THE USE OF COST MINIMISATION ANALYSIS FOR THE APPRAISAL OF HEALTH TECHNOLOGIES

REPORT BY THE NICE DECISION SUPPORT UNIT

10th May 2019

Allan Wailoo, Simon Dixon

NICE Decision Support Unit,
ScHARR,
University of Sheffield,
Regent Court, 30 Regent Street
Sheffield, S1 4DA

Tel (+44) (0)114 222 0734
E-mail dsuadmin@sheffield.ac.uk
Website www.nicedsu.org.uk
Twitter @NICE_DSU

ABOUT THE DECISION SUPPORT UNIT

EXECUTIVE SUMMARY

AIM

The National Institute for Health and Care Excellence (NICE) produces guidance for the NHS based on assessment of the clinical and cost effectiveness of health technologies. In some circumstances, a simpler form of economic analysis is used which focusses solely on the costs of interventions, referred to as Cost Minimisation Analysis (CMA). This report aims to provide information on several issues relating to the use of CMA:

- How and when is CMA used across NICE's programmes?
- Whether the approaches used across different guidance producing programmes are consistent.
- Whether any identified differences are appropriate.

We were asked to make recommendations for the way CMA methods are applied which may inform changes to the Methods Guides that relate to each programme.

METHODS

We reviewed the academic literature on CMA methods. We reviewed formal NICE methods guidelines and documents reporting on specific appraisals where CMA has been used in the Technology Appraisals (TA) programme. Discussions were held with committee members, chairs, academic groups and NICE staff in order to understand the key issues in relation to CMA and to obtain feedback on potential recommendations.

FINDINGS

CMA is not a form of full economic evaluation because, on its own, it does not consider outcomes and is therefore unsuitable in most circumstances. Academic literature has generally focussed on the limited set of circumstances in which CMA is appropriate in the context of economic evaluations alongside clinical trials. In general, the opinion that CMA is inappropriate in most situations has strengthened over the last two decades.

There are two programmes at NICE where CMA is formally considered in place of other forms of economic evaluation. The Medical Technologies Evaluation

Programme (MTEP) exclusively uses a form of CMA in formulating its guidance because it operates a different decision rule to other NICE programmes: technologies are only approved if they are cost saving or have similar costs. The TA Programme has an option for a specific technology to utilise its "Fast–Track Appraisal" (FTA) process on the basis of claims of clinical equivalence to comparators and the use of CMA. We found that three appraisals have been conducted through this fast track option (referred to in NICE's own documentation as the "cost-comparison" case), two of which were for the assessment of "me-too" drugs whose comparators had already been appraised and recommended by NICE.

Other NICE programmes (for example NICE Guidelines) refer to the feasibility of using CMA methods in circumstances where clinical non-inferiority is demonstrated, and this is in-line with guidance from many international HTA bodies, but there are few examples where CMA has been used.

The use of CMA in the case of clinical equivalence, as in the FTA process, is potentially relevant to MTEP, the Diagnostic Appraisal Programme (DAP) and Highly Specialised Technologies (HST). The use of CMA in cases of clinical differences, as in MTEP, is not relevant where different decision rules operate.

The standard use of CMA requires an assessment of clinical non-inferiority. This assessment should begin with consideration of the scientific basis for such a hypothesis. In the case of pharmaceuticals, this requires a consideration of the chemical properties of the competing alternatives, and the implications for health outcomes of any differences. Appraisals of other technology types should consider the scientific and practical properties to understand the rationale for any claim of non-inferiority. The plausibility of the claim may help the interpretation of clinical evidence.

In HTA, interest lies with clinical effectiveness described by multiple outcome measures, including adverse events, rather than a single primary outcome. Those which impact patient outcomes and would inform an economic model require detailed scrutiny. Important outcomes need to have convincing evidence of non-inferiority but making such an assessment with the types of clinical evidence typically generated is not straightforward. There are some parallels with the design of non-inferiority trials to

consider. Where possible, the range of the confidence interval in relation to some non-inferiority margin should be considered. This interpretation requires an assessment of the design of the studies contributing evidence, including their sample sizes. Non–inferiority margins are often constructed in relation to minimally clinically important differences. Clinical and patient experts should contribute to this assessment. Care needs to be taken in interpreting data on patient level versus population level outcomes. Where no degree of non-inferiority is acceptable clinically, there may still be a need to determine an acceptable margin on pragmatic grounds.

Surrogate outcomes, that do not relate directly to patient benefits, must often be considered. This is the case for the assessment of diagnostic technologies where performance is measured in terms of sensitivity and specificity. There is a risk that small differences in diagnostic properties between tests, that may appear to be trivial and/or statistically insignificant, would lead to important differences in estimates of cost effectiveness. In this situation, the appropriate approach is to use cost effectiveness analysis and propagate uncertainty in estimates through the decision model using probabilistic sensitivity analysis. The use of CMA in this setting needs a strong rationale for clinical equivalence and an awareness of often rapidly changing evidence for comparator products.

Cost analysis for standard CMAs should be relatively straightforward since there should only be differences in a limited number of resources covering aspects of acquisition, administration and monitoring. Other costs only become relevant if there are differences in clinical effectiveness, which then negates the use of CMA. Where parameter uncertainty is negligible, total mean costs for a new technology should not exceed the costs of the comparator(s) by any margin.

CMAs conducted within the MTEP process for technologies that do not claim clinical equivalence should be based on cost models constructed using existing standards and time frames that are based on established principles for cost-effectiveness models. The impact of uncertainty in parameter inputs on cost estimates may usefully be described by the use of probabilistic sensitivity analysis. Individual parameter inputs require detailed examination.

Recommendations on clinical effectiveness

1.   Cost minimisation may be used in different circumstances. For all programmes, the option of pursuing CMA when there are plausible claims that the technology in question is clinically equivalent to relevant comparators allows a simpler and potentially faster analysis to be performed. These are a limited set of circumstances and should be applied cautiously.

2.   The principles on which the decision to proceed using CMA in the case of claimed clinical equivalence should be the same for all programmes.

3.   Where the claim of clinical equivalence is made, there must be a detailed consideration of the plausibility of the claim both when making the decision to proceed using CMA, and as part of committee deliberations.

4.   In the case of pharmaceuticals, there should be consideration of the biological plausibility of the claim and the extent to which the mechanisms of action of the new and reference drug differ. Any differences need to also assess how patient outcomes might be affected.

5.   Similarly, for all other technologies, the foundation for any claims of clinical equivalence needs detailed scrutiny from relevant experts in order to assess plausibility. For all programmes, it is important to consider how any differences would be expected to manifest themselves in patient outcomes, particularly those that may be important to an economic model of cost effectiveness.

6.   Where technologies have significantly different adverse event profiles, they should not be considered non-inferior, and thus the cost minimisation route is not appropriate because the implications of these differences in health outcomes measured as QALYs needs to be calculated.

7.   The assessment of clinical non-inferiority requires consideration of both statistical and clinical significance *inter alia*. All outcomes that relate to different aspects of patient benefit should be considered with particular focus on those which are key drivers of an economic model. An assessment of non-inferiority should be required for all important outcomes, not just the primary outcome.

8.   It is useful to consider whether differences in clinical effectiveness plausibly span some non-inferiority margin.

9.   The non-inferiority margin may often be informed by expert patients and clinicians. Independent review groups can help to identify important parameters and

provide information on the assessment of the appropriate margin, informed by relevant experts. Caution is required in the interpretation of individual level differences versus population level differences. Sometimes a pragmatic approach is required to define a non-inferiority margin.

10. Interpretation of results needs to consider the design of studies from which estimates are drawn, including their sample sizes.

11. The impact of differences in intermediate outcomes, such as sensitivity and specificity, needs careful consideration.

RECOMMENDATIONS ON THE ASSESSMENT OF COSTS

12. Where CMA is undertaken because of clinical non–inferiority, there may be very little uncertainty about the cost implications. In this situation, there should be no margin allowed for the preferred estimate of mean cost to be greater than the relevant comparator(s).

13. Where CMA is undertaken and there are claims of differences in clinical benefits, the parameter estimates for individual components of resource use and unit costs and their associated uncertainties should be examined proportionate to their contribution to the cost difference.

14. Probabilistic sensitivity analysis may be particularly useful where sampling uncertainty for key parameters is the primary source of uncertainty.

1. CONTENTS

2. INTRODUCTION

Economic evaluation entails the comparison of the costs and benefits of two or more comparators[1] (Drummond et al. 2005). The aim of such analyses is to support decision makers by providing information on the opportunity costs of alternative funding decisions, given limited resources[2].

In most of its guidance producing programmes, NICE has tended to use cost-utility analysis as the primary method for assessing health technologies. However, in some situations, NICE makes use of "cost minimisation analysis" (CMA), though the use of this terminology in formal NICE documentation is variable.

This project examines the use of CMA methods across NICE's guidance producing programmes. We assess when CMA is used, when it is not and how CMA methods are applied. Where differences are identified, we question whether they are justified given the types of technologies under consideration by the various NICE programmes, the types of evidence they typically receive, and the standing orders of the programmes. We were asked to make recommendations for the way CMA methods are applied which may inform changes to the Methods Guides that relate to each programme.

We report here on reviews of relevant NICE Methods Guides, relevant published literature, international approaches to the use of CMA and discussions with key stakeholders to the various NICE guidance producing programmes. These sources of information are used to highlight a number of issues and recommendations for NICE to consider as methods and processes are updated.

