The Systematic Review Workbook

Gannon University
# Table of Contents

9 steps to creating a Systematic Review ................................................. 3

- **Step 1: Develop a question** ......................................................... 3
- **Step 2: Rationale for review and objectives** ................................. 4
- **Step 3: Inclusion and exclusion criteria** ........................................ 5
- **Step 4: A systematic way to search** ............................................ 6
- **Step 5: Critically appraising the evidence** ..................................... 19
  - CEBM Appraisal forms ............................................................... 20
  - Critical Appraisal Skills Program forms ....................................... 43
  - PEDRO Scale ............................................................................. 88
- **Step 6: Analyzing and synthesizing** ............................................ 91
- **Step 7: Create a flow diagram** .................................................... 93
- **Step 8: Reporting the results** ..................................................... 96
- **Step 9: Take home points** ......................................................... 113

Writing it up! .................................................................................. 114

How to use Medical Subject Headings ............................................. 127

Saving and managing searches in PubMed ..................................... 131
  - Storing a safe search ................................................................ 135
  - Building a search collection ..................................................... 136

Changing display settings ............................................................... 138
Step 1: Develop a Question

What do you need to include when creating your research question?

- **THE POPULATION OF INTEREST**
- **WHAT INTERVENTION WILL BE USED**
- **A COMPARISON OF GROUPS**
- **A MEASURABLE OUTCOME.**

This style of question is sometimes referred to as a P.I.C.O. Question.

Common areas of focus include treatment, diagnosis, etiology, prognosis, or prevention (Bradley & Law, 2014, p. 160).
Step 2: Rationale for review and objectives

Provide reasoning and support for completing a systematic review on this topic.

Be sure to identify the objectives and purpose of the review.

So what? Who cares?
Step 3: Inclusion and exclusion criteria

- Use the P.I.C.O. elements to help define the inclusion and exclusion criteria. Criteria may also include predetermining requirements regarding the level of evidence, the type of evidence, and how current the evidence needs to be for inclusion (Bradley & Law, 2014, p. 160).

- Consider heterogeneity of studies when deciding on inclusion and exclusion criteria. If the differences between studies are many, applying the results of the review can be challenging. Ways to avoid heterogeneity within a systematic review may include setting strict inclusion criteria or incorporating a sub-analysis of the results to allow for more studies to be included (Brown, 2017, p.189). When studies included in a systematic review are very similar, more confidence can be placed in the conclusions that are drawn (Brown, 2017, p.189).
Step 4: A systematic way to search

- Develop a system for searching the evidence and clearly explain this system in your report.

- Track what has been searched to avoid confusion and to prevent any subheadings or key words from being left out.

- In addition to identifying inclusion and exclusion criteria, researchers should include the names of all search terms and databases used during the search.

Selecting the right database

- Selecting the database will be dependent on the field or area that you think will contain the most material related to your project and/or question. (Hissong, Lape, & Bailey, 2015, p.21)
Some databases which are often used to find articles relevant to health professions include:

**Cumulative Index to Nursing and Allied Health (CINAHL)**
This database includes journals, related books, chapters, conferences, audio visuals and educational software. It also includes articles relevant to a wide variety of therapy and rehabilitative studies.

**Cochrane Library**
This database includes a registry of clinical trials as well as man systematic reviews (Brown, 2017, p. 23).

**Medline**
This database includes a large selection of medical related per-reviewed literature.
**PubMed**

This database includes a large resource of biomedical research articles. Medline is included in this database. (Hissong, Lape, & Bailey, 2015, p. 20; Brown, 2017, p. 24).

**Educational Resources Information Center (ERIC)**

This database is sponsored by the United States department of education and includes journal articles and other education-related materials (Hissong, Lape, & Bailey, 2015, p. 20).

**PsychInfo**

This database primarily includes peer-reviewed journal articles (Hissong, Lape, & Bailey, 2015, p. 20).
A good start to selecting terms that will be used while searching for articles is to identify key words that are included in your research question.

It is important to remember that catalogs and databases may use specialized or specific terminology which will impact the search results. As a result, the same search terms may yield different results depending on which database is being searched.

Using a thesaurus or MeSH terms can help with identifying which terms to include in your search title.
MeSH is the U.S. National Library of Medicine's controlled vocabulary used to index journal articles in MEDLINE. Subject analysts assign each article the most specific MeSH terms applicable — typically ten to twelve. MeSH descriptors provide a consistent way to find articles that may have used different terminology for the same concepts” (Pinotti, n.d.).

MeSH terms are organized like a hierarchy with broader terms above terms that are more specific (Ecker & Skelly, 2010, Brown, 2017, p. 24). If you find that the search terms you are using are providing an overwhelming amount of literature, you may need to consider modifying your search terms by using MeSH terms that fall lower on the MeSH term hierarchy. This will help to narrow the search results. Further details on how to use MeSH terms can be found beginning on page 103.
Another way to narrow or broaden your search results is by combining search terms with connector words such as AND, OR, or NOT. These words are sometimes referred to as Boolean operators and should be typed using upper case letters.

The following pages provide a worksheet for help with using Boolean terms in databases such as PubMed and Medline (Jacobs, 2017).
**Boolean Worksheet: A Medline/PubMed Search**

### Step 1: Choose databases that suit your research question

<table>
<thead>
<tr>
<th><strong>Sample Search</strong></th>
<th><strong>Your Search</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sample Research Question:</strong> Does handwashing among healthcare workers reduce hospital acquired infection?</td>
<td><strong>Your Research Question:</strong></td>
</tr>
<tr>
<td><strong>Recommended Databases:</strong></td>
<td><strong>Recommended Databases:</strong></td>
</tr>
<tr>
<td>– CINAHLPlus</td>
<td>– CINAHLPlus</td>
</tr>
<tr>
<td>– <strong>Medline (via PubMed or Ovid)</strong></td>
<td>– Medline (via PubMed or Ovid)</td>
</tr>
<tr>
<td>– PsycINFO</td>
<td>– PsycINFO</td>
</tr>
<tr>
<td>– Cochrane Database of Systematic Reviews</td>
<td>– Cochrane Database of Systematic Reviews</td>
</tr>
<tr>
<td>– JBI (Joanna Briggs Institute) EBPDatabase</td>
<td>– JBI (Joanna Briggs Institute) EBPDatabase</td>
</tr>
<tr>
<td>– Scopus</td>
<td>– Scopus</td>
</tr>
<tr>
<td>– Web of Science</td>
<td>– Web of Science</td>
</tr>
<tr>
<td>– Embase</td>
<td>– Embase</td>
</tr>
<tr>
<td>– ProQuest Central</td>
<td>– ProQuest Central</td>
</tr>
<tr>
<td>– Dissertations and Theses Global</td>
<td>– Dissertations and Theses Global</td>
</tr>
<tr>
<td>– Other</td>
<td>– Other</td>
</tr>
</tbody>
</table>

For the sample search, we will start in PubMed

### Step 2: Perform a simple keyword search

Identify **keywords** that describe the important concepts in your research question. (For a PICO question, select keywords that describe your **patient, problem or population** and your chosen **intervention**)

**Sample Keywords:**

<table>
<thead>
<tr>
<th>Concept 1</th>
<th>Concept 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(Patient or Problem or Population)</strong></td>
<td></td>
</tr>
<tr>
<td>hospital acquired infection</td>
<td></td>
</tr>
<tr>
<td><strong>(Intervention)</strong></td>
<td></td>
</tr>
<tr>
<td>handwashing</td>
<td></td>
</tr>
</tbody>
</table>

Connecting keywords with **AND** in the database search retrieves article citations that contain **both** terms.

©Copyright NYU Libraries 2017
**Step 3: Expand** your search with MeSH terms and other synonyms

From the initial search results, retrieve alternate terms/synonyms/MeSH subject terms for one or both concepts.

Combining these terms with **OR** expands the search to find citations with **ANY** of the added terms (*Remember – OR retrieves MORE!*)

**Sample Synonyms:**

<table>
<thead>
<tr>
<th>Concept 1</th>
<th>Concept 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>hospital acquired infection</td>
<td>handwashing</td>
</tr>
<tr>
<td>OR</td>
<td>OR</td>
</tr>
<tr>
<td>cross infection</td>
<td>hand hygiene</td>
</tr>
<tr>
<td>OR</td>
<td>OR</td>
</tr>
<tr>
<td>nosocomial</td>
<td>hand disinfection</td>
</tr>
</tbody>
</table>

A larger OR’d set of MeSH/keywords/synonyms may be recombined with **AND** to narrow the search

**Your Synonyms:**

<table>
<thead>
<tr>
<th>Your Concept 1</th>
<th>Your Concept 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>_____________</td>
<td>_____________</td>
</tr>
<tr>
<td><strong>OR</strong></td>
<td><strong>OR</strong></td>
</tr>
<tr>
<td>_____________</td>
<td>_____________</td>
</tr>
<tr>
<td><strong>OR</strong></td>
<td><strong>OR</strong></td>
</tr>
<tr>
<td>_____________</td>
<td>_____________</td>
</tr>
</tbody>
</table>

**TIP:** Remember to nest OR’d synonyms with **parentheses** in your database search – e.g. *(handwashing OR hand hygiene OR hand disinfection)*

**Step 4: Add limits to your results**

Most databases allow you to add **“limits” or “filters”** that restrict search results by characteristics such as type of study, type of article, date, time factors, age group, etc.

In your search, look for database **‘filters’, limits’ or ‘refinements’** to narrow your search.

You may want to consider how you can apply the following filters to best suit your research needs:

- **Article Type:** ________________
- **Date:** ________________
- **Language:** ________________
- **Age Group**: ________________

**Pro Tip:** If you are looking for articles about a particular age group of patients, DON’T include terms like ‘pediatric’ or ‘elderly’ in your search. Instead, use the **age group** filter that is available in many databases.

(Jacob, 2017) ©Copyright NYU Libraries 2017
Discovering articles

- Some databases provide a limit or filter feature which can be used to include or exclude citations by age group, language, publication type, date, etc. (Ecker & Skelly, 2010, Brown, 2017, p. 27). If you find that you are lacking articles, try using a major study that includes many participants or studies and look to see if any of the articles sited are relevant to your study (Brown, 2017, p. 29).

- Remember to consider publication bias when searching the evidence for studies as studies that include positive findings are up to four times more likely to be published (Brown, 2017, p.189). While finding unpublished articles can be challenging, including them in your review can decrease the impact of publication bias (Brown, 2017, p.189).
Tracking the process

Keeping track of the databases you select, the combination of search terms that you use, and the literature that you include and exclude is very important.

Every step must be recorded with enough detail so that the search process could be replicated and the persons replicating the search could obtain the same results.

One way to organize this information during the search process is by using charts.

*see example on next page*
<table>
<thead>
<tr>
<th>Brief Citation</th>
<th>Participants/Student design</th>
<th>Independent Variable</th>
<th>Dependent Variable</th>
<th>Important Covariates</th>
<th>Results</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>• N= ___</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Method</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Data source</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Year(s) of study</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Conclusions indicate... ___
- Relevant associations
- What wasn’t considered?
- Any methodological problems?
In addition to charts, some databases, such as PubMed, have built in software that can save your searches and organize them into a collection.

PubMed allows researchers to create a account through My NCBI. By logging in, researchers have the ability to save searches, selected articles, and set up alerts to notify them when new articles are found that meet the search criteria.

PubMed allows users to create a collection through NCBI. When collections are created, users also have the option of setting up notifications that will alert the user when new articles that match the search criteria are published.
Deciding which articles to include

After identifying articles, you will need to review the abstracts to help determine relevance (Hissong, Lape, & Bailey, 2015, p. 21).

One method for organizing which articles to include, and how, is to look for common themes that run through each article and place articles into groups based on those themes. This will also help with the review process (Hissong, Lape, & Bailey, 2015, p. 22).

Be sure to place the articles in a logical sequence so that when it is time to write up the review everything flows smoothly (Hissong, Lape, & Bailey, 2015, p. 22).

Information on the setup and use of these features, though NCBI, is available starting on page 107.
Step 5: Critically Appraising the Evidence

Each article must be critically appraised by a minimum of two reviewers. Reviewers may focus on areas such as methodological strengths and weaknesses, inclusion criteria, bias, level of blinding, missing data, and validity (Bradley & Law, 2014, p.161).

After critically appraising and assigning scores to the articles, reviewers must come together and compare results. Any differences are discussed so that an agreed upon score can be decided. An outside reviewer may step in to offer a 3rd opinion if common ground is not found.

It is important that all methodology and appraisal criteria be reported and clearly explained (Bradley & Law, 2014, p.162).
Appraisal tool resources

The Center for Evidence—Based Medicine provides critical appraisal worksheets for various study types including:

- Systematic Review
- Diagnostic
- Prognosis
- Randomized Controlled Trials
- Qualitative Studies

General worksheets, such as the one provided on the following page, are also accessible through CEBM.

These forms can be accessed by visiting the website below

https://www.cebm.net/2014/06/critical-appraisal/

(CEBM, 2014)
<table>
<thead>
<tr>
<th>Question</th>
<th>Description</th>
<th>Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>P</strong>atients</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>I</strong>ntervention</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>C</strong>omparator</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>O</strong>utcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CER (%)</td>
<td>IER (%)</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Appraisal**

- Randomized
- Ascertainment
- Measures

**Outcomes**

- RDifference: CER - EER
- RRR: RD/CER
- NNT: 1/RD

**Clinical Bottom-line:**

**Further Actions:**
Are the results of the review valid?

What question (PICO) did the systematic review address?

What is best?  Where do I find the information?

The main question being addressed should be clearly stated. The exposure, such as a therapy or diagnostic test, and the outcome(s) of interest will often be expressed in terms of a simple relationship.

The Title, Abstract or final paragraph of the Introduction should clearly state the question. If you still cannot ascertain what the focused question is after reading these sections, search for another paper!

