NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

Briefing paper for methods review working party on choosing comparators

The briefing paper is intended to provide a brief summary of the issues that are proposed for discussion by the Methods Review Working Party to inform an update to the Institute's Guide to Methods of Technology Appraisal. It is not intended to reflect a comprehensive or systematic review of the literature. The views presented in this paper are those of the authors and do not reflect the views of the Institute.

1 Review of the 'Guide to Methods of Technology Appraisal'

The Institute is reviewing the 'Guide to the methods of technology appraisal', which underpins the technology appraisal programme.

The original Methods Guide was published in February 2001, and revised versions were published in 2004 and 2008. The Methods Guide provides an overview of the principles and methods used by the Institute in assessing health technologies. It is a guide for all organisations considering submitting evidence to the technology appraisal programme and describes appraisal methodology.

The revised draft of the Methods Guide will be available for a 3-month public consultation, expected to begin in June 2012. We encourage all interested parties to take part in this consultation.

2 Background

2.1 Relevance of topic to NICE technology appraisals

Clinical and cost effectiveness are relative concepts. A technology cannot be described as “cost effective” *per se*, but is either cost effective (or not) in comparison to some other alternative. It is therefore critical that the additional costs and benefits of a new technology under appraisal by NICE are assessed relative to the appropriate comparator or comparators, to avoid a misleading view of the value of the new technology.

Whilst the choice of comparator can entirely change the assessment of cost effectiveness, there is frequently some judgement to be made about which is the appropriate comparator(s). The purpose of this paper is to highlight and discuss a) the current NICE guidance on the choice of comparator and to consider this in the light of the economic principles that underpin the use of cost effectiveness analysis, b) outline a series of issues that have arisen in past appraisals which collectively demonstrate those situations in which more detail in the Methods Guide may have been advantageous and c) present a number of issues for consideration that arise from a) and b).

2.2 What the current Methods Guide advises with respect to choosing comparators

The 2008 Methods Guide provides only broad guidance as to the selection of appropriate comparator(s). “Routine **and** best practice in the NHS” (emphasis added) is specified throughout. This wording helps to identify the set of potential comparators but does not provide any detail on which from that set should be selected as the basis for calculating the ICER, or if they are to be combined in some way, and if so, how that should be done. Furthermore, the guide does not specify whether “best practice” refers to the option that is most effective or most cost effective. Greater clarity here may help to resolve some of the challenging situations discussed below.

The following quotes exemplify the broad guidance found in the current methods guide:

"Technologies can be considered to be cost effective if their health benefits are greater than the opportunity costs measured in terms of the health benefits associated with programmes that may be displaced to fund the new technology." (Section 1.4.2.)

In relation to the scope, Section 2.2.4 of the Methods Guide states that:

- Relevant comparators are identified, with consideration given specifically to routine and best practice in the NHS (including existing NICE guidance) and to the natural history of the condition without suitable treatment.
- There will often be more than one relevant comparator technology because routine practice may vary across the NHS and because best alternative care may differ from routine NHS practice. For example, this may occur when new technologies are used inconsistently across the NHS.
- Relevant comparator technologies may also include those that do not have a marketing authorisation (or CE mark for medical devices) for the indication defined in the scope but that are used routinely for the indication in the NHS.
- Comparator technologies may include branded and non-proprietary (generic) drugs.
- Sometimes both technology and comparator form part of a treatment sequence, in which case the appraisal may need to compare alternative treatment sequences.

"Relevant comparators for the technology being appraised are those routinely used in the NHS, and therapies regarded as best practice when this differs from routine practice." (section 5. 1.1)

2.3 Guidance from the economic evaluation literature

The 2008 Methods Guide is consistent with standard economic theory to the extent that all relevant comparators are included within the set of technologies that are considered within an evaluation. Most standard texts (see for example Drummond *et al* 2005) describe how the set of potential comparators should be considered alongside the new technology of interest (we refer to this here

as the “decision set” for short). Texts then go on to outline the decision rules that should be implemented in order to identify the optimal choice from each of the comparators included within that set, that is, an incremental analysis. This detail is important because it is possible to calculate a ratio of difference in cost/difference in benefit between *every* pair of technologies in the decision set. There is the potential for such a set of pairwise comparisons to lead to confusion and they may be misleading. Some previous appraisal submissions have failed to include appropriate incremental analyses (see for example retigabine for epilepsy and trastuzumab for HER2 metastatic gastric cancer).The decision rules for incremental analysis are as follows.

