Methodology Checklist 2: Controlled Trials

Study identification (*Include author, title, year of publication, journal title, pages*)		
Guideline topic:	Key Question No:	Reviewer:

Before completing this checklist, consider:

1. Is the paper a **randomised controlled trial** or a **controlled clinical trial**? If in doubt, check the study design algorithm available from SIGN and make sure you have the correct checklist. If it is a **controlled clinical trial** questions 1.2, 1.3, and 1.4 are not relevant, and the study cannot be rated higher than 1+

2. Is the paper relevant to key question? Analyse using PICO (Patient or Population Intervention Comparison Outcome). IF NO REJECT (give reason below). IF YES complete the checklist.

Reason for rejection: 1. Paper not relevant to key question ☐ 2. Other reason ☐ (please specify):

SECTION 1: INTERNAL VALIDITY

	In a well conducted RCT study…	***Does this study do it?***
1.1	The study addresses an appropriate and clearly focused question.[i]	Yes ☐ No ☐ Can't say ☐
1.2	The assignment of subjects to treatment groups is randomised.[ii]	Yes ☐ No ☐ Can't say ☐
1.3	An adequate concealment method is used.[iii]	Yes ☐ No ☐ Can't say ☐
1.4	Subjects and investigators are kept 'blind' about treatment allocation.[iv]	Yes ☐ No ☐ Can't say ☐
1.5	The treatment and control groups are similar at the start of the trial.[v]	Yes ☐ No ☐ Can't say ☐
1.6	The only difference between groups is the treatment under investigation.[vi]	Yes ☐ No ☐ Can't say ☐
1.7	All relevant outcomes are measured in a standard, valid and reliable way.[vii]	Yes ☐ No ☐ Can't say ☐
1.8	What percentage of the individuals or clusters recruited into each treatment arm of the study dropped out before the study was completed?[viii]	
1.9	All the subjects are analysed in the groups to which they were randomly allocated (often referred to as intention to treat analysis).[ix]	Yes ☐ No ☐ Can't say ☐ Does not apply ☐
1.10	Where the study is carried out at more than one site, results are comparable for all sites.[x]	Yes ☐ No ☐ Can't say ☐ Does not apply ☐

SECTION 2: OVERALL ASSESSMENT OF THE STUDY		
2.1	How well was the study done to minimise bias? *Code as follows:*[xi]	High quality (++)☐ Acceptable (+)☐ Unacceptable – reject 0 ☐
2.2	Taking into account clinical considerations, your evaluation of the methodology used, and the statistical power of the study, are you certain that the overall effect is due to the study intervention?	
2.3	Are the results of this study directly applicable to the patient group targeted by this guideline?	
2.4	**Notes.** Summarise the authors' conclusions. Add any comments on your own assessment of the study, and the extent to which it answers your question and mention any areas of uncertainty raised above.	

[i] Unless a clear and well defined question is specified, it will be difficult to assess how well the study has met its objectives or how relevant it is to the question you are trying to answer on the basis of its conclusions.

[ii] Random allocation of patients to receive one or other of the treatments under investigation, or to receive either treatment or placebo, is fundamental to this type of study.

[iii] Allocation concealment refers to the process used to ensure that researchers are unaware which group patients are being allocated to at the time they enter the study. Research has shown that where allocation concealment is inadequate, investigators can overestimate the effect of interventions by up to 40%.

[iv] Blinding refers to the process whereby people are kept unaware of which treatment an individual patient has been receiving when they are assessing the outcome for that patient. It can be carried out up to three levels. Single blinding is where patients are unaware of which treatment they are receiving. In double blind studies neither the clinician nor the patient knows which treatment is being given. In very rare cases studies may be triple blinded, where neither patients, clinicians, nor those conducting the analysis are aware of which patients received which treatment. The higher the level of blinding, the lower the risk of bias in the study.

[v] Patients selected for inclusion in a trial must be as similar as possible. The study should report any significant differences in the composition of the study groups in relation to gender mix, age, stage of disease (if appropriate), social background, ethnic origin, or co-morbid conditions. These factors may be covered by inclusion and exclusion criteria, rather than being reported directly. Failure to address this question, or the use of inappropriate groups, should lead to the study being downgraded.

[vi] If some patients received additional treatment, even if of a minor nature or consisting of advice and counselling rather than a physical intervention, this treatment is a potential confounding factor that may invalidate the results. **If groups were *not* treated equally, the study should be rejected unless no other evidence is available.** If the study is used as evidence it should be treated with caution.

[vii] The primary outcome measures used should be clearly stated in the study. **If the outcome measures are not stated, or the study bases its main conclusions on secondary outcomes, the study should be rejected.** Where outcome measures require any degree of subjectivity, some evidence should be provided that the measures used are reliable and have been validated prior to their use in the study.

[viii] The number of patients that drop out of a study should give concern if the number is very high. Conventionally, a 20% drop out rate is regarded as acceptable, but this may vary. Some regard should be paid to why patients dropped out, as well as how many. It should be noted that the drop out rate may be expected to be higher in studies conducted over a long period of time. A higher drop out rate will normally lead to downgrading, rather than rejection of a study.

[ix] In practice, it is rarely the case that all patients allocated to the intervention group receive the intervention throughout the trial, or that all those in the comparison group do not. Patients may refuse treatment, or contra-indications arise that lead them to be switched to the other group. If the comparability of groups through randomisation is to be maintained, however, patient outcomes must be analysed according to the group to which they were originally allocated irrespective of the treatment they actually received. (This is known as intention to treat analysis.) If it is clear that analysis was not on an intention to treat basis, the study may be rejected. If there is little other evidence available, the study may be included but should be evaluated as if it were a non-randomised cohort study.