In this paper

Yes  No  Unclear

Comment:

F - Is it unlikely that important, relevant studies were missed?

What is best?  Where do I find the information?

The starting point for a comprehensive search for all relevant studies is the major bibliographic databases (eg Medline, Cochrane, EMBASE, etc) but should also include a search of reference lists from relevant studies and contact with experts, particularly to inquire about unpublished studies. The search should not be limited to English language only. The search strategy should include both MESH terms and text words.

The Methods section should describe the search strategy, including the terms used, in some detail. The Results section will outline the number of titles and abstracts reviewed, the number of full-text studies retrieved, and the number of studies excluded together with the reasons for exclusion. This information may be presented in a figure or flow chart.

In this paper

Yes  No  Unclear

Comment:
### A - Were the criteria used to select articles for inclusion appropriate?

**What is best?**

The inclusion or exclusion of studies in a systematic review should be clearly defined *a priori*. The eligibility criteria used should specify the patients, interventions or exposures and outcomes of interest. In many cases the type of study design will also be a key component of the eligibility criteria.

**Where do I find the information?**

The Methods section should describe in detail the inclusion and exclusion criteria. Normally, this will include the study design.

<table>
<thead>
<tr>
<th>In this paper</th>
<th>Yes</th>
<th>No</th>
<th>Unclear</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Comment:**

---

### A - Were the included studies sufficiently valid for the type of question asked?

**What is best?**

The article should describe how the quality of each study was assessed using predetermined quality criteria appropriate to the type of clinical question (e.g., randomization, blinding and completeness of follow-up).

**Where do I find the information?**

The Methods section should describe the assessment of quality and the criteria used. The Results section should provide information on the quality of the individual studies.

<table>
<thead>
<tr>
<th>In this paper</th>
<th>Yes</th>
<th>No</th>
<th>Unclear</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Comment:**
In this paper

Yes ☐ No ☐ Unclear ☐

Comment:

What were the results?

A systematic review provides a summary of the data from the results of a number of individual studies. If the results of the individual studies are similar, a statistical method (called meta-analysis) is used to combine the results from the individual studies and an overall summary estimate is calculated. The meta-analysis gives weighted values to each of the individual studies according to their size. The individual results of the studies need to be expressed in a standard way, such as relative risk, odds ratio or mean difference between the groups. Results are traditionally displayed in a figure called a forest plot, like the one below.

Comparison: 03 Treatment versus Placebo
Outcome: 01 Effect of treatment on mortality

<table>
<thead>
<tr>
<th>Study</th>
<th>Treatment n/N</th>
<th>Control n/N</th>
<th>OR (95%CI Fixed)</th>
<th>Weight %</th>
<th>OR (95%CI Fixed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brown 1998</td>
<td>24 / 472</td>
<td>35 / 459</td>
<td></td>
<td>9.6</td>
<td>0.71[0.42,1.21]</td>
</tr>
<tr>
<td>Geoffrey 1997</td>
<td>120 / 2860</td>
<td>182 / 2638</td>
<td></td>
<td>51.8</td>
<td>0.64[0.51,0.81]</td>
</tr>
<tr>
<td>Mason 1996</td>
<td>56 / 2051</td>
<td>34 / 2030</td>
<td></td>
<td>24.4</td>
<td>0.65[0.46,0.92]</td>
</tr>
<tr>
<td>Peters 2000</td>
<td>5 / 81</td>
<td>4 / 76</td>
<td></td>
<td>1.1</td>
<td>1.22[0.31,4.71]</td>
</tr>
<tr>
<td>Scott 1998</td>
<td>31 / 799</td>
<td>46 / 792</td>
<td></td>
<td>13.1</td>
<td>0.66[0.42,1.06]</td>
</tr>
</tbody>
</table>

Total(95%CI): 286 / 16242 351 / 26237
Test for heterogeneity chi-square=0.82 df=4 p=0.92
Test for overall effect z=4.82 p<0.00001

The forest plot depicted above represents a meta-analysis of five trials that assessed the effects of a hypothetical treatment on mortality. Individual studies are represented by a black square and a horizontal line, which corresponds to the point estimate and 95% confidence interval of the odds ratio. The size of the black square reflects the weight of the study in the meta-analysis. The solid vertical line corresponds to ‘no effect’ of treatment - an odds ratio of 1.0. When the confidence interval includes 1 it indicates that the result is not significant at conventional levels (P>0.05).
The diamond at the bottom represents the combined or pooled odds ratio of all five trials with its 95% confidence interval. In this case, it shows that the treatment reduces mortality by 34% (OR 0.66 95% CI 0.56 to 0.78). Notice that the diamond does not overlap with the ‘no effect’ line (the confidence interval doesn’t include 1) so we can be assured that the pooled OR is statistically significant. The test for overall effect also indicates statistical significance (p<0.0001).

Exploring heterogeneity

Heterogeneity can be assessed using the ‘eyeball’ test or more formally with statistical tests, such as the Cochran Q test. With the ‘eyeball’ test one looks for overlap of the confidence intervals of the trials with the summary estimate. In the example above note that the dotted line running vertically through the combined odds ratio crosses the horizontal lines of all the individual studies indicating that the studies are homogenous. Heterogeneity can also be assessed using the Cochran chi-square (Cochran Q). If Cochran Q is statistically significant there is definite heterogeneity. If Cochran Q is not statistically significant but the ratio of Cochran Q and the degrees of freedom (Q/df) is > 1 there is possible heterogeneity. If Cochran Q is not statistically significant and Q/df is < 1 then heterogeneity is very unlikely. In the example above Q/df is <1 (0.92/4= 0.23) and the p-value is not significant (0.92) indicating no heterogeneity.

*Note: The level of significance for Cochran Q is often set at 0.1 due to the low power of the test to detect heterogeneity.*
Are the results of the study valid?

Was the diagnostic test evaluated in a Representative spectrum of patients (like those in whom it would be used in practice)?

**What is best?**

It is ideal if the diagnostic test is applied to the full spectrum of patients - those with mild, severe, early and late cases of the target disorder. It is also best if the patients are randomly selected or consecutive admissions so that selection bias is minimized.

**Where do I find the information?**

The Methods section should tell you how patients were enrolled and whether they were randomly selected or consecutive admissions. It should also tell you where patients came from and whether they are likely to be representative of the patients in whom the test is to be used.

**In this paper**

Yes [ ]  No [ ]  Unclear [ ]

**Comment:**

Was the reference standard applied regardless of the index test result?

**What is best?**

Ideally both the index test and the reference standard should be carried out on all patients in the study. In some situations where the reference standard is invasive or expensive there may be reservations about subjecting patients with a negative index test result (and thus a low probability of disease) to the reference standard. An alternative reference standard is to follow-up people for an appropriate period of time (dependent on disease in question) to see if they are truly negative.

**Where do I find the information?**

The Methods section should indicate whether or not the reference standard was applied to all patients or if an alternative reference standard (eg follow-up) was applied to those who tested negative on the index test.

**In this paper**

Yes [ ]  No [ ]  Unclear [ ]
Was there an independent, blind comparison between the index test and an appropriate reference (‘gold’) standard of diagnosis?

What is best?
There are two issues here. First, the reference standard should be appropriate - as close to the ‘truth’ as possible. Sometimes there may not be a single reference test that is suitable and a combination of tests may be used to indicate the presence of disease.
Second, the reference standard and the index test being assessed should be applied to each patient independently and blindly. Those who interpreted the results of one test should not be aware of the results of the other test.

Where do I find the information?
The Methods section should have a description of the reference standard used and if you are unsure of whether or not this is an appropriate reference standard you may need to do some background searching in the area.
The Methods section should also describe who conducted the two tests and whether each was conducted independently and blinded to the results of the other.

In this paper

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
<th>Unclear</th>
</tr>
</thead>
</table>

Comment:

What were the results?

Are test characteristics presented?
There are two types of results commonly reported in diagnostic test studies. One concerns the accuracy of the test and is reflected in the sensitivity and specificity. The other concerns how the test performs in the population being tested and is reflected in predictive values (also called post-test probabilities). To explore the meaning of these terms, consider a study in which 1000 elderly people with suspected dementia undergo an index test and a reference standard. The prevalence of dementia in this group is 25%. 240 people tested positive on both the index test and the reference standard and 600 people tested negative on both tests. The first step is to draw a 2 x 2 table as shown below. We are told that the prevalence of dementia is 25% therefore we can fill in the last row of totals - 25% of 1000 people is 250 - so 250 people will have dementia and 750 will be free of dementia. We also know the number of people testing positive and negative on both tests and so we can fill in two more cells of the table.

<table>
<thead>
<tr>
<th>Reference standard</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Index test</th>
<th>+ve</th>
<th>-ve</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ve</td>
<td>240</td>
<td></td>
</tr>
<tr>
<td>-ve</td>
<td></td>
<td>600</td>
</tr>
</tbody>
</table>

| Total      | 250 | 750 | 1000 |

Comment:
Now we are ready to calculate the various measures.

<table>
<thead>
<tr>
<th>What is the measure?</th>
<th>What does it mean?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sensitivity (Sn)</strong> = the proportion of people with the condition who have a positive test result.</td>
<td>The sensitivity tells us how well the test identifies people with the condition. A highly sensitive test will not miss many people.</td>
</tr>
<tr>
<td>In our example the Sn = 240/250 = 0.96</td>
<td>10 people (4%) with dementia were falsely identified as not having it. This means the test is fairly good at identifying people with the condition.</td>
</tr>
<tr>
<td><strong>Specificity (Sp)</strong> = the proportion of people without the condition who have a negative test result.</td>
<td>The specificity tells us how well the test identifies people without the condition. A highly specific test will not falsely identify many people as having the condition.</td>
</tr>
<tr>
<td>In our example the Sp = 600/750 = 0.80</td>
<td>150 people (20%) without dementia were falsely identified as having it. This means the test is only moderately good at identifying people without the condition.</td>
</tr>
<tr>
<td><strong>Positive Predictive Value (PPV)</strong> = the proportion of people with a positive test who have the condition.</td>
<td>This measure tells us how well the test performs in this population. It is dependent on the accuracy of the test (primarily specificity) and the prevalence of the condition.</td>
</tr>
<tr>
<td>In our example the PPV = 240/390 = 0.62</td>
<td>Of the 390 people who had a positive test result, 62% will actually have dementia.</td>
</tr>
</tbody>
</table>
**Negative Predictive Value (NPV)** = the proportion of people with a negative test who do not have the condition.

In our example the NPV = 600/610 = 0.98

This measure tells us how well the test performs in this population. It is dependent on the accuracy of the test and the prevalence of the condition.

Of the 610 people with a -ve test, 98% will not have dementia.

### Applicability of the Results

**Were the methods for performing the test described in sufficient detail to permit replication?**

<table>
<thead>
<tr>
<th>What is best?</th>
<th>Where do I find the information?</th>
</tr>
</thead>
<tbody>
<tr>
<td>The article should have sufficient description of the test to allow its replication and also interpretation of the results.</td>
<td>The Methods section should describe the test in detail.</td>
</tr>
</tbody>
</table>

### In this paper

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
<th>Unclear</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comment:
CRITICAL APPRAISAL OF PROGNOSTIC STUDIES

Are the results of the study valid? (Internal Validity)

1. Was the defined representative sample of patients assembled at a common (usually early) point in the course of their disease?

What is best?

It is preferable if study patients are enrolled at a uniformly early time in the disease usually when disease first becomes manifest. Such groups of patients are called an ‘inception cohort’. Patients should also be representative of the underlying population. Patients from tertiary referral centres may have more advanced disease and poorer prognoses than patients from primary care.

Where do I find the information?

The Methods section should describe the stage at which patients entered the study (eg at the time of first myocardial infarction; Stage 3 breast cancer). The Methods section should also provide information about patient recruitment, whether patients were recruited from primary care or tertiary referral centres.

In this paper

Yes ☐  No ☐  Unclear ☐

Comment:

2. Was patient follow-up sufficiently long and complete?

What is best?

Length of follow-up should be long enough to detect the outcome of interest. This will vary depending on the outcome (eg for pregnancy outcomes, nine months; for cancer, many years). All patients should be followed from the beginning of the study until the outcome of interest or death occurs. Reasons for non follow-up should be provided along with comparison of the demographic and clinical characteristics of the patients who were unavailable and those in whom follow-up was complete.

Where do I find the information?

The Results section should state the median or mean length of follow-up.

The Results section should also provide the number and the reasons for patients being unavailable for follow-up. A comparison of the two groups (those available and those unavailable) may be presented in table form or the authors may simply state in the text whether or not there were differences.
### 3. Were outcome criteria either objective or applied in a ‘blind’ fashion?

**What is best?**
A clear definition of all outcomes should be provided. It is ideal if less objective outcomes are assessed blindly, that is, the individual determining the outcome does not know whether the patient has a potential prognostic factor.

**Where do I find the information?**
The Methods section should provide a clear definition or explicit criteria for each outcome. Whether determination is blinded to prognostic factors will be found in either the Methods or Results sections.

### In this paper

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
<th>Unclear</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Comment:**

### 4. If subgroups with different prognoses are identified, did adjustment for important prognostic factors take place?

**What is best?**
A prognostic factor is a patient characteristic (e.g., age, stage of disease) that predicts the patient’s eventual outcome. The study should adjust for known prognostic factors in the analysis so that results are not distorted.

**Where do I find the information?**
The Results section should identify any prognostic factors and whether or not these have been adjusted for in the analysis. Also look at the tables and figures for evidence of this (e.g., there may be separate survival curves for patients at different stages of disease or for different age groups).

### In this paper

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
<th>Unclear</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Comment:**
What are the results?

How likely are the outcomes over time?

