Cochrane

Cochrane Handbook for Systematic Reviews of Interventions Version 6.5

Technical Supplement to Chapter 20: Integrated full systematic review of economic evidence

Trusted evidence.
Informed decisions.
Better health.

Table of contents

Technical Supplement to Chapter 20: Integrated full systematic review of economic evidence

Pauline Sobiesuo, Patricia Aluko, Erin Graybill, Dawn Craig, Catherine Henderson, Michael Drummond, Edward CF Wilson, Shannon Robalino, Denny John, Pia Johansson, David McDaid, Luke Vale on behalf of the Campbell and Cochrane Economics Methods Group

This Technical Supplement should be cited as: Sobiesuo P, Aluko P, Graybill E, Craig D, Henderson C, Drummond M, Wilson ECF, Robalino S, John D, Johansson P, McDaid D, Vale L. Technical Supplement to Chapter 20: Integrated full systematic review of economic evidence [last updated August 2024]. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.5. Cochrane, 2024. Available from www.training.cochrane.org/handbook.

1 Introduction

There are two optional methodological frameworks for incorporating economic evidence into Cochrane Intervention Reviews (CIRs) (see Chapter 20). These are:

1. an integrated full systematic review of economic evidence, in which additional 'economic' methods procedures are integrated into each stage of the main systematic review of intervention effects; and

2. a brief economic commentary, which provides a 'minimal framework' for incorporating economic evidence.

This supplement to Chapter 20 of the *Cochrane Handbook* focuses on the integrated full systematic review of economic evidence. For guidance on conducting a brief economic commentary see Chapter 20 (Aluko et al 2023). For guidance on deciding when to prioritize and de-prioritize incorporating economic evidence into Cochrane Reviews, see Chapter 20 (Section 20.3.1).

An introduction to core health economics concepts useful when planning and conducting an integrated full systematic review of economic evidence is provided in material from the Campbell and Cochrane Economics Methods Group (CCEMG), available from: http://www.methods.cochrane.org/economics/.

The primary objective of an integrated full systematic review of economic evaluations is to assess impacts of the intervention(s) on resource use, costs, health outcomes (including health-related quality of life) and cost-effectiveness when implemented in different study settings, groups of participants and contexts, at different times. This involves investigating how the impacts of intervention on economic and health outcomes may vary according to specific study conditions and what factors may drive any variations in estimates of these outcomes between studies, populations and settings.

An integrated full systematic review of economic evidence is substantially more resource intensive than the brief economic commentary because it requires integration of ancillary economics methods into each stage of the main systematic review of intervention effects. As such, the choice between developing a brief economic commentary and conducting an integrated full systematic review of economic evidence is often made based on consideration of how much time and expertise is available to develop the health economics component of an intervention review and/or its updates. For an integrated full systematic review of economic evidence, outcomes that can be used to provide an 'economic lens' or interpretation of evidence on effects can be identified and included at the protocol stage. At the review stage, evidence on effectiveness, harm, impact on the use of services and the likely difference in cost of the interventions compared can be brought together in the results and discussion stages. This would make it possible to determine if an intervention could be seen favourably from an economic standpoint.

An important consideration is the availability of specialist support from a health economist to implement the methods outlined in this supplement in Cochrane Protocols and Reviews. We therefore recommend consulting with a health economist with experience of Cochrane Review methods, both: (i) to inform the final decision about whether to conduct an integrated full systematic review of economic evidence; and (ii) to provide advice and support once a decision has been made to proceed.

The CCEMG also seeks to help Cochrane Review author teams identify health economists willing to collaborate on, or provide peer review for, Cochrane Reviews and Protocols (see http://www.methods.cochrane.org/economics/). Author teams are advised to consult with support networks available to them in the first instance.

GRADE guidance on rating certainty of evidence for resource use and costs recommends that important differences in resource use between alternative management strategies should be considered for inclusion, alongside other important outcomes, in 'Summary of findings' tables (Brunetti et al 2013). Consistent with this GRADE guidance, incorporating evidence for resource use into an integrated full systematic review of economic evidence involves three key steps, as follows.

1. Identify (at protocol stage or, if the economics component is added in an update, on a review update form) items of resource use that may differ between alternative management strategies and that are potentially important to decision-makers.

2. Find evidence for these differences in resource use (e.g. from previously published economic evaluations).

3. Make judgements regarding confidence in estimates (for measures of resource use) using the same criteria used for health outcomes (i.e. study limitations, indirectness, imprecision, inconsistency and other considerations – see Chapter 15).

2 Formulating the objectives and eligibility criteria

2.1 Formulating objectives

The research question that will frame an integrated full systematic review of economic evidence can be formulated with close reference to the question(s) that will frame the main review of intervention effects. Research questions addressed by Cochrane Reviews of intervention effects are formulated as objectives (see also Chapter III, Section III.3.3). The main objective of a review is conventionally written as:

"To assess the effects of [*intervention or comparison*] for [*health problem*] for/in [*types of people, disease or problem and setting if specified*]".

The corresponding objective for an integrated full systematic review of economic evidence should therefore be formulated as follows:

"To summarize and assess current evidence on [*secondary economic outcomes, e.g. incremental resource use, utilities (including quality-adjusted life years, capabilities etc.), costs and/or cost-effectiveness*] associated with the use of [*intervention*] versus [*placebo or comparator*] for [*condition and primary health outcome*] in [*population*].

For example, a main objective for a review might be:

"To assess the effects of hospital-based specialist palliative care for adults with advanced illness and their caregivers".

A secondary objective to address economic issues within this review might be:

"To summarize and assess current evidence on the incremental resource use, utilities (including quality-adjusted life years, capabilities etc.), and costs and/or cost-effectiveness of hospital-based specialist palliative care versus usual care for adults with a terminal illness and their unpaid caregivers."

Review authors should avoid formulating objectives that take the form "To assess the cost-effectiveness of…". This is because, for reasons set out in the final paragraph of Chapter 20 (Section 20.2.2), Cochrane systematic reviews of economic evidence cannot be designed to, and therefore should not aim to, produce definitive or generalized claims (statements), findings or conclusions about whether an intervention is, or is not, cost-effective.

2.2 Formulating eligibility criteria

The objective of an integrated full systematic review of economic evidence can be translated into study eligibility criteria using the same PICO framework that is conventionally applied to set study eligibility criteria for the review of intervention effects. As the above example shows, eligibility criteria relating to the population (P), intervention(s) (I), comparator(s) (C) and important health outcome(s) (O) will be identical between the integrated full systematic review of economic evidence and the review of intervention effects. However, for the economic evidence there may be additional relevant evidence about, for example, quality of life or patient satisfaction along with additional information on resource use, costs and efficiency. We discuss these in detail in Section 2.2.2.

When considering additional outcomes of interest for the integrated full review of economic evidence, it is important to check whether some of the economic outcomes are already included as outcomes in the systematic review of intervention effects (see Section 3.1). Therefore, care should be taken not to duplicate outcomes when adding the economic review. However, the level of detail for the specific measures of resource use, costs and/or cost-effectiveness may need to be considered among the economic outcomes (for example, rather than saying that 'resource utilization' is an outcome of interest, the precise types of resource utilization should be specified).

2.2.1 Specifying additional economic study types

We recommend that eligibility criteria related to study types should allow for the inclusion of all types of full economic evaluations: cost-effectiveness analyses (CEAs, including cost-minimization analyses and cost-consequences analyses); cost-utility analyses (CUAs); and cost-benefit analyses (CBAs). For definitions of these types of economic evaluations, see the glossary of terms available on the CCEMG website (see Glossary | Cochrane Economics).

Full economic evaluations compare the costs (resource use) associated with one or more alternative interventions (e.g. intervention X versus comparator Y) with their consequences (health outcomes, effects) in a specific group of patients (population Z). All types of full economic evaluations value resources in the same way (i.e. by applying unit costs to measured quantities of resource use) but differ primarily in the way health outcomes are measured and valued. In CEA, health outcomes are measured in terms of a single natural or clinical unit such as the proportion surviving. In CUA, a broader notion of health is used, such as quality-adjusted life years (QALY) or disability-adjusted life years (DALYs). In CBA, benefits and costs are valued in the same unit, normally monetary. These differences reflect the different aims and viewpoints of different decision-makers or economic questions. Therefore, no type of economic evaluation is considered of 'higher quality'.

However, CBA and CUA may be considered to be more useful for policy-makers or funding organizations because results are reported using metrics such as incremental cost per QALY in a CUA, or cost-benefit ratio (CBR) in a CBA. For CUA, the decision-maker may have decision rules as to what an acceptable cost per QALY is, such as the National Institute for Health and

Care Excellence in England (Dawoud et al 2022, National Institute for Health and Care Excellence 2022). For CBA the decision rule is simpler: if benefits exceed costs, then the experimental intervention is preferred.

A further key preliminary decision is whether the scope of the review will be: (a) limited to trial-based economic evaluations (i.e. those conducted within the framework of studies meeting eligibility criteria for the main review of intervention effects); or (b) extended to encompass model-based economic evaluations (see also Chapter 20, Section 20.2). We recommend that trial- and model-based economic evaluations are treated separately in later stages of the review process. This includes assessments of risk of bias and methodological quality, analysis and synthesis of 'economic outcome' data, presentation of results and (if applicable) rating the certainty of evidence for resource use (cost) outcomes alongside important health outcomes (effects). Boxes 1 and 2 show templates for study design eligibility criteria that are consistent with options (a) and (b). For simplicity, we assume in both cases that study designs eligible for the main systematic review of intervention effects are limited to randomized trials. If this is not the case, these templates can also be adapted to encompass the wider set of non-randomized study designs (for studies of effects) eligible for inclusion. Though randomized trials constitute a key source of data for evidence synthesis, the use of a single randomized trial in an economic evaluation (i.e. trial-based economic evaluation) may only provide partial or insufficient evidence/information to decision-makers (Sculpher et al 2006, Drummond et al 2015). Health economists argue that full economic evaluation should encompass a broader framework where the objectives of a health system subject to resource scarcity can inform the cost-effectiveness of an intervention given existing evidence, and whether more evidence will be needed to support this decision in the future.

Key limitations of economic evaluations conducted as part of randomized trials are the failure to consider all relevant comparators (which typically are not feasible for a trial to do) and the length of trial follow-up, which is typically too short for economic purposes. It is for these reasons that many decision-making organizations rely on model-based economic evaluations to make their decisions. When effects of interventions are less likely to be short-term then it is useful to also consider model-based full economic evaluations for inclusion (as in the eligibility criteria set out in
Box 2.2.a). We return to this issue in Section 2.2.2, where we consider the appropriate *time horizon* for recording outcome data.

Box 2.2.a: Template of study design eligibility criteria for an integrated full systematic review of economic evidence:

(a) scope limited to randomized trial-based economic evaluations
Randomized trials and full economic evaluations (CEAs, CUAs and CBAs) conducted alongside eligible randomized trials.

(b) scope extended to encompass model-based economic evaluations

Randomized trials, full economic evaluations (CEAs, CUAs and CBAs) conducted alongside eligible randomized trials and eligible model-based full economic evaluations.

2.2.2 Specifying which economic outcomes the review includes

Selection of 'economic outcomes' (that is, those sought specifically for the full integrated systematic review of economics evidence) should be guided primarily by consideration of the likely *magnitude* of the expected impacts of interventions on measures of resource use, costs and health effects (and hence cost-effectiveness). Consideration of magnitude involves identifying those items of resource use that are expected to differ between alternative management strategies (experimental intervention and comparator intervention) that are potentially important for decision-making.

Other factors that should also guide the selection of economic outcomes are the *analytic perspective* and *time horizon* that will be adopted for the integrated full systematic review of economic evidence. Deciding on an appropriate *analytic perspective* is challenging in the context of a Cochrane intervention review due to the wide range of potential viewpoints of decision-makers who may wish to use the review findings (e.g. health system, patient/family, societal or third-party payer). It requires consideration of who is likely to bear the incremental costs associated with an intervention, versus comparators, and who is likely to receive the incremental benefits. Some costs (resource use) are relevant from one analytic viewpoint but not from another. For example, the costs of providing informal care, or those associated with lost productivity (due to absence from work), or out-of-pocket costs such as co-payments and other costs such as travel cost and time, may be relevant from a patient or a societal viewpoint but may be excluded when a narrower perspective is adopted, such as that of the healthcare system, an insurance company or single provider (e.g. a hospital or general practice). Similarly, adopting a health and social care system perspective would allow for the inclusion of resource use (costs) incurred in both these sectors but may again exclude the costs of informal care and/or productivity losses.

Given the diversity of end-users of Cochrane Reviews, a pragmatic approach to this issue for an integrated full systematic review of economic evidence is to adopt a societal perspective (which encompasses all perspectives) and report who bears the costs (resource use); for example, 'the health system' (e.g. direct health care costs), 'the social care system' (e.g. direct social care costs), 'patients' (e.g. out-of-pocket expenses) or 'employers and society' (e.g. productivity losses). However, the breadth of economic outcomes reported in a review will inevitably depend on the range of outcomes reported in the studies being reviewed.

