Public Health Monograph Series
No.22
ISSN 1178-7139

LITERATURE SEARCH AND DATA SYNTHESIS METHODS FOR ESTIMATING INPUTS FOR HEALTH ECONOMIC MODELLING

Burden of Disease Epidemiology, Equity and Cost-Effectiveness Programme (BODE^3)

Technical Report: Number 7

Rachel Foster

November 2011
First update: 7 March 2012
Second update: 15 February 2013

A technical report published by the Department of Public Health,
University of Otago, Wellington, New Zealand

ISBN 978-0-473-20313-9

BODE^3 Team*
*** For any queries contact Rachel Foster, BODE^3 Programme, University of Otago, Wellington, New Zealand. Email:** rachel.foster-russell@otago.ac.nz

Acknowledgements

These Standard Procedures were developed as part of the Burden of Disease Epidemiology, Equity and Cost-Effectiveness Programme (BODE^3). The BODE^3 Programme is funded by the Health Research Council of New Zealand. Further details can be found at: www.uow.otago.ac.nz/BODE3-info.html.

Competing Interests

The authors have no competing interests.

Table of Contents

List of Figures

1 Introduction

The Burden of Disease Epidemiology, Equity and Cost-Effectiveness Programme ($BODE^3$) aims to estimate the impact and cost-effectiveness of cancer control and preventive interventions using Markov models, multistate lifetables and microsimulation. Literature reviews will be conducted to inform estimates of the effect size of interventions and other input parameters for modelling purposes such as costs, utilities and prevalence (refer also to the full $BODE^3$ Protocol[1]).

As the first step, literature searches for published review articles will be conducted to identify whether relevant, high-quality systematic reviews exist that can inform effect sizes and other input parameters for $BODE^3$, rather than a de novo review being conducted. This approach is largely pragmatic. Systematic reviews are highly time and resource intensive, and can take a number of months to complete comprehensively. This would not be the best use of the Programme's resources when organisations such as the Cochrane Collaboration, the Campbell Collaboration and other Health Technology Assessment (HTA) bodies have the resources to produce systematic reviews to very high standards. Existing systematic reviews will be updated if necessary.

If there are no suitable systematic reviews, *de novo* reviews will be conducted. Given the time and resources required to do a full systematic review, the $BODE^3$ team will not undertake systematic reviews *per se*, but will conduct literature searches using explicit and systematic methods, including: clear definition of the research question; use of an explicit and documented search strategy; application of pre-determined inclusion and exclusion criteria; documentation of the citations included and excluded with reasons for exclusion, and; critical appraisal techniques to identify the most appropriate and relevant research. It is likely that some interventions investigated in $BODE^3$ will not have been subjected to rigorous evaluation such as randomised controlled trials (RCTs) or well controlled cohort studies. When needed, the $BODE^3$ team will consider less robust levels of evidence and may at times need to involve expert knowledge if existing data are not sufficient. In such cases, broader searches than outlined in these Standard Procedures may be needed. If elicitation of expert knowledge is required, a formal process will be used as appropriate.

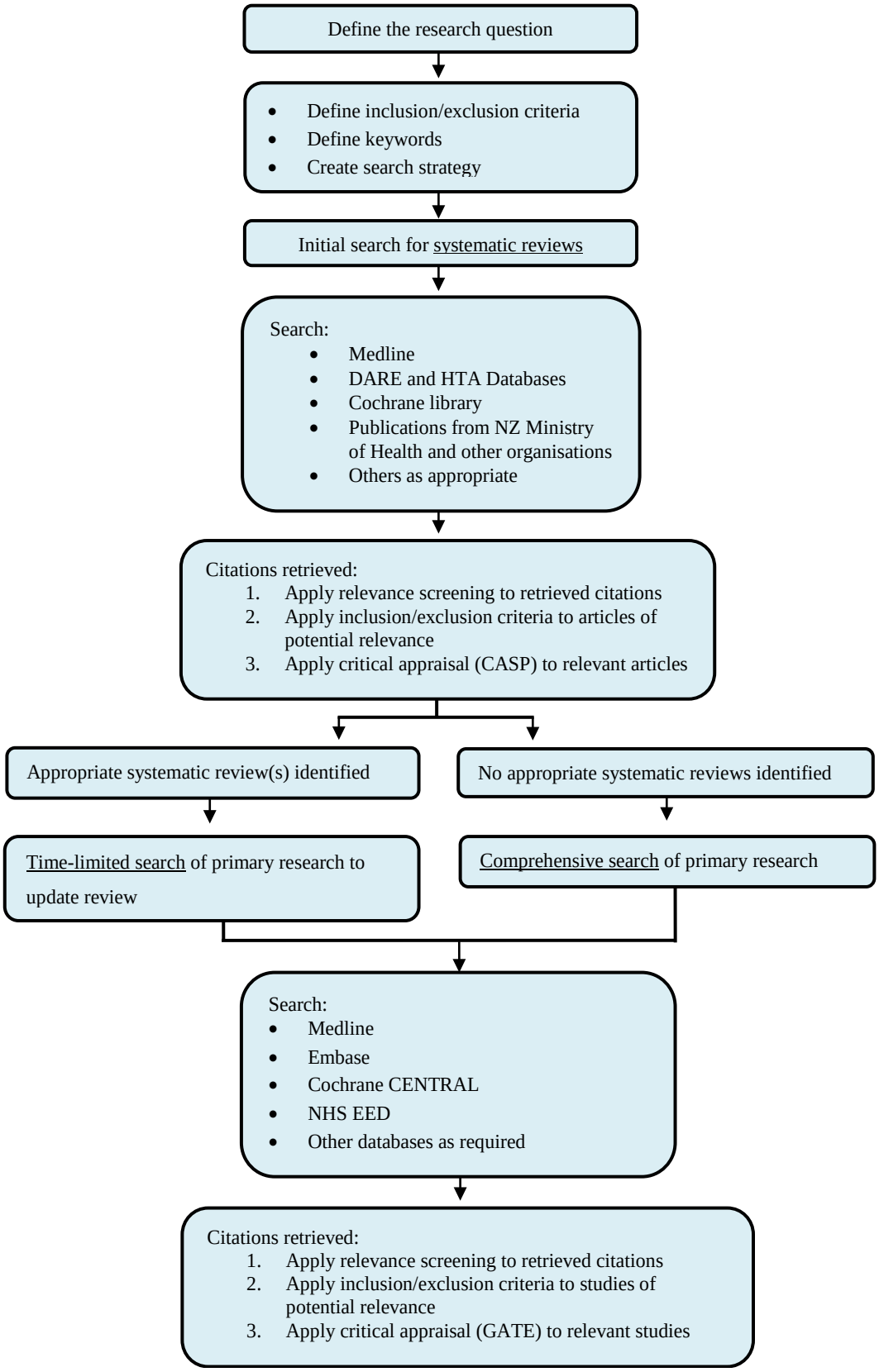
Our aim is to determine the best estimate for input parameters for $BODE^3$ modelling. Thus, the focus is on finding the *best* evidence rather than all available evidence. The best evidence amongst the retrieved literature will be identified by evaluating methodological rigour of relevant papers using a standardised critical appraisal process. Critical appraisal will be applied to both systematic reviews and original research (for *de novo* reviews).

Critical appraisal is a key aspect of this process for several reasons: (i) critical appraisal will determine whether any existing systematic reviews are of sufficient quality to be used to contribute to determination of the effect size or other important parameters; reviews that do not meet quality criteria will be disregarded; (ii) all relevant literature identified for possible inclusion in updating or writing *de novo* systematic reviews will be subjected to critical appraisal to determine their quality and reliability, and the strength of evidence that they provide; where sufficient data exist, papers of lower quality or providing less strong evidence may be disregarded ; (iii) the process of critical appraisal will help identify areas where there is uncertainty around parameters (eg, large variation in reported values, differing of opinions between experts, differential response between patient subgroups etc), and additionally may identify issues that should be raised in the second-filter stage.

This document outlines the standard procedures for undertaking literature searches (and extracting data) for the BODE^3 projects. The methods may prove useful for other groups undertaking health economic modelling.

The search process is outlined in Figure 1.