There are several different ways of reporting outcomes of disease. Often they are reported simply as a rate (e.g., the proportion of people experiencing an event). Expressing prognosis as a rate has some advantages. It is simple, easily communicated and understood and readily committed to memory. Unfortunately, rates convey very little information and there can be important differences in prognosis within similar summary rates. For this reason survival curves are used to estimate survival of a cohort over time. It is a useful method for describing any dichotomous outcome (not just survival) that occurs only once during the follow-up period. The figure below shows the survival curves for three diseases with the same survival rate at 5 years. Notice that the summary rate obscures important differences to patients.

Figure: Five year curves for three different diseases.
How precise are the prognostic estimates?

To determine the precision of the estimates we need to look at the 95% confidence intervals (CI) around the estimate. The narrower the CI, the more useful the estimate. The precision of the estimates depends on the number of observations on which the estimate is based. Since earlier follow-up periods usually include results from more patients than later periods, estimates on the left hand side of the curve are usually more precise. Observations on the right or tail end of the curve are usually based on a very small number of people because of deaths, dropouts and late entrants to the study. Consequently, estimates of survival at the end of the follow-up period are relatively imprecise and can be affected by what happens to only a few people.

Can I apply this valid, important evidence about prognosis to my patient?

The questions that you should ask before you decide to apply the results of the study to your patients are:

- Is my patient so different to those in the study that the results cannot apply?
- Will this evidence make a clinically important impact on my conclusions about what to offer to tell my patients?
Are the results of the trial valid? (Internal Validity)

What question did the study ask?
- Patients
- Intervention
- Comparison
- Outcomes

1a. R- Was the assignment of patients to treatments randomised?

What is best?

Centralised computer randomisation is ideal and often used in multi-centred trials. Smaller trials may use an independent person (e.g., the hospital pharmacy) to ‘police’ the randomization.

Where do I find the information?

The Methods should tell you how patients were allocated to groups and whether or not randomisation was concealed.

In this paper

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
<th>Unclear</th>
</tr>
</thead>
</table>

Comment:

1b. R- Were the groups similar at the start of the trial?

What is best?

If the randomisation process worked (that is, achieved comparable groups) the groups should be similar. The more similar the groups the better it is. There should be some indication of whether differences between groups are statistically significant (i.e., p values).

Where do I find the information?

The Results should have a table of ‘Baseline Characteristics’ comparing the randomized groups on a number of variables that could affect the outcome (i.e., age, risk factors etc). If not, there may be a description of group similarity in the first paragraphs of the Results section.

In this paper

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
<th>Unclear</th>
</tr>
</thead>
</table>

Comment:
### 2a. A - Aside from the allocated treatment, were groups treated equally?

**What is best?**

Apart from the intervention the patients in the different groups should be treated the same, eg additional treatments or tests.

**Where do I find the information?**

Look in the Methods section for the follow-up schedule and permitted additional treatments etc, and in Results for actual use.

<table>
<thead>
<tr>
<th>In this paper</th>
<th>Yes</th>
<th>No</th>
<th>Unclear</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comment:</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 2b. A - Were all patients who entered the trial accounted for? And were they analysed in the groups to which they were randomised?

**What is best?**

Losses to follow-up should be minimal—preferably less than 20%. However, if few patients have the outcome of interest, then even small losses to follow-up can bias the results. Patients should also be analysed in the groups to which they were randomised—‘intention-to-treat analysis’.

**Where do I find the information?**

The Results section should say how many patients were randomised (eg Baseline Characteristics table) and how many patients were actually included in the analysis. You will need to read the results section to clarify the number and reason for losses to follow-up.

<table>
<thead>
<tr>
<th>In this paper</th>
<th>Yes</th>
<th>No</th>
<th>Unclear</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comment:</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 3. M - Were measures objective or were the patients and clinicians kept “blind” to which treatment was being received?

**What is best?**

It is ideal if the study is ‘double-blinded’—that is, both patients and investigators are unaware of treatment allocation. If the outcome is objective (eg death) then blinding is less critical. If the outcome is subjective (eg symptoms or function) then blinding of the outcome assessor is critical.

**Where do I find the information?**

First, look in the Methods section to see if there is some mention of masking of treatments eg placebos with the same appearance or sham therapy. Second, the Methods section should describe how the outcome was assessed and whether the assessor(s) were aware of the patients’ treatment.

<table>
<thead>
<tr>
<th>In this paper</th>
<th>Yes</th>
<th>No</th>
<th>Unclear</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comment:</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In this paper

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
<th>Unclear</th>
</tr>
</thead>
</table>

Comment:

What were the results?

1. How large was the treatment effect?

Most often results are presented as dichotomous outcomes (yes or no outcomes that happen or don’t happen) and can include such outcomes as cancer recurrence, myocardial infarction and death. Consider a study in which 15% (0.15) of the control group died and 10% (0.10) of the treatment group died after 2 years of treatment. The results can be expressed in many ways as shown below.

<table>
<thead>
<tr>
<th>What is the measure?</th>
<th>What does it mean?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Relative Risk (RR)</strong> = risk of the outcome in the treatment group / risk of the outcome in the control group.</td>
<td>The relative risk tells us how many times more likely it is that an event will occur in the treatment group relative to the control group. An RR of 1 means that there is no difference between the two groups thus, the treatment had no effect. An RR &lt; 1 means that the treatment decreases the risk of the outcome. An RR &gt; 1 means that the treatment increased the risk of the outcome. Since the RR &lt; 1, the treatment decreases the risk of death.</td>
</tr>
<tr>
<td><strong>Absolute Risk Reduction (ARR)</strong> = risk of the outcome in the control group - risk of the outcome in the treatment group. This is also known as the absolute risk difference.</td>
<td>The absolute risk reduction tells us the absolute difference in the rates of events between the two groups and gives an indication of the baseline risk and treatment effect. An ARR of 0 means that there is no difference between the two groups - therefore the treatment had no effect. The absolute benefit of treatment is a 5% reduction in the death rate.</td>
</tr>
</tbody>
</table>

In our example the RR = 0.10 / 0.15 = 0.67

In our example the ARR = 0.15 - 0.10 = 0.05 or 5%
Relative Risk Reduction (RRR) = absolute risk reduction / risk of the outcome in the control group. An alternative way to calculate the RRR is to subtract the RR from 1 (e.g. RRR = 1 - RR)

In our example the RRR = 0.05/0.15 = 0.33 or 33%
Or
RRR = 1 - 0.67 = 0.33 or 33%

Number Needed to Treat (NNT) = inverse of the ARR and is calculated as 1 / ARR.

In our example the NNT = 1/ 0.05 = 20

The relative risk reduction is the complement of the RR and is probably the most commonly reported measure of treatment effects. It tells us the reduction in the rate of the outcome in the treatment group relative to that in the control group.

The treatment reduced the risk of death by 33% relative to that occurring in the control group.

The number needed to treat represents the number of patients we need to treat with the experimental therapy in order to prevent 1 bad outcome and incorporates the duration of treatment. Clinical significance can be determined to some extent by looking at the NNTs, but also by weighing the NNTs against any harms or adverse effects (NNHs) of therapy.

We would need to treat 20 people for 2 years in order to prevent 1 death.

2. How precise was the estimate of the treatment effect?

The true risk of the outcome in the population is not known and the best we can do is estimate the true risk based on the sample of patients in the trial. This estimate is called the point estimate. We can gauge how close this estimate is to the true value by looking at the confidence intervals (CI) for each estimate. If the confidence interval is fairly narrow then we can be confident that our point estimate is a precise reflection of the population value. The confidence interval also provides us with information about the statistical significance of the result. If the value corresponding to no effect falls outside the 95% confidence interval then the result is statistically significant at the 0.05 level. If the confidence interval includes the value corresponding to no effect then the results are not statistically significant.

Will the results help me in caring for my patient? (ExternalValidity/Applicability)

The questions that you should ask before you decide to apply the results of the study to your patient are:
• Is my patient so different to those in the study that the results cannot apply?
• Is the treatment feasible in my setting?
• Will the potential benefits of treatment outweigh the potential harms of treatment for my patient?
Was a qualitative approach appropriate?

<table>
<thead>
<tr>
<th>What should I look for?</th>
<th>Where do I find the information?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is the question being asked seeking to further understanding of people's views, opinions and/or experiences in relation to a specific setting/scenario/circumstance?</td>
<td>The Title, Abstract and Introduction/Background should tell you whether a qualitative approach was appropriate for the question being asked.</td>
</tr>
</tbody>
</table>

In this paper

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
<th>Unclear</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comment:

Was the sampling strategy appropriate for the approach?

<table>
<thead>
<tr>
<th>What should I look for?</th>
<th>Where do I find the information?</th>
</tr>
</thead>
<tbody>
<tr>
<td>How were the participants/setting(s) selected? Does the sample include a range of experiences (maximum variation sample), where all relevant ‘variables’ are accounted for, e.g. gender, age, geographical location, severity of condition, social support, socio-economic background, access to services, ethnicity? A convenience sample is seldom a good sampling choice.</td>
<td>The Method should tell you how patients were recruited and selected.</td>
</tr>
</tbody>
</table>

In this paper

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
<th>Unclear</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comment:
What were the data collection methods?

What should I look for?
Are data collection methods described in sufficient detail to allow you to repeat the study? Are they transparent and appropriate? E.g. Interviews are useful to explore individual experience(s); Focus groups are useful to explore views of a particular group or elicit information that is generated during group discussions.

Where do I find the information?
Look in the Methods section for data collection information, including interview guides and field notes.

In this paper

Yes  No  Unclear

Comment:

How were data analysed and how were these checked?

What should I look for?
Was the data analysis approach appropriate for the methodology used? E.g. A grounded theory study needs to include constant comparison. Are the analytical steps explained in detail (are they transparent)? Are the steps to ensure ‘quality control’ described? E.g. Double coding, research team discussion of identified item, respondent validation.

Where do I find the information?
The Methods section should provide sufficient information about how data were analysed.

In this paper

Yes  No  Unclear

Comment:

Is the researcher’s position described?

What should I look for?
It is ideal that the researcher(s) clearly state their position in relation to the research question. For example - their background, gender, and existing knowledge or personal experience of the topic to be researched.

Where do I find the information?
Look in the Methods/Results/Discussion section(s) to see if there is some mention of the researcher’s position as part of the research process.
In this paper

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
<th>Unclear</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comment:

**What were the results?**

**Do the results make sense?**

**What should I look for?**

Do the results answer the question, do they make sense and are they credible? (Credibility). Are the themes/theoretical concepts presented credible and do they relate to the research question?

**What does it mean?**

Look in the findings/results section: Have the authors provided a range of data (quotes) to support their interpretation (themes/theoretical concepts) of data? Are the quotes indexed so they could be traced back to the original data set? For example: patient/participant #2.

Have authors provided ‘negative cases’ i.e. narratives that do not fit the identified themes/theoretical framework. For example where some participants’ experiences differ from the main findings (think outliers!)

Have the authors provided context (background to participant) for quotes in order to interpret meaning? This should be relevant to the findings discussed, for example age and gender/length or severity of condition, socio economic background, educational background, etc.

In this paper

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
<th>Unclear</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comment:
Are the conclusions drawn justified by the results?

What should I look for?
- How well does the analysis explain why people behave in the way they do?
- How comprehensible would this explanation be to a thoughtful participant from the setting (can participants/patients ‘see’ themselves in the interpretation of data)?
- How well does the explanation fit with what we know already and if not why not?

Where do I find the information?
- Look in the Discussion/Conclusion sections of the paper (although some predominately qualitative journals merge findings and discussion).
- Check whether the authors draw on examples of data when providing explanations.
- Look for references to previous research in this area and existing theory and whether these are discussed in relation to findings and explanations offered by authors.
- Does the paper offer a ‘so what’ recommendation?

In this paper

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
<th>Unclear</th>
</tr>
</thead>
</table>

Comment:

Are the finding transferable to other clinical settings?

What should I look for?
This may not be applicable to all studies using qualitative methods (e.g. exploratory, pilot studies). However, research using maximum variation sampling and particularly theoretical sampling needs to demonstrate that the findings are transferable to other settings. E.g. A study aims to explore experiences of breathlessness in COPD and a true theoretical/maximum variation sample has been recruited then the findings are transferrable to other clinical settings with a similar context, E.g. includes a range of illness experiences, age, gender, socio-economic background, illness severity. However if the sample includes only white, middle class men in their 50’s, then this is not maximum variation sampling and cannot be transferred to other settings.

Where do I find the information?
- Check the sampling information in the Methods section. Then compare the sampling strategy mentioned with the actual participant sample recruited in the Findings section. Did the authors recruit the sample they set out to recruit?
- In the Discussion/Conclusion section check whether the authors discuss the transferability of the findings. If not check if the authors have outlined whether the findings are limited to a particular context as part of the limitations of the study.
- True theoretical sampling as described in Grounded Theory Methodology is guided by emerging themes during constant comparative analysis. This is particular to this methodology so does not apply to all other qualitative methodologies. If this methodology is used, steps to illustrate how theoretical sampling has been followed in the research process should be described throughout the Methods section.
In this paper

Yes No Unclear

Comment:

Reference

Critical Appraisal Skills Program is another source that provides critical appraisal forms.

- These forms are provided in pages 20-63 of this workbook.

- Checklists are also available in an electronic, editable format online by vising

https://casp-uk.net/casp-tools-checklists/

Checklists Available include:

- Systematic Reviews p. 20-23
- Randomized Controlled Trials pp. 24-28
- Cohort Studies pp. 29-35
- Case Control Studies pp. 36-41
- Economic Evaluations pp. 42-47
- Diagnostic Studies pp. 48-52
- Qualitative studies pp. 53-58
- Clinical Prediction Rule pp. 59-63

The following documents are free to copy, distribute, and adapt through the Creative Commons License.
CASP Checklist: 10 questions to help you make sense of a Systematic Review

How to use this appraisal tool: Three broad issues need to be considered when appraising a systematic review study:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 10 questions on the following pages are designed to help you think about these issues systematically. The first two questions are screening questions and can be answered quickly. If the answer to both is “yes”, it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a “yes”, “no” or “can’t tell” to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA ‘Users’ guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.