Deciding on an appropriate *time horizon* involves consideration of the period over which important differences in costs (resource use) and effects (health outcomes) between the

intervention(s) and comparator(s) are expected to accrue. Cochrane intervention reviews implicitly establish a time horizon for effects by specifying intermediate and final endpoint measures of health and related outcomes. To ensure that all important differences in costs (resource outputs) and effects are captured, and to protect against biased and potentially misleading conclusions, there is a parallel need to consider whether the same time horizon is applicable when important differences in costs (resource inputs and resource outputs) and effects are considered together. In practice, this means that an integrated full systematic review of economic evidence will often need to consider a longer time horizon (e.g. the patients' whole lifetime) compared with the longest follow-up periods used to measure health outcomes among included studies of effects (e.g. randomized trials). This is especially the case when considering management of long-term conditions or interventions that might have consequences over a whole life course such as, for example, treatments for infants who have had a perinatal stroke. There are circumstances in which shorter follow-ups may be sufficient, such as alternatives to short-term catheterization following surgery for relatively fit adults, where longer-term consequences may be unlikely.

For an integrated full systematic review of economic evidence, economic outcomes should be listed alongside health and related outcomes in the 'Types of outcome measures' part of the 'Criteria for considering studies for this review' section of the protocol. We recommend breaking down specified economic outcome measures to the level of specific items or categories for resource use and costs (e.g. length of hospital stay in days, duration of operation in minutes, number of outpatient attendances, bleeds from secondary prophylaxis at six-month follow-up, number of days off work, direct medical resource use, direct medical costs, overheads like heat, power, light, etc. on healthcare infrastructure, indirect costs such as productivity loss, patient out-of-pocket expenses). The use of general terms (e.g. 'costs', 'resource utilization') to specify economic outcomes should be avoided. While the breaking down of outcome measures forces extraction of economic outcome data from included studies at a granular level (which increases workload), it also ensures that differences between included studies in the components of resource use and costs they measure (in trial-based economic evaluations) or incorporate (in model-based economic evaluations) can be fully considered. These economic outcome measures should be listed in the secondary outcomes section of the systematic review of intervention effects.

A study may still be included if it did not include any of these outcomes, but the fact that it meets the population, intervention and comparator (PIC) criteria but not the outcome (O) criterion of the PICO should be reported as part of the integrated full review of economic evidence. Box 2.2.b shows an example of the additional economic outcomes specified for a full systematic review of economic evidence that is integrated into a Cochrane Review of 'Bone morphogenetic protein (BMP) for fracture healing in adults' (Garrison et al 2010).

Box 2.2.b: Example for economic outcomes: BMP for fracture healing in adults (Garrison et al 2010)

Types of outcome measures

Primary outcomes

1. Time to union
2. Union rate

Secondary outcomes

1. Secondary procedures (any procedures required after initial surgery, specifically those undertaken to promote healing)*
2. Infection*
3. Hardware failure*
4. Clinical response (average change in pain or functional assessment scores such as Short Musculoskeletal Function Assessment)
5. Operative and hospital stay parameters
    a. Operative time*
    b. Operative blood loss
    c. Length of postoperative hospital stay*
6. Other patient outcomes
    a. Employment status before and after treatment
    b. Number and time to return to work (for those patients in employment before treatment)*
7. Donor site appearance (average score/change in donor site appearance)
8. Heterotopic bone formation
9. Immunogenicity (antibody response to BMP or bovine collagen)
10. Any adverse effects
11. Direct medical resource use*
12. Lost or reduced productivity (time off work)*
13. Other non-medical costs (e.g. patient out-of-pocket expenses)*
14. Unit costs associated with direct medical resource use and/or non-medical resource use*
15. Total direct medical costs*
16. Total productivity costs (time off work)*
17. Total other non-medical costs*
18. Incremental cost-effectiveness, cost-utility or cost benefit*

*Economic outcomes (measures of resource use, costs or cost-effectiveness, including clinical outcomes to be used as (proxy) measures of resource use).

A tool to help identify and describe important measures of resource use (resource inputs or resource outputs) for consideration among specified economic outcomes is provided in Box 2.2.c. This tool is adapted from a core set of resource use items (economic outcome measures) for economic evaluations, which was identified by health economists using a Delphi consensus survey (Thorn et al 2017). Additional lists of items of resource use that may be applicable in specific reviews, depending on the types of interventions, comparators and populations being investigated (e.g. emergency care, community care, remote access care), are also available from the same survey study (Thorn et al 2017).

Box 2.2.c: Suggested core and add-on resource use outcome measures (Thorn et al 2017)

Core items

Hospital care

1. Number of hospital admissions (inpatient stay or day case)
2. Length of stay (e.g. dates or number of nights)
3. Number of hospital outpatient appointments

Emergency care

4. Number of visits to accident and emergency
5. Number of admissions to hospital, after accident and emergency

Care at a GP surgery or health clinic or other community setting

6. Number of appointments (and type of professional seen)

Health care at home

7. Number of healthcare professional visits at home (and type of healthcare professional seen)

Medication

8. Name/class and dose/regimen of medication(s) (including period taken for, e.g. dates or number of days)

Additional items

Hospital care

9. Number of operations/procedures undergone (and type of professional seen, e.g. consultant/nurse)
10. Type of operation/procedure undergone (and type of professional seen, e.g. consultant/nurse)

11. Number of imaging scans undergone (e.g. X-ray/MRI)

Emergency care

12. Number of times paramedic care received

Care at a GP surgery or health clinic or other community setting

13. Number of minor surgery/procedures/treatments undergone (and type of professional seen, e.g. nurse)

Health care at home

14. Number of professional visits for help with daily activities (e.g. washing/dressing)
15. Equipment (e.g. wheelchairs/portable oxygen/specialist clothing) or home adaptation (e.g. grab rails/ramp) supplied

Medication

16. Name, number and dose/regimen of prescribed medications (including period taken for (e.g. dates or number of days)

Remote access care

17. Number of real time telephone/computer contacts with health or social care professional (e.g. with GP or telephone helpline (and type of professional contacted, e.g. doctor/nurse/social worker)

Other community care

18. Number of visits to healthcare professional in the community (and type of healthcare professional seen, e.g. dentist, pharmacist, nurse, counsellor, therapist)
19. Number of visits to social care professional in the community (and type of social care professional seen, e.g. social worker/housing worker/drug and alcohol worker)

Residential care

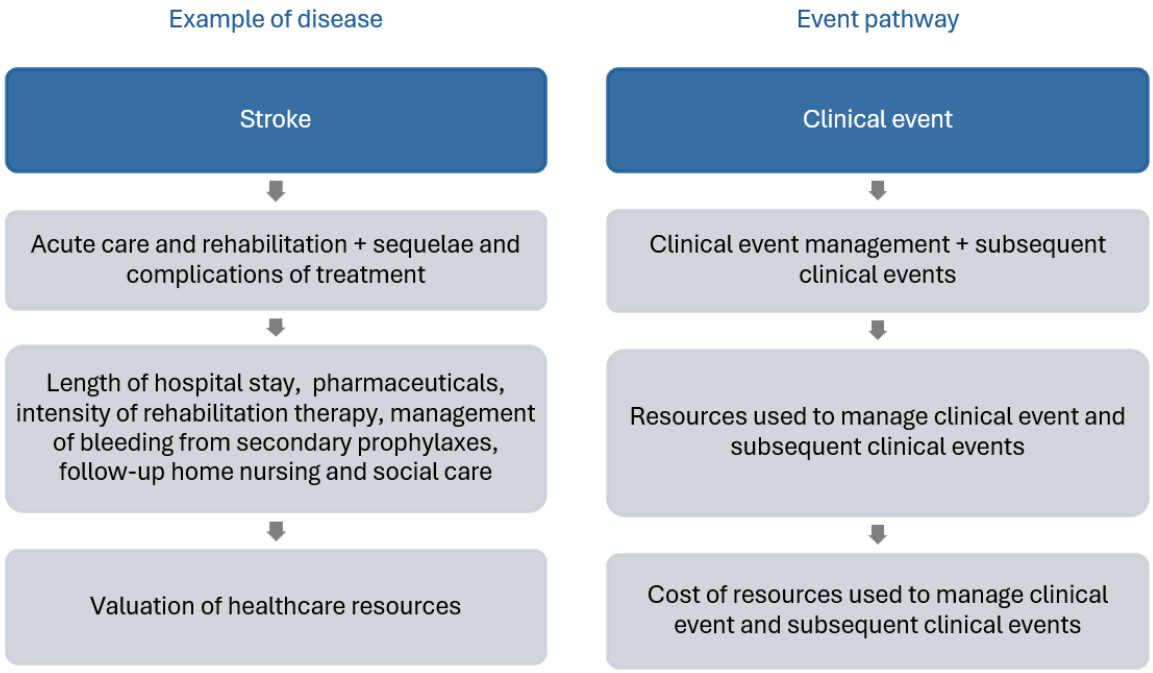
20. Stay in hospice
21. Length of time spent in the hospice
22. Use of short-term respite or rehabilitation care
23. Length of stay in short-term respite or rehabilitation care
24. Living in a nursing home
25. Living in a residential home
26. Living in supported accommodation/sheltered housing

Measures of cost-effectiveness that may be considered in an integrated full systematic review of economic evidence include incremental cost-effectiveness ratios (ICERs), incremental cost-per-QALY and cost-benefit ratios. Note that cost-effectiveness is a composite measure that incorporates both measures of costs (resource use) and effects (health outcomes). (See also the supplementary material on economic evidence available from https://methods.cochrane.org/economics/training-resources).

Clinical event pathways are another useful tool to help identify and describe both the important items of resource use ('economic outcomes') and the important health outcomes. This is conceptually similar to use of logic models to describe theorized causal relationships between an intervention and a set of final health outcomes (Anderson et al 2011, Anderson et al 2013), which may be used when designing a systematic review of effects.

Clinical event pathways describe the main pathways of events that have distinct resource implications or outcome values associated with them. They start at the point of implementation of the interventions or management strategies being compared and continue through subsequent changes in management of patients and outcomes (see also, Chapter 2 of Donaldson et al 2002 (Donaldson et al 2002)). A simplified example of an event pathway for management of acute stroke is shown in Figure 2.2.a. This pathway starts with the clinical event 'acute stroke' and traces the resources needed to manage this event, subsequent clinical events and treatment outcomes of those events along defined patient pathways. Further examples of clinical event pathways, which can be adapted to include identification of important items of resource use at each stage of each pathway (as in the example shown in Figure 2.2.a), can be viewed at: https://pathways.nice.org.uk/.

Figure 2.2.a: Example of an event pathway for management of acute stroke



3 Searching for studies

3.1 Identifying health economic studies to inform an integrated full systematic review of economic evidence

A helpful starting point to inform an integrated full systematic review of economic evidence is to identify cost-of-illness (COI) studies. COI studies give an indication of the size or burden of the problem or condition from an economic perspective. This step helps provide a rationale for the inclusion of an integrated full review of economic evidence within an intervention review. Searching for COI studies does not require the use of systematic search methods. This step is identical to that outlined for a brief economic commentary (see Chapter 20). We recommend that author teams allocate most of the resource available for identifying reports of health economics studies to identifying economic evaluations for inclusion in the review.

3.2 Identifying economic evaluations for the review

This section of the chapter focuses on identifying reports of (full) economic evaluations eligible for consideration in an integrated full systematic review of economic evidence. However, it is important to recognize that relevant economic outcome data (e.g. measures of resource use and/or costs) may also be found in reports of randomized trials and/or other studies of effects meeting the eligibility criteria for consideration in the review of intervention effects of which

the economic review forms part. Detailed guidance on methods for searching for and selecting studies of effects is provided in Chapter 4.

Identifying reports of economic evaluations eligible for consideration in an integrated full systematic review of economic evidence involves using combinations of the following methods:

1. forwards and backwards citation searching from full-text reports of included studies of effects;
2. electronic searching using specialist literature databases;
3. electronic searching using specialist search filters;
4. deploying a specialist machine learning classifier;
5. forwards and backwards citation searching from full-text reports of included economic evaluations; and
6. searching other sources.

We currently recommend that author teams consider using a combination of at least 1 to 3 and 5 and 6. The fourth method is experimental and recommended for use in conjunction with electronic searches and specialist search filters *only* where filtered search yields, and therefore overall screening workloads, remain prohibitively high.

3.2.1 Forwards and backwards citation searching from included studies of effects

Search methods for locating eligible economic evaluations will differ depending on the eligibility criteria set for the integrated full systematic review of economic evidence. However, in all cases, the first procedural step is the same. The first step involves screening full-text reports of eligible studies of effects retrieved for the main systematic review of intervention effectiveness. This screening is to determine which (if any) include reporting of methods and/or results of companion (or 'piggy-back') eligible economic evaluations. This is consistent with the recommendation within Chapter 4 (Section 4.6) that reference lists of included studies should be routinely checked (backward citations). Correspondingly, citations should be tracked forward (forward citations) to identify any further eligible reports. Checking any identified protocols for included studies for descriptions of 'proposed' economic evaluations is also relevant here.

At the end of the screening process, it should be determined whether any of the studies included in the review of clinical effects include reporting of trial-based measures of resource use and/or costs that are specified among target types of economic outcomes (see Chapter 20, Section 20.1.2) but without any formal economic evaluation component. All study reports of economic evaluations and studies of effects with economic outcomes identified by this procedure should also be marked for consideration in the integrated full systematic review of economic evidence.

3.2.2 Electronic searching using specialist literature databases

Several specialist tertiary health economics electronic literature databases have been developed to curate records of health economic evaluations (Box 3.2.a).