Figure 1. Flow diagram of literature search strategies to be used in BODE3



1.1 Search set-up

Literature searching is an iterative process, and will require logic checks along the way. It is often useful to read a few recent review articles in the area to help refine the research question and the search strategy before starting the full literature review search. The TRIP database, which is aimed at answering clinical questions, is excellent for identifying the most recent research, reviews and guidelines for background (www.tripdatabase.com). New Zealand clinical guidelines are available from the New Zealand Guidelines Group (www.nzgg.org.nz/); US and UK guidelines are available from the National Guidelines Clearinghouse (www.guideline.gov/) and NICE (http://guidance.nice.org.uk/), respectively.

1.2 Definition of the research question

This is the essential first step of any literature review. If the research question is not well defined, the literature review may end up off track, collecting data on irrelevant outcomes. The research question should be fully defined using the PICO(T) framework:

- **P**opulation of interest;
- **I**ntervention/Exposure (including setting);
- **C**omparison(s);
- **O**utcomes;
- (**T**imeframe)

The research question will be the same whether searching for existing systematic reviews, updating existing systematic reviews, or creating a *de novo* review. Note that the research question might be of any of the following types:

- What is the best estimate, and uncertainty interval (either/both random and systematic error), for a given input parameter, such as effectiveness of the intervention
- What are the sequelae of a given intervention that are important enough to include in economic decision modelling
- What are the elements to include in costing a given intervention
- What is the change in quality of life resulting from a given intervention.

1.3 Definition of inclusion and exclusion criteria

The systematic approach defined for BODE^3 requires that the same inclusion and exclusion criteria are applied in a consistent manner to all literature identified. The inclusion/exclusion criteria must be explicit and reproducible such that two or more people applying the criteria would select the same papers.

If the search does not provide adequate data, the inclusion and exclusion criteria may need to be loosened, in which case all searches must be re-run applying the new inclusion and exclusion criteria.

Inclusion/exclusion criteria may relate to:

- Study design/methodology
    - What study designs are appropriate for answering the research question?
    - Any indicators of study quality, e.g. include only those with a high level of completion or follow-up?
- The disease/outcome
    - Morbidity or mortality outcomes?
    - Focussed on events or on risk factors?
    - Different levels of disease severity?
- Patient/participant population
    - Limit by age, sex, ethnicity?
    - Exclusion of certain patient groups or conditions?
- Intervention
    - Define scope/specifications
- Treatment setting
    - Inpatient versus outpatient vs community?
- Treatment duration
    - Short-term vs long-term?
- Any other relevant issues

Note 1
Once the research question, inclusion and exclusion criteria and keywords have been defined, the Medical School Library can be asked to help with the search. Do not hesitate to ask them for help if the search is complex or you don't feel confident to do the search.
However, this does not eliminate the more time-consuming aspect of determining which of the citations retrieved are relevant, and allows less flexibility for logic checks and widening and narrowing search terms in response to the results returned.

1.4 Creation of explicit search strategies

Different search strategies often need to be devised for different sources because search platforms differ between databases. The following are general principles that may need to be further tailored.

- Define keywords
    - Use truncation to search for different word endings by using a wildcard character if needed, e.g. “fat*” will pick up fat, fatty, fats etc
    - The wildcard character is often * or $
- Create a search using the keywords and, where possible, mapped Subject Headings or Thesaurus Terms (see Appendix One: Search techniques for search techniques)
- Combine search terms using Boolean operators; OR, AND, NOT
- Further refining the search to meet inclusion and exclusion criteria will provide a more specific search, but this must be balanced against whether possibly relevant papers will be excluded if additional terms are added to the search
    - Searching engines are fairly blunt instruments, and if the search term isn’t in the title, abstract, or keywords the paper may not be captured by the search
- Limit by language, time, study type etc as appropriate.
    - Searches for BODE^3 will be limited to the English language
    - Time limits will depend on the state of knowledge in the area, and when the most recent credible systematic review was published
- Consider any limitations on the type of literature that will be included (media releases, company unpublished data, etc).

The exact terms and operators used must be recorded as part of the review document in sufficient detail to allow replication by others. Complete both the Literature Search Summary Form and Search Results Table; see Appendix Two: Literature search summary forms

The medical library can help with conducting searches, but before approaching them you must have defined the research question, considered appropriate keywords and limits, and created the inclusion and exclusion criteria.

Note 2

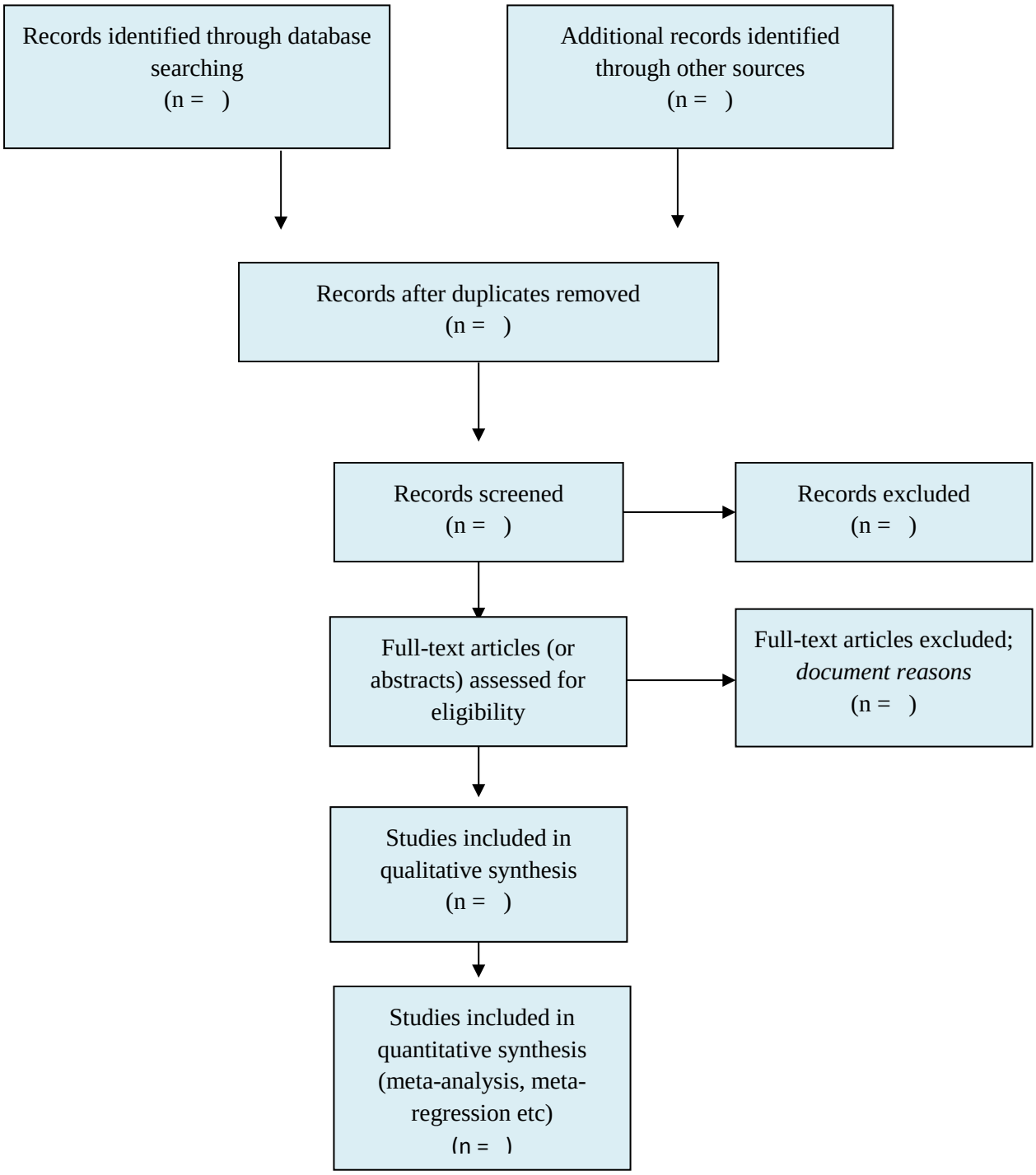
It is recommended that a logic test is carried out at this stage, trialling the search strategy through a quick search of Medline and DARE

- Are the search terms sufficient to address all the important aspects of the research question?
- Does the search produce the expected key articles?
    - If not, reassess the keywords, search strings and limits used
- Should the search be widened or be more focussed?
- If appropriate, trial logical broader or narrower terms to see if substantially different results are produced (in terms of relevant papers identified, not number of hits)

2 Documenting the search

As the searches are carried out, documentation should be adequate to complete the flow diagram below from the PRISMA guidelines for reporting systematic reviews.[2] Template forms to aid documentation are provided in Appendix Two: Literature search summary forms .