3. TYPES OF ECONOMIC EVALUATION

There are three primary forms of economic evaluation: cost-effectiveness analysis (CEA), cost-utility analysis (CUA) and cost-benefit analysis (CBA). Whilst each of these value the costs of competing alternatives in the same way, they differ in terms of how consequences are considered. CEA uses a single measure of effect that is common to both alternatives. Typically, these are some kinds of "natural units" such as life years gained or points of improvement on a clinical outcome scale. CUA combines potentially multiple effects into a single healthy years metric, most commonly the Quality Adjusted Life Year (QALY). It is CUA that has been used as the mainstay for NICE Technology Appraisals since their inception. CBA is an approach that takes a range of different consequences and expresses their value in monetary units, thereby making the benefits commensurate with the costs.

To complicate otherwise simple terminology, some analysts use the term "cost consequences analysis" (CCA) to refer to a variant of CEA where an array of outcomes may be presented alongside costs, leaving decision makers to consider how these should be traded off against each other. Furthermore, some authors do not distinguish between CEA and CUA.

However, what is clear is that CMA is considered only a partial economic evaluation because, on its own, a consideration of costs neglects any consideration of the outcomes. CMA could be considered a full economic evaluation if there is a preliminary step where outcomes are considered and, if they are considered equivalent for different options, then a CMA becomes relevant. Without this preliminary step, CMA is not an appropriate form of economic evaluation. And even with such preliminary considerations of effectiveness, this is not a straightforward issue. There is a requirement to think about what equivalence means, and how uncertainty, including the joint uncertainty between costs and effects, should be taken into account in order to justify the use of CMA. We expand on these considerations using evidence from a systematic review of relevant literature in Section 5.

This report does not directly address the inclusion of biosimilars (or generics) in economic evaluation. NICE is able to apply the same remit and guidance as their

reference products, and as such, implicitly makes the judgement that the products are equally effective[3]. CMA is not applied in these circumstances as biosimilars are not identified as separate comparators, although NICE guidance usually promotes cost minimisation between products in positive guidance by recommending that the least costly product is used.

4. CURRENT NICE METHODS GUIDANCE

4.1. MEDICAL TECHNOLOGIES EVALUATION PROGRAMME (MTEP)

MTEP issues guidance on medical technologies covering medical devices, including diagnostics. Established in 2010, the programme has always used cost minimisation as the key part of its methods.

The current MTEP methods guide[4] sets out the approach in more detail.

MTEP operates a "single technology appraisal" process – this means it evaluates and issues guidance on the single technology that has been notified to NICE. Other technologies could be relevant as comparators. In this approach, the assessment compares to current NHS practice, or range of practices, rather than comparing to similar technologies in a similar class.

The methods guide refers to the use of *"cost consequences analysis"* (CCA) throughout (see for example section 7.3). CCA is a less formal approach to economic evaluation. It treats the benefits of a technology as comprising a range of different effects, each of which may be measured in their own natural units. For example, there may be treatment effects on a primary clinical measure, impacts on hospitalisations or future complications of disease, there may be differences in adverse events of different types and different mechanisms for administering the treatments. The new technology may be better than comparators on some, all or none of these aspects of benefit. In a cost consequences analysis any additional costs are assessed by decision makers in a qualitative manner to determine if the net benefits of the new technology justify additional costs. There is no attempt made to explicitly aggregate these different effects by expressing them in some common metric.

However, despite the terminology and description of the CCA methodology, MTEP methods in fact operate a form of analysis that is closer to CMA than CCA. This is because the decision rules operated by the programme outlined in section 8.2, state that the case for recommending a technology require either that:

- *there is sufficient certainty that the technology has at least equivalent clinical and/or health and social care system benefits compared with current management, and overall uses less resources or*
- *there is sufficient certainty that the technology has significantly greater clinical and/or health and social care system benefits compared with current management, and overall uses similar resources.*

There is therefore no option for the committee to recommend technologies that lie in the "North East quadrant" of the cost effectiveness plane (i.e. those that offer greater benefits but at increased cost), which would require a consideration of the value of the additional benefits, although the phrase "similar resources" may give committees some leeway to recommend technologies with small incremental costs. This is therefore not pure CMA because there does remain open the option to issue positive guidance in the case where costs are considered to be similar, provided that there are positive benefits.

MTG28 "Spectra Optia for automatic red blood cell exchange in patients with sickle cell disease" provides an example of how CMA is typically conducted in the MTEP. Spectra Optia comprises a machine and software for automated red blood exchange for patients needing treatment for sickle cell disease. Sickle red blood cells are replaced by healthy red blood cells according to a user defined software protocol. The comparator is manual blood exchange. Whilst the direct cost of the machine itself is much higher than manual exchange, there are numerous factors that potentially offset that cost: the less frequent need for blood exchange, lower requirement for high cost chelation therapy, improved outcomes like stroke and pain crises. A decision tree cost model that incorporated all of these differences was required in order to estimate the net cost impact of Spectra Optia. The assessment was that this was clearly both clinically effective and cost saving. It is worth noting that, had the cost case indicated that Spectra Optia was likely to be cost incurring, the additional work required to

undertake a cost-utility analysis would have been minimal, since the same health states and events used to generate costs would simply need health utilities attached.

4.2. CLINICAL GUIDELINES

The NICE Guidelines Methods Guide outlines a "reference case", the preferred set of methods and approaches to economic evaluation that should be applied in order to achieve consistency and comparability across pieces of guidance. For interventions funded by the NHS and PSS with health outcomes, the preferred method is cost utility analysis. For interventions funded by non-NHS public sector bodies and that generate either health and non-health outcomes, or with a social care focus, cost-utility analysis is preferred for the base case, but all other types of analysis, including CMA, can be considered. The guide also states that CMA studies may be included in the systematic review of economic studies.

Page 137 of the guide gives a standard description of the CMA approach. It correctly highlights that:

> *"The disadvantage of cost-minimisation analysis is that the health effects of an intervention cannot often be considered equal to those of the status quo."*

Developers of clinical guidelines are often faced with a large number of questions about appropriate management of a disease area. It is rarely possible to conduct comprehensive economic analysis for more than a small number of priority areas. Therefore, an element of pragmatism is required. This means that an estimate of cost can often be sufficient, if an intervention is clearly more effective and associated with fewer adverse events than its comparator (see page 138). The Methods Guide does not set out a requirement for new interventions to be cost saving or cost neutral.

Section 7.7 of the Methods Guide sets out how economic evidence is to be used in formulating guidance and p152 specifically relates to CMA. The methods guide suggests CMA is for those situations where it is known, a priori, that effect differences may be small but cost differences large (presumably in favour of the new intervention). It suggests that ideally, claims of small differences in clinical effect would be established via an equivalence trial, and

13

"[for] this reason, cost-minimisation analysis is only applicable in a relatively small number of cases."

It is worth noting that there are multiple definitions that appear regarding the assessment of clinical benefits (e.g. "small differences", "equivalence") but no detailed instruction on how these should be assessed.

4.3. TECHNOLOGY APPRAISALS

The 2013 Guide to the Methods of Technology Appraisal[5] states that CUA is the reference case approach for TA. However, an Addendum to the Methods Guide (undated), specifically addresses the case of cost comparisons (CCs), which are deemed suitable when a new technology is considered to provide *"similar or greater health benefits than technologies recommended in published NICE technology appraisal guidance for the same indication."*[6]

All other aspects of the methods guide, including the reference case, except those that only pertain to cost-utility analysis, remain relevant in a CC, according to this document. This option allows manufacturers to apply for a fast track appraisal (FTA) which offers a lighter touch review of evidence, and the prospect of a quicker recommendation and funding (within 30 days).

When the fast-track route is sought, the claim of "clinical similarity" needs to be supported by evidence in the manufacturer submission. This evidence is used initially by NICE to determine if it is appropriate for the appraisal to proceed using the FTA process, and then at the appraisal of the evidence by the committee. The committee will recommend a technology that generates similar or greater benefits at a similar or lower cost. Therefore the wording of the TA Methods is very similar to the MTEP methods.

The User Guide for CCs[7] provides further detail on a number of points outlined in the Addendum to the Methods Guide. From this, four points are of note. First, it is stated that cost comparisons should relate to a technology that *"is likely to provide similar or greater health benefits at similar or lower cost than technologies recommended in published NICE technology appraisal guidance for the same indication".* There is potential ambiguity about whether the CC approach described is strictly a CMA: much

depends on the interpretation of the term "similar". One such interpretation is that a technology could generate marginally less health benefit than its comparator (but still be seen as "similar") whilst simultaneously generating marginally higher costs (but still seen as similar). As is the case with the MTEP methods guide, it is worth noting that a CMA would rule out a technology with higher, yet similar, cost.

Second, the user guide states that it *"is acceptable to make a cost-comparison case with only 1 of the comparators in the scope, provided that the selected comparator satisfies all of the following criteria:*

- *It adequately represents the NICE recommended treatments as a whole, both in terms of its costs and effects.*
- *It has a significant market share.*
- *It is recommended in published NICE technology appraisal guidance for the same indication."*

As such, CC does not require a full incremental analysis. In combination with the first deviation from CMA described above, it is therefore theoretically possible for a product to be recommended when it is worse (but similar) and more expensive (but similar) than another product that is less cost-effective that another one of its comparators. It is, however, recognised that the implications of these two deviations from CMA are theoretical; the wider process of scrutiny within FTA may very well prevent such occurrences. For example, the NICE technical team is responsible for determining whether the selection criteria are met, including those relating to the comparator.

Third, the User Guide states that effectiveness within a CC is assessed in the same way as for a Single Technology Appraisal (STA); it should include a full search, critical appraisal and appropriate synthesis. This should relate to all relevant comparators and not just the one selected for the cost-comparison. It is not clear what happens if a network meta-analysis is undertaken that shows the new technology is not statistically different from one comparator, but is statistically worse than another.