- Where only two therapies are in the decision set, the relevant ratio on which to base decisions is the ratio of incremental cost to incremental benefit (the ICER).

- Where there are more than two options:

(adapted from Glick et al. 2007)

1	Rank order therapies in ascending order of either effect or cost
2	Eliminate therapies that are dominated
3	Compute ICERs for each of the remaining adjacent pairs of therapies
4	Eliminate therapies that have a smaller effect but a larger cost effectiveness ratio compared to the next highest ranked therapy (extended dominance)
5	Recalculate the ICERs for each remaining adjacent pair of therapies (steps 4 and 5 may need to be repeated)
6	Select the option with the largest ICER that is less than the maximum willingness to pay (i.e. the cost-effectiveness threshold)

These rules are consistent with the aim of identifying the technology from the decision set with the greatest measure of health benefit and a cost effectiveness ratio that does not exceed the cost effectiveness threshold. In the next section we consider the extent to which this process for identifying the optimal technology can be adopted in NICE Technology Appraisals.

2.4 NICE Technology Appraisals and the scope

In order to consider the relevance of the full incremental analysis as described above, or any other approach to defining appropriate comparators for NICE Technology Appraisals, it is necessary to consider the scope and broad aims of the programme as a whole. Clarity on the following issues will help to provide more detailed guidance, and therefore greater consistency, than that which currently exists in the 2008 guide.

- What is the relevance of current NHS practice when that is not also best practice? Should best practice be defined as the most effective alternative or should it be the most cost effective alternative?

- Should decision rules about appropriate comparators be based on consideration of the set of options that are directly the subject of the specific technology appraisal guidance i.e. the appraised technologies (in which case there is a clear difference between STA and MTA)?

- Alternatively, should the guidance that could be issued via other NICE programmes also be considered relevant in considering appropriate comparators when formulating Technology Appraisal guidance? Are any other ways in which NHS practice could be influenced, beyond those routes open to NICE, relevant when considering which comparator is appropriate?

There are several situations that have arisen in previous appraisals where there is a conflict between the technology that may be considered optimal according to the decision rules that are standard for economic evaluation and the guidance that NICE is able to publish as part of the Technology Appraisal process. The examples discussed below all demonstrate how these conflicts arise as the result of two issues:

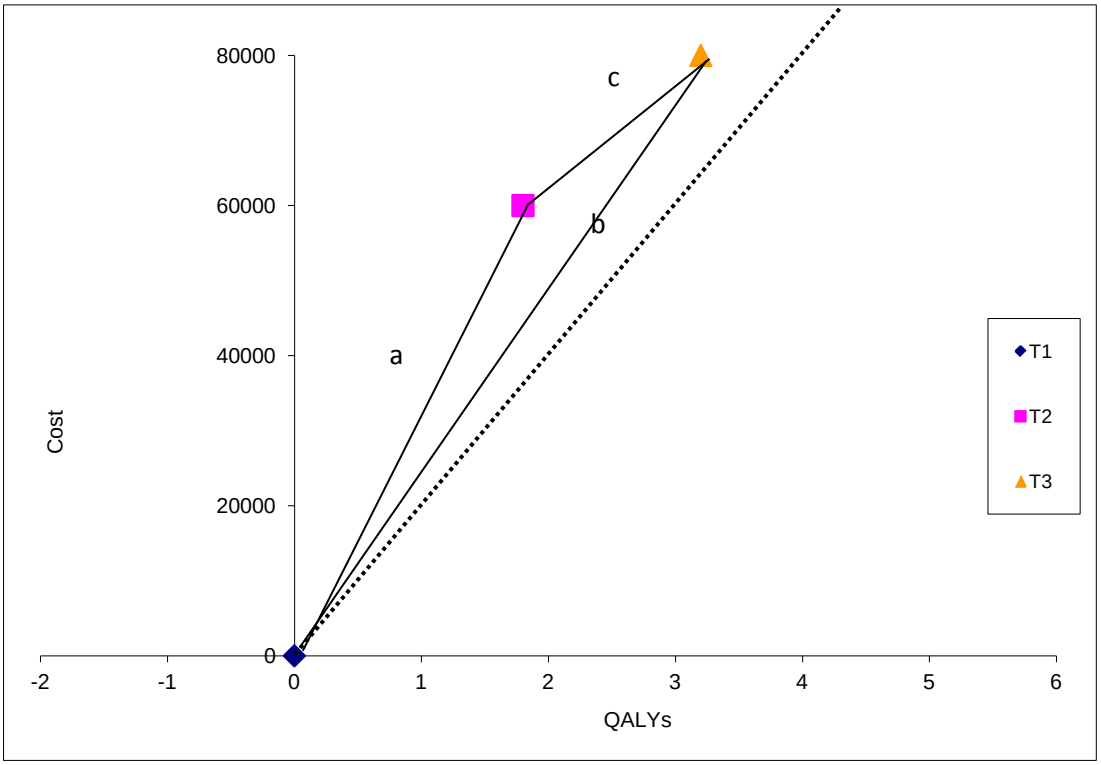
i) the fact that the guidance that NICE may issue as part of a specific Technology Appraisal does not always extend to all potential comparators.