Figure 2: Modified PRISMA Flow Diagram (adapted from the PRISMA statement[2])



3 Searching for existing systematic reviews

As noted previously, BODE[3] plans to use existing systematic reviews whenever an appropriate good quality review is available. The first literature search is to identify what systematic reviews are available. The appropriate sources are described in Appendix Three: Overview of sources.

For the purposes of BODE[3], to be considered a systematic review, the review must state a method that indicates that all studies on the subject (within defined limits) were searched for, and selection of studies was based on a systematic method of inclusion/exclusion.

As previously noted, the aim is to find the best evidence, not all available evidence. Only the sources that index systematic reviews of *high quality* need to be searched. In most cases it will be sufficient to search:

- Medline (and/or PubMed)
    - Note that PubMed has more recent citations available (there is an approx 3 month delay before citations are available on Medline) but it is less well indexed for structured searching
    - Recent (not yet indexed) papers can also be found by searching "Medline pending"
- Centres for Reviews and Dissemination
    - DARE (Database of Abstracts of Reviews of Effects)
    - HTA Database

  The DARE and HTA databases can be searched simultaneously. Note that the HTA database includes NICE and NIHR appraisals, as well as other HTAs
- Cochrane library
- Publications from NZ Ministry of Health[1] and other relevant organisations
- Campbell library if the intervention is within the areas of education, crime and justice, or social welfare.

The search may be broadened if these sources do not identify any suitable systematic reviews. However, the quality of papers identified may be an issue.

[1] MoH publications don't have a search option. Select "Find a publication by title". Manually search (ctl-F9) by keywords and synonyms.

Note 3
Database searches are likely to be imperfect as they are dependent on the quality of indexing, choice of thesaurus terms, and synonyms when free text searching etc. Thus, additional searching may be needed for both systematic reviews and original research.

Snowballing: Reference lists of the most relevant papers should be checked to detect any other important systematic reviews or primary research literature that the database searches have not detected. Snowballing is recommended for BODE^3 searches for both systematic reviews and original research.

Citation links: Additional papers can also be found by looking at papers that have cited a key paper that you have identified. This can be done by calling up the key paper in Web of Knowledge (via the Medical School Library), Google Scholar or Scopus. The citation will show links to papers that have cited the key paper, and links to all papers that were cited by the key paper. This may be particularly useful when searching for more recent research.

Handsearching: The contents pages of relevant journals can be checked to identify additional papers. This is useful when there the topic is likely to be addressed in only a small number of specialised journals. Handsearching will detect papers as soon as they are published, whereas there can be a delay of several months before an article appears in an indexing database. This type of searching is unlikely to be necessary for BODE^3 in most cases as the most important papers should be detected using database searches. and snowballing. However, it may be useful for highly specialised topics.

Review Limits

The “review articles” limit option on Medline may miss important papers, especially meta-analyses, while returning a large number of non-systematic reviews and opinion pieces. Using this option is not advisable for BODE^3 searches. More appropriate options are available under “Additional Limits” (Clinical Queries), as further discussed below.

If there are a large number of reviews on the topic, the most efficient option on Medline is to "**reviews (maximizes specificity)”**, which limits results to meta-analyses and systematic reviews. To test the specificity and sensitivity of the limit options, results from "reviews (maximizes specificity)” were compared to those from the broader limit "reviews (best balance of sensitivity and specificity)" for a literature search on dietary interventions to reduce intake of saturated fats. The broader “balanced” limit option retrieved 302 citations, of which 47 were potentially relevant, and 5 met final criteria for inclusion. The “maximum specificity” option, retrieved 20 citations, of which 10 were potentially relevant and the same 5 as for the broader limit met final criteria for inclusion. One relevant paper was missed by both searches and was identified only by snowballing. Thus, using the "reviews (maximizes specificity)” option appears to greatly reduce time and effort, while still identifying the most important evidence. If the results returned don’t appear to be sufficient, using the broader option "reviews (best balance of sensitivity and specificity)" would be appropriate.

Duplication

Searching on Medline can retrieve a substantial number of duplicate citations. To remove duplicates, type into the keyword search box .dedup and the number of the search, e.g. if you want to remove duplicates from the third search that you have conducted (as listed in the Search History box), you would type .dedup 3 and then hit enter.

Good quality meta-analyses/systematic reviews are also likely to be indexed in more than one database. Where there is duplication of the same citation, the paper should be recorded only once as a retrieved citation. However, if two or more separate papers report the same study/results (e.g. the review has been updated, or an author has produced two or more very similar papers), only the most recent, relevant and/or comprehensive paper should be retained, and the other papers discarded at the stage of applying exclusion criteria.

Note 4

Once the search has been conducted, the following steps are applied to identify which systematic reviews are potentially suitable for determining the effect size or other input parameters for BODE^3 models:

- Relevance screening
- Application of inclusion and exclusion criteria
- Critical appraisal

The number of reviews identified by the search, and the number excluded at each stage and the reasons for their exclusion, must be stated in the Literature Search Summary Form (Appendix Two: Literature search summary forms)

3.1 Relevance screening

Relevance screening is applied to identify all reviews retrieved by the search that are of potential relevance to the research question. Searching is often a blunt instrument and a number of records will be retrieved that are clearly not relevant to the research question. At this stage, the only criterion is possible relevance – no judgment is made regarding value or quality until later stages.

- Screening is based on the record title – referring to abstract if required
- If necessary, export abstracts of selected articles for further assessment
    - In Medline, “Export to Microsoft Word and Save”
        - Select Citation + abstract
        - Select Citation style Ovid Citation
        - Include Link to External Resolver
        - Include URL
        - Include Search History

Record the number of records discarded and retained after the relevance screening on both the Literature Search Summary Form and Search Results Table (Appendix Two: Literature search summary forms).

3.2 Application of inclusion and exclusion criteria

The pre-determined inclusion and exclusion criteria (section 2.2) are then applied to all potentially relevant articles selected by relevance screening

- Based on the abstract and/or full text of the article
- Immediately exclude all articles that are not systematic reviews or meta-analyses
    - Any review that does not provide a methods section describing the search strategy and/or other systematic methods can be excluded
- Exclude duplicate publishing of results
    - If a review has been updated, exclude the earlier version

It is likely that non-systematic reviews or opinion/editorial papers will be found that provide useful background or contextual information. These papers should be counted as "excluded" from the literature search for systematic reviews, but full text copies can be retrieved if they will be useful to the project in other ways, such as writing up papers of our results for publication. They may be particularly useful if there is little literature on the subject and we need to create our own review.

Relevant cost-effectiveness studies and reviews may also be useful.

The number of articles excluded, and the reasons, must be recorded on the Literature Review Summary Form (Appendix Two: Literature search summary forms).

3.3 Critical appraisal: systematic reviews

Reviews identified as suitable in the above steps should be retrieved as full text and subjected to critical appraisal, within the context of what is important to the BODE^3 project. The aim is to ensure that the findings of the review are not invalidated by any significant biases or other significant

weaknesses, *and are of value to the BODE^3 objective of determining effect size or other parameters for modelling.* Systematic reviews that do not meet the key quality criteria should be discarded. Note that systematic reviews produced by authoritative bodies such as the Cochrane Collaboration or the UK NHS National Institute of Health Research (HTA Journal) have already been subjected to multiple appraisals of quality and do not need to undergo critical appraisal.

BODE^3 has chosen to use the CASP Critical Appraisal Tool for Systematic Reviews[3]

- Available from http://www.casp-uk.net/
- Questions 1-7 are the focus
    - Did the review address a clearly focused question?
    - Did the authors look for the appropriate sort of papers?
    - Do you think the important, relevant studies were included?
    - Did the review's authors do enough to assess the quality of the included studies?
    - If the results of the review have been combined, was it reasonable to do so?
    - What are the overall results of the review?
    - How precise are the results?
- Question 8 regarding applicability ("Can the results be applied to the local population?") can be broadly interpreted, e.g. data from a range of countries can be considered applicable for BODE^3 if the context is appropriate
- Question 9 ("Were all important outcomes considered?") should be related to whether the outcomes that are important to BODE^3 have been presented
- Question 10 ("Are the benefits worth the harms and costs?") is not relevant – the BODE^3 analysis itself will address these issues.