©CASP this work is licensed under the Creative Commons Attribution – Non-Commercial-Share A like. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ www.casp-uk.net
### Section A: Are the results of the review valid?

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>Can’t Tell</th>
<th>No</th>
<th>HINT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Did the review address a clearly focused question?</td>
<td></td>
<td></td>
<td></td>
<td>HINT: An issue can be ‘focused’ in terms of</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• the population studied</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• the intervention given</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• the outcome considered</td>
</tr>
<tr>
<td>Comments:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Did the authors look for the right type of papers?</td>
<td></td>
<td></td>
<td></td>
<td>HINT: ‘The best sort of studies’ would</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• address the review’s question</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• have an appropriate study design</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(usually RCTs for papers evaluating interventions)</td>
</tr>
<tr>
<td>Comments:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Do you think all the important, relevant studies were included?</td>
<td></td>
<td></td>
<td></td>
<td>HINT: Look for</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• which bibliographic databases were used</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• follow up from reference lists</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• personal contact with experts</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• unpublished as well as published studies</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• non-English language studies</td>
</tr>
<tr>
<td>Comments:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4. Did the review’s authors do enough to assess quality of the included studies?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Can’t Tell</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**HINT:** The authors need to consider the rigour of the studies they have identified. Lack of rigour may affect the studies’ results (“All that glisters is not gold” Merchant of Venice – Act II Scene 7)

Comments:

5. If the results of the review have been combined, was it reasonable to do so?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Can’t Tell</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**HINT:** Consider whether
- results were similar from study to study
- results of all the included studies are clearly displayed
- results of different studies are similar
- reasons for any variations in results are discussed

Comments:

---

**Section B: What are the results?**

6. What are the overall results of the review?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Can’t Tell</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**HINT:** Consider
- If you are clear about the review’s ‘bottom line’ results
- what these are (numerically if appropriate)
- how were the results expressed (NNT, odds ratio etc.)

Comments:
7. How precise are the results?
   HINT: Look at the confidence intervals, if given

   Comments:

---

Section C: Will the results help locally?

8. Can the results be applied to the local population?
   Yes
   Can’t Tell
   No
   HINT: Consider whether
   • the patients covered by the review could be sufficiently different to your population to cause concern
   • your local setting is likely to differ much from that of the review

   Comments:

9. Were all important outcomes considered?
   Yes
   Can’t Tell
   No
   HINT: Consider whether
   • there is other information you would like to have seen

   Comments:

10. Are the benefits worth the harms and costs?
    Yes
    Can’t Tell
    No
    HINT: Consider
    • even if this is not addressed by the review, what do you think?

   Comments:
CASP Checklist: 11 questions to help you make sense of a Randomised Controlled Trial

How to use this appraisal tool: Three broad issues need to be considered when appraising a trial:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 11 questions on the following pages are designed to help you think about these issues systematically. The first three questions are screening questions and can be answered quickly. If the answer to both is “yes”, it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a “yes”, “no” or “can’t tell” to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA 'Users’ guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.

Referencing: we recommend using the Harvard style citation, i.e.: Critical Appraisal Skills Programme (2018). CASP (insert name of checklist i.e. Randomised Controlled Trial) Checklist. [online] Available at: URL. Accessed: Date Accessed.

©CASP this work is licensed under the Creative Commons Attribution – Non-Commercial-Share A like. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ www.casp-uk.net
### Section A: Are the results of the trial valid?

1. **Did the trial address a clearly focused issue?**
   - **Yes**
   - **Can't Tell**
   - **No**

   **HINT:** An issue can be ‘focused’ in terms of:
   - the population studied
   - the intervention given
   - the comparator given
   - the outcomes considered

   **Comments:**

2. **Was the assignment of patients to treatments randomised?**
   - **Yes**
   - **Can’t Tell**
   - **No**

   **HINT:** Consider:
   - how this was carried out
   - was the allocation sequence concealed from researchers and patients

   **Comments:**

3. **Were all of the patients who entered the trial properly accounted for at its conclusion?**
   - **Yes**
   - **Can’t Tell**
   - **No**

   **HINT:** Consider:
   - was the trial stopped early
   - were patients analysed in the groups to which they were randomised

   **Comments:**

---

**Is it worth continuing?**
4. Were patients, health workers and study personnel 'blind' to treatment?

   Yes [ ]
   Can't Tell [ ]
   No [ ]

Comments:


5. Were the groups similar at the start of the trial

   Yes [ ]
   Can't Tell [ ]
   No [ ]

HINT: Consider
- other factors that might affect the outcome, such as; age, sex, social class

Comments:


6. Aside from the experimental intervention, were the groups treated equally?

   Yes [ ]
   Can't Tell [ ]
   No [ ]

Comments:

Section B: What are the results?
7. How large was the treatment effect?  
HINT: Consider  
• what outcomes were measured  
• Is the primary outcome clearly specified  
• what results were found for each outcome  

Comments:  

8. How precise was the estimate of the treatment effect?  
HINT: Consider  
• what are the confidence limits  

Comments:  

Section C: Will the results help locally?  

9. Can the results be applied to the local population, or in your context?  
Yes  
Can’t Tell  
No  
HINT: Consider whether  
• the patients covered by the trial are similar enough to the patients to whom you will apply this  
• how they differ  

Comments:  

10. Were all clinically important outcomes considered?  
Yes  
Can’t Tell  
No  
HINT: Consider whether  
• there is other information you would like to have seen  
• if not, does this affect the decision  

Comments:
11. Are the benefits worth the harms and costs?

Yes

Can’t Tell

No

HINT: Consider

• even if this is not addressed by the trial, what do you think?

Comments:
CASP Checklist: 12 questions to help you make sense of a Cohort Study

How to use this appraisal tool: Three broad issues need to be considered when appraising a cohort study:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 12 questions on the following pages are designed to help you think about these issues systematically. The first two questions are screening questions and can be answered quickly. If the answer to both is “yes”, it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a “yes”, “no” or “can’t tell” to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA ‘Users’ guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners. For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.

Referencing: we recommend using the Harvard style citation, i.e.: Critical Appraisal Skills Programme (2018). CASP (insert name of checklist i.e. Cohort Study) Checklist. [online] Available at: URL. Accessed: Date Accessed.

©CASP this work is licensed under the Creative Commons Attribution – Non-Commercial-Share A like. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ www.casp-uk.net
Section A: Are the results of the study valid?

1. Did the study address a clearly focused issue?
   - Yes
   - Can’t Tell
   - No

   **HINT:** A question can be ‘focused’ in terms of:
   - the population studied
   - the risk factors studied
   - is it clear whether the study tried to detect a beneficial or harmful effect
   - the outcomes considered

   **Comments:**

2. Was the cohort recruited in an acceptable way?
   - Yes
   - Can’t Tell
   - No

   **HINT:** Look for selection bias which might compromise the generalisability of the findings:
   - was the cohort representative of a defined population
   - was there something special about the cohort
   - was everybody included who should have been

   **Comments:**

Is it worth continuing?
3. Was the exposure accurately measured to minimise bias?

- Yes
- Can’t Tell
- No

HINT: Look for measurement or classification bias:
- did they use subjective or objective measurements
- do the measurements truly reflect what you want them to (have they been validated)
- were all the subjects classified into exposure groups using the same procedure

Comments:

4. Was the outcome accurately measured to minimise bias?

- Yes
- Can’t Tell
- No

HINT: Look for measurement or classification bias:
- did they use subjective or objective measurements
- do the measurements truly reflect what you want them to (have they been validated)
- has a reliable system been established for detecting all the cases (for measuring disease occurrence)
- were the measurement methods similar in the different groups
- were the subjects and/or the outcome assessor blinded to exposure (does this matter)

Comments:
5. (a) Have the authors identified all important confounding factors?

<table>
<thead>
<tr>
<th>Yes</th>
<th>Can’t Tell</th>
<th>No</th>
</tr>
</thead>
</table>

**HINT:**
- list the ones you think might be important, and ones the author missed

Comments:

5. (b) Have they taken account of the confounding factors in the design and/or analysis?

<table>
<thead>
<tr>
<th>Yes</th>
<th>Can’t Tell</th>
<th>No</th>
</tr>
</thead>
</table>

**HINT:**
- look for restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors

Comments:

6. (a) Was the follow up of subjects complete enough?

<table>
<thead>
<tr>
<th>Yes</th>
<th>Can’t Tell</th>
<th>No</th>
</tr>
</thead>
</table>

**HINT:** Consider
- the good or bad effects should have had long enough to reveal themselves
- the persons that are lost to follow-up may have different outcomes than those available for assessment
- in an open or dynamic cohort, was there anything special about the outcome of the people leaving, or the exposure of the people entering the cohort

6. (b) Was the follow up of subjects long enough?

<table>
<thead>
<tr>
<th>Yes</th>
<th>Can’t Tell</th>
<th>No</th>
</tr>
</thead>
</table>

Section B: What are the results?

7. What are the results of this study?

HINT: Consider
- what are the bottom line results
- have they reported the rate or the proportion between the exposed/unexposed, the ratio/rate difference
- how strong is the association between exposure and outcome (RR)
- what is the absolute risk reduction (ARR)

Comments:

8. How precise are the results?

HINT:
- look for the range of the confidence intervals, if given

Comments:
9. Do you believe the results?

|   | Yes |  
|---|-----|---|
|   | Can’t Tell |  
|   | No |  

**HINT:** Consider  
- big effect is hard to ignore  
- can it be due to bias, chance or confounding  
- are the design and methods of this study sufficiently flawed to make the results unreliable  
- Bradford Hills criteria (e.g. time sequence, dose-response gradient, biological plausibility, consistency)

**Comments:**

---

**Section C: Will the results help locally?**

10. Can the results be applied to the local population?

|   | Yes |  
|---|-----|---|
|   | Can’t Tell |  
|   | No |  

**HINT:** Consider whether  
- a cohort study was the appropriate method to answer this question  
- the subjects covered in this study could be sufficiently different from your population to cause concern  
- your local setting is likely to differ much from that of the study  
- you can quantify the local benefits and harms

**Comments:**

---

11. Do the results of this study fit with other available evidence?

|   | Yes |  
|---|-----|---|
|   | Can’t Tell |  
|   | No |  

**Comments:**
12. What are the implications of this study for practice?

Yes

Can’t Tell

No

HINT: Consider
• One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making
• For certain questions, observational studies provide the only evidence
• Recommendations from observational studies are always stronger when supported by other evidence

Comments:
CASP Checklist: 11 questions to help you make sense of a Case Control Study

How to use this appraisal tool: Three broad issues need to be considered when appraising a case control study:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 11 questions on the following pages are designed to help you think about these issues systematically. The first three questions are screening questions and can be answered quickly. If the answer to both is “yes”, it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a “yes”, “no” or “can’t tell” to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA ‘Users’ guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.


©CASP this work is licensed under the Creative Commons Attribution – Non-Commercial-Share A Like. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ www.casp-uk.net
Section A: Are the results of the trial valid?

1. Did the study address a clearly focused issue?

   - Yes
   - Can’t Tell
   - No

   HINT: An issue can be ‘focused’ in terms of:
   - the population studied
   - Whether the study tried to detect a beneficial or harmful effect
   - the risk factors studied

   Comments:

2. Did the authors use an appropriate method to answer their question?

   - Yes
   - Can’t Tell
   - No

   HINT: Consider:
   - Is a case control study an appropriate way of answering the question under the circumstances
   - Did it address the study question

   Comments:
### Is it worth continuing?

<table>
<thead>
<tr>
<th>3. Were the cases recruited in an acceptable way?</th>
<th>Yes</th>
<th>Can’t Tell</th>
<th>No</th>
</tr>
</thead>
</table>

**HINT:** We are looking for selection bias which might compromise validity of the findings
- are the cases defined precisely
- were the cases representative of a defined population (geographically and/or temporally)
- was there an established reliable system for selecting all the cases
- are they incident or prevalent
- is there something special about the cases
- is the time frame of the study relevant to disease/exposure
- was there a sufficient number of cases selected
- was there a power calculation

**Comments:**

<table>
<thead>
<tr>
<th>4. Were the controls selected in an acceptable way?</th>
<th>Yes</th>
<th>Can’t Tell</th>
<th>No</th>
</tr>
</thead>
</table>

**HINT:** We are looking for selection bias which might compromise the generalisability of the findings
- were the controls representative of the defined population (geographically and/or temporally)
- was there something special about the controls
- was the non-response high, could non-respondents be different in any way
- are they matched, population based or randomly selected
- was there a sufficient number of
5. Was the exposure accurately measured to minimise bias?

- Yes
- Can’t Tell
- No

Comments:

HINT: We are looking for measurement, recall or classification bias
- was the exposure clearly defined and accurately measured
- did the authors use subjective or objective measurements
  - do the measures truly reflect what they are supposed to measure (have they been validated)
- were the measurement methods similar in the cases and controls
- did the study incorporate blinding where feasible
  - is the temporal relation correct (does the exposure of interest precede the outcome)

6. (a) Aside from the experimental intervention, were the groups treated equally?

HINT: List the ones you think might be important, that the author may have missed
- genetic
- environmental
- socio-economic

List:

6. (b) Have the authors taken account of the potential confounding factors in the design and/or in their analysis?

- Yes
- Can’t Tell
- No

Comments:

HINT: Look for
- restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors
Section B: What are the results?

7. How large was the treatment effect?

HINT: Consider
- what are the bottom line results
- is the analysis appropriate to the design
- how strong is the association between exposure and outcome (look at the odds ratio)
- are the results adjusted for confounding, and might confounding still explain the association
- has adjustment made a big difference to the OR

Comments:

8. How precise was the estimate of the treatment effect?

HINT: Consider
- size of the p-value
- size of the confidence intervals
- have the authors considered all the important variables
- how was the effect of subjects refusing to participate evaluated

Comments:
9. Do you believe the results?