The User Guide also states that to support the quantitative evidence, manufacturers are required to provide evidence *"on the clinical or biological plausibility of similarities in health benefits between the technology and the comparator(s)".* The nature of this evidence is not described.

Fourth, the User Guide makes it clear that a CC is not just a comparison of acquisition costs. Costs should include the acquisition cost of the product over the treatment period and resources directly related to its administration, for example, hospital appointments and monitoring. In addition, the guidance states that:

> *"If there are relevant differences in health outcomes that affect resource use (for example, managing adverse events), the time horizon must be long enough to capture these. Substantial differences between technologies in costs directly relating to health outcomes (such as adverse events) indicate that the intervention and comparator(s) may not provide similar overall health benefits, so any such cost difference must be clearly justified." (p31-32)*

Another set of costs that could be relevant is highlighted in internal documents, namely, costs associated with any impact on subsequent treatments. The meaning here is not stated. It could be that this is intended to relate to impacts on sequencing, rather than any argument relating to the duration of later treatments caused by improved health relating to the new product.

A related issue to which categories of cost should be included is how the inclusion of a relevant cost is determined. For example, which side-effects should be included; just those that are 'statistically significant'? This issue has wider implications for the way in which effectiveness is included within the cost-comparison, as many costs will be directly related to effectiveness (e.g. length of treatment and progression free survival in many cancers); are these costs calculated on the 'statistically insignificant' difference in outcomes, or on the assumption of equivalence?

4.3.1. Application of CMA in Technology Appraisals

We are aware of three published TAs where the CMA approach has been formally applied as part of the FTA process: Guselkumab for psoriasis (TA 521), Golimumab for non-radiographic axial spondlyloarthritis, (TA 497) and Aflibercept for choroidal neovascularisation (TA 486).

TA521

In the case of TA521, the proposal was for the use of guselkumab in a population consistent with NICE's previous recommendations for biologics, albeit narrower than the marketing authorisation. Guselkumab is a monoclonal antibody (mAb) to the interleukin (IL)-23 protein. This differs from the TNF-alpha inhibitors (adalimumab, etanercept, infliximab). Other comparator mAbs were ustekinumab (IL-12 and IL-23), secukinumab and ixekizumab (IL-17A). Guselkumab is described as *"a novel biological therapy that selectively targets IL-23 and is the first-in-class selective IL-23 inhibitor approved to treat moderate-to-severe plaque psoriasis."*[8]

There was direct and indirect trial evidence that guselkumab is more effective than TNF-alpha inhibitors and ustekinumab. Specifically, there was direct evidence from 2 trials that guselkumab was associated with a statistically significant improvement in PASI75 response rates against adalimumab. A network meta-analysis of 45 trials suggested statistically significant improvements in PASI75 response rates compared to adalimumab, etanercept, infliximab and ustekinumab. There was also a statistically significant improvement compared to secukinumab but the committee felt "the difference might not be clinically meaningful", though it is unclear what this means in the context of an outcome that relates to patient response. There was no statistically significant difference to ixekizumab. The committee felt that results relating to PASI100 were broadly similar and that the safety profile was similar to other biologics.

The company presented a cost-comparison analysis that modelled the total costs of guselkumab, adalimumab and ustekinumab treatment. The cost case against adalimumab and ustekinumab showed (in the ERG analysis) that guselkumab was more costly, presumably because there is a higher stopping rate in the less effective therapies. The committee stated that these were not acceptable comparators for the cost-comparison case, since they do not provide similar clinical benefits. The committee therefore decided that the comparison for the cost analysis should be secukinumab and ixekizumab. This is important because it appears that there are two comparators against which guselkumab is both more costly and more effective and therefore a cost utility analysis would ordinarily be required. However, the committee did not compare against these options because the comparators that they did compare to (secukinumab and ixekizumab) had already been deemed cost effective in previous

NICE appraisals and were expected to be in widespread NHS use either at the time or in the near future.

The evidence also suggested similar health benefits to two other NICE recommended drug treatments, against which, guselkumab has either the same or lower costs.

TA497

Golimumab is a TNF inhibitor. TA497 compared golimumab to three other TNF inhibitor biologic drugs, approved in previous NICE guidance, for treating severe non-radiographic axial spondyloarthritis: adalimumab, etanercept and certolizumab pegol. A single RCT showed superior outcomes for golimumab vs placebo, and those trial results were then used in a network meta-analysis with the trials of the other three biologics. The published guidance reports that the committee considered the results showed that golimumab was "similar" to the other three for some outcomes and statistically significantly superior for others, leading to an assessment that the clinical effectiveness of golimumab was likely to be similar to those of the comparators. There was no formal analysis of adverse events reported in the guidance. It is reported that the adverse event profile is well established for both golimumab and the comparators and the committee believed them to be similar.

For the cost case, it was accepted that all costs, other than drug acquisition costs, would be identical. For the drug costs, golimumab was shown to be the same as adalimumab, lower than etanercept, and higher than certolizumab pegol in the first year, but lower in subsequent years. Overall, costs were believed to be similar.

TA486

TA486 considered that there was a single relevant comparator, ranibizumab, already recommended by NICE, to aflibercept for the treatment of visual impairment because of myopic choroidal neovascularisation. Both of these drugs are antivascular endothelial growth factor (VEGF) therapies.
An indirect comparison of the two drugs showed that there were no statistically significant differences between them in terms of clinical effectiveness. Whether the point estimates favoured one or the other depended on the retreatment criteria for

ranibizumab. The committee also heard from clinical experts who stated that aflibercept is considered slightly more effective than ranibizumab,

Regarding adverse events, the committee refer to direct evidence from a trial of patients with wet age related macular degeneration which showed them to be similar. It was not possible to perform an indirect comparison but the rates and types appear similar.

For costs and resource use, the evidence and clinical opinion about the number of injections required was considered. These were judged to be equal. Overall costs, taking into account the patient access schemes, meant aflibercept was similar or lower than those of ranibizumab.

In these three examples from the FTA process, there is some variation in the types of drugs and their comparators that have been considered. In TA497 and TA486, it is clear that the new drug under appraisal can be considered a "me-too" drug to the comparators – they are chemically related or have an identical mechanism of action. In the case of TA497, this is true for all three of the NICE recommended drugs treated as comparators: all are TNF inhibitors. In TA486, there is a single comparator that is also another VEGF therapy. The case of TA521 is different in this respect and more complex. Of the six NICE approved comparators to Guselkumab, 3 were TNF-alpha inhibitors and the other three block different target receptors (ustekinumab targets both IL-12 and IL-23). In this example, guselkumab is a first-in-class therapy as "the first biologic to selectively block IL-23"[9].

7 other appraisals that had used CMA but were not formally designated as FTA's were identified by NICE staff: TA131, TA183, TA184, TA310, TA397, TA425 and TA426. See Appendix Table 1 for more details of these appraisals.

Key issues that arise from these examples of CMA are that:
- In most of these cases, cost utility models were constructed and the results of these models were presented to committee. The reasons for subsequently conducting CMA, or restricting the considerations of the committee to an analysis of costs alone, was either because of a lack of data to allow the CUA to be considered robust, or that the committee's assessment of the clinical effectiveness evidence led them to conclude that the technologies and

comparators were similar. This is different to the use of CMA in the FTA process that is underpinned by a decision that CMA is warranted *a priori.*

- In TA184, the only comparison where a cost comparison was used as the basis for decision making was for one subgroup of patients where the comparator was the combination of cyclophosphamide, doxorubicin and vincristine (CAV). Here, the committee judged the clinical effectiveness to be similar between CAV and topetecan. There was no cost effectiveness model provided for this comparison, and topetecan was considered to be more costly. The committee stated that the required effectiveness difference required for this to be considered cost effective was unlikely to be achieved.

- Two examples are partial reappraisals of one of three tyrosine kinase inhibitors, in two slightly differing conditions, as part of the Cancer Drugs Fund reappraisals in 2016. The three drugs had been appraised using standard methods previously in TA425 and TA426. In both cases, full cost utility models had been considered as part of the initial appraisals but for the reappraisals CMA was used based on the conclusions of clinical equivalence between dasatinib and nilotinib from the previous appraisals. There was no new clinical evidence to change those judgements by the committee.

5. INTERNATIONAL APPROACHES

5.1. CANADIAN AGENCY FOR DRUGS AND TECHNOLOGIES IN HEALTH (CADTH)

CADTH and the associated pan-Canadian Oncology Drug Review (pCODR) are sometimes identified as allowing CMA. In its 3rd Edition Guidelines (2006), it states that "A cost-minimization analysis (CMA) is appropriate as the Reference Case when the evidence shows that the important patient outcomes of the intervention and comparators are essentially equivalent".[10] Consequently, CMA was used in several HTAs, with each use of the method requiring an explicit justification, for example for the pCODR appraisal of axitinib for metastatic Renal Cell Carcinoma, it is stated that:

> *"In the absence of head to head data, in order to support equivalent efficacy of axitinib and everolimus, the manufacturer conducted three different approaches for indirect treatment comparisons: a side by side comparison, the Bucher fixed effect model, a Bayesian fixed-effect model, and a simulated treatment comparison. Based on the pCODR Clinical Guidance Report, conclusions*

drawn from such indirect comparisons are not as robust as conclusions based on direct, head-to-head trial data and there are some serious limitations which need to be considered when interpreting the results of indirect treatment comparisons. However, the pCODR Clinical Guidance Panel concluded that the clinical effects of axitinib and everolimus appear similar and it is on this clinical basis that a cost minimization analysis was considered justifiable." (p3)11

However, the 4th Edition of the Guidelines adopts a different position:

"In the reference case, the economic evaluation should be a cost-utility analysis (CUA) with outcomes expressed as quality-adjusted life-years (QALYs). Any departure from this approach should be clearly justified." (p17)[12]

"A cost-minimization analysis (CMA) is a costing exercise and not a formal economic evaluation. As such, a CMA is not an appropriate reference case analysis. A CUA remains the appropriate approach, even where convincing evidence is available to show that important outcomes are similar, as it allows for the analysis of the uncertainty in incremental effect (through probabilistic analysis), facilitating the necessary comparison across all technologies." (p17)[12]

As such, the most recent guidance sets out a more limited role for CMA within CADTH. Non-reference case analyses using CMA are presumably allowed, but no example of this has been identified. What use these additional analyses would be is unclear as a CUA will always be available to decision makers in the reference case.