Further details to help answer the questions are provided on the CASP appraisal form.
CASP also provides critical appraisal tools for economic evaluations, RCTs, case-control studies, cohort studies, qualitative studies and diagnostic studies (see www.casp-uk.net).

The CASP critical appraisal tools were developed by the Critical Appraisals Skills Programme at the Public Health Resource Unit, Oxford. They are based on the guidelines produced by the Evidence-Based Medicine (EBM) Working Group. The EBM guidelines are widely accepted but are too detailed in themselves for our needs. Thus, the CASP tool combined with the expertise within the research group represents an acceptable compromise. This tool has been recommended by other groups (e.g. the Canadian National Collaborating Centre for Methods and Tools[4]).

For the purposes of BODE^3, an additional level of selection can be included if appropriate. A systematic review may be discarded on the grounds that a more recent, more comprehensive and/or better performed systematic review is available.

The decision whether a paper is of sufficient quality and relevance to be used in BODE^3 is based on a qualitative appraisal of the issues identified by applying the CASP tool, rather than the assignment of a summary score. An explicit decision was made to not use a tool that produced summary scores because the use of such scoring has been criticised,[5] and is explicitly discouraged in Cochrane reviews.[6] The weighting applied to different scale items to calculate summary scores does not have a good theoretical underpinning and is inconsistent between tools. Better transparency is achieved with qualitative checklists.

While it would be ideal to test for inter-rater reliability, this would not be feasible given the large number of interventions that will be investigated in the BODE^3 programme. The risk of subjectivity associated with only one person carrying out the search, selection and appraisal procedures are likely to be low because of the high level of expertise in epidemiology and literature review within the team. Furthermore, the aim is to identify the *best* evidence, which should be relatively easily identifiable in most cases. If there is any ambiguity over whether a paper should be included in the BODE^3 analysis, a second opinion will be sought from another team member. The critical appraisal forms will be retained in case they need to be referred to in the future. The number of, and reasons for, discarded papers will be stated on the Literature Review Summary Form (Appendix Two: Literature search summary forms).

4 Searching for primary research literature

Searches to identify original research literature will be needed if:

1. an appropriate systematic review has been identified but needs updating, or
2. no appropriate systematic review has been identified and a *de novo* review is required.

The approach is the same for these two scenarios with one important difference. Updating requires only a time-limited search, while a *de novo* review requires a comprehensive search within the boundaries outlined in this document. Snowballing should also be applied (see section 4, Note 3).

Relevance screening and inclusion/exclusion criteria are the same whether searching for a systematic review, primary research to update a review, or primary research to conduct a *de novo* review. The only exception is criteria relating to study type (i.e. whether the search aims to identify systematic reviews or original research).

Note that if a full systematic review is going to be written for publication, the PRISMA statement should be adhered to.[2]

4.1 Updating (time-limited search)

Once appropriate high-quality systematic reviews have been identified, it may be appropriate to update the evidence. The decision regarding whether this is required will depend on:

- whether the systematic review is likely to be out of date
- whether there are any gaps in the evidence, and
- whether there have been any significant trials published post-dating the systematic review.

It will not always be necessary, or appropriate, to update systematic reviews.

If updating, the same keyword search strategies should be used as previously defined (section 2.3), but with time limits to include 2 years before publication of the systematic review, up until the present. Pre-dating the systematic review by 2 years allows for any lag between when a review is written and when it is published. When recording your search strategy, include what time limits were applied.

Relevance screening and application of the pre-determined inclusion/exclusion criteria (as per section 2.2) should be rigorous. Good quality studies that add to the findings so as to reduce the uncertainty

around any given estimate are of value, as are those that negate the existing findings. Studies that are of low methodological quality, or those that investigate additional mechanisms or subgroups that are not relevant to the particular BODE^3 objective, can be discarded. The hierarchy of evidence can be considered and only higher levels of evidence included (see section 5.4). The number and reasons for papers being excluded should be recorded on the Literature Search Summary Form (Appendix Two: Literature search summary forms).

4.2 *De novo* review (comprehensive search)

Searching for a *de novo* review differs only in that there is no time limit applied to the search unless required for pragmatic reasons (e.g. may only want more recent data). The search is also likely to be more inclusive, as all original research studies that meet relevance screening and inclusion/exclusion criteria will be initially included. However if there is a large body of evidence, papers may be selected according to a hierarchy of evidence (see section 5.4).

4.3 Sources

Keeping in mind that the aim is to find the *best* evidence and not *all* the evidence, it will generally be sufficient to search the following for original research articles, whether for updating or *de novo* review:

- Cochrane Central Register of Controlled Trials (CENTRAL)
    - Along with other data, CENTRAL includes all controlled trials (RCTs and quasi RCTS) indexed on Medline and Embase, but the latter databases still need to be searched for other types of trials
- Medline (and/or PubMed)
    - Note that doing an additional search on "Medline pending" will find papers recently published but not yet indexed
- EMBASE

Searching the above 3 sources is in line with that recommended by Cochrane.[6] However, controlled trials will not necessarily have been performed for all interventions to be modelled in BODE^3. Alternative sources are Scopus and Google Scholar. These search beyond journals (e.g. professional bodies, universities, reputable websites etc) and have a broader focus (e.g. social sciences as well as health). These sites will be useful when Medline/Embase searches return insufficient evidence, but can also be too broad and return many irrelevant results. There is no filtering of quality of the information.

Snowballing and citation linking may also be an important way to ensure that all key research has been identified (Section 4, Note 3).

Inclusion of unpublished data is important for an unbiased and fully comprehensive systematic review. For BODE^3, there may not be adequate time to comprehensively search for unpublished data. However, if unpublished data can be readily attained, it should be included. If possible, abstracts from the major conferences in the area should be included. Note that Cochrane CENTRAL includes some unpublished information on controlled trials.

4.4 Hierarchy of evidence

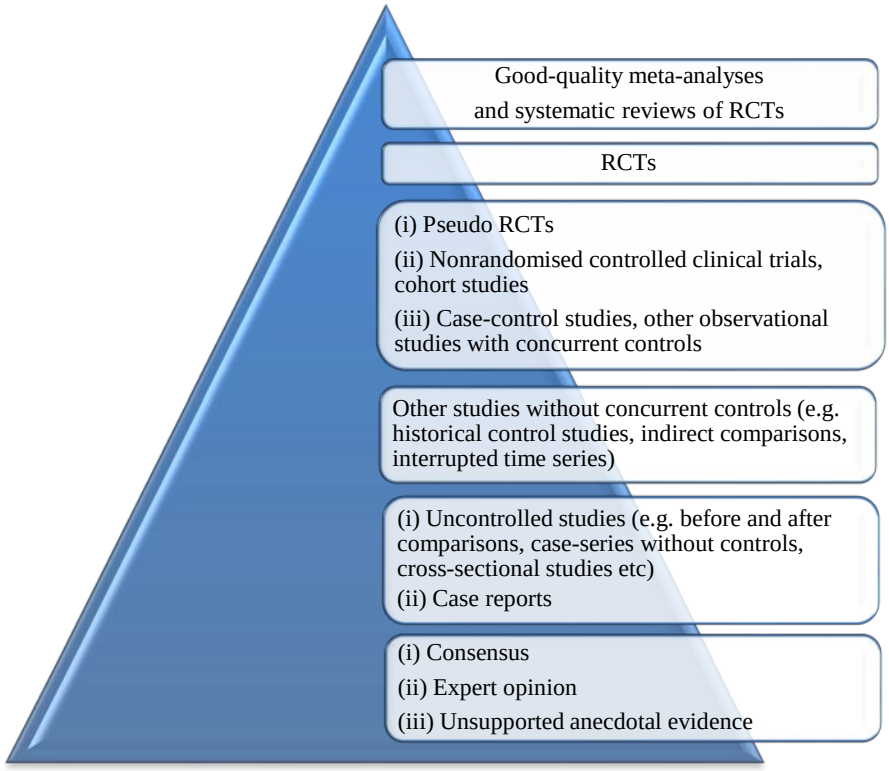
Where there is sufficient evidence, a hierarchy can be applied to include only papers representing the highest levels of evidence (see also Table 2: Approach to classification of strength of evidence in BODE^3 in the full Protocol[1]). For instance, if several RCTs are available, then lower quality papers such as historical control studies or uncontrolled studies could be excluded.