Box 3.2.a: Specialist health economics electronic literature databases

The Tufts Medical Center Cost-Effectiveness Analysis Registry

- Published online by the Center for the Evaluation of Value and Risk in Health at http://healtheconomics.tuftsmedicalcenter.org/cear4/Home.aspx (regularly updated but there is a lag of 12 to 18 months).
- The objective of the CEA Registry project is to create a single electronic source comprising all the information contained in a detailed database of cost-utility and cost-effectiveness analyses.
- A searchable database of bibliographic details of cost-utility analyses, which can display extended information about cost-utility ratios (cost-per-QALY), preference weights and the methodology of the underlying analyses (1976 onwards).
- Inclusion criteria are that the study must: report a cost-per-QALY value that can be converted into $/QALY value; be published in a peer-reviewed journal; be reported in the English language.
- Database includes a subjective overall score for methodological quality (1 low to 7 high).
- Public access to a select set of data and paid premium access.

The Global Health Cost Effectiveness Analysis (GHCEA) Registry

- Published online by the Center for the Evaluation of Value and Risk in Health at http://healtheconomics.tuftsmedicalcenter.org/ghcearegistry/
- Compiles articles utilizing the “cost-per-DALY averted” metric to measure the efficacy of health interventions.
- Includes published English-language studies that report original cost-per-DALY assessments from 2000 to 2015.
- Published articles summarised in the Registry undergo a formalized review protocol.

Econlit

- Published by the American Economic Association – see https://www.aeaweb.org/econlit/ for details (subscription only via universities and libraries - updated monthly).
- Contains bibliographic records, abstracts and links to full-text articles in over 750 indexed titles in mainstream economics, including peer-reviewed journals, books, book reviews, collective volume articles (such as conference proceedings and collected essay volumes), working papers (including the Cambridge University

Press Abstracts of Working Papers in Economics) and dissertations covering a range of health and social science topics (1969 onwards).

Paediatric Economic Database Evaluation (PEDE)

- Published online at http://pede.ccb.sickkids.ca/pede/search.jsp (free access - updated periodically, lag of 12 to 18 months).
- Developed using funds provided by the Canadian Agency for Drugs and Technologies in Health (CADTH, formerly the Coordinating Office for Health Technology Assessment, CCHOTA) and The Hospital for Sick Children Research Institute, Toronto, Canada and from in-kind support from the Canadian Institute of Health Economics.
- Contains over 2700 bibliographic records (with added information) of published full economic evaluations of healthcare interventions, programmes, services and processes directed at the paediatric population, including neonates, infants, children or adolescents less than 19 years (1980 onwards).
- Specifically excludes cost analyses, cost descriptions and cost-of-illness studies.

NHS Economic Evaluation Database (NHS EED)

- Published online by the Centre for Reviews and Dissemination (CRD), University of York at http://www.crd.york.ac.uk/CRDWeb.
- Contains > 17,000 records of full economic evaluations, published in all languages (1992 to 2014), including > 9500 quality-assessed structured abstract records.
- NHS EED ceased updating at the end of 2014 and is currently available only as a closed database.

As the information in Box 3.2.a shows, these specialist databases vary substantively in their scope, eligibility criteria and coverage of published reports of health economic evaluations (see also (Pitt et al 2016) for a recent bibliometric analysis with a global perspective). However, the information also illustrates why searches of specialist databases alone are unlikely to be sufficient to enable comprehensive searching for published reports of eligible economic evaluations in a Cochrane Review.

3.2.3 Electronic searching using specialist search filters

Appendices A to F reproduce search filters originally developed to identify published reports of (trial- and model-based) economic evaluations for potential inclusion on NHS EED among records indexed in the MEDLINE (Ovid SP), EMBASE (Ovid SP), PsycINFO (Ovid SP), PubMed (NIH) and CINAHL (EBSCO) databases. With the dual need to locate reports of eligible studies of effects for the review of intervention effects, author teams will need to plan carefully whether and how these electronic search filters are applied in their workflow. It is also

important to highlight that these (and other) 'economic evaluation search filters' are typically highly sensitive with low precision (Glanville et al 2009a, Glanville et al 2009b).

The use of 'study design' search filters may not be necessary when overall search yields using just topic terms are low enough for all retrieved records to be screened within available resources. This equally applies in the case of 'economic evaluation search filters.' However, when yields from searches conducted for the main review of intervention effects are too large to manage, *all* electronic records retrieved from the five electronic literature databases listed above can be filtered using the corresponding 'economic evaluation search filter' in order to help identify relevant reports of economic evaluations. At a minimum, we recommend applying these search filters to retrieved records published from 1 January 2015 onwards, when search yields are prohibitively high. This date is specifically recommended as the NHS EED has not been updated since the end of December 2014 (see Section 3.2.2).

Before applying an 'economic evaluation search filter', it is important to check whether a randomized trial or other study design search filter is already incorporated into the corresponding search strategy developed for the review of intervention effects. If it is, then the randomized trial or other study design filter may need to be removed from the search strategy before appending the 'economic evaluation filter' (that is, unless the eligibility criteria for the integrated full systematic review of economic evidence are limited to trial-based economic evaluations conducted alongside included studies of effects). In principle, these 'economic evaluation search filters' can also be adapted for use in other electronic literature databases. However, we suggest Cochrane author teams consider seeking advice and/or peer review from the CCEMG to help with the adaption.

3.2.4 Deploying a specialist machine learning classifier

When conducting an integrated full systematic review of economic evidence, Cochrane authors may consider deploying an Economic Evaluations classifier (See Box 3.2.b) within the study identification workflow to prioritize those records most likely to report an economic evaluation for screening (i.e. the review author would screen in order of 'probability score' – highest to lowest).

Box 3.2.b: Machine learning classifiers

Machine learning classifiers 'learn' to distinguish between one class of records (the 'positive class') and another (the 'negative class') based on the text contained in the titles and abstracts of those records. Specifically, they assign a 'probability score' to each record retrieved by electronic searches, which represents the likelihood (based on the machine learning model) that it belongs to the 'positive class'. For example, a randomized trial classifier assigns a 'probability score' that reflects how likely a given record does, or does not, report a randomized trial.

Evidence for Policy and Practice Information (EPPI) Reviewer is an online systematic review software application freely available to Cochrane authors to support all stages (or specific stages) of their Cochrane intervention reviews (see https://community.cochrane.org/help/tools-and-software/eppi-reviewer for details). EPPI Reviewer includes a suite of tools designed to support searching for studies and study selection, including user interfaces (UIs) for record management, title-abstract and full-text screening, and machine learning classification (Thomas and Brunton 2007).

EPPI Reviewer includes an Economic Evaluations classifier built by applying machine learning to NHS EED screening data. This Economic Evaluations classifier is relatively new and has therefore yet to be formally validated using 'gold-standard' external datasets (early evaluation and validation work is currently in progress but has not yet been completed). As such, no detailed guidance or recommendations are provided for Cochrane authors about how the Economic Evaluations classifier should be used. However, as evidence accumulates from using this classifier, recommendations will be formulated.

Currently, the Economic Evaluations classifier can be regarded as an optional, experimental tool that should only be considered for use in conjunction with specialist electronic search filters. Cochrane author teams who wish to use the Economic Evaluations classifier should contact the CCEMG in the first instance. The CCEMG will coordinate with the EPPI Centre to support deployment of this Economic Evaluations classifier in integrated full systematic reviews of economic evaluations within EPPI Reviewer. For author teams with access to the Cochrane Register of Studies (CRS) system, the same Economic Evaluations classifier is available here too.

3.2.5 Forwards and backwards citation searching from full-text reports of included economic evaluations

The same recommended methods of 'forwards citation searching' and 'backward citations searching' are applicable to full-text reports of economic evaluations (see Chapter 4 and Section 4.S1 of the technical supplement to that chapter).

3.2.6 Searching other sources

Other potential sources of health economic evaluations include technology assessment databases (e.g. the international HTA database maintained by the International Network of Agencies for Health Technology Assessment, INAHTA), websites of HTA agencies (e.g. NICE in England and CADTH in Canada) and grey literature (e.g. conferences such as the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) and Health Technology Assessment international (HTAi); the Research Papers in Economics (RePEC) working papers collection). Other sources of guidance and links to further resources on searching for health economic evaluations in the grey literature are the HTAi Vortal (Virtual Portal) Etext chapter on Health Economics Information and preprint servers such as medRxiv. The Vortal also provides useful general guidance on search methods, including sources to search and designing search strategies, with access to linked resources.

4 Selecting studies and collecting data

4.1 Records screening and studies selection

Procedures for screening and selecting economic evaluations for inclusion in an integrated full systematic review of economic evidence should follow those applied in the systematic review of intervention effects (see Chapter 4, Section 4.6). In line with general methodological expectations for Cochrane intervention reviews, this should involve (at least) two people working independently to determine the final decision, based on full-text assessment, that an economic evaluation meets the eligibility criteria set for the review and following a pre-specified process for resolving disagreements. Reasons for excluding reports of economic evaluations at the full-text screening stage should also be reported in the 'Characteristics of excluded studies' section.

As in the systematic review of intervention effects, linked study reports (including separate reports of trial-based economic evaluations that link to reports of companion studies of effects) should be treated as belonging to the same study. Once the selected studies are assembled, the set of eligible economic evaluations (linked with companion trials in the case of trial-based economic evaluations) should be accepted for inclusion, along with those included studies of effects that do not include a formal economic evaluation component, but which do report relevant 'economic outcome' measures. These studies collectively represent the set of eligible studies that will be considered by the integrated review of economic evidence.

4.2 Classification of studies

The next stage is to classify (code) included economic evaluations based on their analytic framework and type. The first step is to separate trial- from model-based economic evaluations (if the latter are eligible) and then, within each set, classify the type of economic evaluation that was conducted and reported (e.g. CEA, CUA or CBA). Specialist input may be required here and classification of the type of economic evaluation should not rely on design labels used by study authors, as these are known to be inconsistently applied (Zarnke et al 1997). It is also recommended to classify (and code), at an early stage, the source of effects data used in full economic evaluations. For trial-based economic evaluations, this source will typically be the corresponding randomized trial. For model-based economic evaluation, the source(s) of data inputs for effect size parameters may either be a single study of effects (e.g. a randomized trial) or a summary estimate from a meta-analysis in a systematic review that has synthesized the results of multiple studies of effects.

5 Data extraction

Specific data extraction requirements for an integrated full systematic review of economic evidence will need to be determined on a case-by-case basis. In general, two types of data will

need to be extracted (Brazier et al 2010b): details of the characteristics of included studies and details of their results. As well as being needed to inform assessments of methodological quality (see Section 6.2), data on the characteristics of each included study are also needed to inform investigation of potential sources of variation in study-level estimates of resource use, costs and/or cost-effectiveness (i.e. economic outcomes). Data collection requirements will therefore invariably include details of the PICO characteristics of each study, as well as its design and methods (e.g. details of the analytic framework, type of economic evaluation, year of study, geographic and organizational setting, methods and data sources for resource use, unit cost, source of health state utility values and effects data (for benefits and harms); analytic perspective and time horizon; details of study participant selection and recruitment methods, attrition and losses to follow-up). It is also important to collect information on the price year and currency used to calculate estimates of costs. This process should be followed for both trial- and model-based economic evaluations (if included). In addition, for models it would be necessary to collect details of the model structure and assumptions (see Section 6.2). A challenge that might be faced is that some economic evaluations are very complex so the amount of data that need to be extracted for a study may be substantial. We currently recommend collecting comprehensive data, although further research is needed to help guide review authors so that they focus on the key determinants of incremental costs, incremental effects and incremental cost-effectiveness.

Specific data on study characteristics that, if applicable, may need to be collected from included cost-utility analyses (CUAs) include: the disease-related health states measured (e.g. established osteoporosis, hip, vertebral or wrist shoulder fracture, along with the time period since any osteoporotic incident); health state description system (e.g. EQ-5D, or vignettes); and the valuation technique (e.g. standard gamble, time trade-off, visual analogue scale).

For results, study-level estimates of resource use should be extracted at the level of specific items for both intervention and comparator groups. The type and quantity of each resource used should be extracted in natural units (e.g. length of hospital stay in days, duration of operation in minutes, number of outpatient attendances at six-month follow-up, number of days of work). Wherever possible, study-level measures of incremental resource use and costs should be collected at the ‘per patient' level (i.e. resource use per patient, cost per patient). If reported, estimates of unit costs should be extracted separately, as well as estimates of costs for each group.

If reported, data required to compute a point estimate and a standard error or confidence interval should be extracted for measures of (incremental) resource use, costs and cost-effectiveness, along with numbers of participants in each group for trial-based economic evaluations. Similarly, from each included study that reports estimates of health state utility values, it will be necessary to collect descriptive statistics on the utility values, including sample size, means, medians and standard errors (or data needed to compute these) (Brazier et al 2010b).

It is also useful to collect details of any sensitivity analyses undertaken and their results (i.e. information regarding the impact of varying assumptions on the magnitude and direction of results). This can lead to challenges if a very large number of sensitivity analyses is performed. However, we suggest that data are extracted on all of them.

NHS EED structured abstract records for trial-based and model-based economic evaluations can provide useful templates for the design of data collection forms for use in an integrated full review of economic evidence. Indeed, if a full economic evaluation already has a corresponding NHS EED structured abstract, this may substantively reduce the need for researchers to undertake further data collection from the corresponding full-text study report (although checking for data accuracy and consistency between these two sources is still advised). Structured abstract records can be browsed at https://www.crd.york.ac.uk/CRDWeb/ by inputting basic keyword terms into the search box (title) and ticking boxes to restrict search results to NHS EED and CRD assessed economic evaluation (full abstract) records (inputting an asterisk into the search box (all fields) with the same restrictions returns all structured abstracts archived in the database).