5.2. AUSTRALIAN PHARMACEUTICAL BENEFITS ADVISORY COMMITTEE (PBAC)

PBAC provides guidelines for submissions and includes a section on the circumstances when CMA may be used and the preferred methods associated with applying the approach[13].

The guidance stresses the importance of non-inferiority in both clinical effectiveness and safety for CMA to be considered appropriate. This is based on the assessment of clinical evidence as for a standard assessment. No specific reference is made to the evidence requirements to support this claim. CMA can be used where the new technology has a safety profile that is non-inferior or superior to the comparator. Claims of non-inferiority require justification, though it is not stated whether this requires an explicit statement of the chemical similarities of the drug and its

comparators, or if this claim can be supported solely based on the clinical evidence. However, Taylor et al suggest that for PBAC these cases are for drugs expected to be non-inferior due to being in the "same therapeutic class or biosimilars"[14]. They cite examples of the use of CMA by PBAC in the case of a vaccine for the prevention of diphtheria, tetanus and pertussis and an infliximab biosimilar[15].

The guidance also requires clarity on the doses that provide clinical equivalence.
The cost approach is then standard for CMA. Costs of administration and safety management profiles must be presented and any other costs of cost offsets summarised.

5.3. ISPOR SUMMARY OF HTA GUIDELINES

According to the ISPOR summary of pharmacoeconomic guidelines around the world, there are 10 other sets of guidelines that permit the use of CMA methods in certain circumstances (see

Appendix Table 2). This is not a count of country specific guidelines, because these may refer to regionally agreed documents. Nor does the ISPOR summary provide sufficient detail to understand the circumstances when CMA may be used or how frequently that occurs

6. REVIEW OF ACADEMIC LITERATURE

6.1. METHODS

The aim of the literature review is to identify and critically appraise publications that examine the validity of CMA in the assessment of health technologies. Experience of previous 'methodological reviews' has shown that simple searching on keywords or MeSH search tags is difficult as the classification systems are not specific enough. Conversely, searching for a set of related terms within the title and abstract can sometimes produce too many records. Consequently, we undertook a targeted search strategy.

Records were to be screened by reading title and abstract (if available), with those papers classified as being papers undertaking a detailed examination of the CMA methods being retrieved for full text review. At this point, all papers found to meet the original classification were summarised and critically assessed.

The strategy was to first search for all English language papers in Medline, Embase and Google Scholar with "cost minimi?ation" within the title and abstract. To this we added all papers that had cited the seminal paper by Briggs & O'Brien (2001). Finally, we excluded duplicates and papers before 2001. The results from this were as follows:

- Medline - 582 records
- Embase - 1131 records
- Citation search in Google Scholar - 392
- Total number of records - 2105
- Removal of pre-2001 records and duplicate titles - 1304 unique records

The first 101 of the 1304 records was screened and failed to identify a single methods paper. Given the low yield of appropriate papers, a review of all papers was abandoned and replaced by a more targeted review of papers focusing purely on the 372 papers that cited the Briggs and O'Brien paper. This is not thought to introduce a significant selection bias in favour of papers agreeing with their position as anyone suggesting a contrary position would still be expected to make reference to this paper. However, to mitigate this potential bias, we also undertook two additional searches

which limited the original search to papers with "reimbursement" or "health technology assessment" in the title or abstract; these produced 65 and 25 papers respectively.

6.2. RESULTS

The citation search identified 9 methods papers, whilst the reimbursement and health technology assessment searches identified 0 and 1 methods papers, respectively. Two papers by O'Brien and Briggs are excluded as they effectively repeat the same argument, including using the same case study. The paper by Bosmans is excluded as it examines notion of economic non-inferiority in equivalence trials, which is not relevant. The remaining papers, plus two additional papers known to the authors of this report, plus the original paper by Briggs and O'Brien are summarised below:

Briggs and O'Brien (2001)[16]

The main conclusion from this work is that cost-effectiveness analysis should estimate the joint density of cost and effect differences rather than undertake sequential hypothesis tests of costs and effects using arbitrary type I error rates. As such, point estimates of incremental costs and effects should be used, together with the associated uncertainty relating to the joint distribution.

However, a proviso is added to this following the presentation of a case study which is used to suggest that CMAs can be conducted alongside equivalence trials; the reason being that sample sizes within such trials are much greater and so 'lack of a difference' is determined with greater certainty.

One further point of interest is mentioned in passing within the paper; the more comprehensively effectiveness is defined, the less likely equivalence will be established. The reasoning behind this is that the variability of narrowly defined outcome measures for a given health problem are generally less than those of broadly defined outcome measures for the same problem. This is important for economic evaluation as the QALY falls into the latter category, and as such, following their line of thought, equivalence of QALYs is unlikely. As such, an implication of the Briggs and O'Brien argument is that CMA is generally only relevant to trials with a single outcome measure of interest.

We will return to this issue later, but it is of note that the view that evidence derived from equivalence trials is sufficient to avoid the need for a cost-effectiveness study is no longer generally accepted by health economists. It is also of note that both 'significance level' and 'equivalence margin' are both subjectively determined, and as such, equivalence trials are subject to the same criticism of arbitrariness that Briggs and O'Brien direct toward superiority trials.

Johnston et al (2003)[17]

This paper is supportive of the Briggs and O'Brien argument. Much of the paper by Johnston and colleagues focusses on a case study that showed superiority of a new intervention based on an intermediate end-point, but equivalence for life-years gained. Their key message – the danger of using intermediate outcomes in economic evaluation which is exacerbated by publication bias – is not relevant to our review. However, their discussion of cost-minimisation implies an important practical problem of CMA; with studies that include multiple end-points, it is unlikely that equivalence is achieved (or powered) for all endpoints, and as such, CMA is rarely legitimate.

Span et al (2006)[18]

This paper is supportive of the Briggs and O'Brien argument, but makes one minor amendment; the legitimate use of CMA should be extended to use with trials designed to prove non-inferiority. The reasoning being that "because non-inferiority covers a broader range of possible outcomes (not only equal, but also possibly better)".

Stewart et al (2010)[19]

This paper is critical of the Briggs and O'Brien argument that equivalence trials are required before CMA should be considered; this criticism is not direct, but implied by their examination of the use of CEA in the evaluation of biosimilar pharmaceuticals. They note that generic pharmaceuticals are assumed to be equivalent to the index product and are essentially evaluated via a CMA restricted to acquisition price by most HTA agencies. In contrast, they report that biosimilars are treated as sufficiently different products and are therefore evaluated using full submission, including CEA, by most HTA agencies. Whilst they recognise that biosimilars can be different chemically and require effectiveness data, bioequivalence for generics is commonly based on the pharmacokinetic measures with equivalence defined by wide ranges.

Their criticism is based on their view that the legitimate use of CMA requires three criteria to be met; *"Firstly, an a priori expectation that the treatments should perform equally…secondly, pharmacodynamics and pharmacokinetic evidence that treatments are clinically equivalent…[and]…thirdly, …evidence from adequately designed and powered equivalence or non-inferiority studies".*

Their basic argument is that for biosimilars, *"in the absence of compelling reasons….HTA bodies should accept CMA as the basis of the cost-effectiveness deliberations."* However, they rein back on this as a blanket proposal by arguing that this should only apply to studies where comparative data with the reference product are available, whilst in other cases threshold analysis could be used to explore relative efficacy. For example, they suggest that *"the HTA authority should define an acceptable threshold for treatment effect. For example, if a 30% difference in treatment effect crosses the threshold, the HTA body could reasonably approve the biosimilar. If a 5% difference crossed the threshold, the HTA body might reasonably reject the biosimilar".*

Haycox (2010)[20]

This chapter of a book is supportive of the Briggs and O'Brien position around the use of CMA in relation to Equivalence Trials (ETs). It does, however, provide explicit guidance in relation to the legitimate use of CMAs. Firstly, ETs provide the best evidence, but even with these, the decision maker should consider the extent to which the primary outcome measure fully captures the benefits of the intervention. If other benefits are important, additional comparisons of clinical equivalence may be required. Secondly, decision makers should recognise that failure to show superiority is not the same as equivalence. Thirdly, if non-inferiority is to be used to justify a CMA, proof should be available that demonstrates that non-inferiority is an acceptable approximation of therapeutic equivalence. Finally, the decision maker should be open to use their judgment as to the quality and relevance of the evidence in relation to therapeutic equivalence.

Dakin and Wordsworth (2013)[21]

This paper examined three aspects of CMA in decision making; (i) the prevalence of CMA before and after the Briggs and O'Brien paper, (ii) the impact of using CMA in two trials where there were non-statistically significant effects, and (iii) the impact of

using CMA in ten simulated datasets with varying incremental costs, effects and associated standard errors. In the third set of analyses, the impact of using CMA was assessed in terms of producing the wrong decision, the error probability and the impact on value of information (VOI). The impact on VOI was raised as an issue as CMA is not only based on the premise of zero incremental effect, but also zero uncertainty relating to this estimate of incremental effect; this interpretation is a development to the original Briggs and O'Brien idea.

