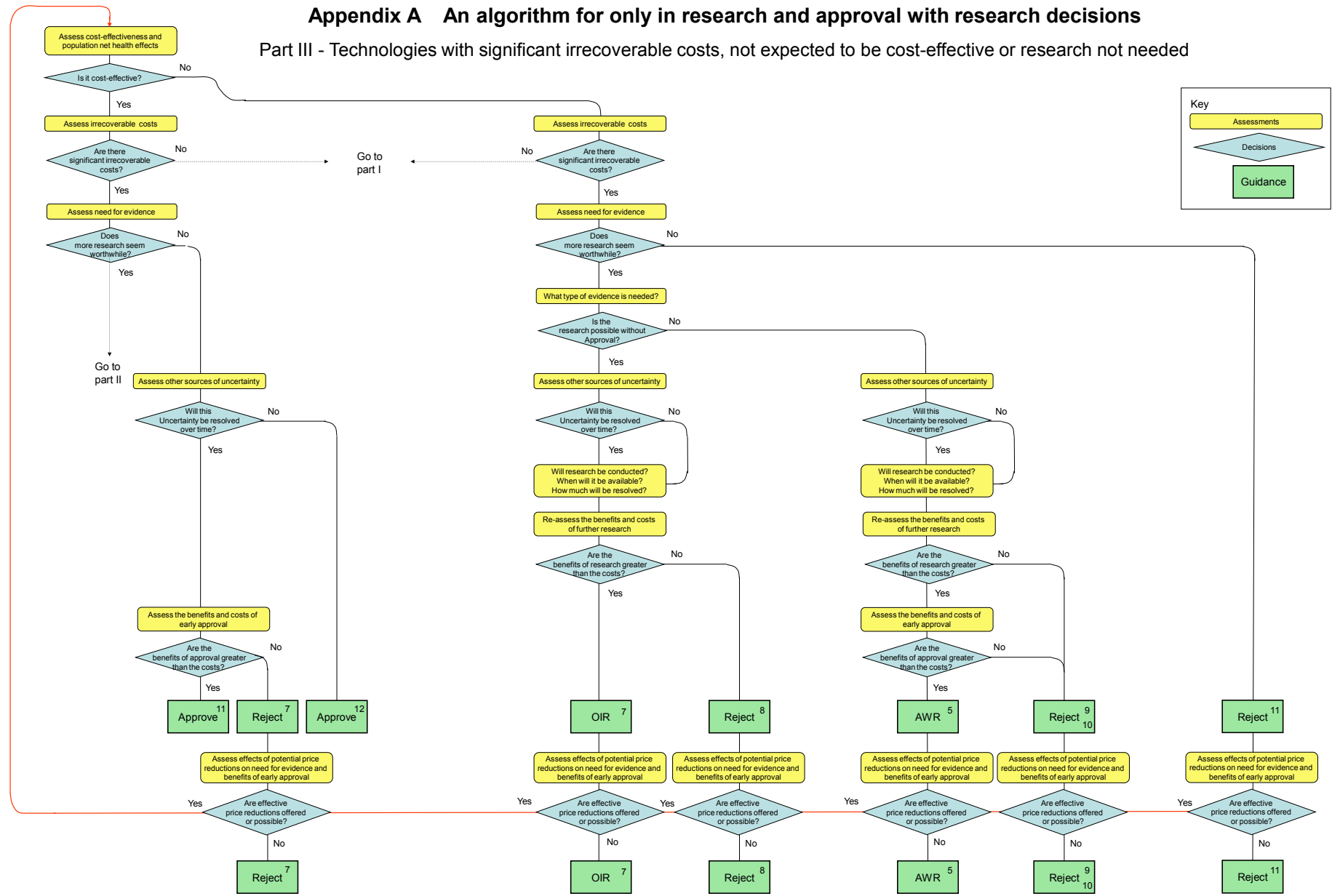|	
Assessments	(yellow box)
Decisions	(blue diamond)
Guidance	(green box)

Appendix A   An algorithm for only in research and approval with research decisions

Part III - Technologies with significant irrecoverable costs, not expected to be cost-effective or research not needed