Box 4.2.a shows the template outline used in the NHS EED structured abstract data collection. This template outline can be used for data collection in the review. In addition, some example data collection forms used in an integrated full systematic review of economic evidence in a Cochrane Review of BMP for fracture healing in adults are presented in Appendix 3 of Garrison et al (Garrison et al 2010).

Box 4.2.a: NHS EED structured abstract template outline (structure)*

- Summary
- Type of economic evaluation
- Author's objective
- Interventions
- Location/setting
- Method
- Analytical approach
- Effectiveness data
- Monetary benefit and utility valuations
- Measure of benefit
- Cost data
- Analysis of uncertainty
- Results
- Conclusions

* As an outline this does not have the level of detail that covers all the issues covered in this document, but all these are relevant to include and fit within these broad headers.

6 Addressing risk of bias

The recommended procedure for addressing risk of bias among included economic evaluations is similar to that applied to the main systematic review. This assessment has two main stages, shown in Figure 6.1.a and Figure 6.1.b.

1. Assessment of risk of bias in the single primary study, or among the body of primary studies, that produced the estimates of relative intervention effects used as data inputs to each included economic evaluation.
2. Assessment of the overall methodological quality of each included economic evaluation.

The choice of applicable tools to use in each stage of this procedure depends on whether included economic evaluations are trial-based, model-based or both.

The first stage of this assessment may be able to draw heavily on study-based risk of bias judgements and supporting information being assembled for the main review of intervention effects. This is because in the case of trial- and model-based economic evaluations, effects (and possibly resource use data) may come from a single study. In the case of a model, effects may come from several studies (see Figure 6.1.a) included in the review of effects. However, some data, for example some measures of resource use, may be available but not form part of the main review of intervention effects. In this situation further study-based risk of bias judgements will need to be made for some bias domains (i.e. some risk of bias judgements are outcome-specific).

6.1 Addressing risk of bias in underlying studies of effects

Three tools are discussed in the *Cochrane Handbook* for assessing the risk of bias in the single primary study, or body of primary studies, used to generate the estimate(s) of relative intervention effects used in each included economic evaluation. The appropriateness of these tools depends on whether the economic evaluation is trial-based or model-based (see also supplementary material on economic evidence available from https://methods.cochrane.org/economics/). Two of these tools – for assessing risk of bias in a randomized trial (RoB 2; see Chapter 8) and in a non-randomized study of an intervention (ROBINS-I; see Chapter 25) – are applicable to both trial-based and model-based economic evaluations. The third tool, ROBIS (see Chapter V), assesses risk of bias in a systematic review rather than a primary study. ROBIS is recommended here exclusively for application to model-based economic evaluations (see Section 6.1.3).

6.1.1 Cochrane Risk of Bias tool (RoB 2)

The second version of the Cochrane risk of bias tool for randomized trials (RoB 2) is recommended for assessing risk of bias in estimates of effect size (for beneficial and adverse effects) that are used as data inputs to an included trial- or model-based economic evaluation when the source of those data inputs is a single randomized trial. Similarly, when a trial-based economic evaluation derives its estimates of impacts on resource use and/or health state utility values (HSUVs) (if applicable) from data collected prospectively among participants in the companion trial (single randomized trial), the RoB 2 tool should be used to assess risk of bias in estimates of effect size for these economic outcomes.

6.1.2 The Risk of Bias in Non-randomized Studies - of Interventions (ROBINS-I) assessment tool

The Risk of Bias In Non-randomized Studies of Interventions (ROBINS-I) tool is recommended for assessing risk of bias in estimates of effect size (for beneficial and adverse effects and possibly for resource use and/or HSUVs – see also Section 6.1.1) that are used as data inputs to an included trial- or model-based economic evaluation when the source of those data inputs is a single non-randomized study. It should be used only when: (a) the companion review of intervention effects includes non-randomized studies (see also Chapter 25); and (b) matching eligibility criteria are specified for the integrated full systematic review of economic evidence (i.e. the review will consider evidence from economic evaluations conducted alongside a non-randomized study of effects, as well as those conducted alongside an randomized trial (Sterne et al 2016).

6.1.3 ROBIS assessment tool

The Risk Of Bias In Systematic reviews (ROBIS) assessment tool is recommended for assessing the risk of bias in the summary effect sizes (for beneficial and adverse effects and possibly for resource use and/or HSUVs – see also Section 6.1.1) used as data inputs to an included model-based economic evaluation (Whiting et al 2016) when the source of those data inputs is one or more meta-analyses from a published systematic review. Model-based economic evaluations (especially those developed for health technology assessments (HTAs)) are increasingly developed in conjunction with systematic review workflows. Therefore, they often use summary effect sizes from meta-analyses as data inputs to inform values of corresponding effect size parameters in the mathematical model (Briggs et al 2012).

The ROBIS authors describe the tool as “designed for completion in three phases: (1) assess relevance (optional), (2) identify concerns with the review process, and (3) judge risk of bias. For a Cochrane Review, it is recommended to complete phases 2 and 3 only. Phase 2 covers four domains through which bias may be introduced into a systematic review: study eligibility criteria, identification and selection of studies, data collection and study appraisal, and synthesis and findings. Phase 3 assesses the overall risk of bias in the interpretation of review findings and whether this considered limitations identified in any of the phase 2 domains” (Whiting et al 2016).

In addition to the use of the above tools, we suggest drawing on any summary of risk of bias judgements reported by the authors of those systematic reviews and meta-analyses used in model-based economic evaluations to inform an assessment of overall study limitations (risk of bias) in the body of evidence. The original systematic review authors' ratings of overall study limitations (risk of bias) may be available to inform this assessment if the original systematic review publication includes a GRADE 'Summary of findings' table (coded as 'no serious limitations', 'serious limitations' or 'very serious limitations' see Chapter 15). Additionally, if the original systematic review authors identified serious or very serious limitations, this should be reported in a footnote.

In model-based economic evaluations, which typically incorporate data inputs obtained from multiple different sources, it is also possible that estimates of some measures of resource use (e.g. length of stay or re-intervention rates for a comparator intervention) were derived from administrative databases (Cooper et al 2005, Cooper et al 2010, Coyle et al 2010). This issue is addressed in the second stage of our recommended, two-stage assessment (see Section 6.2), as the choice of data inputs for resource use (and unit cost) parameters primarily relates to the issue of applicability (external validity) as opposed to the issue of bias (internal validity).

In addition, because model-based economic evaluations are frequently designed to predict cost-effectiveness (e.g. cost per QALY) over an extended time horizon (up to 'patient's lifetime'), such model-based analyses typically require estimates of utility values (HSUVs) for several different health states that patients may experience over the specified time horizon. Such data are typically beyond the scope of investigation in a single randomized trial, so may come from multiple studies that collected these data from different populations (e.g. utilities for early disease events, later disease events and palliative care) using different health state description systems (e.g. EQ-5D, vignettes) and valuation techniques (e.g. standard gamble, time trade-off, visual analogue scale) (Brazier et al 2010b). For practical reasons, it may not be feasible to assess risk of bias for all utilities (if this is applicable) in studies that provide data utilized in a model-based economic evaluation. In these circumstances, we recommend limiting risk of bias assessments using the tools described above to those studies of effects and/or economic evaluations that included measurement of the subset of health state utility values that are both incorporated into an included model-based CUA and specified among eligible outcome measures in the Cochrane Review. When no measures of health state utilities values (HSUVs) are specified among eligible outcome measures, this risk of bias assessment need not be conducted.

Figure 6.1.a: First stage - assessing risks of bias in the underlying study of effect

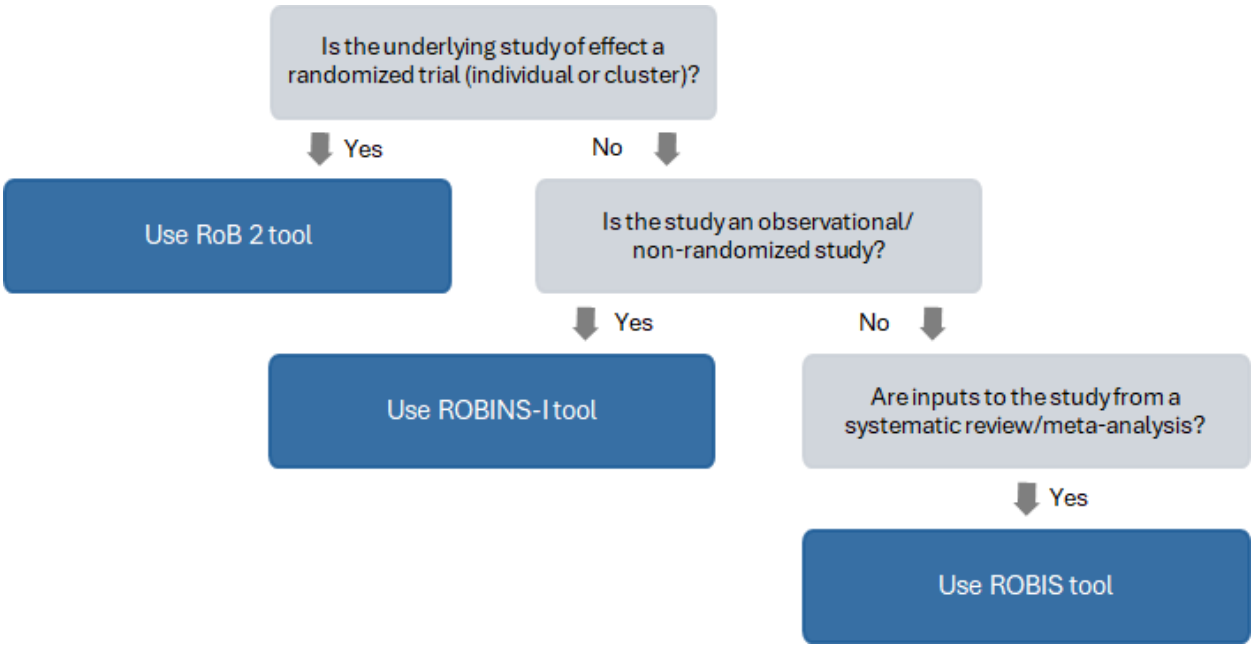
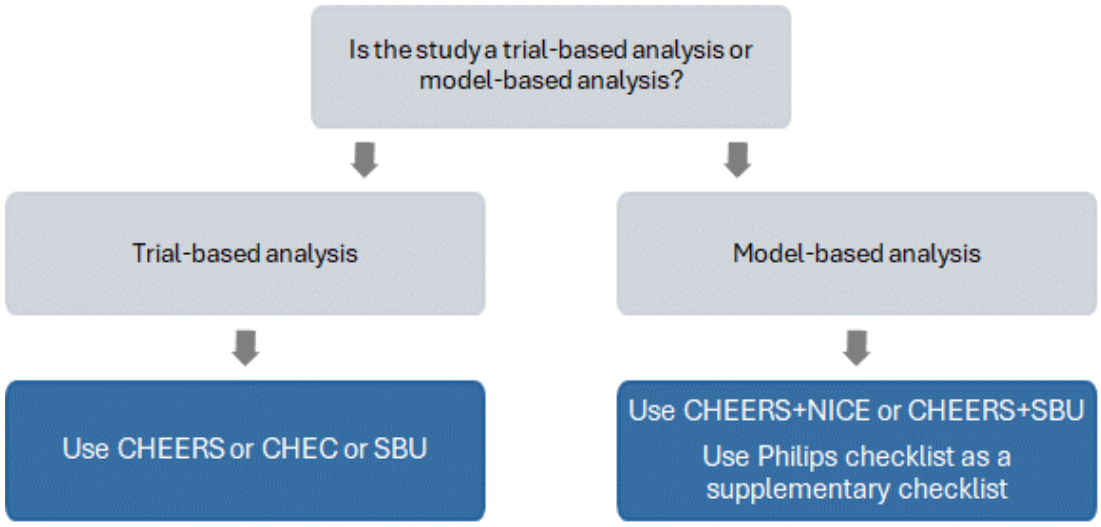


Figure 6.1.b: Second stage - assessing methodological quality of economic evaluation studies



6.2 Assessing the overall methodological quality of each included economic evaluation

This second stage of the assessment involves applying tools designed to inform critical appraisal of the overall quality of included trial- and model-based economic evaluations. Applying these tools involves a series of complex judgements and is likely to require specialist health economics expertise.

The overall methodological quality of an economic evaluation is not exclusively determined by the internal validity of its data inputs; it is also determined by the judgements concerning, for example, in a model-based analysis, the relevance of its mathematical structuring of the decision problem.

The core objective of this second stage is therefore to assess the applicability of the scope of each economic evaluation (e.g. what costs and effects were included in the trial-based evaluation, or whether the structure used to describe the decision problem in a model-based evaluation is appropriate to address the decision problem the economic evaluation is trying to address). Here the emphasis is on applicability and relevance (external validity) as key factors in determining the overall methodological quality of an economic evaluation.

Many critical appraisal tools have been developed to inform assessments of the overall quality of economic evaluation studies. For included trial-based economic evaluations, we currently recommend the combined use of two tools: the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist (Husereau et al 2013, Husereau et al 2022) and the Consensus Health Economic Criteria (CHEC) list (Evers et al 2005). For included model-based economic evaluations, we currently suggest the combined use of two tools (the CHEERS statement and the NICE 'study limitation' checklist), supplemented by reference to a third checklist, the Philips checklist (Philips et al 2004). CHEERS focuses on reporting quality, whilst CHEC and the Philips tools focus on methodological quality.