A conceptual model of the hierarchy of evidence is shown in Figure 2. Note that this is a guide only, and will depend entirely on the quality of the study and the certainty of the estimate. For instance, a poorly conducted RCT may rank lower than a well conducted cohort study. Furthermore, where specific study types are placed within the pyramid varies depends on the type of research question being asked. In BODE^3, the highest level of evidence available will vary according to the type of intervention, and it is important to consider what study type(s) are most suitable to the research question. It is acknowledged that BODE^3 may investigate some interventions that do not have a strong evidence base, and that some parameters will have to be sought through consensus and/or expert knowledge using a systematic approach, and then subjected to uncertainty analyses.

Furthermore, while the hierarchy of evidence is likely to be applicable where the model input parameter of interest is effect size, there are likely to be other model parameters within BODE^3 for which this approach to grading evidence will not be relevant. For instance, representativeness may be more important for prevalence data, and validity of the measurement tool may be more important for evidence around quality of life and utilities.

Figure 3: Hierarchy of evidence pyramid

Note this is a guide only and individual studies may move up or down the pyramid depending on their quality and their appropriateness for a specific research question



4.5 Critical appraisal: original research

Original research papers that are identified for possible inclusion in updating or conducting a *de novo* review are subjected to critical appraisal to determine whether the research findings are likely to be valid.

In line with the recommendations of the New Zealand pharmaceutical purchasing agency PHARMAC,[7] for critical appraisal of effectiveness studies BODE^3 has chosen to use the GATE (Graphic Appraisal Tool for Epidemiology) method developed by the Effective Practice, Informatics & Quality Improvement (EPIQ) group at the University of Auckland.[8] As with CASP, the GATE tool is based on the JAMA evidence-based medicine (EBM) guidelines. For assessing original research, GATE has two advantages over CASP; GATE is designed to allow critical appraisal of any

type of epidemiological study, and the assessment points relating to possible bias are more detailed. GATE has been praised for its simplicity and clarity by a working group from the Clinical Effectiveness Forum of the Royal College of Physicians[2].[9] An amended version of GATE is used by the New Zealand Guidelines Group (NZGG). In validation tests of the GATE tool for systematic reviews by the NZGG, some issues with individual items were identified, but the overall quality assessment for each paper had good inter-rater reliability.[10]

Of most importance for critically appraising studies of effectiveness for BODE^3 is the validity assessment section of GATE. Excel templates to carry out the appraisal are provided at the following website: http://www.fmhs.auckland.ac.nz/soph/depts/epi/epiq/ebp.aspx

For BODE^3, the relevant sections are those relating directly to appraisal of the study (Step 3):

- Step 3(a): Hanging the study on the GATE
    - Page 2 of Excel worksheet
    - Complete only briefly, and only with the information relevant to BODE^3
    - Note that entering numbers into the GATE calculator is inaccurate for case-control studies because incidence in the population is unknown, and the calculation approximates an odds ratio as being the same as a relative risk.
- Step 3(b): Assess study quality
    - RAMMbo
        - **R**ecruitment
        - **A**llocation
        - **M**aintenance within allocated categories
        - **M**easurement of outcomes: **b**lind or **o**bjective
    - Page 3 of Excel worksheet
    - Complete in detail
    - Identify any issues of particular relevance to modelling for BODE^3

Step 1 (Ask a clinical question) is already covered by the "PICOT" framework defined here as part of the literature search (section 2.1). Step 2 (Accessing the best evidence) is not required as we are interested in appraising the validity of only the single study of interest. Step 3 (see above). Step 4 of

[2] The working group from the Clinical Effectiveness Forum of the Royal College of Physicians, in consultation with NICE, carried out a review of grading systems for use in developing evidence-based guidelines.[9] Because of its simplicity and clarity, GATE was chosen for appraisal along with three other major grading systems: the Scottish Intercollegiate Guidelines Network (SIGN), the Grading of Recommendations Assessment, Development and Evaluation (GRADE) and the National Service Framework for long-term neurological conditions critical appraisal tool (NSF-LTC). GATE was concluded to be an excellent tool for teaching critical appraisal, but not suitable for use in developing guidelines because it doesn't grade papers. As noted earlier in this protocol, BODE^3 has explicitly chosen to avoid tools that assign a summary score to papers.

applying the evidence (The "X factor") is not relevant because the data are being appraised as an input into the $BODE^3$ model, not as an outcome in its own right. Step 5 (Auditing) is not required.

More information on how to complete the GATE appraisal can be obtained from Jackson et al. (2006)[8] and Martin et al. (2006)[11]. The latter in particular works through the steps very clearly.

Note that studies of model parameters other than effect size may be less suitable for evaluation by GATE, e.g. studies of costs, resource use, utilities, prevalence etc. The best way to evaluate these studies will need to be considered on a case-by-case basis, and with consideration of how the evidence fits with the principles of $BODE^3$ as outlined in the protocols.[1,12] CASP provides a critical appraisal tool for economic evaluations (www.casp-uk.net/wp-content/uploads/2011/11/CASP_Economic_Evaluation_Checklist_14oct10.pdf) that may be able to be adapted for other related studies.

5 Economic papers and other evidence for model parameters

It is important to identify relevant economic papers, as well as searching for systematic reviews and original research relating to the effect size and other model parameters. These papers are important to understanding the issues surrounding costs and resource utilisation for the intervention, for example what are the cost drivers and where the most important cost differences between comparators arise. However, we will largely obtain actual costs for the BODE^3 models through our own activity costing processes and/or use of HealthTracker, and disability weights from the Global Burden of Disease.[1] There may be cases where we use resource utilisation data directly from other cost-effectiveness analyses.

To obtain relevant economic papers, the NHS EED (NHS Economic Evaluation Database) should be searched. It is available from the Centre for Reviews and Dissemination (see Appendix Three: Overview of sources), and comprehensively indexes economic evaluations of healthcare, with a focus on those relevant to the UK NHS. Critical abstracts commenting on the overall reliability and generalisability of the analysis are provided for many of the citations. This database is comprehensive for health economic evaluations, but does not comprehensively cover methodological papers or reviews of cost-effectiveness analyses.

Data and studies on quality of life and utilities are also important, and may require additional specific searches in Medline, Embase, NHS EED etc.

There are many other model parameters that evidence may be required for. Each will require a suitable search strategy and consideration of the most appropriate sources, e.g. surveys for prevalence and trends over time, extraction of cost and/or resource utilisation data from surveys or databases.

6 Data extraction and computation

Data should be extracted from retrieved articles in a systematic manner. This may be best done with tabulation. What data are to be extracted will depend on the research question (see section 2.1).

All information of importance to estimating the effect size for BODE^3 modelling (including uncertainty), and how this varies by age, sex and ethnicity, should be extracted, along with any *key* information relevant to understanding the quality of the evidence. Along with tabulation, the overall findings (and relevant issues) should be briefly summarised in a 1-2 page "primer" to help inform modelling structure, and to aid in writing up papers for publication on the BODE^3 analysis of the intervention.

The exact information to be abstracted may vary for different interventions and/or the research question, but the following lists provide a loose framework for extracting data on effect size and/or sequelae.

Data from systematic reviews

- Aim of the systematic review
- Year data collection completed if available; if not, year of publication
- Population studied
- Very brief description of intervention(s) and comparator(s)
- Types of studies included
- *Important* inclusion and exclusion criteria
    - Note only those that will affect: applicability to the BODE^3 modelled patient population, or; combinability with other papers on the subject
- Meta-analysis method (if applicable)
- Number of studies included
- Total number of subjects included
- Follow-up (if important to modelling the intervention)
- Countries studies conducted in (if important in terms of generalisability to the NZ population)
- Results
    - By arm if applicable
    - Include only outcomes relevant to modelling the effect size
        - e.g. RRs, ORs, HRs, risk differences, etc for health states and mortality rather than "soft" or intermediate outcomes
    - Include 95% CIs and p values

    - Include results by age, sex and ethnicity if possible
- Safety data
    - See below for original research papers
    - If safety data are not provided, may need to identify safety issues from other sources
- Very brief comments on validity of the paper
    - Heterogeneity
    - Bias
    - Confounding
    - Contextual comments

Data from original research papers

- Aim of study
- Date study conducted if available; if not, year of publication
- Country of study
- Brief description of patient population
    - Age range
    - Presence of risk factors and/or disease (including severity)
    - Other *important* prognostic factors, including previous treatments (e.g. first-line vs second-line chemotherapy etc)
- Brief description of intervention(s) and comparator(s), as appropriate:
    - Dosage and formulation, or method, intensity and timing of intervention delivery
    - Duration of intervention and/or follow-up
    - Setting
- Study design
- Number of subjects
- *Important* inclusion and exclusion criteria
    - Note only those that will affect: applicability to the $BODE^3$ modelled patient population, or; combinability with other papers on the subject
- Efficacy/effectiveness results
    - By arm if applicable
    - Include only outcomes relevant to modelling the effect size
        - e.g. RRs, ORs, HRs for health states and mortality rather than "soft" or intermediate outcomes
    - Include 95% CIs and p values
    - Include results by age, sex and ethnicity if possible
- Adverse event (AE)/safety data

    - Number withdrawn from treatment due to AEs
    - Incidence of significant AEs that would have implications for modelling because they cause sequelae, affect the health state of the patient, or require treatment that incurs costs
    - For chemotherapy, generally only include grade 3 or 4 events
- Cost or resource utilisation data if relevant
- Brief comments on validity of the paper
    - Bias (information bias, measurement bias, selection bias)
    - Confounding
    - Contextual comments

  If there is evidence of problematic confounding or bias, consider documenting information necessary for qualitative bias analyses.