Appendix B

Table B1   Types and categories of guidance from using the checklist

Assessment	1	2	3	4	5	6	7	Guidance
1	Yes	No	Yes	Yes	Yes/No	Yes	-	AWR 1
2	Yes	No	Yes	Yes	Yes/No	No	-	Approve 1
3	Yes	No	Yes	No	Yes/No	Yes	Yes	Approve 2
4	Yes	No	Yes	No	Yes/No	Yes	No	OIR 1
5	Yes	No	Yes	No	Yes/No	No	-	Approve3
6	Yes	No	No	-	-	-	-	Approve 4
7	No	No	Yes	Yes	Yes/No	Yes	-	OIR 2
8	No	No	Yes	Yes	Yes/No	No	-	Reject 1
9	No	No	Yes	No	Yes/No	Yes	Yes	AWR 2
10	No	No	Yes	No	Yes/No	Yes	No	Reject 2
11	No	No	Yes	No	Yes/No	No	-	Reject 3
12	No	No	No	-	-	-	-	Reject 4
13	Yes	Yes	Yes	Yes	Yes	Yes	Yes	AWR 3
14	Yes	Yes	Yes	Yes	Yes	Yes	No	OIR 3
15	Yes	Yes	Yes	Yes	Yes	No	Yes	Approve 5
16	Yes	Yes	Yes	Yes	Yes	No	No	Reject 5
17	Yes	Yes	Yes	Yes	No	Yes	Yes	AWR 4
18	Yes	Yes	Yes	Yes	No	Yes	No	OIR 4
19	Yes	Yes	Yes	Yes	No	No	-	Approve 6
20	Yes	Yes	Yes	No	Yes	Yes	Yes	Approve 7
21	Yes	Yes	Yes	No	Yes	Yes	No	OIR 5
22	Yes	Yes	Yes	No	Yes	No	Yes	Approve 8
23	Yes	Yes	Yes	No	Yes	No	No	Reject 6
24	Yes	Yes	Yes	No	No	Yes	Yes	Approve 9
25	Yes	Yes	Yes	No	No	Yes	No	OIR 6
26	Yes	Yes	Yes	No	No	No	-	Approve 10
27	Yes	Yes	No	n/a	Yes	n/a	Yes	Approve 11
28	Yes	Yes	No	n/a	Yes	n/a	No	Reject 7
29	Yes	Yes	No	n/a	No	-	-	Approve 12
30	No	Yes	Yes	Yes	Yes/No	Yes	-	OIR 7
31	No	Yes	Yes	Yes	Yes/No	No	-	Reject 8
32	No	Yes	Yes	No	Yes/No	Yes	Yes	AWR 5
33	No	Yes	Yes	No	Yes/No	Yes	No	Reject 9
34	No	Yes	Yes	No	Yes/No	No	-	Reject 10
35	No	Yes	No	-	-	-	-	Reject 11

Part I of the algorithm (Assessments 1–12)

Part II of the algorithm (Assessments 13–26)