ii) One or more of the comparator technologies in NHS use is not cost effective

Figure 1 illustrates a situation where there are three technologies that form the decision set and an assumed threshold value of £20,000 per QALY gained (represented by the dashed line). Standard decision rules would conclude that the optimal technology is T1. This is because neither options T2 nor T3 have an ICER compared to T1 that is below the threshold of £20k (the gradients of lines *a* and *b* are steeper than the dashed line). It is also the case that T2 would be excluded on the basis of extended dominance, that is, there is a combination of T1 and T3 that would cost less and generate more QALYs than T2.

It should be noted that the gradient of line *c*, the cost effectiveness ratio of T3 compared to T2, is less steep than the dashed threshold line i.e. the ICER is less than £20,000 per QALY gained. If T1 is not an option, then T3 is preferable to T2.

Figure 1: Cost effectiveness plane where a comparator is not cost effective



Given the current NICE process, there are situations whereby the existence of T1 may be deemed by some to be irrelevant. In these situations, it may be argued that "*c*" does represent the appropriate comparison for the problem at hand since this represents the differences in costs and benefits that will occur in the NHS depending on whether NICE Technology Appraisal guidance is positive or negative.

As previously mentioned, if the "best" alternative practice is defined in terms of clinical rather than cost effectiveness then this situation may arise (T2 is more effective than T1). Likewise, if the focus is on current NHS practice then it is feasible that this is T2 rather than T1. This problem may be more acute in the Single Technology Appraisal (STA) process, where the focus is T3 as the new technology, since the only guidance possible is either to recommend or not recommend T3. No recommendations will be made directly about T1 or T2 within the appraisal since these lie outside the remit. In the Multiple Technology Appraisal (MTA) process the problem will occur if neither T1 or T2 are among the technologies specified in the remit of the appraisal i.e. it is not possible to issue guidance on T1 or T2 as part of the appraisal. If T1 alone is included then the recommendation is likely to be for T1 only because T3 is not cost effective relative to it. If T2 alone is included then the appraisal might not recommend either T2 or T3 which would only leave T1 despite their being no formal NICE guidance on it.

It is therefore clear that T3 can be recommended in those situations where there is a focus on current NHS activity as the comparator, where the comparator is defined as "best practice" in terms of clinical effectiveness rather than cost effectiveness and where the optimal strategy is to be chosen only from those which NICE may directly issue guidance on within an appraisal (the appraised technologies) . This latter point requires that the broader set of activities in which NICE, or the NHS in general, may engage in order to influence NHS practice are not considered relevant to the issue of comparators in an appraisal.

There are several reasons why T2 may be current NHS practice, despite the fact that it is cost ineffective compared to T1. For example

- T2 has not been appraised by NICE. This can include the possibility that relevant comparators emerge or are only licensed after the point at which a scope is produced for a particular appraisal.
- T2 may represent off-label use for the specific indication in question. This does not currently rule it out as a comparator in either the STA or MTA process but it does mean that NICE would normally not be able to issue guidance on its use in the NHS as part of the Technology Appraisals Programme. However, it is worth noting that NICE Clinical Guidelines can make recommendations regarding off-label use (NICE Guidelines Manual p.110). This situation is most common in paediatrics, although it has also featured in a number of non-paediatric technology appraisals. This can make it difficult to ascertain whether a treatment really does represent routine practice in the NHS and therefore can increase the debate as to whether a treatment is an appropriate comparator in accordance with the strict definition given in the current Methods Guide.
- T2 may have been appraised by NICE but the technology has been adopted in the NHS contrary to NICE guidance. One previous example of this situation relates to the appraisal of natalizumab for Multiple Sclerosis (MS). Within this STA, it was accepted that current NHS treatment for these patients is beta interferon or glatiramer acetate, provided by the Department of Health supported “Risk Sharing Scheme”, which permits patients to continue to receive these treatments despite the fact that NICE did not recommend them for NHS use on the basis of their cost effectiveness.
- A similar issue is likely to arise in relation to the Cancer Drugs Fund (CDF). The CDF aims to ensure that drugs which have been deemed cost ineffective and are therefore not recommended by NICE are still made available to NHS patients in England only. It is administered regionally around Strategic Health Authority established panels and is intended to be a temporary measure until the expiry of the Pharmaceutical Price Regulation Scheme (PPRS) at the end of 2013. This temporary nature of