Yes

No

HINT: Consider

• big effect is hard to ignore!
• Can it be due to chance, bias, or confounding
• are the design and methods of this study sufficiently flawed to make the results unreliable
• consider Bradford Hills criteria (e.g. time sequence, does-response gradient, strength, biological plausibility)

Comments:

Section C: Will the results help locally?

10. Can the results be applied to the local population?

Yes

Can’t Tell

No

HINT: Consider whether

• the subjects covered in the study could be sufficiently different from your population to cause concern
• your local setting is likely to differ much from that of the study
• can you quantify the local benefits and harms

Comments:

11. Do the results of this study fit with other available evidence?

Yes

Can’t Tell

No

HINT: Consider

• all the available evidence from RCT’s Systematic Reviews, Cohort Studies, and Case Control Studies as well, for consistency

Comments:

Remember One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making. However, for certain questions observational studies provide the only evidence. Recommendations from observational studies are always stronger when supported by other evidence.
CASP Checklist: 12 questions to help you make sense of an Economic Evaluation

How to use this appraisal tool: Three broad issues need to be considered when appraising an economic evaluation study:

- Is the economic evaluation valid? (Section A)
- How were consequences and costs assessed and compared? (Section B)
- Will the results help in purchasing for local people? (Section C)

The 12 questions on the following pages are designed to help you think about these issues systematically. The first two questions are screening questions and can be answered quickly. If the answer to both is “yes”, it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a “yes”, “no” or “can’t tell” to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA 'Users' guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.


©CASP this work is licensed under the Creative Commons Attribution – Non-Commercial-Share A like. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ www.casp-uk.net
### Section A: Is the economic evaluation valid?

1. **Was a well-defined question posed?**
   - **Yes**
   - **Can’t Tell**
   - **No**

   **HINT:** Is it clear what the authors are trying to achieve?
   - What is the perspective?
   - How many options are compared?
   - Are both costs and consequences considered?
   - What is the time horizon?

2. **Was a comprehensive description of the competing alternatives given?**
   - **Yes**
   - **Can’t Tell**
   - **No**

   **HINT:** Is there a clear decision tree (or similar given):
   - Can you tell who did what, to whom, where and how often?

---

**Is it worth continuing?**
3. Does the paper provide evidence that the programme would be effective? (i.e. would the programme do more good than harm?)

Yes
Can’t Tell
No

HINT: Consider:
• if an RCT or systematic review was used;
  if not, consider how strong the evidence was (economic evaluations frequently have to integrate different types of knowledge stemming from different study designs)

Comments:

4. Were the effects of the intervention identified, measured and valued appropriately?

Yes
Can’t Tell
No

HINT: Effects can be measured in natural units (e.g. years of life) or more complex units (e.g. years adjusted for quality of life such as QALYS) or monetary equivalents of the benefit gained (e.g. $)

Comments:

Section B: How were consequences and costs assessed and compared?

5. Were all important and relevant resources required, and health outcome costs for each alternative identified, measured in appropriate units and valued credibly?

Yes
Can’t Tell
No

HINT: Identified?
• remember the perspective being taken
  HINT: measured accurately?
• appropriate units may be hours of nursing time, number of physician visits, years-of-life gained etc.
  HINT: valued credibly?
• are the values realistic
• how have they been derived
• have opportunity costs been considered
6. Were costs and consequences adjusted for different times at which they occurred (discounting)?

Yes  
Can’t Tell  
No

Comments:

7. What were the results of the evaluation?

Yes  
Can’t Tell  
No

HINT: Consider:  
- what is the bottom line  
- what units were used (e.g. cost/life year gained, cost/QALY, net benefit)

Comments:

8. Was an incremental analysis of the consequences and cost of alternatives performed?

Yes  
Can’t Tell  
No

Comments:
9. Was an adequate sensitivity analysis performed?

- Yes
- Can’t Tell
- No

**HINT:** Consider
- if all the main areas of uncertainty were considered by changing the estimate of the variable *and*
- looking at how this would change the result of the economic evaluation

**Comments:**

---

**Section C: Will the results help in purchasing for local people?**

10. Is the programme likely to be equally effective in your context or setting?

- Yes
- Can’t Tell
- No

**HINT:** Consider whether
- the patients covered by the review could be sufficiently different to your population to cause concern
- your local setting is likely to differ much from that of the review

**Comments:**

---

11. Are the costs translatable to your setting?

- Yes
- Can’t Tell
- No

**Comments:**
12. Is it worth doing in your setting?

- Yes
- Can’t Tell
- No

Comments:
CASP Checklist: 12 questions to help you make sense of a Diagnostic Test study

How to use this appraisal tool: Three broad issues need to be considered when appraising a trial:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 12 questions on the following pages are designed to help you think about these issues systematically. The first three questions are screening questions and can be answered quickly. If the answer to both is “yes”, it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a “yes”, “no” or “can’t tell” to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA ‘Users’ guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.

Referencing: we recommend using the Harvard style citation, i.e.: Critical Appraisal Skills Programme (2018). CASP (insert name of checklist i.e. Diagnostic Test Study) Checklist. [online] Available at: URL. Accessed: Date Accessed.

©CASP this work is licensed under the Creative Commons Attribution – Non-Commercial-Share A like. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ www.casp-uk.net
Section A: Are the results of the trial valid?

1. Was there a clear question for the study to address?  
   - Yes  
   - Can’t Tell  
   - No  
   
   HINT: A question should include information about  
   - the population  
   - the test  
   - the setting  
   - the outcomes

Comments:

2. Was there a comparison with an appropriate reference standard?  
   - Yes  
   - Can’t Tell  
   - No  
   
   HINT: Is this reference test(s) the best available indicator in the circumstances

Comments:

Is it worth continuing?

3. Did all patients get the diagnostic test and reference standard?  
   - Yes  
   - Can’t Tell  
   - No  
   
   HINT: Consider  
   - were both received regardless of the results of the test of interest  
   - Check the 2x2 table (verification bias)

Comments:
4. Could the results of the test have been influenced by the results of the reference standard?

Yes
Can’t Tell
No

HINT: Consider
• was there blinding
• were the tests performed independently
• review bias

Comments:

5. Is the disease status of the tested population clearly described?

Yes
Can’t Tell
No

HINT: Consider
• presenting symptoms
• disease stage of severity
• co-morbidity
• differential diagnoses (spectrum bias)

Comments:

6. Were the methods for performing the test described in sufficient detail?

Yes
Can’t Tell
No

HINT: Consider
• was a protocol followed

Comments:

Section B: What are the results?
7. What are the results?

HINT: Consider
- are the sensitivity and specificity and/or likelihood ratios presented
- are the results presented in such a way that we can work them out

Comments:

8. How sure are we about the results?

HINT: Consider
- could they have occurred by chance
- are there confidence limits
- what are they

Comments:

Section C: Will the results help locally?

Consider whether you are primarily interested in the impact on a population or individual level

9. Can the results be applied to your patients/the population of interest?

Yes
Can’t Tell
No

HINT: Do you think your patients/population are so different from those in the study that the results cannot be applied, such as age, sex, ethnicity and spectrum bias

Comments:

10. Can the test be applied to your patient or population of interest?

Yes
Can’t Tell
No

HINT: Consider
- resources and opportunity costs
- level and availability of expertise required to interpret the tests
- current practice and availability of services
11. Were all outcomes important to the individual or population considered?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Can’t Tell</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>HINT: Consider</td>
<td>• will the knowledge of the test result improve patient wellbeing</td>
<td>• will the knowledge of the test result lead to a change in patient management</td>
<td></td>
</tr>
</tbody>
</table>

12. What would be the impact of using this test on your patients/population?

Comments:
CASP Checklist: 10 questions to help you make sense of a Qualitative research

How to use this appraisal tool: Three broad issues need to be considered when appraising a qualitative study:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 10 questions on the following pages are designed to help you think about these issues systematically. The first two questions are screening questions and can be answered quickly. If the answer to both is “yes”, it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a “yes”, “no” or “can’t tell” to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA ‘Users’ guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.

Referencing: we recommend using the Harvard style citation, i.e.: Critical Appraisal Skills Programme (2018). CASP (insert name of checklist i.e. Qualitative) Checklist. [online] Available at: URL. Accessed: Date Accessed.

©CASP this work is licensed under the Creative Commons Attribution – Non-Commercial-Share A like. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ www.casp-uk.net
### Section A: Are the results valid?

1. Was there a clear statement of the aims of the research?
   - **Yes**
   - **Can’t Tell**
   - **No**

   **HINT:** Consider
   - what was the goal of the research
   - why it was thought important
   - its relevance

**Comments:**

2. Is a qualitative methodology appropriate?
   - **Yes**
   - **Can’t Tell**
   - **No**

   **HINT:** Consider
   - If the research seeks to interpret or illuminate the actions and/or subjective experiences of research participants
   - Is qualitative research the right methodology for addressing the research goal

**Comments:**

### Is it worth continuing?

3. Was the research design appropriate to address the aims of the research?
   - **Yes**
   - **Can’t Tell**
   - **No**

   **HINT:** Consider
   - if the researcher has justified the research design (e.g., have they discussed how they decided which method to use)

**Comments:**
4. Was the recruitment strategy appropriate to the aims of the research?  

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Can’t Tell</th>
<th>No</th>
</tr>
</thead>
</table>

HINT: Consider
- If the researcher has explained how the participants were selected
- If they explained why the participants they selected were the most appropriate to provide access to the type of knowledge sought by the study
- If there are any discussions around recruitment (e.g. why some people chose not to take part)

Comments:

5. Was the data collected in a way that addressed the research issue?  

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Can’t Tell</th>
<th>No</th>
</tr>
</thead>
</table>

HINT: Consider
- If the setting for the data collection was justified
- If it is clear how data were collected (e.g. focus group, semi-structured interview etc.)
- If the researcher has justified the methods chosen
- If the researcher has made the methods explicit (e.g. for interview method, is there an indication of how interviews are conducted, or did they use a topic guide)
- If methods were modified during the study. If so, has the researcher explained how and why
- If the form of data is clear (e.g. tape recordings, video material, notes etc.)
- If the researcher has discussed saturation of data

Comments:
6. Has the relationship between researcher and participants been adequately considered? Yes  Can’t Tell  No

HINT: Consider
• If the researcher critically examined their own role, potential bias and influence during (a) formulation of the research questions (b) data collection, including sample recruitment and choice of location
• How the researcher responded to events during the study and whether they considered the implications of any changes in the research design

Comments:

Section B: What are the results?

7. Have ethical issues been taken into consideration? Yes  Can’t Tell  No

HINT: Consider
• If there are sufficient details of how the research was explained to participants for the reader to assess whether ethical standards were maintained
• If the researcher has discussed issues informed consent or confidentiality or how they have handled the effects of the study on the participants during and after the study
• If approval has been sought from the ethics committee

Comments:
8. Was the data analysis sufficiently rigorous?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Can’t Tell</th>
<th>No</th>
</tr>
</thead>
</table>

HINT: Consider
- If there is an in-depth description of the analysis process
- If thematic analysis is used. If so, is it clear how the categories/themes were derived from the data
- Whether the researcher explains how the data presented were selected from the original sample to demonstrate the analysis process
- If sufficient data are presented to support the findings
- To what extent contradictory data are taken into account
- Whether the researcher critically examined their own role, potential bias and influence during analysis and selection of data for presentation

Comments:

9. Is there a clear statement of findings?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Can’t Tell</th>
<th>No</th>
</tr>
</thead>
</table>

HINT: Consider whether
- If the findings are explicit
- If there is adequate discussion of the evidence both for and against the researcher’s arguments
- If the researcher has discussed the credibility of their findings (e.g. triangulation, respondent validation, more than one analyst)
- If the findings are discussed in relation to the original research question

Comments:
Section C: Will the results help locally?

10. How valuable is the research?

HINT: Consider

- If the researcher discusses the contribution the study makes to existing knowledge or understanding (e.g., do they consider the findings in relation to current practice or policy, or relevant research-based literature)?
- If they identify new areas where research is necessary.
- If the researchers have discussed whether or how the findings can be transferred to other populations or considered other ways the research may be useful.

Comments:
CASP Checklist: 11 questions to help you evaluate a clinical prediction rule

How to use this appraisal tool: Three broad issues need to be considered when appraising a clinical prediction rule study:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 10 questions on the following pages are designed to help you think about these issues systematically. The first two questions are screening questions and can be answered quickly. If the answer to both is “yes”, it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a “yes”, “no” or “can’t tell” to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA ‘Users’ guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

This material has been developed by CASP España (CASPe) http://redcaspe.org it was translated into English and tested by the Critical Appraisal Skills Programme, Oxford, UK (CASP)


©CASP this work is licensed under the Creative Commons Attribution – Non-Commercial-Share A like. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/www.casp-uk.net
Section A: Are the results of the study valid?