6.1 Quantitative estimation of effect size

For input into modelling, BODE^3 will often require a quantitative estimate (with confidence intervals) of the effect size, such as RR, HR or OR. Thus, a quantitative estimate should be derived from an existing meta-analysis (updated if necessary) or calculated *de novo* from a review if possible (see section 5.2).

However, a meta-analysis will not be appropriate if there are not sufficient data from 2 or more independent studies, if there is significant heterogeneity between the studies, or if there is likely substantial systematic error in one or more of the studies (in which case such studies should be deleted or subjected to quantitative bias analysis before inclusion in the meta-analysis). Outcomes must have been measured in a similar and combinable manner, and there must not be substantial heterogeneity arising from clinical and/or methodological differences between studies.[6] Looking at the individual study results graphically (e.g. forest plot) to identify where heterogeneity is arising can be useful; if confidence intervals have poor overlap, there is likely heterogeneity. It may be possible to exclude problematic studies to achieve a better result. However, if there is *true* variation arising from differences in response between patient subgroups, analysis within the subgroups may prove useful; BODE^3 aims, in particular, to capture heterogeneity by sex, age, ethnicity and deprivation.

If a meta-analysis cannot be performed, effect size data may have to be taken from individual studies, possibly in combination with expert opinion and critical appraisal of the validity of the estimate and quantitative bias analysis. Uncertainty around effectiveness estimates will be needed as part of BODE^3 modelling; uncertainty around a single study estimate is likely to be greater than that from a

meta-analysis of a number of studies, assuming that random error is the main (or sole) source of error to consider, or that study-specific systematic biases 'cancel out' across studies.

Of note, link models may also be required to convert the outcome from clinical trials into the outcome best operationalized for the BODE^3 models; this is covered in the full BODE^3 Protocol.[1]

6.2 Meta-analysis

Meta-analysis calculates a summary (pooled) estimate of the effect size of the intervention based on the weighted average of results of individual studies. It is a two-step process: (i) calculation of the summary statistic (e.g. OR or RR) of the individual study, and: (ii) application of weighting (e.g. inverse of variance) that takes into account the size of the study and the rate of events. More precise estimates are given more weight. (Note that within the BODE^3 programme we may add a third step – quantitative bias analysis – to individual studies before meta-analysis as described above.

Both the point estimate of the (pooled) treatment effect and the confidence interval are calculated, along with a measure of the heterogeneity such as the Q test (chi-squared; X^2) or the I^2 test.[6] A statistically significant chi-squared test indicates that there is heterogeneity; however, a non-significant result does not necessarily means that there is not heterogeneity. The I^2 test may be easier to interpret, although thresholds are somewhat arbitrary. As a rough guide, the level of heterogeneity indicated by I^2 values are as follows: <25% small/insubstantial; 25-50% = moderate; >50% = sufficiently large to be important, and; >75% = substantial.

Meta-analysis can be based on a fixed-effect or random effects model.[6] In a fixed-effect model, it is assumed that the true effect of the intervention (in both magnitude and direction) is the same value in every study (that is, fixed across studies). The result can be viewed as a 'typical intervention effect' from the studies included in the analysis. This model assumes that the observed differences among study results are due solely to chance, and that there is no statistical heterogeneity. In contrast, the random-effects model assumes that the effects being estimated in the different studies follow some kind of distribution (e.g. normal, beta, gamma). The centre of this distribution describes the average of the effects, while its width describes the degree of heterogeneity. It is likely that meta-analyses performed for BODE^3 will use a random-effects model in most cases.

References

1. Blakely T, Foster R, Wilson N, Bode3 Team. *Burden of disease epidemiology, equity and cost-effectiveness (bode3) study protocol. Version 1.0*. Burden of Disease Epidemiology, Equity and Cost-Effectiveness Programme Technical Report: Number 3. Wellington: Department of Public Health, University of Otago, Wellington, 2011. http://www.otago.ac.nz/wellington/otago019355.pdf.
2. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: The prisma statement. *BMJ* 2009;339:b2535.
3. Critical appraisal skills programme (casp): 10 questions to help you make sense of reviews. England: Public Health Resource Unit, 2006.
4. Ciliska D, Thomas H, Buffett C. *A compendium of critical appraisal tools for public health practice*. Hamilton, Ontario: National Collaborating Centre for Methods and Tools, 2008.
5. Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: A systematic review and annotated bibliography. *Int J Epidemiol* 2007;36(3):666-76.
6. *Cochrane handbook for systematic reviews of interventions*. The Cochrane Collaboration, 2009 www.cochrane-handbook.org.
7. *Prescription for pharmacoeconomic analysis: Methods for cost-utility analysis*. Wellington: PHARMAC, 2007. http://www.pharmac.govt.nz/2007/06/19/PFPAFinal.pdf.
8. Jackson R, Ameratunga S, Broad J, Connor J, Lethaby A, Robb G, et al. The gate frame: Critical appraisal with pictures. *Evid Based Med* 2006;11(2):35-8.
9. Baker A, Young K, Potter J, Madan I. A review of grading systems for evidence-based guidelines produced by medical specialties. *Clin Med* 2010;10(4):358-63.
10. Fitzgerald A, Coop C. Validation and modification of the graphical appraisal tool for epidemiology (gate) for appraising systematic reviews in evidence-based guideline development. *Health Outcome Research in Medicine* 2011;2:e51-e59.
11. Martin A, Srihari V. Geometrically evident: Framing studies using the graphic appraisal tool for epidemiology (gate). *J Am Acad Child Adolesc Psychiatry* 2006;45(12):1521-6.
12. Foster R, Blakely T, Wilson N, O'Dea D. *Protocol for direct costing of health sector interventions for economic modelling (including event pathways)*. Burden of Disease Epidemiology, Equity and Cost-Effectiveness Programme, Technical Report No. 12. Wellington: Department of Public Health, University of Otago, Wellington, 2012.

Appendix One: Search techniques

Medline and Embase

Access through the Medical School Library, under Databases.

Search terms

Select search by Keyword.

Medline and Embase (on Ovid from the Medical School Library) provide an option to "Map Term to Subject Headings" (MeSH; also called Thesaurus Terms). Alternative spelling and synonyms for the keyword will be accounted for with the subject heading indexing. Searching this way is recommended. Note that mapping search terms to MeSH subject headings works only if you key in a single keyword or phrase.

Select the relevant subject heading(s), and also the term as a keyword. It is a useful logic check to look at the search tree to ensure you have selected the most appropriate terms (click on the underlined subject heading). If the subject headings that appear are not what you expected, try changing your keyword slightly.

Consider exploding or focussing subject headings.