the scheme may distinguish these treatments from those provided through other means in the NHS.

In each of these situations, it can be argued, and indeed has been in previous appraisals (see for example Natalizumab for MS), that the “theory of the second best” becomes relevant. Essentially this accepts that efficiency within the limited set of NHS options that NICE Technology Appraisals can influence is the goal of the appraisal. Current NHS practice may be cost ineffective, but if this is not something that NICE Appraisal guidance is able to advise on, then further departures from an inefficient situation may be warranted. A broader view of the remit of the Technology Appraisals Committee, for example that includes as part of its considerations the range of NICE activities that may, at some point in the future, allow a much broader set of guidance to be issued, would lead to a different conclusion. Indeed, a view that considers a full range of NHS activities including disinvestment, implementation and research may be one which provides a rational framework for the consideration of costs, benefits and their associated uncertainties.

One important implication of this view, if accepted for all the various situations highlighted above, is that the guidance that emerges from the MTA process may sometimes be very different from that which would emerge from the STA process. The scope of an STA is limited to issuing guidance on the use of the new technology, whereas an MTA would seek to issue guidance about all technologies in the decision set in many, though not all the examples above.

In each of these situations it is also worth noting that the patient group could be perceived as having already benefitted more than other groups since a non cost effective therapy is available to them. To issue positive guidance for another new technology on the basis of a comparison to a cost ineffective alternative may be seen to exacerbate an already unfair situation.

There are practical issues that must be considered if the view is taken that “best” practice rather than current NHS practice is the relevant comparator. Particularly important is the potential for comparators to emerge after the scope for an appraisal has been finalised. Such comparators may have

gained licensing approval in the interim, or even been through the Appraisal Process. This means that at the point of Committee consideration and guidance production, the appropriate comparators may have changed, be at the point of changing, or be subject to consideration at the same point in time as the technology in the scope in preparation. Indeed, it may be reasonable to assume that such a new comparator would become future NHS practice. In such situations, there are obvious challenges to the submitting manufacturers in terms of access to data on clinical effectiveness (and other parameter values) relating to the new comparator, as well as the time constraints of the appraisal process.

In many settings standard NHS practice may be clear. However, there are situations in which standard care will vary substantially and a number of different comparators may each comprise a significant proportion of current care. There have been situations where it has been argued that the additional costs and benefits of the new technology should be calculated against some form of “average” costs and benefits associated with the mix of current approaches, sometimes referred to as a “blended comparator”. This could be seen as an appropriate approach if, as described above, the goal of a NICE appraisal is considered to be restricted to identifying whether a single new technology is efficient compared to current NHS practice as a whole. The approach also assumes that the displacement of existing practices will occur in the same proportion as in current use. It should also be clarified that where different comparators can be identified for identifiable patient groups then these should be dealt with as separate subgroups. Furthermore, there are several practical issues to be considered even if this goal is considered appropriate. These include:

How should the “average” be determined? Should weights be applied to each of the comparator technologies according to their estimated NHS use? Where should estimates of use come from?

Accepting a blended comparator approach will require at least some NHS practitioners to switch away from their current treatment approach to a new

NICE recommended approach that is relatively cost ineffective and may even be less clinically effective. At the extreme this could entail switching to a less clinically effective option.

This situation arose in relation to the appraisal of lapatinib for the treatment of women with previously treated or metastatic breast cancer. In this case, the manufacturer argued that lapatinib was cost effective compared to trastuzumab-containing treatment regimens and that these were in widespread NHS use. Lapatinib did not appear likely to be cost effective compared to other potential comparators. The manufacturer presented a "blended comparator", which was comprised of a weighted average of the costs and benefits of three treatments, including trastuzumab, where weights were estimated from market research data. The NICE Decision Support Unit (DSU) report on this appraisal argued against the concept of the blended comparator and the Appraisal Committee also adopted this view which was upheld at appeal. However, it should be noted that there are several issues specific to this appraisal that the Appraisal Committee considered pertinent and that may make it inappropriate to infer that the concept of the "blended comparator" was rejected in principle. In particular, a forthcoming NICE guideline regarding the use of trastuzumab containing therapies, that trastuzumab was being used in an unlicensed indication in this situation and the lack of evidence of the magnitude of treatment effect were considered relevant factors.