The Swedish Agency for HTAs and Assessment of Social Services (SBU) (https://www.sbu.se/en/method/) have also developed checklists for assessing the quality of both trial-based studies and model-based economic evaluations (see Section 6.2.5). An alternative to using a combination of CHEERS 2022 and CHEC or a combination of CHEERS 2022 and Philips would be to combine the CHEERS and SBU checklists to assess the methodological quality of both trial- and model-based studies, as SBU has a separate checklist for trial-based and model-based studies. Research is needed to see if this recommendation can be simplified to allow the use of one checklist in some circumstances.

Note that all these tools include items that address the internal and external validity of trial- and/or model-based economic evaluations. However, to a large extent, the internal validity items can be completed by drawing on risk of bias assessments made in the first stage of our recommended two-stage process.

6.2.1 The Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement

This CHEERS statement represents an attempt to consolidate and update previous health economic evaluation guideline efforts into one reporting guideline. The CHEERS 2022 checklist includes two 'tracks' (with several common items) for trial- and model-based economic evaluations and can therefore be applied to inform appraisal of the overall quality of reporting in both designs.

Like many other similar checklists, CHEERS focuses on the quality of the reporting of economic evaluations and not on methodological quality *per se*. Therefore, while reporting checklists of this kind can usefully *inform* the overall assessment of methodological quality by directing authors' attention to all the relevant issues in study reports, further judgements will be required to assess the quality of each study's methodological conduct (i.e. what was done). For example, CHEERS includes an item on 'Estimating resource use and costs', which (for model-based economic evaluations) recommends that authors reporting economic evaluations should "Describe approaches and data sources used to estimate resource use associated with model health states" (see page 3 of Husereau et al (2013) (Husereau et al 2013)). In this case, the review author is directed to identify this description in the study report but will need to make a further judgement about whether the approaches and data sources used were appropriate, given the economic evaluation question and specific decision context at hand (as well as how these relate to the review question at hand, including the specified list of 'economic outcomes').

6.2.2 The CHEC checklist

The Consensus Health Economic Criteria (CHEC) checklist (reproduced in Appendix G) is applicable for the assessment of the overall methodological quality of trial-based economic evaluations (Evers et al 2005) and directly addresses issues of methodological quality (i.e. what was done). The checklist is not appropriate for cost-analyses, nor is it suitable for model-based evaluations, but it can be used for economic evaluation based upon a variety of different study designs (e.g. randomized trials, cohort studies and case-control studies). The 19 items on the checklist are each answered with a 'Yes' or 'No' response. A 'Yes' response is given if a review author thinks the study has sufficiently addressed a given question. Where there is insufficient information for a given question, they provide a 'No' response. A brief explanation of each question covered in the checklist is provided by an assessment guide to help review authors (Maastricht University Care and Public Health Research Institute).

6.2.3 The NICE 'Study limitations' checklist

In general, appraising the methodological quality of a model-based economic evaluation based on published reports of the study can be a conceptually challenging and time-intensive exercise, regardless of the level of expertise and experience of the review authors. Overall, the quality of models depends on:

1. its fitness for purpose: relevance to the underlying research question;

2. the methods by which data inputs to the model are combined (both of which relate to the quality of the model structure);

3. the quality of the source studies from which data are taken; and

4. the model's internal and external consistency, which can broadly be interpreted as the internal and external validity of the model.

These broad dimensions of quality are covered in further detail in an Encyclopaedia of Health Economics entry on Quality Assessment in Modelling in Decision Analytic Models for Economic Evaluation (Shemilt et al 2012, Shemilt et al 2014).

The NICE 'Study limitations' checklist is the 'Study limitations (methodological quality)' section of the 'Methodology checklist: economic evaluations', which appears in Appendix G of the English National Institute for Health and Care Excellence (NICE) 'Guidelines Manual'. The NICE 'Study limitations' checklist is a combined, abridged version of the CHEC Criteria List (Evers et al 2005) (Section 6.2.2) and the Philips Checklist (Philips et al 2004) (Section 6.2.4).

6.2.4 Philips checklist

The Philips checklist is recommended by NICE for quality assessment in decision-analytic models in HTA. This checklist facilitates a more granular assessment of the methodological quality of model-based economic evaluations but is much more time-consuming to complete compared with the NICE 'Study limitations' checklist, the CHEERS tool or the SBU checklist.

We therefore recommend the Philips checklist as a supplementary tool for use in the assessment of the overall methodological quality of model-based economic evaluations. A 2004 review of good practice guidelines in decision-analytic modelling in HTA (Philips et al 2004) identified four dimensions of quality pertaining to data: identification, modelling, incorporation and assessment of uncertainty (which map onto broader dimensions 2 and 3). The checklist consists of four dimensions. Each of these dimensions of 'data quality' has a corresponding set of attributes of good practice, or factors that determine quality, and each attribute refers to the quality of processes used in the identification, appraisal, selection or use of data at different stages of the model development process (Shemilt et al 2012). A key message here is that general expertise in health economics, and specific expertise in model-based economic evaluation (including familiarity with key concepts and methods), is likely to be required to inform a full critical appraisal of model quality informed by this checklist.

6.2.5 The Swedish Agency for Health Technology Assessment and Assessment of Social Services (SBU) checklist

The SBU has produced checklists for model-based and trial-based economic evaluations. The checklists are identical apart from model-specific items included in the former. The 'Quality of economic analysis' section is made up of four parts: model structure; costs and effects; interpretation of results; sensitivity analysis and discounting. Previous checklists (Philips et al 2004, Evers et al 2005, Brunetti et al 2010, Drummond et al 2015) influenced the development

of these checklists. The checklists have been reproduced in Appendix H and Appendix I to suit the Cochrane context, where we recommend the use of the checklist together with the CHEERS checklist.

6.3 Summary assessment of study limitations (risk of bias)

It may be challenging to distil the large number of complex judgements that feed into completing the two-stage process described above into an overall summary rating of study limitations for each included economic evaluation, particularly for model-based economic evaluations. From a procedural perspective, this rating will primarily be needed to inform GRADE assessments of study limitations in the body of evidence that underlie each (summary) estimate of resource use ('economic outcome') that is to be included in a 'Summary of findings' table (see Chapter 15). In these cases, ratings of 'no serious limitations', 'serious limitations' or 'very serious limitations' may be applied.

Apart from those economic evaluations contributing output data to (summary) estimates of differences in resource use, we advise that it is unnecessary to attempt an overall summary rating of study limitations for each included economic evaluation. Instead, we advise incorporating text into the 'Risk of bias in included studies' part of the 'Results' section of the review to summarize the results of the two-stage assessment process. This summary should distinguish between each stage of this process and highlight key issues of bias and quality that emerged from the assessment across the whole body of studies included in the integrated review of economic evidence. It should also summarize the results of assessments for trial- and model-based evaluations separately. This commentary should be supplemented by presentation of risk of bias assessments for underlying studies of effects and presentation of completed checklists in appendices. For example, Appendix 4 in a Cochrane Review of BMP for fracture healing in adults presents checklists completed to inform assessments of methodological quality of economic evaluations (Garrison et al 2010).

6.4 Equity considerations in an integrated full systematic review of economic evidence

Health equity is a key issue on policy agendas, most often associated with the aim of achieving universal health coverage (Ottersen and Norheim 2014, Cotlear et al 2015) and reducing inequality in health (Marmot et al 2012). To further enhance the usefulness of reviews in decision-making and policy, health equity should be considered even at the initial stages of a review. Review authors should explicitly consider the relevance of equity to their review when formulating their review questions at the title and protocol stages to help guide the design of their methods (including methods to identify and appraise evidence related to equity; see Chapter 16, Section 16.2). Health inequity may be experienced across characteristics defined by PROGRESS-Plus (Place of residence, Race/ethnicity/culture/language, Occupation, Gender/sex, Religion, Education, Socio-economic status, Social capital and other factors ('Plus') such as sexual orientation, age and disability (Welch et al 2019).

Whilst economic evaluations are usually conducted with the aim of improving the total health of a population given the available resources (either technical efficiency or allocative efficiency (see glossary of economic terms, Glossary | Cochrane Economics)), they may not consider the impact of the interventions on equity. Healthcare decision-makers often have to consider multiple policy goals simultaneously (e.g. a goal of maximizing health and reducing health inequality (Östlin et al 2011)). However, typically there is a trade-off between these goals. The most effective intervention may not be the most efficient and an effective or efficient intervention may either increase or reduce health inequalities. In recent years, work has been conducted on incorporating equity into economic evaluations (Asaria et al 2016, Verguet et al 2016, Cookson et al 2017, Boujaoude et al 2018, Dukhanin et al 2018, Lal et al 2018). In addition, the Guidance on Priority Setting in health care (GPS-Health) (Norheim et al 2014) checklist was developed to include equity criteria in cost-effectiveness analysis. The checklist explicitly captures comprehensive equity dimensions relevant for priority-setting and decision-making in health care and can be integrated into economic evaluations.

Given that the methods and tools to incorporate equity into an integrated full systematic review of economic evidence are not yet formally developed, at this stage we advise review authors to consider, and explicitly elaborate on the relevance of, considering equity at the title and protocol stages.

It should be noted that none of the checklists recommended in Section 6.2 for assessing the overall methodological quality of each included economic evaluation explicitly include equity items. For an integrated full systematic review of economic evidence (when equity is relevant to the review), in addition to the methods outlined in Chapter 16, review authors might consider developing a checklist based on the GPS-Health checklist and some of the recent equity methods to appraise equity considerations of included economic evaluation studies.

7 Analysing and presenting results

The aim of an integrated full systematic review of economic evidence is not to provide any single, summary estimate of the intervention effect on each economic outcome but to help end-users understand key economic trade-offs between the alternative courses of action (i.e. intervention(s) and comparator(s)) under consideration (Anderson and Shemilt 2010a)). Therefore, the focus is on explaining how interventions produce outcomes of value to patients, health systems and societies and on why such outcomes, including measures of resource use and costs, differ between studies and settings. However, the extent to which this goal can be realised in practice will depend on the types and numbers of identified studies contributing usable economic outcome data to the synthesis.

The guidance in this document therefore places emphasis on the tabulation of the characteristics and results of included health economic evaluation studies, supplemented by a narrative synthesis of quantitative results. This narrative synthesis should focus on describing and summarizing the principal findings of included studies. Additionally, in some

circumstances, meta-analysis of resource use or cost data may be considered as an analytic method for identifying statistical heterogeneity in estimates of economic and health outcomes between studies and settings. In principle, this meta-analytic framework can be extended to investigate factors that may explain any such heterogeneity (see also Chapter 10). These options are described in further detail in the sections below.

7.1 Presenting results in tables

The key characteristics of each study accepted for inclusion into an integrated full systematic review of economic evidence should be presented in the 'Characteristics of included studies' tables, based on collected data (see Chapter 9, Section 9.3). This may be supplemented by presentation of completed data collection forms for each study in an appendix to the published full review.

When presenting tables that summarize the results of included studies (either in the main 'Results' section or in appendices), as in all methods of analysis and presentation of results, we recommend treating trial- and model-based economic evaluations separately. Point estimates (means) for resource use, health state utility values and/or costs should be presented along with relevant measures of precision, such as standard errors, bootstrapped 95% credible intervals or interquartile ranges (IQRs) for both the target intervention and each of its comparators (including sample sizes), as well as the descriptive statistics for *incremental* measures. It is also important to state the currency and price year alongside estimates of (incremental) costs and/or (incremental) cost-effectiveness (if reported).

Cost estimates should be presented in the original currency and price year. Each estimate can also be converted into a common currency and price year using the CCEMG-EPPI Centre Cost Converter (an online tool developed and maintained by members of the CCEMG). The tool is underpinned by international exchange rate data based on purchasing power parities (PPPs) and gross domestic product (GDP) deflator indices, which are preferred for converting cost estimates to a common target currency and price year (for example, as opposed to pure exchange rate conversion). A sensible target currency and price year to use in a Cochrane intervention review might be US Dollars for the most recent price year for which data are available (note that 1 US Dollar is equivalent to 1 International Dollar in a given price year (Shemilt et al 2010)).

7.2 Narrative synthesis of quantitative results

A narrative summary of the main characteristics and results of included economics evaluations should be developed to supplement, and provide a commentary on, tabulated results. This can be placed in the Results section, alongside and placed in the context of the narrative summary of the (synthesized) quantitative results of included studies of effects (see Chapter 12). Again, the narrative summary should distinguish between trial- and model-based economic evaluations. The central aim of a narrative synthesis is to make explicit the extent to which

estimates collected from multiple studies are heterogeneous between studies. This can be achieved by:

- summarizing variation in the design of studies and methods for assessing economic outcomes;
- describing observed patterns of relative effect sizes for resource use, costs, health state utility values and health outcomes, and patterns of cost-effectiveness results, among included studies; and
- offering potential explanations for inconsistent results between studies and settings (see Section 7.3).

Other features of good practice specific to a narrative synthesis of the quantitative results of studies included in an integrated full systematic review of economic evidence include:

- reporting the currency and price year alongside estimates of costs extracted from included studies;
- converting cost estimates extracted from reports of each included study to a common currency and price year and presenting converted estimates; and
- highlighting the key features of sensitivity analyses undertaken within included studies and the consistency of results obtained when key parameters are varied in sensitivity analysis both within and among included studies.