- Explode retrieves results using the selected term and all of its more specific terms
    - This is recommended where the subject could be indexed in a number of ways
        - E.g. "cardiovascular diseases" without exploding will miss some of the papers that are indexed as coronary disease
    - Important : if you do not select the explode option, the papers indexed with more specific terms will not be included in the search; indexers are instructed to use the most specific term possible, and do not necessarily include both the general and more specific term
    - Exploding will reduce specificity, so this must be balanced against the risk of missing papers if exploding is not used
- Focus limits the search to those documents in which your subject heading has been indexed as a major topic of the article
    - This may miss some possibly relevant papers but should have greater specificity
- You can manually select broader and narrower terms to optimise the balance of capture versus specificity
    - Click on the underlined subject heading term to show the search tree

- E.g. “cancer” maps to the subject heading “neoplasms”, but you may wish to manually select the most relevant subheadings, e.g. neoplastic malignant disease, metastasis, tumor and tumor recurrence.
- It is always advisable to also include the general term (i.e. the higher level heading) – some papers will be indexed only as “neoplasms” without a more specific term
- This produces an intermediate number of hits between the Explode and Focus options
- This is likely the optimal strategy, but requires understanding of the medical terms listed

If a more comprehensive search is needed, a second search can be conducted using only free text keywords. Consider UK and US spelling, synonyms, and use of truncations with wildcard characters. Then combine the first (MeSH) and second (free text keywords) searches using OR, and remove duplicates. This combined search will have good sensitivity but poor specificity. The free text search should pick up any papers that have not been indexed properly.

When searching free text, the following commands can be useful:

- * or $ can be used as wildcard characters, e.g. child* will find child or children
- The wildcard can also be used within a word, e.g. an*esthetic will find both anaesthetic and anesthetic
- adj will find words that are adjacent within the specified number of words, e.g. liver adj3 cancer will find all instances where liver and cancer occur within 3 words of each other so will pick up “liver and bowel cancer” as well as “liver cancer”

Limits

Warning – use of the “review articles” limit option in Medline on the basic search or advanced search is not recommended for searching for BODE[3]. This may miss important papers, especially meta-analyses, while returning a large number of non-systematic reviews and opinion pieces. Instead the following is recommended:

- Select Limits
- Open “Additional Limits”
  - Use the review options within the Clinical Queries
    - limit to "reviews (maximizes sensitivity)" – produces a large number of hits with more irrelevant papers, but should capture everything
    - limit to "reviews (maximizes specificity)” – produces a small number of hits but limits to systematic reviews and meta-analyses

        - this appears to have sufficient sensitivity for most searches for BODE^3, but a less specific search should be done if this does not return adequate results
    - limit to "reviews (best balance of sensitivity and specificity)" – produces better results than the simple “review articles” limit

Removing duplicates

In the keyword search box, type in ..dedup and the number of the search to remove duplicates (e.g. ..dedup 3).

Cochrane

Cochrane has only limited search functionality. Because of this, it is recommended that both a free text search and a MeSH search is done for BODE^3, as outlined below.

To combine searches of different terms and to use MESH terms, the option “Go directly to Search History” must be used when searching in Cochrane.

Select which databases to search (Cochrane Reviews, DARE, CENTRAL, Methodology Register, HTA database and/or NHS EED)

- Conduct an advanced search using free text with “Go directly to Search History” selected.
- Conduct a second search by clicking on the “MeSH Search” option
    - Click on “Go to MeSH trees”
    - Choose whether to explode or use the specific term only
    - Select “Go directly to Search History”
    - View Results
- Combine the 2 searches in the Search History page by typing in the search numbers (e.g. #1) in the search box with Boolean operators, e.g. #1 OR #2

Appendix Two: Literature search summary forms

Literature Search Summary Form: Master	
Completed by:	**Date:**
1. Define research question	
State research question	
P (population) =	
I (intervention/exposure and setting) =	
C (Comparisons) =	
O (outcome) =	
T (timeframe) =	
2. Define Inclusion and Exclusion Criteria	
Inclusion	Exclusion
Systematic reviews	
All systematic reviews relevant to the research question published from XXXX onwards	Opinion pieces, non-systematic reviews
Other specific inclusion criteria	*Other specific exclusion criteria*
Systematic review update	
All epidemiological, clinical and/or economic original research relevant to the research question published in the 2 years prior to, and all years subsequent to, the systematic review	Non-clinical research papers Descriptive and/or mechanistic papers
All other specific inclusion and exclusion criteria as defined above	
De novo Review	

All epidemiological, clinical and economic original research relevant to the research question	Non-clinical research papers Descriptive and/or mechanistic papers
All other specific inclusion and exclusion criteria as defined above	

3. Devise explicit search strategy	
Keywords	Limits

4. List sources used (delete any not applicable)		
Enter search strings into the separate Search Results Table		
Systematic reviews		
• Medline	• Cochrane library	• Other (specify)
• DARE and HTA databases	• Ministry of Health Publications	
Updating/de novo review		
• Medline	• Cochrane CENTRAL	
• Embase	• Other (specify)	
Economic papers		
• NHS EED	• Other (specify)	

5. Summary of search (from Search Results Table)	
Systematic reviews	
Number of records retrieved from databases	
Number of additional records received from other sources	

Total number of records retrieved after removal of duplicates	
Number of records excluded by relevance screening	
Number of articles excluded by exclusion criteria:	
o Unsuitable article type, e.g. non-systematic review or editorial	
o Unsuitable study population to address our research question	
o Unsuitable intervention/exposure to address our research question	
o Unsuitable outcome to address our research question	
o Updated or repetition of other similar review by same authors	
o Other (specify)	
Number of articles excluded by critical appraisal	
Reasons:	
Total number of systematic review articles retained for inclusion	
Updating/de novo review	
Number of records retrieved from databases	
Number of additional records received from other sources	
Total number of records retrieved after removal of duplicates	
Number of records excluded by relevance screening	
Number of articles excluded by exclusion criteria:	
o Unsuitable study design to address our research question	
o Unsuitable study population to address our research question	
o Unsuitable intervention/exposure to address our research question	

o Unsuitable outcome to address our research question	
o Updated or duplicate publication of same study	
o Other (specify)	
Number of articles excluded by critical appraisal	
Reasons:	
Total number of original research articles retained for inclusion	

Search Results Table: MASTER										
Database (date)	Search string (i) copy and paste string from electronic search, including any limits (ii) delete any sources below that were not searched (iii)add any additional sources searched	No. of papers								
		Initial no. retrieved	Duplicates[a]	Total no. excl duplicates	Relevance screening of records		Application of Inclusion/Exclusion Criteria to paper		Critical appraisal (final selection)	
					Exclude	Retain	Exclude	Retain	Exclude	**Final**
Systematic reviews										
NZ MoH										
Cochrane Library										
DARE/HTA										
Medline										
Snowballing										
Other (specify)										
Total										

Updating/De novo review										
Medline										
Embase										
CENTRAL										
Snowballing										
Other (specify)										
Total										

a Refers to duplicate detection of the same citation, either within the same database, or in different databases; ensure that one citation is retained, and other duplicates discarded. This does not include duplicate publishing where the same or very similar results have been published more than once in different papers, which would be excluded after application of inclusion/exclusion criteria.

EXAMPLE: Summary of Literature Search: Saturated Fats (questions 1-3)	
Conducted by: Rachel Foster **DATE:** 20/6/11	
1. Define research question	
How does a reduction in intake of dietary saturated fats affect the relative risks for cardiovascular disease events in adults in developed countries	
P (population) = Adults in developed countries	
I (intervention/exposure) = (i) Comparison of CVD events by high versus low saturated fat; (ii) Dietary intervention that reduces saturated fat intake, including replacement with PUFAs or other foods	
C (Comparisons) = (i) Low dietary saturated fat intake; (ii) No dietary intervention	
O (outcome) = Cardiovascular disease events	
T (timeframe) = n/a	
2. Define Inclusion and Exclusion Criteria	
Systematic reviews	
Inclusion	Exclusion
All systematic reviews relevant to the research question published from 2000 onwards	Opinion pieces, non-systematic reviews and non-clinical research papers Descriptive and/or mechanistic papers
Exposure = Intervention to reduce dietary saturated fat intake	Outcome cardiovascular risk rather than cardiovascular events
Outcome: cardiovascular events	High-risk populations
Study conducted in developed countries (OECD criteria)	Children
	Multifactorial interventions
	Supplementation with healthy oils (e.g. omega-6 or omega-3) rather than primary reduction in saturated fat
	Due to the large amount of data, systematic reviews that covered a range of dietary components, rather than focussing on fats, were also excluded

3. Devise explicit search strategy	
Keywords	Limits
Dietary cholesterol	English language
Saturated fat (maps to dietary fats)	Human
Cardiovascular disease	2000-present