1. Is the CPR clearly defined?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Can’t Tell</th>
<th>No</th>
</tr>
</thead>
</table>

HINT: • is the type of patients to whom the CPR will be applied clearly defined  
      • are the variables included in the rule clearly defined  
      • is the outcome relevant and is it clinically reasonable (the outcome can be expressed as a probability or as a course of action)

**Comments:**

2. Did the population from which the rule was derived include an appropriate spectrum of patients?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Can’t Tell</th>
<th>No</th>
</tr>
</thead>
</table>

HINT: Consider  
      • Is it adequate the way the patients were selected  
      • The spectrum of patient, to whom the rule will apply, is represented well

**Comments:**

3. Was the rule validated in a different group of patients?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Can’t Tell</th>
<th>No</th>
</tr>
</thead>
</table>

HINT: • it’s not good enough that the rule had a good performance on the patient group used to derive it. The rule should be validated in a different set of patients  
      • the validation was done in a group of patients similar to the one used to derive it

**Comments:**

**Is it worth continuing?**
4. Were the predictor variables and the outcome evaluated in a blinded fashion?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Can’t Tell</th>
<th>No</th>
</tr>
</thead>
</table>

HINT:
- did people evaluating the outcome know the predictor variables
- did people evaluating the predictor variables know the outcome

Comments:

5. Were the predictor variables and the outcome evaluates in the whole sample selected initially?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Can’t Tell</th>
<th>No</th>
</tr>
</thead>
</table>

HINT:
- are exclusions and drop outs well described and do the authors discuss the reasons for them
- sometimes the outcome cannot be measured in the same way in all patients

Comments:

6. Are the statistical methods used to construct and validate the rule clearly described?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Can’t Tell</th>
<th>No</th>
</tr>
</thead>
</table>

HINT:
- were all important variables included and the positivity criteria explained
- is the statistical method adequately described
- was the reliability of the rule considered

Comments:

Section B: What are the results?
7. Can the performance of the rule be calculated?

HINT:
- Performance results can be presented as: Sens, Sp, +LR, -LR, ROC curve, calibration curves etc.
- Sensitivity = \( \frac{a}{a+c} \)
- Specificity = \( \frac{d}{b+d} \)
- LR+ = \( \frac{a}{a+c} \) / \( 1 - \frac{b}{b+d} \)
- LR- = \( \frac{b}{b+d} \) / \( 1 - \frac{a}{a+c} \)

<table>
<thead>
<tr>
<th>Rule+</th>
<th>Outcome+</th>
<th>Rule-</th>
<th>Outcome-</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a</td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comments:

8. How precise was the estimate of the treatment effect?

HINT: Think about
- the sample size and the number of variables included in the CPR
- is the rule robust, has there been any attempt to refine it

Comments:

Section C: Will the results help locally? Are the findings applicable to the scenario?
9. Would the prediction rule be reliable and the results interpretable if used for your patient?  

Yes  
Can’t Tell  
No

**HINT:** Consider  
- is your setting too different from that of the study

**Comments:**

---

10. Is the rule acceptable in your case?  

Yes  
Can’t Tell  
No

**HINT:** Consider  
- the ease of use and the availability of the rule and the costs  
- if the rule is reasonable from a clinical point of view

**Comments:**

---

11. Would the results of the rule modify your decision about the management of the patient, or the information you can give to him/her?  

Yes  
Can’t Tell  
No

**HINT:** Consider  
- in addition to your opinion, might there be studies analysing the impact (in monetary terms or health results) of the rule  
- if nothing will change, the rule is at best useless in terms of benefit to the patients  
- how the initial estimation has changed after applying the rule, and the effect it has had on the action threshold

**Comments:**
The PEDro Scale is another resource that can be used to help with critically appraising evidence.

Additional tools provided by PEDro can be assessed by visiting the PEDro website: https://www.pedro.org.au/

The scale is provided on the following page.
### PEDro scale

| 1. eligibility criteria were specified | no □ yes  □ where: |
| 2. subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | no □ yes  □ where: |
| 3. allocation was concealed | no □ yes  □ where: |
| 4. the groups were similar at baseline regarding the most important prognostic indicators | no □ yes  □ where: |
| 5. there was blinding of all subjects | no □ yes  □ where: |
| 6. there was blinding of all therapists who administered the therapy | no □ yes  □ where: |
| 7. there was blinding of all assessors who measured at least one key outcome | no □ yes  □ where: |
| 8. measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups | no □ yes  □ where: |
| 9. all subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention to treat” | no □ yes  □ where: |
| 10. the results of between-group statistical comparisons are reported for at least one key outcome | no □ yes  □ where: |
| 11. the study provides both point measures and measures of variability for at least one key outcome | no □ yes  □ where: |

The PEDro scale is based on the Delphi list developed by Verhagen and colleagues at the Department of Epidemiology, University of Maastricht (Verhagen AP et al (1998). The Delphi list: a criteria list for quality assessment of randomised clinical trials for conducting systematic reviews developed by Delphi consensus. Journal of Clinical Epidemiology, 51(12):1235-41). The list is based on "expert consensus" not, for the most part, on empirical data. Two additional items not on the Delphi list (PEDro scale items 8 and 10) have been included in the PEDro scale. As more empirical data comes to hand it may become possible to "weight" scale items so that the PEDro score reflects the importance of individual scale items.

The purpose of the PEDro scale is to help the users of the PEDro database rapidly identify which of the known or suspected randomised clinical trials (ie RCTs or CCTs) archived on the PEDro database are likely to be internally valid (criteria 2-9), and could have sufficient statistical information to make their results interpretable (criteria 10-11). An additional criterion (criterion 1) that relates to the external validity (or “generalisability” or “applicability” of the trial) has been retained so that the Delphi list is complete, but this criterion will not be used to calculate the PEDro score reported on the PEDro web site.

The PEDro scale should not be used as a measure of the “validity” of a study’s conclusions. In particular, we caution users of the PEDro scale that studies which show significant treatment effects and which score highly on the PEDro scale do not necessarily provide evidence that the treatment is clinically useful. Additional considerations include whether the treatment effect was big enough to be clinically worthwhile, whether the positive effects of the treatment outweigh its negative effects, and the cost-effectiveness of the treatment. The scale should not be used to compare the “quality” of trials performed in different areas of therapy, primarily because it is not possible to satisfy all scale items in some areas of physiotherapy practice.
Notes on administration of the PEDro scale:

All criteria **Points are only awarded when a criterion is clearly satisfied.** If on a literal reading of the trial report it is possible that a criterion was not satisfied, a point should not be awarded for that criterion.

Criterion 1 This criterion is satisfied if the report describes the source of subjects and a list of criteria used to determine who was eligible to participate in the study.

Criterion 2 A study is considered to have used random allocation if the report states that allocation was random. The precise method of randomisation need not be specified. Procedures such as coin-tossing and dice-rolling should be considered random. Quasi-randomisation allocation procedures such as allocation by hospital record number or birth date, or alternation, do not satisfy this criterion.

Criterion 3 **Concealed allocation** means that the person who determined if a subject was eligible for inclusion in the trial was unaware, when this decision was made, of which group the subject would be allocated to. A point is awarded for this criterion, even if it is not stated that allocation was concealed, when the report states that allocation was by sealed opaque envelopes or that allocation involved contacting the holder of the allocation schedule who was “off-site”.

Criterion 4 At a minimum, in studies of therapeutic interventions, the report must describe at least one measure of the severity of the condition being treated and at least one (different) key outcome measure at baseline. The rater must be satisfied that the groups’ outcomes would not be expected to differ, on the basis of baseline differences in prognostic variables alone, by a clinically significant amount. This criterion is satisfied even if only baseline data of study completers are presented.

Criteria 4, 7-11 **Key outcomes** are those outcomes which provide the primary measure of the effectiveness (or lack of effectiveness) of the therapy. In most studies, more than one variable is used as an outcome measure.

Criterion 5-7 **Blinding** means the person in question (subject, therapist or assessor) did not know which group the subject had been allocated to. In addition, subjects and therapists are only considered to be “blind” if it could be expected that they would have been unable to distinguish between the treatments applied to different groups. In trials in which key outcomes are self-reported (eg, visual analogue scale, pain diary), the assessor is considered to be blind if the subject was blind.

Criterion 8 This criterion is only satisfied if the report explicitly states both the number of subjects initially allocated to groups and the number of subjects from whom key outcome measures were obtained. In trials in which outcomes are measured at several points in time, a key outcome must have been measured in more than 85% of subjects at one of those points in time.

Criterion 9 An intention to treat analysis means that, where subjects did not receive treatment (or the control condition) as allocated, and where measures of outcomes were available, the analysis was performed as if subjects received the treatment (or control condition) they were allocated to. This criterion is satisfied, even if there is no mention of analysis by intention to treat, if the report explicitly states that all subjects received treatment or control conditions as allocated.

Criterion 10 A between-group statistical comparison involves statistical comparison of one group with another. Depending on the design of the study, this may involve comparison of two or more treatments, or comparison of treatment with a control condition. The analysis may be a simple comparison of outcomes measured after the treatment was administered, or a comparison of the change in one group with the change in another (when a factorial analysis of variance has been used to analyse the data, the latter is often reported as a group × time interaction). The comparison may be in the form hypothesis testing (which provides a “p” value, describing the probability that the groups differed only by chance) or in the form of an estimate (for example, the mean or median difference, or a difference in proportions, or number needed to treat, or a relative risk or hazard ratio) and its confidence interval.

Criterion 11 A point measure is a measure of the size of the treatment effect. The treatment effect may be described as a difference in group outcomes, or as the outcome in (each of) all groups. Measures of variability include standard deviations, standard errors, confidence intervals, interquartile ranges (or other quantile ranges), and ranges. Point measures and/or measures of variability may be provided graphically (for example, SDs may be given as error bars in a Figure) as long as it is clear what is being graphed (for example, as long as it is clear whether error bars represent SDs or SEs). Where outcomes are categorical, this criterion is considered to have been met if the number of subjects in each category is given for each group.
Step 6: Analyzing and Synthesizing

Qualitative Study

- Check to see how similar the articles are in the areas of population, methodology, and outcome. If the studies are similar enough, then they may be pooled.

- Thematic Synthesis - A thematic synthesis involves using a systematic process to identify themes within individual studies and then find other studies where similar themes emerged.
Quantitative study

- Test the studies for homogeneity. Studies that have significant differences, especially those with differing outcomes or populations, may need to be removed due to a lack of compatibility with other studies.

- During the synthesis process try to pull out the number of participants, participant demographics, and the outcomes of each study. Extracted data may be included through table presentation.

- Meta-Analysis - A meta-analysis is a type of systematic review that results from using statistical methods to pool similar quantitative studies. You will need to calculate the effect size of each study and obtain an overall effect size with the pooled data. In some cases, studies might be weighed so that studies having a larger sample size hold more weight in determining the overall effect size (Brown, 2017). Forest plots are often included in a meta-analysis to present the data in a graphic form (Bradley & Law, 2014, p. 162).

- Odds Ratio and Relative Risk are often identified within a Meta-Analysis.
Step 7: Create a Flow Diagram

- You MUST show a flow diagram that illustrates how you ended up with the articles that you are left with. The flow diagram includes databases searched, inclusion/exclusion criteria, number of articles retrieved, number of articles excluded, number of articles excluded after screening, number of articles excluded during data extraction, and the final number of articles included.

- The PRISMA flow diagram generator can be used to create a flow chart. The generator offers 10 different chart formats and can be found at the site lister below.

- [http://prisma.thetacollaborative.ca/](http://prisma.thetacollaborative.ca/)
Records identified through database searches (PubMed, CINAHL, PsychINFO) (n = 118)

Records identified through other sources (n = 9)

Records left after duplicates removed (n = 126)

Records Screened (n = 126)

Full-text article assessed for eligibility (n = 28)

Studies included in qualitative synthesis (n = 12)

Records excluded (n = 98)

Full-text articles excluded with reasons,
  (n = 5)
  Reason #1 (n = 3)
  Reason #2 (n = 8)
  Reason #3

All studies included in qualitative synthesis (n = 12)
Example flow diagram produced through the PRISMA flow Diagram Generator

Generator can be accessed at: http://prisma.thetacollaborative.ca/
Step 8: Reporting the Results

This includes write up of the entire process. Explanations should provide enough information for the study to be replicated.

It is important that the reviewers include any potential bias that exists within the systematic review and within the original trials of studies used.

If your review is in the form of a meta-analysis, you should develop a forest plot to display the effect sizes of each study.
Data extraction and management worksheets
General Remarks

This template of a data extraction form is intended to help you to start developing your own data extraction form. It certainly **has to be adapted** to your specific question. Delete unnecessary information and include all information important for your field.