7.3 Meta-analysis of resource use and cost data

Meta-analysis (see Chapter 10) of study-level estimates of resource use and cost may be considered but should not be done to obtain (weighted average) estimates of the 'true' sizes of resource use or cost differences between the alternative courses of action. This is because included economic evaluations will likely be set in different countries, jurisdictions or time points and any such summary estimates are unlikely to be applicable to any specific decision-making context (Anderson 2010, Drummond 2010) (see Chapter 12, Section 12.1). In cases where it appears that differences between countries or jurisdictions are leading to differences in the economic outcomes, a key contribution of the review would be to highlight these differences so that readers of the review can appreciate how cross-country differences can impact on resource use, cost and cost-effectiveness.

Nonetheless, judicious application of meta-analysis methods may be useful if its explicit purpose is to identify and investigate statistical heterogeneity in effect sizes for resource use and/or cost outcomes. Since these are often measured as continuous outcomes, review authors should be familiar with issues specific to effect sizes for this type of outcome data (Chapter 10, Section 10.5). Any such investigations of heterogeneity should be based on *a priori* hypotheses about specific factors expected to drive systematic variation in study-level

differences between the intervention(s) and its comparator(s). Data on these specific factors will need to have been collected as study-level variables in the data collection stage (see Chapter 5, Section 5.4). They can then be incorporated as covariates into a meta-regression analysis designed to explore whether these factors (individually and/or collectively) are associated with resource use and/or cost outcomes. The core principles and methods of meta-regression analysis described in Chapter 10 (Section 10.11.4) are also applicable here. Furthermore, guidance and support are likely to be needed from an experienced statistician, along with specialist input from a health economist, to design and implement meta-regression analysis to support the analysis of economic outcomes in a Cochrane intervention review.

In general, we recommend that it may, in principle, be appropriate to combine study-level estimates of relative treatment effects for measures of *resource use* using meta-analysis, provided these outcome data are considered by the authors to represent a common metric (i.e. to have a common 'meaning') between included studies, or if they can be converted into a common metric prior to analysis (e.g. operative time in minutes, length of hospital stay in days, revisional surgical procedure) (Brunetti et al 2013). For example, it is only sensible to combine study-level effect sizes for a measure of the quantity of a drug provided to patients if the same drug regimen (dose, administration method and treatment duration) was implemented consistently among included studies.

Meta-analysis of study-level measures of relative intervention effects for measures of *costs* may, in principle, be feasible and useful only under the following more limited set of conditions.

- Study-level estimates of costs should incorporate sufficiently similar cost components (i.e. a similar 'bundle' of resource inputs have been incorporated into the estimated cost) between studies, for both intervention and control conditions.
- If the above condition is met, cost estimates for each comparison group (intervention(s) and comparator(s)) should be converted to a common currency and price year.
- Statistical distributions of cost data (a continuous outcome) are often skewed at study-level. If so, this may need to be taken into account in any statistical analysis (see Chapter 10, Section 10.5).

Applications of meta-analysis to costs and outcome data have been explored in individual studies (Anderson and Shemilt 2010b), but further evaluation of this approach in systematic reviews of interventions, in conjunction with meta-regression analysis, is needed (see Box 7.3.a).

Box 7.3.a: Application of meta-analysis

- Meta-analysis and meta-regression analysis have been applied to synthesize estimates of health state utility values (HSUVs) collected from multiple studies (Peasgood and Brazier 2015, Brazier et al 2019), including investigation of variation by the different methods (e.g. an elicitation or preference-based measure) used to obtain values between studies and settings (Brazier et al 2010a, Peasgood and Brazier 2015).
- A possible synthesis approach is to apply strict eligibility criteria within the included studies to reduce heterogeneity, such as limiting HSUVs to studies with the same preference-based measure and population. This method is appropriate only when sufficient studies meet the stricter eligibility criteria (Peasgood and Brazier 2015). For example, a study on health states of osteoporosis conditions excluded all HSUVs not collected with EQ-5D (the measure preferred by NICE in England) and combined data from nine studies to estimate mean HSUVs (Brazier et al 2019).
- Gyllensten et al (2022) (Gyllensten et al 2022) conducted a meta-analysis on the costs associated with screening for retinopathy of prematurity (ROP), resource use among people who develop ROP and lifetime costs (for treatment and follow-up). The authors reported that comparisons between studies were challenging due to a lack of detail on cost and resource use, although quality assessment of studies indicated low risk of bias.
- Some studies have used multilevel modelling techniques, which take into consideration that several economic evaluations provide multiple estimates of resource use, costs and effects in the form of a base case analysis along with a series of sensitivity and/or subgroup analysis. The conclusions of such an analysis might be biased if sensitivity and subgroup analyses are selectively reported among included studies or this approach might not be useful for small sets of included studies (Shemilt et al 2014).
- Meta-analysis with subgroup analysis has been conducted to compare effect sizes for primary health outcomes among included studies of effects with a concurrent economic evaluation (i.e. included trial-based economic evaluations), with effect sizes among included studies of effects *without* a concurrent economic evaluation (see, for example, Gilbody et al (2007)(Gilbody et al 2007)). The results of such an analysis (which requires at least two study-level effect sizes within each subgroup) may highlight similarity, or uncover differences, between these subgroups (differences may also suggest the presence of reporting bias – see also Section 7.4).

Furthermore, in recent times there has been growing use of comparative efficiency research methodology (Crespo et al 2014) for the statistical combination of study-level estimates of cost-effectiveness extracted from multiple economic evaluations (i.e. meta-analysis of economic evaluation studies (MAEEs)) (Crespo et al 2014, Bagepally et al 2022). Although this

may not be relevant in an integrated review of economic evidence in a Cochrane Review whose purpose is not to generate a single estimate, MAEEs can be useful for evidence-informed healthcare decisions and policies (Veettil et al 2022), especially in settings with limited economic evaluations due to resource constraints. The World Health Organization (WHO) Immunization and Vaccine-related Implementation Research Advisory Committee (IVIR-AC) has approved MAEEs to demonstrate the cost-effectiveness of vaccines (World Health Organisation 2021). Some studies that have conducted MAEEs include Bagepally et al 2019 and 2020 (Bagepally et al 2019, Bagepally et al 2020), Syeed et al (2023) (Syeed et al 2023), Udayachalerm et al (2022) (Udayachalerm et al 2022) and Veettil et al (2023) (Veettil et al 2023). Large heterogeneity is observed in MAEEs and there is currently no consensus on a standardized approach to handle this heterogeneity (Veettil et al 2022). Further research is needed to explore the relevance of conducting a MAEE in an integrated review of economic evidence in a Cochrane Review.

In conclusion, if quantitative analyses of 'economic outcome' data are conducted, using any of the methods described above (or any other method) as part of a Cochrane intervention review, then a narrative summary of these results should be included in the corresponding 'Results' section of the published full review.

7.4 Addressing reporting biases

Methods for addressing publication bias covered in Chapter 13 can be applied, with the same caveats, in an integrated full systematic review of economic evidence.

As described in Section 7.3, meta-analyses within subgroups to compare effect sizes for primary health outcomes among included studies of effects with and without a concurrent economic evaluation can be informative, as differences between subgroups may indicate the presence of publication bias (see, for example, (Gilbody et al 2007)) widely recognized that commercial and other pressures may affect the funding of studies and reporting of the results of studies that focus on the economic value of healthcare interventions (Drummond et al 1992). Despite this, relatively little research attention has been focused on the issue of publication and related biases in economic evaluation studies, compared with coverage of this issue with respect to primary studies of effects. However, some studies have examined reporting bias using systematic reviews and research synthesis methods.

Bell et al (2006)(Bell et al 2006), in their systematic review of published cost-effectiveness studies in health care, found that studies sponsored by industry were more likely to report ratios that fall beneath, and cluster around, commonly proposed cost-effectiveness acceptability thresholds when compared with studies sponsored by non-industry source. Miners and colleagues conducted a similar study to compare evidence on cost-effectiveness submitted to the National Institute for Health and Clinical (now Care) Excellence (NICE) by manufacturers of the relevant healthcare technologies and by contracted university-based assessment groups respectively (Miners et al 2005). They found that estimated incremental cost-effectiveness ratios (ICERs) submitted by manufacturers were, on average, lower than

those provided by the assessment groups for the same technology; that is, manufacturers' estimates of cost-effectiveness were more favourable than the assessment groups. A similar finding was found by Friedberg and colleagues (Friedberg et al 1999).

In order to examine outcome reporting bias and dissemination bias in economic evaluation studies, Thorn et al (2013)(Thorn et al 2013) identified and selected 100 randomized trials (starting from year 2000 and ending before 2008) with a planned concurrent trial-based economic evaluation that had been registered on the International Randomised Controlled Trial Number (ISRCTN) database. Only 43 of these trials had published their planned economic evaluation results by the year 2010 as compared with the corresponding studies of effects, of which 70 had published their results.

A common theme of discussion among authors of these methodology review studies is the proposal that reporting, or publication, biases are likely to be instrumental in the observed patterns of results. The general hypothesis is that sponsors, authors or journal editors, consciously or unconsciously, do not publish economic analyses with results that suggest an intervention may be economically unattractive. If true, this would suggest that health economic evaluations from studies with equivocal or negative findings are less likely to be published.

7.5 Interpreting results and drawing conclusions

There are clearly many potential resource allocation decisions that could be informed by the findings of a Cochrane intervention review, which are intended for an international audience. However, most published economic evaluations are still dominated by authors from high-resource settings, and hence the relevance of economic evidence in intervention reviews is likely to be limited for lower-resource settings. Decision-makers in different constituencies will need to consider a range of contextual factors, many of which are specific to the target health system (and related systems) and population(s).

In this context, it is not feasible to interpret the results of a Cochrane integrated full systematic review of economic evidence to make general claims or draw definitive conclusions about the cost-effectiveness of an intervention, or whether a treatment option should be adopted or rejected. Whilst in these circumstances review findings can inform only some aspects of any policy decision, it can still help to refine an economic discussion and to set this in an international context (Gilbody and Petticrew 1999).

Interpretation of the findings of an integrated full systematic review of economic evidence needs to be informed, as far as possible, by full assessment of the body of evidence that underpins each finding. It should also seek to place the economic evidence in the context of the evidence of effects. To draw out similarities and inconsistencies between findings, for example, are estimates of effect used in economic evaluations similar to those found in the review of effects?

For (synthesized) estimates of relative intervention effects for resource use and costs, full assessment of the body of evidence can be undertaken using the GRADE system (and, if the outcome is judged sufficiently important to decision-making, GRADE certainty ratings can be presented in 'Summary of findings' tables (see Chapter 14). The overall GRADE approach to rating economic evidence is summarized in the introduction section and we recommend consulting Brozek et al (2021) (Brozek et al 2021) and Brunetti et al (2013) (Brunetti et al 2013)for detailed guidance. Presently, the focus of this guidance is on how to rate the certainty of (bodies of) evidence for (summary) estimates of resource use and effects among included studies (i.e. assessment of two important 'data inputs' that are combined within each included economic evaluation to estimate incremental cost-effectiveness). In part, this rating should draw on the overall rating of study limitations among included economic evaluations, based on authors' assessments of risk of bias and overall methodological quality (see Section 6).

We do not *currently* recommend applying GRADE to rate the certainty of evidence for the final 'output(s)' of trial-based or model-based economic evaluations (i.e. estimates of incremental costs and/or incremental cost-effectiveness). However, the GRADE Working Group has developed new guidance to assess the certainty of evidence for model-based economic evaluations in collaboration with members of CCEMG (Brozek et al (2021)(Brozek et al 2021)). The new GRADE approach will consider the credibility of the model and the certainty of the input evidence as drivers of the overall quality of the outputs (i.e. ICERs) and would be assessed in conjunction with the other GRADE domains (i.e. inconsistency, indirectness, imprecision). In the interim, we recommend that any commentary on published estimate(s) of incremental cost-effectiveness among included economic evaluations, which is presented in the 'Discussion' section of the review, should fully consider and comment on the overall methodological quality of those evaluations (see Section 6.2), as well as avoid any generalized claims or conclusions.

Finally, if an integrated full review of economic evidence identifies few or no eligible economic evaluations, this can usefully highlight a lack of economic evidence, which future research may need to address. If review authors judge there is a need for new economic evaluation studies to be conducted, this can be highlighted in the 'Implications for research' part of the 'Authors' conclusions' section of the review.

8 References

Aluko P, Graybill E, Craig D, Henderson C, Drummond M, Wilson EC, Robalino S, Vale L, Campbell, Group CEM. Economic evidence. In: Higgins JPT TJ, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editor. *Cochrane Handbook for Systematic Reviews of Interventions*2023. p. 507-523. http://www.training.cochrane.org/handbook.

Anderson LM, Petticrew M, Rehfuess E, Armstrong R, Ueffing E, Baker P, Francis D, Tugwell P. Using logic models to capture complexity in systematic reviews. *Research Synthesis Methods* 2011; **2**: 33-42.

Anderson LM, Petticrew M, Chandler J, Grimshaw J, Tugwell P, O'Neill J, Welch V, Squires J, Churchill R, Shemilt I. Introducing a series of methodological articles on considering complexity in systematic reviews of interventions. *Journal of Clinical Epidemiology* 2013; **66**: 1205-1208.

Anderson R, Shemilt I. The role of economic perspectives and evidence in systematic review. *Evidence-based Decisions and Economics: Health Care, Social Welfare, Education and Criminal Justice* 2010a: 23-42.