The "blended comparator" has been raised in other appraisals, including that of azacitidine for the treatment of myelodysplastic syndromes, chronic myelomonocytic leukaemia and acute myeloid leukaemia. Following appeal the committee did accept a blended comparator comprising of best supportive care, low dose chemotherapy and standard dose chemotherapy. The ACD is clear however that several appraisal specific issues led the committee to accept this comparator as the basis for decision making only in this specific instance rather than accepting this as a general decision rule. In particular, the Committee heard that the populations for each of the comparator conventional care regimens could not be clearly defined.

Further examples of potential differences between full incremental and pairwise cost effectiveness ratios

There are many technologies which could form part of a sequence of treatments for individual patients. In these situations, the fundamental principles of economic evaluation still apply. Namely, the costs and benefits of each feasible alternative sequence should be compared in an incremental fashion. This approach considers each sequence of treatments, including a sequence that excludes the new technology entirely, as if they were separate individual treatments and it is correct to do so because they are mutually exclusive: patients can only receive one sequence.

An example of an appraisal where this issue was debated was tocilizumab for the treatment of rheumatoid arthritis. Here, the manufacturer compared a number of different treatment sequences that included tocilizumab in a pairwise fashion to a sequence that excluded tocilizumab (current treatment). Since all generated ICERS were approximately equal it was argued that guidance should permit tocilizumab in any position in the sequence, including as first-line treatment. A full incremental analysis gave very different results and suggested the optimal sequence was one where tocilizumab is used as a second-line treatment within the sequence.

There may also be differences between the incremental and pairwise approaches when consideration of different patient subgroups or strategies for using a technology are considered. For example, it is often the case that separate subgroups of patients can be identified distinguished by those that are naive to currently available treatment and those that are not. For the naive group comparisons can be made between the new therapy and both current care and “do nothing”. Within the licensed indication for a new therapy it is possible to consider a range of different uses of that therapy. NICE appraisals often consider starting and stopping rules for example. As with the use of therapies in a sequence of treatments, these strategies can be considered mutually exclusive and therefore an incremental analysis may be appropriate.

3 Proposed issues for discussion

After consideration of the developments in this methodological area, the current Methods Guide and the requirements of the Institute's Technology Appraisal Programme, it is proposed that the following key areas are discussed by the Methods Guide Review Working Party.

1. What are the general principles that should govern the selection of comparators? Specifically, should NICE Technology Appraisals focus on comparing to "best" or "standard NHS" practice? If "best" should this be defined in terms of clinical or cost effectiveness?

2. Should the Appraisals committee consider only the narrow set of options that can be influenced directly by its guidance or should a broader view be taken? If the latter, what boundaries should be set e.g. the set of NICE activities as a whole, the broader NHS?

3. Should NICE appraisals ever consider the relevant ICER for a new technology to be that compared to a technology that itself is not cost effective, though in use in the NHS? If so, in which circumstances?

    - Comparator recommended for use by DoH despite NICE guidance
    - Comparator available via Cancer Drugs Fund
    - Comparator not been appraised by NICE (does it matter why not appraised? Too new, not licensed, other?)

4. In which circumstances, if any, is it appropriate to consider the comparator to be a "blend" of other options?

5. Should the identification of comparators (during the scoping stage) be focussed on providing a 'protocol' for the appraisal (i.e. a binding list of comparators to be used in the appraisal) or as more of a source of information about all possible options to be defined during the appraisal?

- If the latter approach is taken, should clear guidance on choosing options for the Committee be given?

4 References

Drummond, M.F., Sculpher, M.J., Torrance, G.W., O'Brien, B.J., Stoddart, G.L. (2005) Methods for the Economic Evaluation of health care programmes, Oxford: Oxford University Press.

Glick, H.A., Doshi, J., Sonnad, S.S., Polsky, D. (2007) Economic Evaluation in Clinical Trials, Oxford: Oxford University Press.

NICE (2009) The Guidelines Manual

5 Author/s

This briefing paper has been prepared by Allan Wailoo, Rebecca Trowman and Andrew Stevens.