- It is advisable to use one data-extraction form for one study, so that one data-extraction form may contain the information gained from several publications on the same trial.
- If several different trials are mentioned in one publication, the data of each should be extracted in a separate data extraction form.
- Fill in **every** field as it must be obvious from the form if a certain information is missing or uninterpretable (versus forgotten to extract)
- Extract all information that you will need for further analysis (e.g. subgroup analysis) and which allow you to classify or group several studies with common features (e.g. study quality, protocol of intervention)
- Specify which information is unclear or name conflicting details in order to avoid duplication of effort
- Extraction of statistics: extract all information on variables on location and variability, standard error, confidence interval and p-values. Extract exact figures of p-values (instead of "[not] significant") and add niveau of confidence (95 or 99%)

**Abbreviations:**

<table>
<thead>
<tr>
<th>REF</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>Identification</td>
</tr>
<tr>
<td>NR</td>
<td>Not reported</td>
</tr>
<tr>
<td>IN</td>
<td>Included</td>
</tr>
<tr>
<td>EX</td>
<td>excluded</td>
</tr>
<tr>
<td>DB</td>
<td>Database</td>
</tr>
<tr>
<td>STUDY ELIGIBILITY FORM</td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td></td>
</tr>
<tr>
<td><strong>FACTORS</strong></td>
<td><strong>ASSESSMENT</strong></td>
</tr>
<tr>
<td><strong>TYPE OF STUDY</strong></td>
<td></td>
</tr>
<tr>
<td>1. Is the study described as randomized?</td>
<td>Yes</td>
</tr>
<tr>
<td>NB. Please answer “No” if the study is a crossover or quasi-randomized trial.</td>
<td>↓</td>
</tr>
<tr>
<td>Exclude</td>
<td></td>
</tr>
<tr>
<td><strong>PARTICIPANTS</strong></td>
<td></td>
</tr>
<tr>
<td>2. Were participants diagnosed as patients with disease of interest?</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>↓</td>
</tr>
<tr>
<td>Exclude</td>
<td></td>
</tr>
<tr>
<td>3. Were participants of the prespecified age?</td>
<td>Yes</td>
</tr>
<tr>
<td>NB: Please answer „Yes”, if mix age participants i.e. both &gt;18 years and &lt; 18 years are included and state it as comments. No: If only &lt; 18 years.</td>
<td>↓</td>
</tr>
<tr>
<td>Exclude</td>
<td>Subgroups available?</td>
</tr>
<tr>
<td><strong>INTERVENTIONS</strong></td>
<td></td>
</tr>
<tr>
<td>4. Were comparison groups treated with prespecified intervention in one group and control intervention in other group?</td>
<td>Yes</td>
</tr>
<tr>
<td>NB: study can have 3 arms e.g. CT arm, CT+RT (CMT) arm or RT arm, if so please cross „Yes” and state it as comments.</td>
<td>↓</td>
</tr>
<tr>
<td>Exclude</td>
<td></td>
</tr>
<tr>
<td><strong>OUTCOMES</strong></td>
<td></td>
</tr>
<tr>
<td>5. Did the study report prespecified outcomes?</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>↓</td>
</tr>
<tr>
<td>Exclude</td>
<td></td>
</tr>
<tr>
<td><strong>FINAL DECISION</strong></td>
<td></td>
</tr>
<tr>
<td>1 X “No” = EXCLUDE</td>
<td></td>
</tr>
<tr>
<td>1 X “Unclear” = UNCLEAR</td>
<td></td>
</tr>
</tbody>
</table>
## Systematic Review on “Intervention X in patients Y”

### Organisational Aspects

<table>
<thead>
<tr>
<th>REF ID</th>
<th>Reviewer, Date</th>
<th>Checked by</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Author, Year**

**Journal/Source**

**Country of origin**

**Publication type**

- Fulltext
- Abstract
- Book chapter
- Internal progress report
- Other (please specify)

**Other relevant publications in DE-form**

**Fate**

- Decision pending
- Check references
- Use for discussion
- EX without listing
- EX with listing
- Other (please specify)

**Notes / Short description**

### Reasons for Exclusion of Study from Review (Please Specify according to protocol)

**Methods**

- No RCT
- Inadequate concealment of allocation
- Other

**Patients**

- Different disease
- Stage
- Pretreatment schedule
- Age
- Subgroups available?

**Outcomes**

- No clinically relevant outcomes assessed
- No data for relevant subgroup extractable

**Other**

- Duplicate publication
- Other

NONE: Included

### Current Status: (Name of reviewer + Date)

**Question to clinician**

**Question to author**

**Status verified with study investigators or sponsors:** Yes / No

Enter name of the source (e.g. PI, sponsor, etc.)

**Contact address:**


<table>
<thead>
<tr>
<th><strong>STUDY INTERVENTION BASICS</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Disease(s)/stage(s) studied</strong></td>
</tr>
<tr>
<td><strong>Category of treatment investigated</strong></td>
</tr>
</tbody>
</table>
| First line therapy □ / Consolidation therapy □ / Salvage therapy □  
Other: |
| **Inclusion criteria** |
| **Exclusion criteria** |
| **Specials:** |
| **Experimental Intervention** |
| *If more than two, please specify/add further rows* |
| **Intervention Control** |
| **Type of control** |
| Active □ / Placebo □ / Active + placebo □ / No therapy □ |
| **Additional treatment** |
| Balanced between treatment arms? Y / N |
| **Compliance** |
| Evaluated? Y / N |
| **Planed treatment in case of failure/as long-term treatment?** |
| □ Infection related mortality  
□ Infection incidence  
□ Neutropenia incidence  
□ Neutropenia duration  
□ Treatment-related mortality  
□ Response  
□ Overall survival  
□ Event-free survival  
□ Progression-free survival  
□ Adverse events  
□ Quality of life  
□ Other (please specify) |
<p>| <strong>Outcomes assessed</strong> |
| <strong>Treatment arms comparable?</strong> |
| Significant differences between arms: |
| <strong>Subgroup evaluated</strong> |
| <em>(extractable data for these subgroups)</em> |
| <strong>Confounders</strong> |
| <em>(were confounders mentioned? A priori / a posteriori? Which? Multivariate analysis?)</em> |</p>
<table>
<thead>
<tr>
<th>TRIAL CHARACTERISTICS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sample size</strong></td>
</tr>
<tr>
<td><strong>Number of excluded patients</strong></td>
</tr>
<tr>
<td><strong>Recruitment method</strong></td>
</tr>
<tr>
<td><strong>Setting</strong></td>
</tr>
<tr>
<td><strong>Location of trial</strong></td>
</tr>
<tr>
<td><strong>Dates of Recruitment</strong></td>
</tr>
<tr>
<td><strong>Trial Design</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Length of follow-up</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Funding</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Conflict of interest statement</strong></td>
</tr>
<tr>
<td><strong>Number of groups</strong></td>
</tr>
<tr>
<td><strong>Flow diagram?</strong></td>
</tr>
<tr>
<td><strong>Method of randomisation</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Method of concealment of allocation</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Blinding</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Primary study aims</strong></td>
</tr>
<tr>
<td><strong>Secondary study aims</strong></td>
</tr>
</tbody>
</table>
**Outcomes**

- Statistically significant for primary end-point: Yes [ ] / No [ ] / Enter p-value: __________
- Statistically significant for secondary end-point: Yes [ ] / No [ ] / Enter p-value: __________

If outcome was NOT statistically significant is it because due to:

- [ ] Evidence of absence of treatment effect (true negative study) (i.e. clearly defined based on primary outcomes in the trial), or
- [ ] No evidence for absence of treatment effect (i.e. inconclusive, or low-powered study) (i.e. not clearly defined or not based on primary outcomes in the trial)
- [ ] Unclear

**Power calculation?**

- No [ ] / Yes [ ] (expected effect: _______

- Expected difference on primary outcome: Yes [ ] / No [ ] / Enter Value: _______

- Alpha (α) pre-specified: Yes [ ] / No [ ] / Enter value: _______

- Beta error (β) pre-specified: Yes [ ] / No [ ] / Enter value: _______

- Calculated sample size: Yes [ ] / No [ ] / Enter value: _______

- Sample Size achieved? Yes [ ] / No [ ]

**Statistical methods**

- ITT [ ] as treated [ ] per protocol [ ] unclear [ ]

**Analysis (+ definition)**

- Definition: [ ] available and acceptable [ ] not available

**Stopping rules**

**Drop outs stated**

- No [ ] / Yes [ ]

---

**BASELINE CHARACTERISTICS OF PATIENTS**

<table>
<thead>
<tr>
<th></th>
<th>Experimental Arm</th>
<th>Control Arm</th>
<th>Others</th>
<th>Notes: p-values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overall comment</strong></td>
<td>No significant differences</td>
<td></td>
<td></td>
<td>Reports how to transform units, all ± values = means</td>
</tr>
<tr>
<td><strong>Number of patients</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>mean/±</strong></td>
<td>±</td>
<td>±</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>median/±</strong></td>
<td>±</td>
<td>±</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>No. %</strong></td>
<td>NR</td>
<td>NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>No. %</strong></td>
<td>Male:</td>
<td>Male:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Female:</td>
<td>Female:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Definition of Diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extent of disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organ involvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Additional diagnoses in group</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staging system</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Status of patients at Rdx</td>
<td>e.g. untreated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Previous treatment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concurrent conditions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Considered as high risk patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Considered as low risk patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laboratory parameter (UNITs)</td>
<td>--</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cytogenetics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BASELINE CHARACTERIZATION OF PATIENTS (continued)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Source</strong></td>
</tr>
<tr>
<td><strong>Information</strong></td>
</tr>
<tr>
<td>Important prognostic factor A</td>
</tr>
<tr>
<td>Info 1</td>
</tr>
</tbody>
</table>
# Treatment Details according to STUDY PROTOCOL (as planned)

<table>
<thead>
<tr>
<th>Experimental Arm</th>
<th>Control Arm</th>
<th>Others</th>
<th>Notes;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary intervention (Medication, dosage, administration)</td>
<td>The form can be adapted to list expected medication and or schedules in order to reduce the amount of necessary writing.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Timing of treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration of treatment (days, cycles)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Important treatment information</td>
<td>e.g. bone marrow or peripheral blood stem cells</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment specials</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supportive treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

# Patient flow according to PUBLICATION (as it really happened)

<table>
<thead>
<tr>
<th>Experimental Arm</th>
<th>Control Arm</th>
<th>Others</th>
<th>Notes; p-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of patients screened</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of patients recruited</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of patients allocated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of patients evaluated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of patients receiving planned treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reasons for not receiving treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of drop-outs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reasons for drop-outs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outcome</td>
<td>Primary</td>
<td>Secondary</td>
<td>not defined</td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>-----------</td>
<td>-------------</td>
</tr>
</tbody>
</table>

**Definition of outcome** *(Check definitions carefully and compare to definitions of outcome you have specified in your protocol for the meta-analysis)*

**Timing of assessment**

**Statistics**

**Length of follow-up**

<table>
<thead>
<tr>
<th>No. of patients evaluated for this outcome</th>
<th>All randomised</th>
<th>Unclear</th>
<th>Less [%]</th>
</tr>
</thead>
</table>

**Reasons for drop-out**

**Reasons for exclusion**

**Source of information**

### Dichotomous data

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Time</th>
<th>Intervention group</th>
<th>Control Arm</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Observed events</td>
<td>Sample size</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Observed events</td>
<td>Sample size</td>
<td></td>
</tr>
</tbody>
</table>

Source [text, p_____] [figure No.____] [table No.____]

Expert statistical attention needed? [Y] / [N]

### Continuous data

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Time</th>
<th>Intervention group</th>
<th>Experimental Arm</th>
<th>Notes, p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Sample size</td>
<td>Mean/mean change (incl. Range)</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sample size</td>
<td>Mean/mean change (incl. Range)</td>
<td>Standard Deviation</td>
</tr>
</tbody>
</table>
## Survival probabilities

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Time</th>
<th>Patients at risk</th>
<th>Intervention group (incl. CI)</th>
<th>Patients at risk</th>
<th>Control Arm</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rate [%] of patients alive / SE / Sample size / %CI</td>
<td>Rate [%] of patients alive / SE / Sample size / %CI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PFS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Source** [text, p_____] [figure No._____] [table No._____]
### Calculation of Hazard ratio for e.g. Death (table derived from i and ii)

<table>
<thead>
<tr>
<th></th>
<th>Arm 1 (CSF)</th>
<th>Arm 2 (Control)</th>
<th>Arm 3</th>
<th>Arm 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomization ratio</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patients randomised</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patients analysed</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed Deaths</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logrank expected</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>events</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HR (CI 95% or</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>standard error or</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>variance from Cox)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logrank variance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logrank O-E</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>test statistik (&amp;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>test used, 1 or 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sided?)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advantage to control</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>or research?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kaplan-Meier curves</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>or Actuarial curves?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numbers at risk</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>reported?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follow up details</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Estimates for Death:

<table>
<thead>
<tr>
<th></th>
<th>HR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower 95% CI</td>
<td></td>
</tr>
<tr>
<td>Upper 95% CI</td>
<td></td>
</tr>
<tr>
<td>ln(HR)</td>
<td></td>
</tr>
<tr>
<td>se(ln(HR))</td>
<td></td>
</tr>
<tr>
<td>Variance</td>
<td></td>
</tr>
<tr>
<td>O-E</td>
<td></td>
</tr>
</tbody>
</table>

### Definition of death:
## METHODOLOGICAL QUALITY - OVERVIEW

<table>
<thead>
<tr>
<th>REF ID</th>
<th>Author, Year</th>
<th>Journal/Surce</th>
<th>Study ID</th>
<th>Publication type</th>
<th>Fulltext</th>
<th>Abstract</th>
<th>Other (please specify)</th>
<th>Randomization</th>
<th>Treatment allocation</th>
<th>Similarity of groups</th>
<th>Implementation of blinding</th>
<th>Transparent patient flow?</th>
<th>Completeness of trial</th>
<th>ITT (less than 15% loss)</th>
<th>Loss to follow up symmetric in both arms?</th>
<th>Different drop-out rates for different endpoints?</th>
<th>Treatment preference (see below)</th>
<th>Type of primary end-point</th>
<th>Summarized validity:</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>yes</td>
<td>unclear</td>
<td>no</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Randomization:
- **Yes**: random numbers, etc.
- **No**: patient number, day of week, etc.
- **Unclear**: method not stated

### Allocation concealment:
- **Yes**: central
- **No**: alternate, etc.
- **Unclear**: not stated

### Similarity of groups:
Were the participant characteristics at baseline similar in both groups regarding the most important prognostic factors?

### Blinding:
Was the treatment allocation masked at the outcome assessments/to data managers?

### Transparency:
Were withdrawals, drop-outs and patients lost to follow-up stated for each group? (Yes if there were no drop-outs, withdrawals etc.)