Their work showed that the proportion of published economic evaluations that were CMAs reduced form 8.4% in 1999 to 0.5% in 2009 and the rate of decline accelerated after the publication of the Briggs and O'Brien paper. Their reanalysis of two trials showed that CMA generated the wrong cost-effectiveness conclusion in the first study and seriously underestimated the VOI in both studies (which arguably, should have led to an 'only in research' conclusion in the second study, thereby suggesting that the conclusion of that study was wrong, too).

The simulations produced several important results, with two useful rules of thumb being generated, whilst noting that in all other situations "CEA is necessary to inform decisions about current resource allocation and future research". The rules of thumb are:

- CMA is only appropriate "where the difference in costs is sufficiently large that no plausible difference in efficacy could change the conclusions or uncertainty estimates".
- CMA is also appropriate "where one intervention is significantly more effective and significantly less costly".

The authors also note that "bias within uncertainty estimates from CMA will also be negligible in these cases because error probabilities and VOI will approach zero".

The authors then highlight one proviso; noting that additional costs are required to undertake a CEA, they suggest that "it may, therefore, be appropriate to consider whether the reduced risk of bias is worth the additional research cost of conducting CEA". In these situations, Dakin and Wordsworth suggest that the problems associated with using CMA could potentially be mitigated by undertaking sensitivity analyses based on simplified CEA model.

In summary, this work highlights several important issues. First, the impact of using CMA is two-fold; bias in the estimate of incremental net benefit (INB) and bias in the expected value of perfect information (EVPI), with both of these being of importance to funding decisions. Second, non-significant differences in effects can lead to incorrect funding decisions when using CMA. Thirdly, more restrictive rules of thumb may be of use to decision makers relating to the size of incremental effect, anticipated size of incremental cost, uncertainty of both and the cost of further research.

Eckerman (2017)[22]

This discussion of CMA is limited to a sub-section within a chapter of a book. It is supportive of the notion of CEA in <u>all</u> situations in order to "avoid partial analysis biases". The general argument is that the preferred treatment in a CMA is rarely dominant; small incremental health losses, together with cost savings place the intervention in the south-west quadrant of the cost-effectiveness plane, which demonstrates that a trade-off is required. In fact, Eckerman takes this argument further, claiming that in many circumstances CMA could recommend an intervention that is dominated by the alternative if the ('statistically insignificant') health losses have costs associated with them that have not been included in the CMA:

> *"That is, effects ignored are usually also expected to have cost implications for the health system where they have associated treatment, such as hospitalisation for various forms of morbidity or treatment of side effects. Hence, an intervention which on the face of a cost minimisation analysis has lower direct cost and 'equivalent effect' can easily represent worse effects, but also higher health system costs in appropriately including the cost of treating such effects". (p64)*

Eckerman also highlights one potential problem with CMA not highlighted elsewhere, namely that its use can provide an incentive for manufacturers to under-power studies in order to produce 'statistical insignificance', thereby giving them the excuse to undertake a CMA.

Russo et al (2018)[23]

This paper uses a case study examining the use of oral anticoagulants for the prevention of stroke in atrial fibrillation. This is then used to argue that the CEA, which identifies apixiban as the most cost-effective treatment, is misleading as the underlying

clinical evidence shows no statistically significant differences in efficacy and the economic modelling show no statistically significant differences in costs or QALYs. As such the paper is a reiteration of pre-Briggs approach of sequential hypothesis testing, without any consideration of the expected value approach that has been widely accepted by health economists since then.

Nuitjen (2004)[24] and O'Hagan et al (2005)[25]

This refers to a pair of papers. First a proposed analytical framework by Nuijten, which was followed by a critique by O'Hagan and colleagues. Consideration of these papers is considered relevant here as the proposed framework is, in essence, a CMA approach, whilst the critique highlights several problems that are not directly addressed in other papers.

The Nuijten approach extends the idea of using non-statistically significant differences in effect to indicate the appropriateness of CMA, but additionally considering non-statistically significant differences in cost. These two ranges are then used to identify an 'area of indifference' on the cost-effectiveness plane. Nuitjen then extends this further by specifying 13 zones within the cost-effectiveness plane which he uses to assess decision uncertainty for an exemplar cost-effectiveness analysis. O'Hagan and colleagues highlight several problems including:

- It requires decision makers to specify thresholds for meaningful differences in costs and effects. They further note that because the opportunity cost of a making the incorrect decision is dependent on the size of the patient population it relates to, then these thresholds and the 'area of indifference' will be decision specific.

- A full cost-effectiveness analysis, with PSA, is required to apply the decision framework, thereby undermining the need for reanalysis within another framework.

- 'Areas of indifference' undermine the correct decision rule, which is "to choose the treatment with largest expected net benefit conditional on a threshold value for a unit of health gain".

- 'Areas of indifference' undermine the correct basis for identifying future research, which is the use of the VOI framework.

6.3. SUMMARY OF REVIEW FINDINGS

The review shows that there is almost universal support for the view that CMA is an inappropriate method of economic evaluation as it can lead to incorrect conclusions compared to those based on expected net benefit. Whilst Briggs and O'Brien[16] and several other authors[17,18,20] have focused on economic evaluations alongside controlled trials, the same conclusion is relevant to HTAs. In addition, HTA brings further practical problems of using CMA (e.g. the need for equivalence across all outcome measures and all comparators).

As noted earlier, the argument against CMA has actually strengthened since 2001 among health economists by arguing that there are no situations where CMA is appropriate.[21,25] This is partly due to a growing recognition of the irrelevance of inference to funding decisions,[26] but also a greater appreciation that CMA will also undermine recommendations for future research priorities by biasing estimates of uncertainty and VOI.[21] A further consequence of the bias introduced to VOI analysis that has yet to be recognised, is that health economic analyses to support only in research (OIR), approval with research (AWR) and managed entry agreements (MEA), will also be undermined.[27,28] In practical terms, it should be recognised, however, that further research recommendations are rarely made on the basis of VOI and the proposed quantitative methods for identifying possibilities for OIR, AWR and MEA are not used, either.

However, practical considerations have been identified by some authors which has led them to provide advice relating to situations where CMA is less likely to provide misleading conclusions. In relation to the decision making context faced by NICE, the most relevant appear to be:

• The decision makers should consider the extent to which the primary outcome measure fully captures the benefits of the intervention.[20]

• The need for a strong a priori expectation that the treatments should perform equally.[19]

• Where the difference in costs is sufficiently large that no plausible difference in efficacy could change the conclusions or uncertainty estimates.[21]

• Where one intervention is significantly more effective and significantly less costly.[21]

• When supplemented with sensitivity analyses in order to assess the impact of non-zero incremental effects.[19,21]

Stewart also highlighted the topic of biosimilars as a special case, arguing that CMA should be adopted for this class of drug.[19] Whilst his argument was partly based on the need for consistency with the use of CMA for evaluating generic products, this was aligned to a certain degree with generalizable criteria such as the need for a strong a priori expectation that the treatments should perform equally.

Other, potentially useful issues were also highlighted by the literature. Firstly, the opportunity cost of making an incorrect decision due to the use of CMA is related to the size of the patient population. As such, there needs to be much greater certainty about equivalence in appraisals for larger patient populations. Secondly, the cost of the research needed to undertake a CEA compared to a CMA needs to be considered; large costs, combined with strong evidence of equivalence and cost savings may suggest that a CEA is not worthwhile.

We would add one further point for consideration. It is widely recognised that the quality of research can influence the speed and level of diffusion of new technologies. If an appraisal's recommendations are considered to be flawed, due to a partial analysis, it is possible that implementation of the recommendations is lower than if a fuller consideration of costs and health outcomes was undertaken. In such situations, any gain from undertaking a partial analysis are likely to be overwhelmed by the QALY losses of reduced implementation.

7. CONSULTATION

Based on the evidence described in the preceding sections, we consulted with a large number of UK-based expert advisors with expertise in the TA, MTEP and Diagnostics Programmes at NICE. These included committee members, NICE staff and academic groups. The purpose of these discussions was not to generate a further body of research evidence, but instead, to test out the issues and ideas raised from the reviews for theory, practice and policy.

Although largely anecdotal, it is worth noting that the expert advisors were generally supportive of the use of CMA in the circumstances where clinical effectiveness was claimed to be equivalent, as in the current FTA process, provided that its use were

reserved for specific circumstances and was applied with caution. For MTEP, there was a feeling that current methods work reasonably well, but that the need for 'similar cost' restricts the scope of the programme's work; some cost-increasing technologies can be identified as being clearly cost-effective without the need for a full appraisal estimating cost per QALY. For FTA, experts felt that it was a reasonable simplification of the TA process, but that it was only appropriate to assume CMA was likely to be the appropriate method a priori in a very small number of circumstances, notably the case of "me-too" drugs, and needed to be used cautiously.

These points, and further views expressed by experts, are referred to below in the discussion section.

8. DISCUSSION

Based on the evidence reviews and the consultation with experts a series of issues were raised that are examined below. Whilst many of these issues were raised during consultation, they are not intended to reflect the views of those experts, either collectively or as individuals.

There are two broad situations in which CMA may be considered appropriate.

i)   The appraisal is for a technology that is considered to be clinically non-inferior to the relevant comparator(s)

ii)   The appraisal is of a technology that differs from relevant comparator(s) in terms of clinical effectiveness but there is an expectation that those differences will lead to a net cost saving (where clinical effectiveness is superior).