Part III of the algorithm (Assessments 27–35)
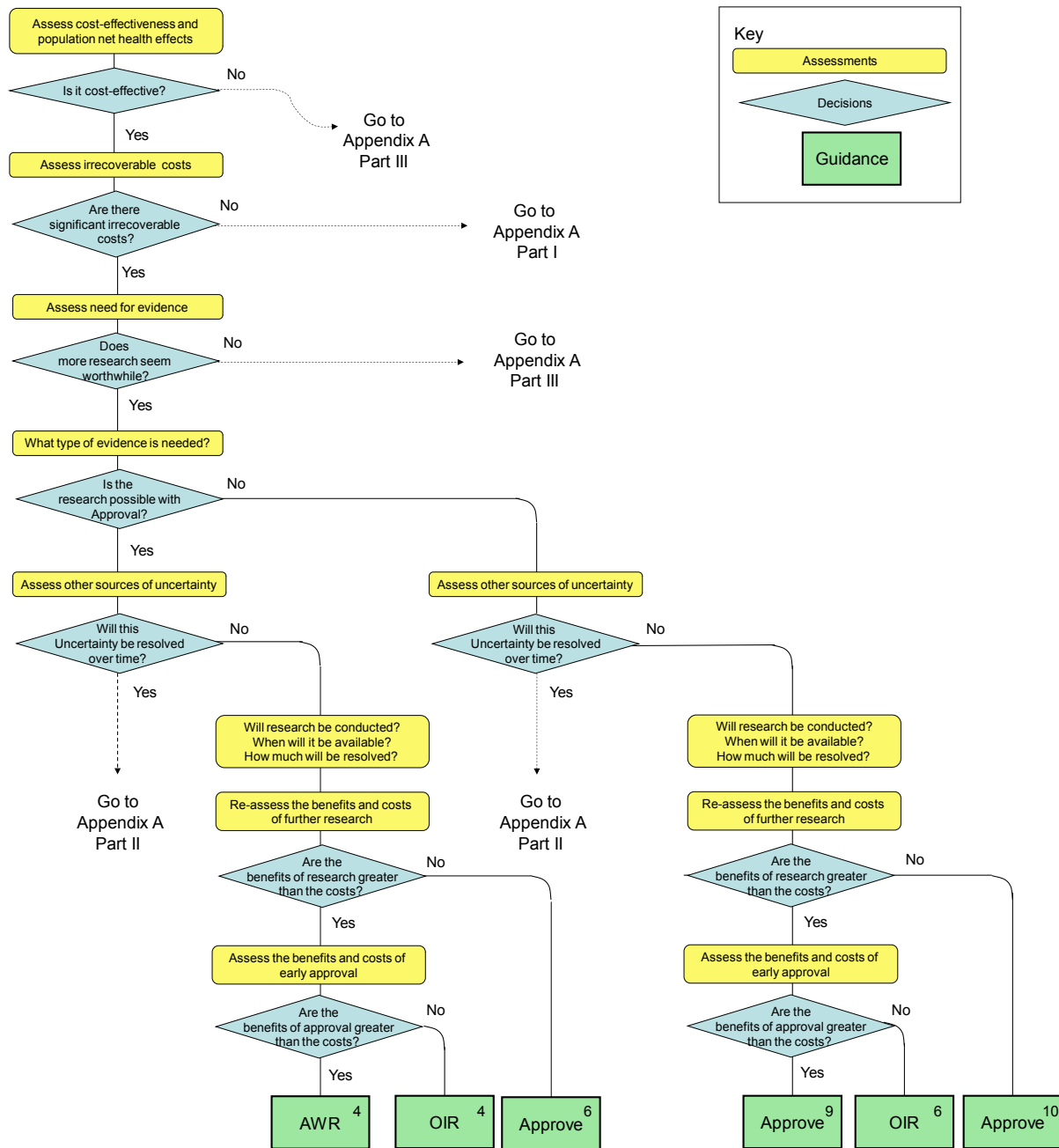
Figure B1   Possible pathways for EECP