Anderson R, Shemilt I. The role of economic perspectives and evidence in systematic review. *Evidence-based decisions and economics: health care, social welfare, education and criminal justice 2nd ed Oxford, UK: Wiley-Blackwell* 2010b: 23-42.

Anderson R. Systematic reviews of economic evaluations: utility or futility? *Health Economics* 2010; **19**: 350-364.

Asaria M, Griffin S, Cookson R. Distributional cost-effectiveness analysis: a tutorial. *Medical Decision Making* 2016; **36**: 8-19.

Bagepally BS, Gurav YK, Anothaisintawee T, Youngkong S, Chaikledkaew U, Thakkinstian A. Cost utility of sodium-glucose cotransporter 2 inhibitors in the treatment of metformin monotherapy failed type 2 diabetes patients: a systematic review and meta-analysis. *Value in Health* 2019; **22**: 1458-1469.

Bagepally BS, Gurav YK, Anothaisintawee T, Youngkong S, Chaiyakunapruk N, McEvoy M, Attia J, Thakkinstian A. Glucagon-like peptide 1 agonists for treatment of patients with type 2 diabetes who fail metformin monotherapy: systematic review and meta-analysis of economic evaluation studies. *BMJ Open Diabetes Research and Care* 2020; **8**: e001020.

Bagepally BS, Chaikledkaew U, Chaiyakunapruk N, Attia J, Thakkinstian A. Meta-analysis of economic evaluation studies: data harmonisation and methodological issues. *BMC Health Services Research* 2022; **22**: 202.

Bell CM, Urbach DR, Ray JG, Bayoumi A, Rosen AB, Greenberg D, Neumann PJ. Bias in published cost effectiveness studies: systematic review. *BMJ* 2006; **332**: 699-703.

Boujaoude M-A, Mirelman AJ, Dalziel K, Carvalho N. Accounting for equity considerations in cost-effectiveness analysis: a systematic review of rotavirus vaccine in low-and middle-income countries. *Cost Effectiveness and Resource Allocation* 2018; **16**: 18.

Brazier J, Papaioannou D, Cantrell A, Paisley S, Herrmann K. *Identifying and reviewing health state utility values for populating decision models*: Wiley-Blackwell, Oxford; 2010a.

Brazier J, Papaioannou D, Cantrell A, Paisley S, Herrmann K. *Identifying and reviewing health state utility values for populating decision models*. Second ed. Ian Shemit MM, Luke Vale, Kevin Marsh and Cam Donaldson, editor: Wiley-Blackwell, Oxford; 2010b.

Brazier J, Ara R, Azzabi I, Busschbach J, Chevrou-Séverac H, Crawford B, Cruz L, Karnon J, Lloyd A, Paisley S. Identification, review, and use of health state utilities in cost-effectiveness models: an ISPOR good practices for outcomes research task force report. *Value in Health* 2019; **22**: 267-275.

Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group–6. *Medical Decision Making* 2012; **32**: 722-732.

Brito V, Ciapponi A, Kwong J. Factor Xa inhibitors for acute coronary syndromes. *Cochrane Database of Systematic Reviews* 2011.

Brozek JL, Canelo-Aybar C, Akl EA, Bowen JM, Bucher J, Chiu WA, Cronin M, Djulbegovic B, Falavigna M, Guyatt GH. GRADE Guidelines 30: the GRADE approach to assessing the certainty of modeled evidence—An overview in the context of health decision-making. *Journal of Clinical Epidemiology* 2021; **129**: 138-150.

Brunetti M, Ruiz F, Lord J, Pregno S, Oxman AD. Grading economic evidence. In: Shemilt I MM, Vale L, Marsh K, Donaldson C, editor. *Evidenced Based Decisions and Economics*. Oxford: Wiley-Blackwell; 2010.

Brunetti M, Shemilt I, Pregno S, Vale L, Oxman AD, Lord J, Sisk J, Ruiz F, Hill S, Guyatt GH. GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. *Journal of Clinical Epidemiology* 2013; **66**: 140-150.

Cookson R, Mirelman AJ, Griffin S, Asaria M, Dawkins B, Norheim OF, Verguet S, Culyer AJ. Using cost-effectiveness analysis to address health equity concerns. *Value in Health* 2017; **20**: 206-212.

Cooper JD, Smith KJ, Ritchey AK. A cost-effectiveness analysis of coagulation testing prior to tonsillectomy and adenoidectomy in children. *Pediatric Blood & Cancer* 2010; **55**: 1153-1159.

Cooper N, Coyle D, Abrams K, Mugford M, Sutton A. Use of evidence in decision models: an appraisal of health technology assessments in the UK since 1997. *Journal of Health Services Research & Policy* 2005; **10**: 245-250.

Cotlear D, Nagpal S, Smith O, Tandon A, Cortez R. *Going universal: how 24 developing countries are implementing universal health coverage from the bottom up*: The World Bank; 2015.

Coyle D, Lee KM, Cooper NJ. Use of evidence in decision models. *Evidence-based Decisions and Economics: Health Care, Social Welfare, Education and Criminal Justice*. 2nd ed. London: Wiley: BMJ; 2010. p. 106-113.

Crespo C, Monleon A, Díaz W, Ríos M. Comparative efficiency research (COMER): meta-analysis of cost-effectiveness studies. *BMC Medical Research Methodology* 2014; **14**: 1-9.

Dawoud D, Lamb A, Moore A, Bregman C, Rupniewska E, Paling T, Wolfram V, Lovett RE, Dent R. Capturing what matters: updating NICE methods guidance on measuring and valuing health. *Quality of Life Research* 2022; **31**: 2167-2173.

Donaldson C, Mugford M, Vale L. From effectiveness to efficiency: an introduction to evidence-based health economics. *Evidence-based Health Economics: From Effectiveness to Efficiency in Systematic Review* 2002.

Drummond M. Evidence-based decisions and economics: an agenda for research. *Evidence-based Decisions and Economics: Health Care, Social Welfare, Education and Criminal Justice* 2010: 181-182.

Drummond MF, Bloom BS, Carrin G, Hillman AL, Hutchings HC, Knill-Jones RP, De Pouvourville G, Torfs K. Issues in the cross-national assessment of health technology. *International Journal of Technology Assessment in Health Care* 1992; **8**: 670-682.

Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*: Oxford University Press; 2015.

Dukhanin V, Searle A, Zwerling A, Dowdy DW, Taylor HA, Merritt MW. Integrating social justice concerns into economic evaluation for healthcare and public health: A systematic review. *Social Science & Medicine* 2018; **198**: 27-35.

Evers S, Goossens M, De Vet H, Van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. *International Journal of Technology Assessment in Health Care* 2005; **21**: 240-245.

Friedberg M, Saffran B, Stinson TJ, Nelson W, Bennett CL. Evaluation of conflict of interest in economic analyses of new drugs used in oncology. *JAMA* 1999; **282**: 1453-1457.

Garrison KR, Shemilt I, Donell S, Ryder JJ, Mugford M, Harvey I, Song F, Alt V. Bone morphogenetic protein (BMP) for fracture healing in adults. *Cochrane Database of Systematic Reviews* 2010.

Gilbody S, Bower P, Sutton AJ. Randomized trials with concurrent economic evaluations reported unrepresentatively large clinical effect sizes. *Journal of Clinical Epidemiology* 2007; **60**: 781-786.

Gilbody SM, Petticrew M. Rational decision-making in mental health: the role of systematic reviews. *The Journal of Mental Health Policy and Economics* 1999; **2**: 99-106.

Glanville J, Fleetwood K, Yellowlees A, Kaunelis D, Mensinkai S. Development and testing of search filters to identify economic evaluations in MEDLINE and EMBASE. *Ottawa: Canadian Agency for Drugs and Technologies in Health* 2009a.

Glanville J, Kaunelis D, Mensinkai S. How well do search filters perform in identifying economic evaluations in MEDLINE and EMBASE. *International Journal of Technology Assessment in Health Care* 2009b; **25**: 522-529.

Gyllensten H, Humayun J, Sjöbom U, Hellström A, Löfqvist C. Costs associated with retinopathy of prematurity: a systematic review and meta-analysis. *BMJ Open* 2022; **12**: e057864.

Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, Augustovski F, Briggs AH, Mauskopf J, Loder E. Consolidated health economic evaluation reporting standards (CHEERS) statement. *Cost Effectiveness and Resource Allocation* 2013; **11**: 6.

Husereau D, Drummond M, Augustovski F, de Bekker-Grob E, Briggs AH, Carswell C, Caulley L, Chaiyakunapruk N, Greenberg D, Loder E. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations. *International Journal of Technology Assessment in Health Care* 2022; **38**.

Lal A, Moodie M, Peeters A, Carter R. Inclusion of equity in economic analyses of public health policies: systematic review and future directions. *Australian and New Zealand Journal of Public Health* 2018; **42**: 207-213.

Maastricht University Care and Public Health Research Institute. CHEC list: Consensus Health Economic Criteria, Assessment instruction. https://www.maastrichtuniversity.nl/research/caphri/our-research/creating-value-based-health-care/chec-list-consensus-health-economic.

Marmot M, Allen J, Bell R, Bloomer E, Goldblatt P. WHO European review of social determinants of health and the health divide. *The Lancet* 2012; **380**: 1011-1029.

Miners A, Garau M, Fidan D, Fischer A. Comparing estimates of cost effectiveness submitted to the National Institute for Clinical Excellence (NICE) by different organisations: retrospective study. *BMJ* 2005; **330**: 65.

National Institute for Health and Care Excellence. NICE health technology evaluations: the manual 2022. https://www.nice.org.uk/process/pmg36/chapter/introduction-to-health-technology-evaluation.

Norheim OF, Baltussen R, Johri M, Chisholm D, Nord E, Brock D, Carlsson P, Cookson R, Daniels N, Danis M. Guidance on priority setting in health care (GPS-Health): the inclusion of equity criteria not captured by cost-effectiveness analysis. *Cost Effectiveness and Resource Allocation* 2014; **12**: 18.

Östlin P, Schrecker T, Sadana R, Bonnefoy J, Gilson L, Hertzman C, Kelly MP, Kjellstrom T, Labonté R, Lundberg O. Priorities for research on equity and health: towards an equity-focused health research agenda. *PLOS Medicine* 2011; **8**: e1001115.

Ottersen T, Norheim OF. Making fair choices on the path to universal health coverage. SciELO Public Health; 2014.

Peasgood T, Brazier J. Is meta-analysis for utility values appropriate given the potential impact different elicitation methods have on values? *Pharmacoeconomics* 2015; **33**: 1101-1105.

Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, Woolacott N, Glanville J. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. 2004.

Pitt C, Goodman C, Hanson K. Economic evaluation in global perspective: a bibliometric analysis of the recent literature. *Health Economics* 2016; **25**: 9-28.

Sculpher MJ, Claxton K, Drummond M, McCabe C. Whither trial-based economic evaluation for health care decision making? *Health Economics* 2006; **15**: 677-687.

Shemilt I, Thomas J, Morciano M. A web-based tool for adjusting costs to a specific target currency and price year. *Evidence & Policy: A Journal of Research, Debate and Practice* 2010; **6**: 51-59.

Shemilt I, Wilson E, Vale L. Quality Assessment in Modeling in Decision Analytic Models for Economic Evaluation. Elsevier; 2012.

Shemilt I, Wilson E, Vale L. Quality Assessment in Modeling in Decision Analytic Models for Economic Evaluation. 2014.

Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, Henry D, Altman DG, Ansari MT, Boutron I. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016; **355**: i4919.

Syeed MS, Ghule P, Le LM, Veettil SK, Horn EK, Perdrizet J, Wasserman M, Thakkinstian A, Chaiyakunapruk N. Pneumococcal vaccination in children: a systematic review and meta-analysis of cost-effectiveness studies. *Value in Health* 2023; **26**: 598-611.

Thomas J, Brunton J. EPPI-Reviewer: software for research synthesis. 2007.

Thorn JC, Noble SM, Hollingworth W. Timely and complete publication of economic evaluations alongside randomized controlled trials. *Pharmacoeconomics* 2013; **31**: 77-85.

Thorn JC, Brookes ST, Ridyard C, Riley R, Hughes DA, Wordsworth S, Noble SM, Thornton G, Hollingworth W. Core items for a standardized resource use measure (ISRUM): expert Delphi consensus survey. *Value in Health* 2017; **21**: 640-649.

Udayachalerm S, Renouard MG, Anothaisintawee T, Thakkinstian A, Veettil SK, Chaiyakunapruk N. Incremental net monetary benefit of herpes zoster vaccination: a systematic review and meta-analysis of cost-effectiveness evidence. *Journal of Medical Economics* 2022; **25**: 26-37.

Veettil SK, Syeed MS, Noviyan R, Thakkinstian A, Chaiyakunapruk N. Does meta-analysis of economic evaluations have the potential to play a role in healthcare decision-making in the United States? *Journal of Medical Economics* 2022; **25**: 750-754.

Veettil SK, Vincent V, Shufelt T, Behan E, Syeed MS, Thakkinstian A, Young DC, Chaiyakunapruk N. Incremental net monetary benefit of biologic therapies in moderate to severe asthma: a systematic review and meta-analysis of economic evaluation studies. *Journal of Asthma* 2023: 1-13.

Verguet S, Kim JJ, Jamison DT. Extended cost-effectiveness analysis for health policy assessment: a tutorial. *Pharmacoeconomics* 2016; **34**: 913-923.