The use of CMA in situation i) is potentially of relevance for both TA, MTEP, DAP and HST. This is the situation in which CMA has most frequently been considered in previous literature and aligns with the use of CMA in the current TA FTA process.

The use of CMA in situation ii) is not relevant to any NICE programme except MTEP. The use of CMA where there are differences in health benefits between technologies is only appropriate because of a difference in the decision rules for producing positive

guidance between MTEP versus other programmes. MTEP requires the use of cost minimisation methods because technologies must be cost saving or cost neutral to receive positive guidance. All other programmes may recommend cost-incurring technologies by balancing additional health benefits against additional costs using cost effectiveness analysis.

8.1. THE ASSESSMENT OF CLINICAL NON-INFERIORITY

The primary route through which CMA is used within the TA programme is via FTA. At the time of writing, NICE has seen only three technologies appraised in this way since its launch. There is some ambiguity about the selection criteria here, and differences of views about what the criteria ought to be.

The methods guide states that CMA is an acceptable approach where the new technology *"is likely to provide similar or greater health benefits at similar or lower cost than technologies recommended".* There are two ways in which this could be demonstrated. The first is to establish this through review of the evidence of clinical effectiveness and safety. However, before the appraisal takes place this relevant literature will not have been fully assembled or interrogated. The second is for there to be a clear clinical or pharmacological reason why similarity of effect would be expected. In practice, 2/3 of the FTAs have been based on so-called "me-too" drugs where such a hypothesis could be made from the outset. The experts we consulted gave differing accounts of the extent to which the third example, guselkumab, could be considered plausibly similar to some of the comparator, NICE recommended products. Certainly, guselkumab was marketed as a novel therapy but it is unclear the extent to which differences in the mechanism of action might be expected to result in differences in patient outcomes compared to other biologic therapies.

For many of the experts we consulted, it was felt that the existence of an underlying, plausible hypothesis for clinical equivalence should be a required aspect for the use of CMA within the FTA process. This requirement would lower the risk of falsely concluding a technology was non-inferior, given the uncertainties in empirical evidence. If this approach were to be adopted, this would mean that "first in class" technologies, which guselkumab for psoriasis in TA521 claims to be, could not be

33

routed to this part of the programme. That does not rule out the possibility that a decision may be based on cost minimisation at the subsequent appraisal should the review of clinical evidence be deemed to warrant it, but it would not prejudge the likelihood of CMA being appropriate. Not all experts held such strong views and, in the context of drug appraisals, some felt that the scientific rationale for how therapies worked could often be misleading. Nevertheless, it is important, in both drug and non–drug appraisals to consider the plausibility and underlying scientific rationale for any claims of non–inferiority. This assessment should be used to help interpret the evidence of clinical effectiveness.

Much of the academic literature on the use of CMA has focussed on decision rules linked to the clinical trials setting. Where an economic evaluation alongside a clinical trial is planned, access to patient level data on a range of outcomes and resource use variables can be analysed simultaneously and used to assess the appropriateness or otherwise of conducting CMA.

Whilst the principles remain equally relevant, translating them into practice in the HTA setting is not straightforward. In the ideal scenario, a head-to-head trial of equivalence would have been conducted but this situation is uncommon. Even where such studies exist, other evidence is also relevant. In current practice, committee members rightly take into account a range of different considerations in making the judgement of clinical non-inferiority, just as they do when judging clinical superiority and its magnitude. These include issues of statistical significance, clinical significance and cover a range of different outcome measures from the clinical studies that have been conducted. With a range of outcomes, even the typical equivalence trial may not be sufficient to demonstrate that competing health technologies are sufficiently similar on all relevant domains to warrant CMA. These judgements are influenced by prior beliefs about the technologies drawn from the pharmacological basis of the treatments (in the case of drugs) or other clinical or technical issues that determine how the technology works, and must draw heavily on the input from clinical or other relevant experts.

The difficulty of these judgements is perhaps less standardised in relation to the assessment of adverse events. The experts we consulted outlined different accounts of how a similar AE profile is assessed. However, they all required a consideration of

the types of AEs referred to, the magnitude of their severity and implications for patients. Committees can draw on various experts to help provide relevant information on these matters, but there is no formal decision rule that determines how these issues can be balanced. It is possible to conceive of a more formal means of undertaking these assessments, for example by explicitly categorising AEs into severity categories, but this itself raises many additional challenges and there was no appetite for such an approach. As with many aspects of the decision making process, and not unique to the issue of assessing clinical non-inferiority, the subjective views of the committee in making these judgements was seen as a key committee function.

Several of our experts felt that the assessment of clinical non-inferiority was one of the core functions of committee and the risk from the FTA process, or similar variants adopted for other NICE programmes, is that the role of the committee in making this assessment is diminished. For example, committees frequently spend a lot of time examining the robustness of key model parameters that have been derived from network meta-analyses. Forgoing a CUA on the basis of such analyses requires detailed consideration.

The assessment needs to consider a range of outcomes, with particular focus on those relevant to different aspects of patient outcome, including adverse events. It is also important to consider how evidence in relation to outcomes relates to those that would feature in a cost effectiveness analysis. Experts were clear that any judgement of clinical non-inferiority should be made with reference to the main drivers of the economic model. In the case of multiple outcome measures being appropriate, experts had mixed views on whether they would accept potential trade-offs between some outcomes being superior, whilst others were inferior. In the absence of a formal cost effectiveness study that quantifies the net health effect into a single metric, there are clear dangers and a lack of transparency from making this an issue of judgement.

In the case of a non-inferiority trial, we would usually conclude that a new treatment is non–inferior to another if the range of the 95% confidence interval (though the 95% level is itself arbitrary and not universally applied in the design of non-inferiority studies) of the primary outcome does not exceed (i.e. the new treatment is considered worse than) a pre-specified, non-inferiority margin. An equivalence trial is similar in

that the specification of an equivalence margin is required. If the confidence interval based on the observed data lies within the equivalence range, then the evidence supports the hypothesis that the treatments are considered therapeutically equivalent. These concepts are not straightforward to translate to the HTA setting. Interpretation of the evidence needs to recognise the multiple outcomes considered and the nature of the studies from which estimates are drawn. Studies, either alone or in combination, will not have been explicitly powered for the assessment of many outcomes of interest. As sample sizes increase, confidence intervals will shrink.

Even in the case of statistically significant differences in relevant outcomes, it may be the case that the difference is considered too small to be of clinical significance. The non-inferiority / equivalence margin in a trial specifically designed for such assessments is usually related to this concept. The assessment of clinical non-inferiority would benefit from consideration of this issue. This type of information could be provided by the independent review groups, drawing on appropriate experts, and further explored within the committee discussion with their appointed expert advisors. Clinical and patient experts can often be informative in making judgements about what constitutes a minimally important difference. It is important to avoid confusion between those outcomes that have relevance at the individual patient level (for example the difference between scores on a pain scale) versus those that are relevant only at the population level (for example, the proportion of patients achieving a clinical response or death). For some outcomes, no degree of non-inferiority is acceptable from a clinical perspective. However, in this situation it would never be possible to design a study that demonstrates non-inferiority as it would require an infinite sample size. In this situation, there may still be a requirement for committees to determine an acceptable margin on pragmatic grounds.

Some outcomes have a complex, surrogate relationship with patient outcomes, including those that drive cost effectiveness considerations. This is the case for the assessment of diagnostic technologies where performance is measured in terms of sensitivity and specificity. These are intermediate outcomes in relation to their impacts on health outcomes of interest. The impact of differences between technologies depends on the treatments available for patients requiring them, the adverse events of those treatments and the long term sequelae of misdiagnosis. It is therefore unclear

how one can specify an acceptable margin for non–inferiority, or what the role of standard measures of statistical significance should play here. There is a risk that differences between tests that may appear to be trivial and/or statistically insignificant, would in fact lead to important differences in cost effectiveness analysis estimates. In this situation, the appropriate approach is to use cost utility analysis and propagate uncertainty in estimates through the decision model using probabilistic sensitivity analysis. This is illustrated in the case of DG25, High-throughput non-invasive prenatal testing (NIPD) for fetal RHD genotype. In this appraisal, the assessment was to determine the value of NIPD as a potential means of reducing the use of inappropriate routine anti-D prophylaxis in rhesus-D negative pregnant women. Whilst non-invasive NIPD was considered to have very high sensitivity and specificity (Sensitivity was 0.998 (95% CI 0.992 to 0.999), specificity was 0.942 (95% CI 0.920 to 0.959) in the preferred results) it is not a perfect test. In particular, to assess non–inferiority would require an assessment of the impact of failing to provide anti-D to a mother on the basis of an incorrect test result, and the risks to subsequent pregnancies, albeit in a small number of cases. Since the appraisal was conducted using cost utility methods, this was shown to have a relatively small impact on QALY differences (-0.46 QALYs per 100,000 pregnancies). It is through conducting the modelling that the impact on health outcomes is demonstrated, allowing the calculation of cost effectiveness to inform decision making. In this situation, the risks of having conducted CMA without the explicit assessment of how differences in diagnostic performance would impact patients are clear.

In most cases, diagnostics assessed by NICE, either through MTEP or DAP, are not evaluating technologies with equivalent diagnostic properties. Many are assessing a range of similar tests against current practice, and, whilst those tests may themselves be similar, they do differ in subtle but important ways. It may be the case that new tests become available that are similar in nature to those that have already been assessed by NICE. In this situation the same principles described above, relevant for all technology types, apply. In addition to this, and unlike the typical case for pharmaceuticals, the evidence base for diagnostic technologies often continues to mature after tests come into routine practice. For drug technologies, new randomised evidence is much more sparse after approval due to ethical restrictions and the lack of incentive from potential study funders. In the case of diagnostics, the changed

nature of the evidence for a comparator would also need to be taken into account as part of any appraisal of a new diagnostic technology, and this may negate the option of performing a CMA.