Figure B2   Possible pathways for CLOP

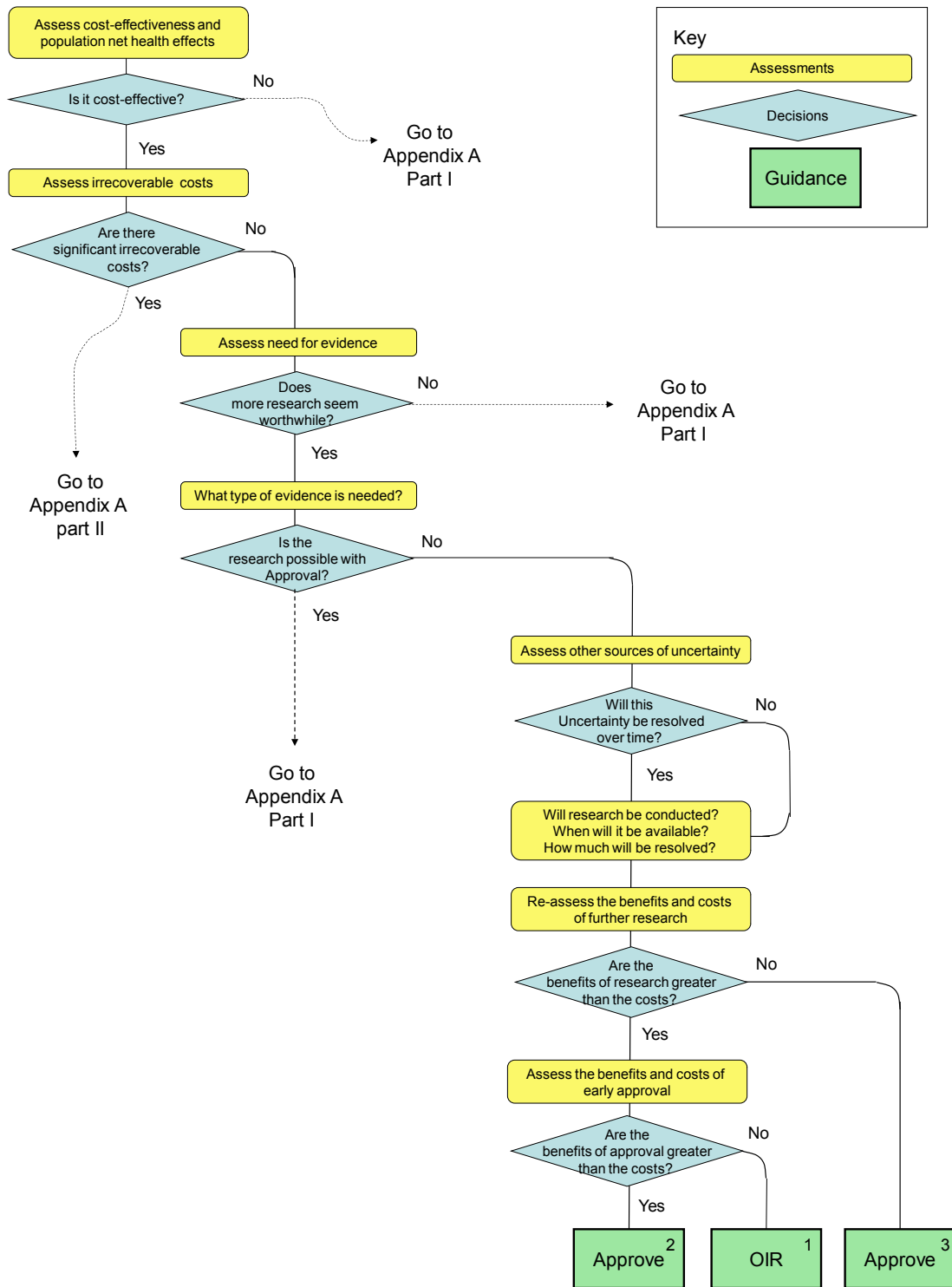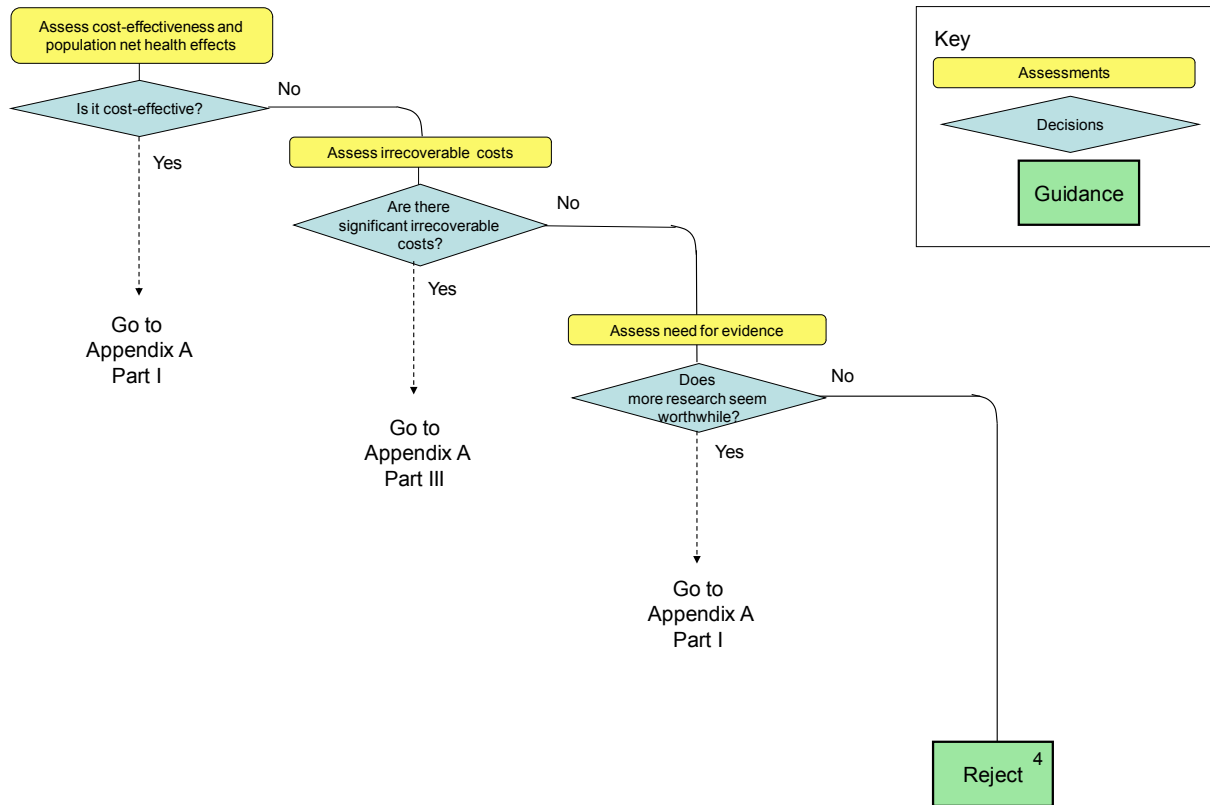