### Completeness:
If transparent, drop-out rate per study < 15%?), if asymmetric, please specify in comments

### ITT:
Did the analysis include an ITT analysis and were there less than 10% of patients excluded in each group? Comment if appropriate definition of ITT
**Treatment preference:**

1. Standard treatment highly preferred
2. Standard preferred to innovation
3. About equal, innovation a disappointment
4. About equal, innovation a success
5. Innovation preferred to standard
6. Experimental treatment highly preferred

**Type of endpoint:**

- Hard e.g. mortality, survival

---


Sample Tables
<table>
<thead>
<tr>
<th>Study</th>
<th>ED type</th>
<th>Annual visit volume</th>
<th>Sample size</th>
<th>Duration</th>
<th>Study design</th>
<th>Intervention/s</th>
<th>Main outcome measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bartlett (2002)</td>
<td>Multi-Center, Academic and Community Hospitals (17)</td>
<td>20,000-60,000</td>
<td>n/a</td>
<td>10 months</td>
<td>Time-series</td>
<td>1. Operational projects to reduce length of stay and turnaround times for supporting services. 2. Clinical projects to reduce time to treatment and introduce new care pathway rules.</td>
<td>1. ED LOS. 2. Patient satisfaction. 3. Team achievement.</td>
</tr>
<tr>
<td>Baumann (2006)</td>
<td>Single-Center, Academic Hospital – Urban</td>
<td>52,000</td>
<td>Interventions 61 Controls 182</td>
<td>1 week</td>
<td>Quasi-Experimental</td>
<td>Triage exam prior to ED admission performed by attending physician, RN &amp; technician team; compared with usual RN triage.</td>
<td>1. ED LOS. 2. Wait time before seeing a medical provider. 3. Daily LWBS rate (leaving without being seen) during week of intervention vs. 2 weeks before intervention.</td>
</tr>
<tr>
<td>Baumlin (2010)</td>
<td>Single-Center, Academic Hospital – Urban</td>
<td>76,930</td>
<td>Interventions 601 LOS; 108 xray; 29 CT scan; 271 lab Controls 508 LOS; 60 xray; 40 CT scan; 121 lab</td>
<td>48 months</td>
<td>Time-series; random sampling</td>
<td>ED throughout redesign processes including patient registration, test order entry, lab/imaging results retrievals &amp; implement Integrated ED Information System.</td>
<td>1. ED LOS. 2. Door-to-doctor time: Triage time to when the attending physician signed up for the patient. 3. Doctor-to-disposition time: First doctor-patient contact to time of disposition decision. 4. Disposition-to-discharge time: Time of disposition decision to when the patient physically left the ED. 5. Laboratory, x-ray, and CT scan TAT: Time from test order entry to when the results were available.</td>
</tr>
<tr>
<td>Bond (2001)</td>
<td>Single-Center, Community Hospital - Urban</td>
<td>68,000</td>
<td>Interventions 200 Controls 200</td>
<td>1 month pre; 1 month post</td>
<td>Time-series; random sampling</td>
<td>Triage exam prior to ED admission performed by physician &amp; RN on non-urgent patients from 8am to midnight.</td>
<td>1. ED LOS. 2. Wait time before seeing a medical provider.</td>
</tr>
</tbody>
</table>
Step 9: Take Home Points

- Discuss what it means for clinical practice and what the take home points are regarding the intervention.

- It is also important to address study limitations and suggestions for future research.
Writing it Up!
Title:

- Identify the report as a Systematic review.

- The title should reflect the content of the paper (Shamseer et al., 2015).

- If registered, provide registry name and number.

- Identify protocol authors and include the name, institutional affiliation, and e-mail address. Authors should be listed in order of contribution. A physical mailing address should be provided for the corresponding author (Shamseer et al., 2015).
Abstract:

- Summary of the review that includes enough detail so that relevance to the evidence-based question can be determined.

- The format style and number of characters allotted for this section will be determined by either the citation style that you are following, or by the journal that you are submitting to.
Introduction:

- Provide the background information on the topic of focus.

- Present the purpose of the study followed by a statement of the research question and/or hypothesis.

- Provide rationale and a strong justification for the need of the study review.
Methods:

- This section may be broken down into subsections.

- State the inclusion and exclusion criteria.

- Operational definitions of all key terms should be included within this section (Hissong, Lape, & Bailey, 2015, p.163).

- Describe the way in which studies were included for the review. Be as specific as possible and include any databases searched and key terms that were used to locate studies.

- Explain who collected the data.

- Include a description of any mechanisms used to manage records (Shamseer et al., 2015).
Methods (Cont.)

► Overall, this section should include enough detail to where a reader could replicate the study selection process and obtain the same articles.

► Try to pull out the number of participants, participant demographics, and the outcomes of each study. List and define all variables and outcomes of interest (Shamseer et al., 2015).

► State plans for assessing meta-bias (Shamseer et al., 2015).

► Determine and define how the strength of the review will be assessed (Shamseer et al., 2015).
Results

- This section should begin with the explanation of how many studies were included in the review.

- A table including each of the studies used in the review should be included. This allows for each study to be summarized and displayed openly for interpretation by the reader.

- In addition to the table, each study included in the review should be summarized and interpreted in narrative form within the results section.

- An overall statement, or remarks about the evidence, should be included after synthetization of the studies has taken place. The synthesis may include positive findings, negative findings, or both.

- Be sure that all results are presented factually without subjective interpretation.
The Cochrane Collaboration recommends that forest plots be included in systematic reviews as a way of visualizing the results of a systematic review.

**Forest Plot**

The Cochrane’s free reference manager (RevMan), can be used to help produce forest plots.

Information on how to download the RevMan program can be found at [https://community.cochrane.org/help/tools-and-software/revman-5](https://community.cochrane.org/help/tools-and-software/revman-5)
Discussion

- Restate the purpose of the study and relate the results to the original hypothesis or answer the research question.

- This section includes a summary of the results.

- Researcher may provide speculation of why certain findings were obtained.

- Provide citations from existing literature to support the results.

- Acknowledge limitations of the study.

- Clinical recommendations and implications for practice based on the results should be included.
References & Acknowledgement

References:

- Should include all studies that were used in the review.

- Be sure to format references based on the style required by the journal you are submitting to.

Acknowledgements:

- Contributors to the study, who are not authors of the study, can be mentioned here.

- Identify any conflicts of interest (if any).

- If there is a funding source, they are included in this section.
The PRISMA Checklist, provided on the following pages, can be used as a helpful outline to guide you through the process of writing up your systematic review.
<table>
<thead>
<tr>
<th>Section/topic</th>
<th>#</th>
<th>Checklist item</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TITLE</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>1</td>
<td>Identify the report as a systematic review, meta-analysis, or both.</td>
</tr>
<tr>
<td><strong>ABSTRACT</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structured summary</td>
<td>2</td>
<td>Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.</td>
</tr>
<tr>
<td><strong>INTRODUCTION</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rationale</td>
<td>3</td>
<td>Describe the rationale for the review in the context of what is already known.</td>
</tr>
<tr>
<td>Objectives</td>
<td>4</td>
<td>Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).</td>
</tr>
<tr>
<td><strong>METHODS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protocol and registration</td>
<td>5</td>
<td>Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.</td>
</tr>
<tr>
<td>Eligibility criteria</td>
<td>6</td>
<td>Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.</td>
</tr>
<tr>
<td>Information sources</td>
<td>7</td>
<td>Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.</td>
</tr>
<tr>
<td>Search</td>
<td>8</td>
<td>Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.</td>
</tr>
<tr>
<td>Study selection</td>
<td>9</td>
<td>State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).</td>
</tr>
<tr>
<td>Data collection process</td>
<td>10</td>
<td>Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.</td>
</tr>
<tr>
<td>Data items</td>
<td>11</td>
<td>List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.</td>
</tr>
<tr>
<td>Risk of bias in individual studies</td>
<td>12</td>
<td>Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.</td>
</tr>
<tr>
<td>Summary measures</td>
<td>13</td>
<td>State the principal summary measures (e.g., risk ratio, difference in means).</td>
</tr>
<tr>
<td>Synthesis of results</td>
<td>14</td>
<td>Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., $I^2$) for each meta-analysis.</td>
</tr>
<tr>
<td>Section/topic</td>
<td>#</td>
<td>Checklist item</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>---</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Risk of bias across studies</td>
<td>15</td>
<td>Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).</td>
</tr>
<tr>
<td>Additional analyses</td>
<td>16</td>
<td>Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.</td>
</tr>
<tr>
<td><strong>RESULTS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study selection</td>
<td>17</td>
<td>Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.</td>
</tr>
<tr>
<td>Study characteristics</td>
<td>18</td>
<td>For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.</td>
</tr>
<tr>
<td>Risk of bias within studies</td>
<td>19</td>
<td>Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).</td>
</tr>
<tr>
<td>Results of individual studies</td>
<td>20</td>
<td>For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.</td>
</tr>
<tr>
<td>Synthesis of results</td>
<td>21</td>
<td>Present results of each meta-analysis done, including confidence intervals and measures of consistency.</td>
</tr>
<tr>
<td>Risk of bias across studies</td>
<td>22</td>
<td>Present results of any assessment of risk of bias across studies (see Item 15).</td>
</tr>
<tr>
<td>Additional analysis</td>
<td>23</td>
<td>Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).</td>
</tr>
<tr>
<td><strong>DISCUSSION</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Summary of evidence</td>
<td>24</td>
<td>Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).</td>
</tr>
<tr>
<td>Limitations</td>
<td>25</td>
<td>Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).</td>
</tr>
<tr>
<td>Conclusions</td>
<td>26</td>
<td>Provide a general interpretation of the results in the context of other evidence, and implications for future research.</td>
</tr>
<tr>
<td><strong>FUNDING</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Funding</td>
<td>27</td>
<td>Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.</td>
</tr>
</tbody>
</table>


For more information, visit: [www.prisma-statement.org](http://www.prisma-statement.org).
How to use the Medical Subject Headings (MeSH) Database
Go to the PubMed home page.

Under “More Resources” click the “MeSH Database.”

Let’s say you want to find articles about problems associated with giving flu vaccines to individuals with egg allergies. You may first search for “flu vaccine” which returns the MeSH term, *Influenza Vaccines*.

The page will then provide a definition of the term *Influenza Vaccines* along with subheadings and other information.

In the MeSh database, you will see a hierarchically organize list of terms with more specific terms being place below broader terms. For this example, the terms would be displayed as follows. *(see next page)*
Subheadings may be used to narrow the search and can be found underneath the definition of the selected term. Continuing with the example, because we are interested in the negative effects of the flu vaccine, we will click the subheading “adverse effects.” Next, you will click the button “Add to Search Builder.”

To further limit the articles on egg allergies we can search MeSH again to find the appropriate MeSH term for “Egg Allergies.” Type in “Egg Allergies” in the search box at the top of the page. For this example, the database displayed the MeSH term *egg hypersensitivity*.
Next, we can add this search term to our search by clicking the “add to search builder” button on the right side of the page.

After adding MeSH terms to your search builder, you can click “Search PubMed” and see which target citations are presented.

Keep in mind that using MeSH terms this way will limit your search to a subset of PubMed citations. If you have further questions, you can use the “write to home desk” feature. This can be found on the bottom right of the MeSH database homepage.

All information provided was derived from Use MeSH to Build a Better PubMed Query (National Center for Biotechnology Information [NCBI], 2013).
Saving and Managing Searches in PubMed
Begin by visiting the PubMed home page at [https://www.pubmed.gov](https://www.pubmed.gov)

Let’s use Adolescent COPD for this example.

Step 1: Create a search query.

Enter “adolescent AND (COPD OR “chronic obstructive pulmonary disease”) into the search box.

Step 2: Click the search button.

Step 3: Click the “create alert” button which is located beneath the search box. A pop-up window will appear prompting you to sign in to your NCBI account.
Step 4: Save the search name.

- The saved search name may be up to 100 characters long and will become the search line of your e-mail. On this page, Search terms can also be edited. Always remember to click “test search terms” before saving the search terms. This will tell you how many results are found using the currently included search terms.

Step 5: Set up e-mail result updates.

- This page is also where you will indicate if you would like to receive e-mail updates from the new results that are found using the search terms. (Check “Yes, please” to receive e-mails, or “No, thanks”) You also have the option of changing which e-mail you would like the results to be sent to, however, only 1 e-mail address can be associated with a NCBI account. *You may want to add nih.gov to your e-mail “safe list”
Step 6: Set up the frequency of the e-mail.
▶ Under “schedule”
▶ E-mails can be set up monthly, weekly, or daily.

Step 7: Select the format and number of included items.
▶ Options: Summary, Abstract, Medline
▶ What to expect in an e-mail?
▶ All e-mails will include the name of the saved search, the date sent, the PubMed search terms, the ability to view the complete results in PubMed, the ability to edit the search settings, the option to unsubscribe, and the PubMed Results.

Step 8: Click SAVE!
▶ This will immediately bring you back to the PubMed search page and a check mark will appear stating “Your search was saved”
Storing a safe search through NCBI.

- When you are logged in to PubMed, click the “my NCBI” link located at the top right corner next to your username.

- You will see the section “Saved Searches.”

- Here you can see how many matching citations have been made available since you last checked your results. This is found by looking under the “What’s New” column. You will also see how many days have passed since you last viewed the search results.
How to build a search collection

- Go to the My NCBI homepage. From the homepage you will see that your account will come with 3 default collections titled “Favorites”, “My Bibliography”, and “Other Citations.”

- You can search directly through PubMed or click on the “What’s new” column under your “Saved searches”.

- This will bring you to all the citations that are included under the saved search. From here, you can click the box next to each citation that you wish to include in your collection.

- Next, click “Send to.”

- You will be provided with a drop-down box to allow you to choose destination. Click “Collections” and then click “Add to collections.”
A window will appear and can choose “create new collection” or “append to an existing collection.” If you are creating a new collection, you would enter the name of the new collection here.

Collections can be set to public or private.

If a collection is set to public, a URL is provided that will bring you to the collection. This URL can then be shared with others.

To set a collection to public. Click the “My NCBI” to go to the NCBI homepage. Click the “private” or “public” link found under the setting/sharing column corresponding to the collection.

This will bring you to the edit page. From this page you can modify the collection name, change the status to private or public, retrieve the public URL, and retrieve the HTML. The URL will no longer work if the collection is changed back to private.
Changing the display settings

Once you are in a collection, click the “Display settings” link that is found in the top left corner. Click the down arrow, here you can sort by “date,” “author,” or “title.”

All information provided was derived from NCBI Minute: Automate PubMed Searches & Save Citation Collections with My NCBI (Bailey, 2017).
References

- Bailey, S. H. (2017). NCBI Minute: Automate PubMed Searches & Save Citation Collections with My NCBI. National Center for Biotechnology Information. Retrieved from https://www.youtube.com/watch?v=kXY5P4C_2l4