Some consultees expressed the view that the independent review groups should consistently provide a stronger steer to the committee about evidence of clinical non-inferiority, particularly where this relied on the use of network meta-analysis (NMA).

Uncertainty needs to be assessed using appropriate methods, regardless of its source. The impact of the degree of uncertainty in the assessment of clinical effectiveness was a source of variation in the responses of consultees. For some, substantial uncertainty undermined the rationale for CMA and should be an automatic trigger to a full appraisal. Others felt that they would be prepared to accept a substantial degree of uncertainty within the CMA assessment, but would expect a greater degree of cost savings through lower prices.

The choice of comparator also impacted the assessment of uncertainty in the views of experts we consulted. In the current FTA process, whilst there may be several relevant comparators, it seems that the interpretation of the methods guide addendum is that the new technology only needs to be shown to be cost saving and equally effective against one of them to be recommended (see for example TA521), provided that comparator adequately represents the full comparator set. Consideration should be given to clarifying the requirements for the comparator in a CMA case, particularly in other programmes where it is more likely that relevant comparators have not been previously appraised by NICE.

8.2. THE ASSESSMENT OF COSTS

In scenario i) described above, where the technology under appraisal is claimed to be clinically equivalent to its comparators, the only relevant costs relate to the acquisition costs (prescribing costs of drug technologies), administration and monitoring of treatments. In this situation, there is likely to be very little uncertainty associated with the assessment of costs. The unit costs of technologies should usually be known with complete certainty. There may, be variation in best practice for monitoring patients or

some other aspects of patient management. However, the source of any existing uncertainty is not likely to arise from study sampling uncertainty. Therefore, in this situation, it could be argued that there should be no margin allowed for the preferred estimate of mean cost to be greater than the relevant comparator(s).

Where the technology claims to deliver the same health benefits as comparators, but changes the way the therapy is delivered, then the cost assessment should still be relatively straightforward. We would expect this type of claim to be unlikely for drug technologies but may be associated with drug delivery devices, or technologies which impact on the management of patients, perhaps substituting staff type or time at particular points in the process, for example. Costs comprise acquisition costs, administration and monitoring, part of which will be different owing to the delivery aspect. In this situation, cost differences will usually only accrue for the period in which the technology is applied to the patient. There could be scenarios where some aspects of monitoring continue long after the cessation of the therapy in question. The calculation is also likely to be quite straightforward because in many situations any differences in costs will simply be equivalent in each time period.

Scenario ii) above describes the use of CMA in the MTEP setting - where the therapy is clinically superior, and typically has higher acquisition costs, but these are expected to be more than offset by the reduced NHS costs from improved outcomes and/or adverse events. In this situation, the analysis needs to model the total cost associated with the care pathway with and without the technology, typically using sampled data (from a trial or observational study) and to extrapolate as required to a time frame over which these differences continue to occur and have cost implications. The analysis should reflect the extent to which the health technologies in question continue to lead to differences in clinical events (for example the rate of strokes) and the entirety of the costs of those events (lifetime costs in the example of stroke). Health benefits feature in the model only to the extent that they manifest themselves in cost differences.

Many parameter values will contribute to the overall assessment of cost difference and will be subject to sampling uncertainty. Here, it is feasible for the mean cost estimate to show the technology to be cost-incurring, but this may be solely due to sampling uncertainty in those parameter estimates. The use of probabilistic sensitivity analysis

to propagate the uncertainty in the cost model inputs, with results presented as the probability of being cost-incurring, may be a particularly insightful method of conveying the impact of uncertainty. This is referred to in the MTEP methods guide (see section 7.3.2) but is not a requirement. In any case, there needs to be a detailed examination of the evidence and associated uncertainty associated with the individual parameter input values that collectively inform the cost estimate, with particular emphasis on those that contribute most to the cost model. This assessment should be based on clinical and operational plausibility, together with an examination of all relevant data and the sources from which those data were derived. In addition, it should be noted that the resolution of uncertainties in parameters which drive the cost assessment, yet are unlikely to alter the assessment of clinical superiority, may justifiably lead to a recommendation that a technology is used only in research or is approved with further research. The MTEP methods guide (sections 8.2.2. and 8.2.3) explicitly allow for these types of recommendations, though in practice their use has been restricted to the resolution of other types of uncertainties.

It is worth noting here that the complexity of the evidence review and economic modelling required to undertake this type of analysis is often similar to that required to undertake a full cost effectiveness analysis. Indeed, some of our experts pointed out that, even in the case of CMA used in the case of claimed clinical non–inferiority, the workload is neither substantially less, nor quicker to produce, than that required in the case of a full cost effectiveness analysis. This is of relevance for those situations where a technology is found to be cost-incurring. Current or future NICE processes that switch to the use of a cost effectiveness assessment coupled with appropriate decision rules for adoption in this situation can be designed around the existing cost analysis. Indeed, in some cases within the MTEP programme, manufacturers have submitted full CEA models with any assessment of the impact on quality of life stripped out.

9. RECOMMENDATIONS

Recommendations on clinical effectiveness

1. Cost minimisation may be used in different circumstances. For all programmes, the option of pursuing CMA when there are plausible claims that the technology in question is clinically equivalent to relevant comparators allows a simpler and

potentially faster analysis to be performed. These are a limited set of circumstances and should be applied cautiously.

2. The principles on which the decision to proceed using CMA in these circumstances should be the same for all programmes. The use of these principles may mean that CMA is even more rarely employed in the assessment of certain types of technologies than for others.

3. Where the claim of clinical equivalence is made, there must be a detailed consideration of the plausibility of the claim both when making the decision to proceed using CMA, and as part of any committee considerations.

4. In the case of pharmaceuticals, there should be consideration of the biological plausibility of the claim and the extent to which the mechanisms of action of the new and reference drug differ should be considered. Where any differences are identified, additional consideration is required to assess how patient outcomes might be affected.

5. Similarly, for all other technologies, the foundation for any claims of clinical equivalence needs detailed scrutiny from relevant experts in order to assess plausibility. For all programmes, it is important to consider how any differences would be expected to manifest themselves in patient outcomes, particularly those that may be important to an economic model of cost effectiveness.

6. Where technologies have significantly different adverse event profiles, they should not be considered non-inferior, and thus the cost minimisation route is not appropriate because the implications of these differences in health outcomes measured as QALYs needs to be calculated.

7. The assessment of clinical non-inferiority requires consideration of both statistical and clinical significance *inter alia*. All outcomes that relate to different aspects of patient benefit should be considered with particular focus on those drivers of an economic model. An assessment of non-inferiority should be required for all important outcomes, not just the primary outcome.

8. It is useful to consider whether differences in clinical effectiveness plausibly span some non-inferiority margin.

9. The non-inferiority margin may often be informed by expert patients and clinicians. Independent review groups can help to identify important parameters and provide information on the assessment of the appropriate margin, informed by relevant experts. Caution is required in the interpretation of individual level differences

versus population level differences. Sometimes a pragmatic approach is required to define a non-inferiority margin.

10. Interpretation of results needs to consider the design of studies from which estimates are drawn, including their sample sizes.

11. The impact of differences in intermediate outcomes, such as sensitivity and specificity, needs careful consideration in order to establish their impact on final health outcomes.

Recommendations on the assessment of costs

12. Where CMA is undertaken because of clinical non–inferiority, there may be very little uncertainty about the cost implications. In this situation, there should be no margin allowed for the preferred estimate of mean cost to be greater than the relevant comparator(s).

13. Where CMA is undertaken and there are claims of differences in clinical benefits, the parameter estimates for individual components of resource use and unit costs and their associated uncertainties should be examined proportionate to their contribution to the cost difference.

14. Probabilistic sensitivity analysis may be particularly useful where sampling uncertainty for key parameters is particularly important.

10. REFERENCES

[1] Drummond M, Sculpher M, Torrance G et al (2005) Methods for the Economic Evaluation of Health Care Programmes, 3rd Edition, Oxford University Press, Oxford.

[2] Palmer S, Raftery J. (1999) "Opportunity costs", BMJ, Vol.318:1551-2.

[3] NICE (undated) "NICE's biosimilars position statement" available at https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/biosimilars-statement.pdf (last accessed 10/5/19)

[4] NICE (2017) "Medical Technologies Evaluation Programme Methods Guide" available at https://www.nice.org.uk/process/pmg33/chapter/introduction (last accessed 10/5/19)

[5] NICE (2013) "Guide to the Methods of Technology Appraisal" available at: https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781 (last accessed 10/5/19)

[6] NICE (undated) "Cost comparison Addendum to the Guide to the methods of technology appraisal" available at: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/methods-guide-addendum-cost-comparison.pdf (last accessed 10/5/19)

[7] NICE (2017) "User guide for the cost comparison company evidence submission template, 1st April 2017" available at https://www.nice.org.uk/process/pmg32/resources/user-guide-for-the-cost-comparison-company-evidence-submission-template-pdf-72286772526277 (last accessed 10/5/19)

[8] Machado A, Torres T. (2018) "Guselkumab for the Treatment of Psoriasis", BioDrugs. Vol.32(2):119-128

[9] See https://www.tremfya.com/what-is-tremfya last accessed 13th March 2019

[10] CADTH (2006) "Guidelines for the Economic Evaluation of Health Technologies", Canada. Ottawa: Canadian Agency for Drugs and Technologies in Health.