Welch VA, Petkovic J, Jull J, Hartling L, Klassen T, Kristjansson E, Pardo JP, Petticrew M, Stott DJ, Thomson D, Erin U, Katrina W, Camila Y, Peter T. Equity and specific populations. In: Higgins JT CJ, Cumpston M, Li T, Page M, Welch V, editor. *Cochrane Handbook for Systematic Reviews of Interventions*. 6.0 ed: Wiley; 2019. p. 433-449. www.training.cochrane.org/handbook.

Whiting P, Savović J, Higgins JP, Caldwell DM, Reeves BC, Shea B, Davies P, Kleijnen J, Churchill R. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. *Journal of Clinical Epidemiology* 2016; **69**: 225-234.

World Health Organisation. Weekly epidemiological record, No 17, 96:133-144: WHO; 2021. https://www.who.int/publications/i/item/weekly-epidemiological-record-no-17-2021-96-133-144 / https://iris.who.int/bitstream/handle/10665/341059/WER9617-eng-fre.pdf?sequence=1.

Zarnke KB, Levine MA, O'Brien BJ. Cost-benefit analyses in the health-care literature: don't judge a study by its label. *Journal of Clinical Epidemiology* 1997; **50**: 813-822.

9 Appendices

Appendix A: Example of NHS EED search filter sample adapted from EMBASE

The example below shows an extract from an EMBASE search designed to identify studies of effects in a Cochrane Review of Factor Xa inhibitors for acute coronary syndromes (Brito et al 2011), followed by the corresponding search strategy adapted for execution in the NHS EED search interface.

EMBASE:

1. (fondaparinux or idraparinux or Arixtra or otamixaban or Razaxaban or Fonadaparin or Dx 9065$).mp.
2. xa inhibit$.mp.
3. 10a inhibit$.mp.
4. xa antagonist$.mp.
5. 10a antagonist$.mp.
6. xa block$.mp.
7. factor x inhibit$.mp.
8. Fxa inhibit$.mp.
9. vaso flux.mp.

NHS EED:

fondaparin* OR idraparinux OR arixtra OR otamixaban OR ((xa OR 10a) AND (inhibit* OR antagonist* OR block*)) OR ("factor x" NEAR inhibit*) OR (fxa NEAR inhibitor*) OR "vaso flux" OR razaxaban OR "dx 9065"

Appendix B: NHS EED search filter for MEDLINE using OvidSP

1 Economics/

2 exp "costs and cost analysis"/

3 Economics, Dental/

4 exp economics, hospital/

5 Economics, Medical/

6 Economics, Nursing/

7 Economics, Pharmaceutical/

8 (economic$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$).ti,ab.

9 (expenditure$ not energy).ti,ab.

10 value for money.ti,ab.

11 budget$.ti,ab.

12 or/1-11

13 ((energy or oxygen) adj cost).ti,ab.

14 (metabolic adj cost).ti,ab.

15 ((energy or oxygen) adj expenditure).ti,ab.

16 or/13-15

17 12 not 16

18 letter.pt.

19 editorial.pt.

20 historical article.pt.

21 or/18-20

22 17 not 21

23 exp animals/ not humans/

24 22 not 23

25 bmj.jn.

26 "cochrane database of systematic reviews".jn.

27 health technology assessment winchester england.jn.

28 or/25-27

29 24 not 28

30 limit 29 to yr="2014 -Current"

Appendix C: NHS EED search filter for EMBASE using OvidSP

1. Health Economics/

2. exp Economic Evaluation/

3. exp Health Care Cost/

4. pharmacoeconomics/

5. 1 or 2 or 3 or 4

6. (econom$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$).ti,ab.

7. (expenditure$ not energy).ti,ab.

8. (value adj2 money).ti,ab.

9. budget$.ti,ab.

10. 6 or 7 or 8 or 9

11. 5 or 10

12. letter.pt.

13. editorial.pt.

14. note.pt.

15. 12 or 13 or 14

16. 11 not 15

17. (metabolic adj cost).ti,ab.

18. ((energy or oxygen) adj cost).ti,ab.

19. ((energy or oxygen) adj expenditure).ti,ab.

20. 17 or 18 or 19

21. 16 not 20

22. animal/

23. exp animal experiment/

24. nonhuman/

25. (rat or rats or mouse or mice or hamster or hamsters or animal or animals or dog or dogs

or cat or cats or bovine or sheep).ti,ab,sh.

26. 22 or 23 or 24 or 25

27. exp human/

28. human experiment/

29. 27 or 28

30. 26 not (26 and 29)

31. 21 not 30

32. 0959-8146.is.

33. (1469-493X or 1366-5278).is.

34. 1756-1833.en.

35. 32 or 33 or 34

36. 31 not 35

37. Conference abstract.pt.

38. 36 not 37

39. limit 38 to yr="2010 -Current"

Appendix D: NHS EED search filter for PsycINFO using OvidSP

1. "costs and cost analysis"/

2. "Cost Containment"/

3. (economic adj2 evaluation$).ti,ab.

4. (economic adj2 analy$).ti,ab.

5. (economic adj2 (study or studies)).ti,ab.

6. (cost adj2 evaluation$).ti,ab.

7. (cost adj2 analy$).ti,ab.

8. (cost adj2 (study or studies)).ti,ab.

9. (cost adj2 effective$).ti,ab.

10. (cost adj2 benefit$).ti,ab.

11. (cost adj2 utili$).ti,ab.

12. (cost adj2 minimi$).ti,ab.

13. (cost adj2 consequence$).ti,ab.

14. (cost adj2 comparison$).ti,ab.

15. (cost adj2 identificat$).ti,ab.

16. (pharmacoeconomic$ or pharmaco-economic$).ti,ab.

17. or/1-16

18. (task adj2 cost$).ti,ab,id.

19. (switch$ adj2 cost$).ti,ab,id.

20. (metabolic adj cost).ti,ab,id.

21. ((energy or oxygen) adj cost).ti,ab,id.

22. ((energy or oxygen) adj expenditure).ti,ab,id.

23. or/18-22

24. (animal or animals or rat or rats or mouse or mice or hamster or hamsters or dog or dogs or cat or cats or bovine or sheep or ovine or pig or pigs).ab,ti,id,de.

25. editorial.dt.

26. letter.dt.

27. dissertation abstract.pt.

28. or/24-27

29. (0003-4819 or 0003-9926 or 0959-8146 or 0098-7484 or 0140-6736 or 0028-4793 or 1469-493X).is.

30. 17 not (23 or 28 or 29)

31. limit 30 to yr="2010 -Current"

Appendix E: NHS EED search filter for PubMed

#1 economic evaluation*[ti]

#2 economic analy*[ti]

#3 cost analy*[ti]

#4 cost effectiveness[ti]

#5 cost benefit*[ti]

#6 cost utilit*[ti]

#7 (#1 OR #2 OR #3 OR #4 OR #5 OR #6)

Appendix F: NHS EED search filter for CINAHL using EBSCO

S1 MH "Economics+"

S2 MH "Financial Management+"

S3 MH "Financial Support+"

S4 MH "Financing, Organized+"

S5 MH "Business+"

S6 S2 OR S3 or S4 OR S5

S7 S1 NOT S6

S8 MH "Health Resource Allocation"

S9 MH "Health Resource Utilization"

S10 S8 OR S9

S11 S7 OR S10

S12 TI (cost or costs or economic* or pharmacoeconomic* or price* or pricing*) OR AB (cost or costs or economic* or pharmacoeconomic* or price* or pricing*)

S13 S11 OR S12

S14 PT editorial

S15 PT letter

S16 PT commentary

S17 S14 or S15 or S16

S18 S13 NOT S17

S19 MH "Animal Studies"

S20 (ZT "doctoral dissertation") or (ZT "masters thesis")

S21 S18 NOT (S19 OR S20)

S22 PY 2009-

S23 S21 AND S22

Appendix G: CHEC checklist

	Item	Yes	No
1.	Is the study population clearly described?	☒	☒
2.	Are competing alternatives clearly described?	☒	☒
3.	Is a well-defined research question posed in answerable form?	☒	☒
4.	Is the economic study design appropriate to the stated objective?	☒	☒
5.	Is the chosen time horizon appropriate to include relevant costs and consequences?	☒	☒
6.	Is the actual perspective chosen appropriate?	☒	☒
7.	Are all important and relevant costs for each alternative identified?	☒	☒
8.	Are all costs measured appropriately in physical units?	☒	☒
9.	Are costs valued appropriately?	☒	☒
10.	Are all important and relevant outcomes for each alternative identified?	☒	☒
11.	Are all outcomes measured appropriately?	☒	☒
12.	Are outcomes valued appropriately?	☒	☒
13.	Is an incremental analysis of costs and outcomes of alternatives performed?	☒	☒
14.	Are all future costs and outcomes discounted appropriately?	☒	☒
15.	Are all important variables, whose values are uncertain, appropriately subjected to sensitivity analysis?	☒	☒
16.	Do the conclusions follow from the data reported?	☒	☒
17.	Does the study discuss the generalizability of the results to other settings and patient/client groups?	☒	☒

18.	Does the article indicate that there is no potential conflict of interest of study researcher(s) and funder(s)?	☒	☒
19.	Are ethical and distributional issues discussed appropriately?	☒	☒

Appendix H: SBU checklist for assessing methodological quality of model-based economic evaluations

'Quality of the economic analysis'					
1 Choice of analysis	Yes	No	Unclear	Not applicable	Comments
a) Is the type of economic analysis justified in relation to the research questions?					
2 Model structure	Yes	No	Unclear	Not applicable	Comments
a) Is the model structure appropriate for the specific research question and the specific health condition?					
b) Is the model structure, including the underlying assumptions, transparent?					
c) Is the external validity of the model explored?					
d) Is the time horizon sufficient to reflect all-important differences in costs and effects?					
e) Markov models: Is the model cycle length motivated by the research question?					
3 Costs and effects	Yes	No	Unclear	Not applicable	Comments
a) Have all relevant outcomes been identified (including side effects)?					
b) Is the data on treatment effects taken from the best possible sources?					
c) Is the difference in treatment effects, which determines the model outcomes, statistically significant?					
d) Are appropriate methods used to extrapolate treatment					

effects over the chosen time horizon?					
e) Has the study considered compliance?					
f) Are the quality-of-life weights from the best possible sources?					
g) Given the perspective of the analysis, have all relevant costs been identified (including those due to side effects)?					
h) Is the data on resource use (e.g. number of social worker visits, number of hospital care days) from the best possible sources?					
i) Are the unit costs from the best possible sources?					
4 Interpretation of results	Yes	No	Unclear	Not applicable	Comments
a) Was an incremental analysis of both costs and outcomes conducted (or is it possible to calculate)?					
b) Are appropriate statistical methods used?					
c) Are the conclusions consistent with the reported results?					
5 Sensitivity analysis					
a) Are all important variables explored in sensitivity analyses?					
b) Is the uncertainty in the result explored using probabilistic sensitivity analysis?					
c) Is the result insensitive to changes in examined variables?					
6 Discounting (for studies with a time horizon exceeding 1 year)	Yes	No	Unclear	Not applicable	Comments
a) Are costs discounted appropriately?					

b) Are outcomes discounted appropriately?					
7 Potential conflicts of interests	Yes	No	Unclear	Not applicable	Comments
a) Is there a low risk that the conflict of interest declared by the authors may have influenced the study results					
b) Is there a low risk that a sponsor with an economic interest in the outcome may have influenced the study results?					
c) Is there a low risk of conflict of interest from other sources (e.g. the authors have developed the intervention					

Appendix I: SBU checklist for assessing quality of trial-based economic evaluation studies

'Quality of the economic analysis'					
1 Choice of analysis and reporting of results	Yes	No	Unclear	Not applicable	Comments
a) Is the type of economic analysis justified in relation to the research question(s)?					
b) Was an incremental analysis of both costs and outcomes performed (or is it possible to calculate)?					
c) Are appropriate statistical methods used?					
d) Are the conclusions consistent with the reported results?					
e) Is the time horizon sufficient to reflect all important differences in costs and effects?					
2 Costs and effects	Yes	No	Unclear	Not applicable	Comments
a) Is the difference in outcomes between the alternatives statistically significant?					
b) Has the study considered compliance?					
c) Is the proportion of missing data (costs and outcomes) acceptable?					
d) Have all relevant outcomes been identified (including side effects)?					
e) Are the outcomes quantified appropriately?					
f) If the outcome measure is QALYs, are the quality-of-life weights valued appropriately?					
g) Given the perspective of the analysis, have all relevant costs					

been identified (including those due to side effects)?					
h) Is the resource use quantified appropriately in physical units (e.g. number of social worker visits, number of hospital care days)?					
i) Are the unit costs valued appropriately?					
3 Sensitivity analysis	Yes	No	Unclear	Not applicable	Comments
a) Are all important variables explored in sensitivity analyses?					
b) Is the uncertainty in the result explored using probabilistic sensitivity analysis?					
c) Is the result insensitive to changes in examined variables?					
4 Discounting (for studies with a time horizon exceeding 1 year)	Yes	No	Unclear	Not applicable	Comments
a) Are costs discounted appropriately?					
b) Are outcomes discounted appropriately?					
Conflicts of interest					
a) Is there a low risk that the conflict of interest declared by the authors may have influenced the study results					
b) Is there a low risk that a sponsor with an economic interest in the outcome may have influenced the study results?					
c) Is there a low risk of conflict of interest from other sources (e.g. the authors have developed the intervention					