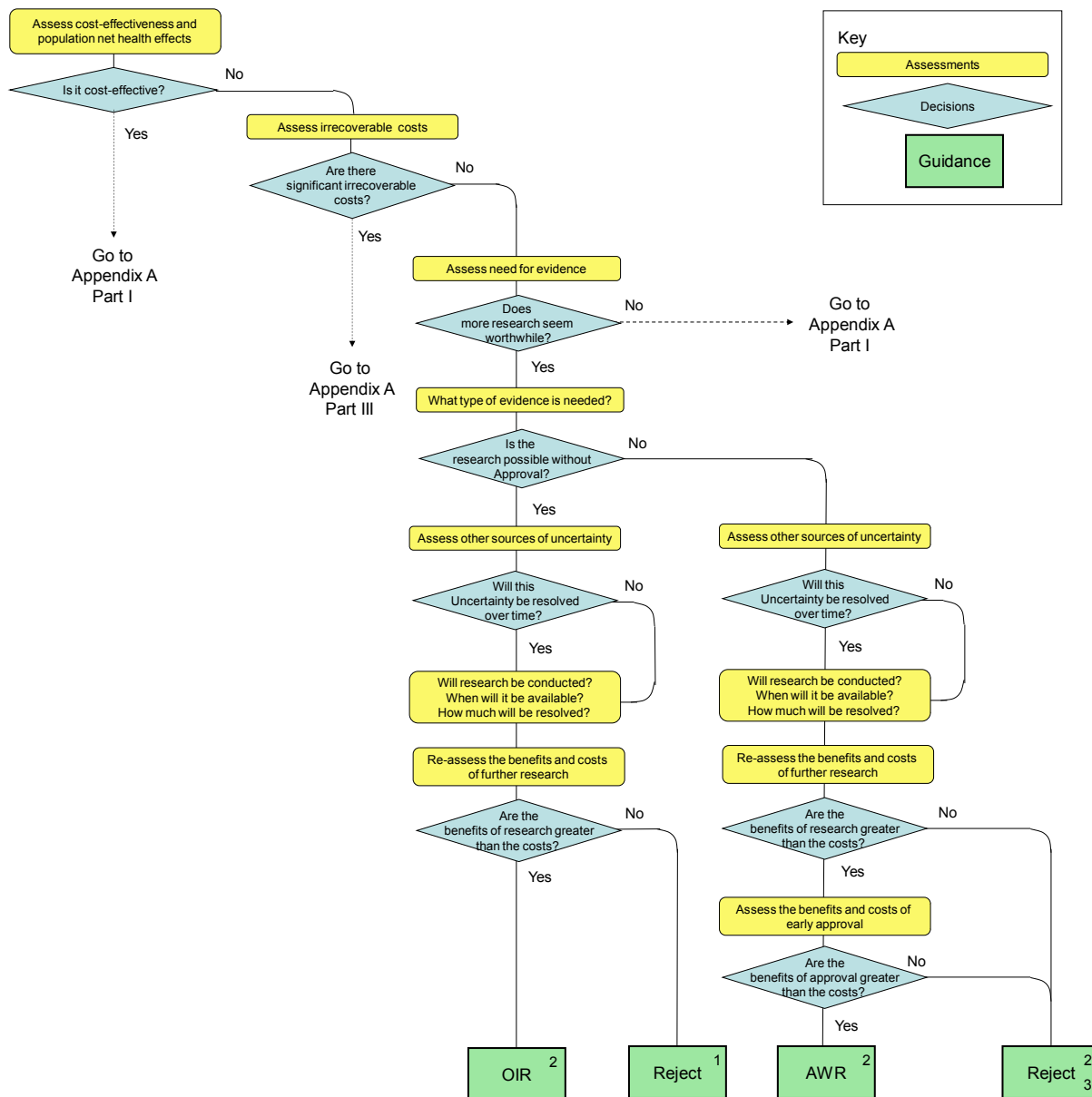**Figure B3   Possible pathways for OMAL**

Figure B 4   Possible pathways for PsA