[11] pCODR (2013) "Axitinib (Inlyta) for metastatic Renal Cell Carcinoma", Toronto: pan-Canadian Oncology Drug Review.

[12] CADTH (2017) "Guidelines for the Economic Evaluation of Health Technologies", Canada. Ottawa: Canadian Agency for Drugs and Technologies in Health.

[13] PBAC submission guidelines. https://pbac.pbs.gov.au/information/about-the-guidelines.html last accessed 14th March 2019

[14] Taylor C, Jan S. (2017) "Economic evaluation of medicines", Aust Prescr 2017;40:76–8

[15] Pharmaceutical Benefits Scheme. Recommendations made by the PBAC July 2015. Canberra: Department of Health; 2015. http://www.pbs.gov.au/info/industry/listing/elements/pbac-meetings/pbac-outcomes/pbac-outcomes-2015-07 (last accessed 10/5/19)

[16] Briggs AH, O'Brien BJ. (2001) "The death of cost-minimization analysis?", Health Economics, Vol.10(2):179-184.

[17] Johnston K, Gray A, Moher M, Yudkin P, et al. (2003) "Reporting the cost-effectiveness of interventions with nonsignificant effect differences: example from a trial of secondary prevention of coronary heart disease", Int J Technol Assess Health Care, Vol.19(3):476-89.

[18] Span M, TenVergert E, Hilst CVD, et al. (2006) "Noninferiority testing in cost-minimization studies: Practical issues concerning power analysis", Int J Technol Assess Health Care, Vol.22(2):261-6.

[19] Stewart A, Aubrey P, Belsey J. (2010) "Addressing the health technology assessment of biosimilar pharmaceuticals", Curr Med Res Opin, Vol.26(9):2119-2126.

[20] Haycox A. (2009) "What is cost-minimization analysis?", in Arnold R (ed) Pharmacoeconomics: From theory to practice. Boca Raton, FL: CRC Press.

[21] Dakin H, Wordsworth S. (2013) "Cost-minimisation analysis versus cost-effectiveness analysis, revisited", Health Economics, Vol.22(1):22-34.

[22] Eckermann S. (2017) "Avoiding Frankenstein's Monster and Partial Analysis Problems: Robustly Synthesising, Translating and Extrapolating Evidence", in Eckermann S (ed) Health Economics from Theory to Practice. Optimally Informing Joint Decisions of Research, Reimbursement and Regulation with Health System Budget Constraints and Community Objectives, Adis: Cambridge.

[23] Russo P, Favato G, Rosano G, Staniscia T, Romano F. (2018) "Adequate use of economic evaluations based on the endpoints of a registrative clinical trial: cost-minimization analysis and the return of the living dead", SSRN Electronic Journal.

[24] Nuijten MJ.(2004) "Incorporation of statistical uncertainty in health economic modelling studies using second-order Monte Carlo simulations", Pharmacoeconomics. Vol.22(12):759-769.

[25] O'Hagan A, McCabe C, Akehurst R, et al. (2005) "Incorporation of uncertainty in health economic modelling studies", Pharmacoeconomics. Vol.23(6):529-536.

[26] Claxton K. (1999) "The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies", Journal of Health Economics, Vol.18(3):341-364.

[27] Walker S, Sculpher M, Claxton K, Palmer S. (2012) "Coverage with evidence development, only in research, risk sharing or patient access scheme? A framework for coverage decisions", University of York: Centre for Health Economics.

[28] Grimm SE, Strong M, Brennan A, Wailoo AJ. (2017) "The HTA Risk Analysis Chart: Visualising the Need for and Potential Value of Managed Entry Agreements in Health Technology Assessment", Pharmacoeconomics, Vol.35(12):1287-1296.

11. APPENDICES

Appendix Table 1: Summary of Technology Appraisals that used Cost Minimisation analysis but were not Fast-Track Appraisals.

	Technology	Patient group	Date of issue	Clinical effectiveness	Adverse events	Costs
TA131	Inhaled corticosteroids	Children with Asthma	Nov-07	No clinically relevant differences	not appropriate to distinguish based on AEs	A decision would be necessary based on cost comparisons
TA183	Topetecan	recurrent Stage IVB cervical cancer	Oct-09			ICERS were calculated. This isn't a cost minimisation example
TA184	Topetecan	relapsed small-cell lung cancer	Nov-09	No stat sig differences in main outcome measures. The Committee concluded that intravenous topotecan might have some benefits over CAV in terms of symptomatic relief, but these were difficult to confirm or quantify on the basis of current evidence. Some symptomatic gains conceivable but CAV more convenient for patients.		ICERs for most comparisons (against other comparators for specific patient groups). vs the combination of cyclophosphamide, doxorubicin and vincristine (CAV) a cost model was produced because of absence of data. Assumed equivalence of outcomes. With threshold analysis of magnitude of gain required to be cost effective.
TA258	Erlotinib	Non small cell lung cancer	Jun-12	Insufficient evidence to suggest a difference in PFS or OS.	AEs mild and the committee took this to mean that a choice of treatments would be valuable.	ICERS were presented. The model was resestimated with equal effectiveness so that differences in costs came from the drug, its administration and the costs of the PAS schemes.
TA310	Afatinib	non-small-cell lung cancer	Apr-14	The Committee concluded that on balance afatinib is likely to have similar clinical efficacy to erlotinib and gefitinib	Higher rate of diarrhoea and rash, but felt these were well managed in practice. So	ICERS calculated. But committee felt the ICERs were not reliable because of the structural problems with the model.

					overall, they felt similar toxicity.	
TA397	Belimumab	Lupus	Jun-16	Compared to standard care belimumab superior. No data to allow comparison to rituximab.	No data to compare to rituximab	Cost effectiveness was calculated. ERG calculated a cost model.
TA425	Dasatinib, nilotinib and high dose imatinib	chronic or accelerated-phase Philadelphia-chromosome-positive chronic myeloid leukaemia	Dec-16	Dasatinib and nilotinib superior to imatinib. Insufficient evidence to distinguish between dasatinib and nilotinib in terms of clinical effectiveness.	All 3 drugs well tolerated.	This appraisal was a Cancer Drugs Fund partial reconsideration of dasatinib, one of three tyrosine kinase inhibitors. ICERS were calculated in the original appraisal. This was a CMA between dasatinib and nilotinib, only for the Cancer Drugs reappraisal. Committee agreed CMA was appropriate.
						As part of the CDF reappraisal, the company submitted a cost minimisation analysis of dasatinib. Committee agreed.
TA426	Dasatinib, nilotinib and imatinib	untreated, chronic-phase Philadelphia-chromosome-positive chronic myeloid leukaemia	Dec-16	Dasatinib and nilotinib superior to imatinib. Insufficient evidence to distinguish between dasatinib and nilotinib in terms of clinical effectiveness. No new evidence changed that view.	All 3 drugs well tolerated.	This appraisal was a Cancer Drugs Fund partial reconsideration of dasatinib, following the original appraisals consideration of three tyrosine kinase inhibitors. ICERS were calculated in the original appraisal. However, the committee decided it was important to have an additional drug available (imatinib) as an option given it had a longer term record of safety and efficacy. CMA was considered appropriate for dasatinib compared to nivolumab because they were "sufficiently similar". Against imatinib, if dasatinib cost saving then it would be considered to dominate.

Appendix Table 2: Pharmacoeconomics guidelines that include the use of cost Minimisation Analysis, from ISPOR Pharmacoeconomics Guidelines Website

Region	
Baltic (Latvia, Lithuania, Estonia)	Any one of CMA, CEA, CUA. Need justification.
France	Any one of CMA, CEA, CUA, CBA, and CCA. The choice must be justified.
Portugal	Any scientific recognised economic evaluation technique can be used such as CMA, CEA, CUA, CBA
Sweden	CUA, CEA, CMA
Taiwan	The most appropriate method that can reflect the purpose of the study
Brazil	Cost-consequence Analysis, CMA, CEA, CUA, CBA
Cuba	CBA, CEA, CMA, CUA
México	Cost-effectiveness, Cost-utility, Cost-benefit and Cost minimization
Egypt	Any of CMA, CEA and CUA considered.
Slovenia	The pharmacoeconomic analysis must include treatment outcomes of other studies, presented by systematic literature review or meta-analysis. The following types of analysis can be used: Cost-Effectiveness Analysis (CEA), Cost-Utility Analysis (CUA), Cost-Minimization Analysis (CMA) and Cost Analysis.
MERCOSUR (Argentina, Brazil, Paraguay, Uruguay)	There is no preference, the analytical technique should be justified and adequate to answer the research question.

Appendix Table 3: Classification of papers by the four searches

Category	Description	Sample of full search	Briggs and O'Brien citation search	HTA search	Reimburse-ment search
$$methods	Papers with the primary of focus of assessing the methods of economic evaluation	Number of records 0	Number of records 9	Number of records 1	Number of records 0
$$drug	Papers reporting the evaluation of drugs, with or without, non-drugs as comparators	Number of records 38	Number of records 26	Number of records 2	Number of records 15
$$nondrug	Papers reporting the evaluation of non-drug interventions, exclusively	Number of records 33	Number of records 97	Number of records 4	Number of records 25
$$other	Other included papers in which CMA forms a substantive part	Number of records 9	Number of records 61	Number of records 7	Number of records 16
$$excluded	Papers which do not provide sufficient information to categorise as above, are non-English language or do not relate to CMA within health care.	Number of records 21	Number of records 179	Number of records 11	Number of records 9
	Total	101	372	25	65