Search results Table: EXAMPLE Saturated fats										
Database (date)	Search string	Initial retrieve	Dupl	Total unique	Relevance screening		Application of Inclusion/Exclusion criteria		Critical appraisal (final selection)	
					Exclude	Retain	Exclude	Retain	Exclude	Final
Medline (30/6/11)	1 exp Cardiovascular Diseases/co, di, dh, dt, ec, ep, eh, et, mo, nu, pa, pp, pc, rh, th [Complications, Diagnosis, Diet Therapy, Drug Therapy, Economics, Epidemiology, Ethnology, Etiology, Mortality, Nursing, Pathology, Physiopathology, Prevention & Control, Rehabilitation, Therapy] (682790) 2 dietary fats/ or butter/ or exp cholesterol, dietary/ (19613) 3 1 and 2 (2956) 4 limit 3 to (english language and humans and yr="2000 -Current") (1270)									
	5 limit 4 to **"reviews (maximizes specificity)"** (22) 6 remove duplicates from 7 (20)	20	1	19	9	10	5	5	2	3
Ministry of Health (13/5/11)	Manual search: diet, nutrition, fat	17	0	17	16	1	1	0	0	0
Cochrane Library	Dietary fat OR saturated fat AND cardiovascular disease	16	0	16	14	2	1	1	0	1

(13/5/11)										
DARE and HTA (1/11/11)	1. (saturated fat*):TI OR (cholesterol):TI OR (dietary fat*):TI OR (fatty acids):TI IN DARE, HTA FROM 2000 TO 2011 AND 2. (cardiovascular) OR (coronary) OR (ischaemic heart) IN DARE, HTA FROM 2000 TO 2011	37	3	34	32	2	2	0	0	0
Snowballing		1	0	1	0	1	0	1	0	1
Total		**91**	**4**	**87**	**71**	**16**	**9**	**7**	**2**	**5**

Appendix Three: Overview of sources

A. Systematic reviews/guidelines/Health Technology Assessments (HTAs)

Centre for Reviews and Dissemination (CRD)

The Centre for Reviews and Dissemination is a department of the University of York and is part of the National Institute for Health Research (NIHR). It is one of the largest groups in the world engaged exclusively in evidence synthesis in the health field. They produce the DARE, NHS EED and HTA databases (in collaboration with the Cochrane library).

Available via http://www.crd.york.ac.uk/crdweb/ or via the Cochrane Library.

- DARE: Database of Abstracts of Reviews of Effects
    - Provides structured abstracts of previously published systematic reviews and meta-analyses, with commentary on the strengths and weaknesses of the review/analysis
    - 15,000 abstracts of systematic reviews including over 9,000 quality assessed reviews and details of all Cochrane reviews and protocols and Campbell reviews
- HTA database
    - details of over 10,000 completed and ongoing health technology assessments from around the world
    - includes NICE and NIHR appraisals, HTAs from CADTH, and other HTAs

Cochrane Library Systematic Reviews

- Highly respected systematic reviews of research in healthcare and health policy; only trials that meet certain quality criteria are included and methods aim to reduce the impact of bias.
- Available via Medical School Library databases

Campbell Library

- Quality-defined systematic reviews on the effects of interventions within the areas of education, crime and justice, and social welfare.
- Include unpublished data to limit publication bias.
- Freely available at http://www.campbellcollaboration.org/library.php

National Institute of Health Research (NIHR) Health Technology Assessment programme (UK NHS)

- Considered a world leader in HTA

- Independent comprehensive assessments of the effectiveness, costs and broader impact of healthcare treatments and tests
- Analyses inform the work of NICE
  - NIHR HTA undertakes assessments and provides this information to NICE; NICE undertakes appraisals and provides guidance to the NHS
- Published in the journal "Health Technology Assessment"
- www.hta.ac.uk
- Also indexed in the HTA database

Canadian Agency for Drugs and Technologies in Health (CADTH)

- Previously Canadian Coordinating Office for Health Technology Assessment (CCOHTA)
- Provides Canada's health care decision makers with credible, impartial advice and evidence-based information about the effectiveness and efficiency of drugs and other health technologies
- www.cadth.ca
- Also indexed in the HTA database

Other sources that may be more useful to provide background and context (e.g. current gold standards for treatment, treatment pathways, validation etc) include:

- Clinical evidence (BMJ)
  - *Clinical Evidence* systematic reviews summarise the current state of knowledge and uncertainty about the prevention and treatment of clinical conditions, based on thorough searches and appraisal of the literature. It describes the best available evidence from systematic reviews, RCTs, and observational studies where appropriate, and highlights gaps in evidence.
  - Available via Medical Library databases

- NICE (UK National Institute for Health and Clinical Excellence)
  - National guidance to ensure that practice in local health communities is in line with the best available evidence of effectiveness and cost effectiveness.
  - www.nice.org.uk
  - Also indexed in the HTA database

- Pharmac (NZ) Technology Assessment Reports (TARs)
  - Cost-utility analyses of pharmaceuticals
  - TARs comprehensively report effectiveness, safety and cost-effectiveness.

- Most are confidential; only a limited number are publically released
- www.pharmac.govt.nz

- Australia
  - Similar to Pharmac – most HTAs are confidential
  - PBAC = drugs; MSAC = medical services

B. Key Indexing databases

These indexing databases are available from the Medical Library, except for NHS EED (available from CRD at http://www.crd.york.ac.uk/crdweb/).

NHS Economic Evaluation Database (NHS EED)

- Very comprehensive - 24,000 abstracts of health economics papers including over 11,000 quality-assessed economic evaluations
- Focuses on cost analyses of interest to the UK NHS, so does not cover methodological issues comprehensively
- Does not include *reviews* of cost-effectiveness analyses

Medline

- The US National Library of Medicine's premier bibliographic database providing information from the fields of medicine, nursing, dentistry, veterinary medicine, allied health and pre-clinical sciences.
- Journals are required to meet very strict criteria in order to be indexed in Medline (quality of content, importance, originality, quality of editorial work and production etc). This means that the indexed articles are generally of a high standard, but the coverage is more restricted.
  - Only a quarter of the journals who apply are accepted
  - About 5,300 titles are currently indexed

Embase

- major biomedical and pharmaceutical database indexing over 7,500 international journals
- more inclusive than Medline and has pharmaceutical focus (covers 2000 journals not covered by Medline), but included journals may be of more variable quality
- includes conference abstracts

Cochrane Central Register of Controlled Trials (CENTRAL)

- Most comprehensive database of *controlled* clinical trials

- Citations sourced from Medline, Embase, hand searching of conference proceedings, and other less easily attained sources.
- Over 600,000 records of randomised trials
- Available through the Cochrane Library

Scopus

- Scopus is the largest abstract and citation database of peer-reviewed literature and quality web sources
- Covers 25 million abstracts from over 14,000 titles across 4,000 publishers
- Covers Scientific, Technical, Medical and Social Sciences literature
- Provides a broad, but not always specific search

Google Scholar

- Broadly searches scholarly literature across many disciplines and sources: peer-reviewed papers, theses, books, abstracts and articles, from academic publishers, professional societies, preprint repositories, universities and other scholarly organizations.
- Provides a broad, but not always specific search

Additional indexing databases

Other more specific indexing databases may be useful if they are a good fit for the research question, and Medline and Embase have not provided adequate coverage. For instance, the AMED and/or CINAHL databases may be useful *additional* sources for interventions around palliative care. All are available through the Medical School Library.

- AMED (Allied and Complementary Medicine)
  - produced by the Health Care Information Service of the British Library
  - covers a selection of journals in complementary medicine, palliative care, and several professions allied to medicine.

- CINAHL
  - CINAHL is the claimed to be authoritative resource for nursing and allied health

- PsycINFO
  - provides citations with abstracts to the scholarly literature in the psychological, social, behavioral, and health sciences.

- REHABDATA
    - Disability and rehabilitation database produced by the National Rehabilitation Information Center. It describes over 65,000 documents covering physical, mental, and psychiatric disabilities, independent living, vocational rehabilitation, special education, assistive technology, law, employment, and other issues as they relate to people with disabilities.

D. Databases of clinical trials

Databases of clinical trials are important for identifying unpublished data.

- ClinicalTrials.gov
    - Run by the US NIH in collaboration with the FDA
    - Holds records for federally and privately supported clinical trials for a wide range of diseases and conditions
    - Contains approx 108,000 trials sponsored by the National Institutes of Health, other federal agencies, and private industry.
    - Covers 174 countries
    - Available at http://clinicaltrials.gov/