[x] In multi-site studies, confidence in the results should be increased if it can be shown that similar results were obtained at the different participating centres.

[xi] Rate the overall methodological quality of the study, using the following as a guide: **High quality** (++): Majority of criteria met. Little or no risk of bias. Results unlikely to be changed by further research. **Acceptable** (+): Most criteria met. Some flaws in the study with an associated risk of bias, Conclusions may change in the light of further studies. **Low quality** (0): Either most criteria not met, or significant flaws relating to key aspects of study design. Conclusions likely to change in the light of further studies.

SIGN

Methodology Checklist 4: Case-control studies

Study identification (*Include author, title, year of publication, journal title, pages*)

Guideline topic:	Key Question No:	Reviewer:

Before completing this checklist, consider:

1. Is the paper really a case-control study? If in doubt, check the study design algorithm available from SIGN and make sure you have the correct checklist.

2. Is the paper relevant to key question? Analyse using PICO (Patient or Population Intervention Comparison Outcome). IF NO REJECT (give reason below). IF YES complete the checklist.

Reason for rejection: Reason for rejection: 1. Paper not relevant to key question ☐ 2. Other reason ☐ (please specify):

Section 1: Internal validity

In an well conducted case control study:		***Does this study do it?***	
1.1	The study addresses an appropriate and clearly focused question.	Yes Can't say	No
Selection of subjects			
1.2	The cases and controls are taken from comparable populations.	Yes	No
		Can't say	
1.3	The same exclusion criteria are used for both cases and controls.	Yes Can't say	No
1.4	What percentage of each group (cases and controls) participated in the study?	Cases: Controls:	
1.5	Comparison is made between participants and non-participants to establish their similarities or differences.	Yes Can't say	No
1.6	Cases are clearly defined and differentiated from controls.	Yes Can't say	No
1.7	It is clearly established that controls are non-cases.	Yes Can't say	No
ASSESSMENT			
1.8	Measures will have been taken to prevent knowledge of primary exposure influencing case ascertainment.	Yes Can't say	No Does not apply
1.9	Exposure status is measured in a standard, valid and reliable way.	Yes Can't say	No
CONFOUNDING			
1.10	The main potential confounders are identified and taken into account in the design and analysis.	Yes Can't say	No
STATISTICAL ANALYSIS			

1.11	Confidence intervals are provided.	Yes	No
Section 2: OVERALL ASSESSMENT OF THE STUDY			
2.1	How well was the study done to minimise the risk of bias or confounding?	High quality (++) □ Acceptable (+) □ Unacceptable – reject 0 □	
2.2	Taking into account clinical considerations, your evaluation of the methodology used, and the statistical power of the study, do you think there is clear evidence of an association between exposure and outcome?	Yes Can't say	No
2.3	Are the results of this study directly applicable to the patient group targeted by this guideline?	Yes	No
2.4	**Notes.** Summarise the authors conclusions. Add any comments on your own assessment of the study, and the extent to which it answers your question and mention any areas of uncertainty raised above..		

SIGN

Methodology Checklist 3: Cohort studies

Study identification (*Include author, title, year of publication, journal title, pages*)

Guideline topic:	Key Question No:	Reviewer:

Before completing this checklist, consider:

1. Is the paper really a cohort study? If in doubt, check the study design algorithm available from SIGN and make sure you have the correct checklist.
2. Is the paper relevant to key question? Analyse using PICO (Patient or Population Intervention Comparison Outcome). IF NO REJECT (give reason below). IF YES complete the checklist..

Reason for rejection: 1. Paper not relevant to key question ☐ 2. Other reason ☐ (please specify):

Please note that a retrospective study (ie a database or chart study) cannot be rated higher than +.

Section 1: Internal validity

	In a well conducted cohort study:	***Does this study do it?***	
1.1	The study addresses an appropriate and clearly focused question.	Yes □ Can't say □	No □
Selection of subjects			
1.2	The two groups being studied are selected from source populations that are comparable in all respects other than the factor under investigation.	Yes □ Can't say □	No □ Does not apply □
1.3	The study indicates how many of the people asked to take part did so, in each of the groups being studied.	Yes □	No □ Does not apply □
1.4	The likelihood that some eligible subjects might have the outcome at the time of enrolment is assessed and taken into account in the analysis.	Yes □ Can't say □	No □ Does not apply □
1.5	What percentage of individuals or clusters recruited into each arm of the study dropped out before the study was completed.		
1.6	Comparison is made between full participants and those lost to follow up, by exposure status.	Yes □ Can't say □	No □ Does not apply □

ASSESSMENT			
1.7	The outcomes are clearly defined.	Yes □ Can't say □	No □
1.8	The assessment of outcome is made blind to exposure status. If the study is retrospective this may not be applicable.	Yes □ Can't say □	No □ Does not apply □
1.9	Where blinding was not possible, there is some recognition that knowledge of exposure status could have influenced the assessment of outcome.	Yes □ Can't say □	No □ □
1.10	The method of assessment of exposure is reliable.	Yes □ Can't say □	No □
1.11	Evidence from other sources is used to demonstrate that the method of outcome assessment is valid and reliable.	Yes □ Can't say □	No □ Does not apply□
1.12	Exposure level or prognostic factor is assessed more than once.	Yes □ Can't say □	No □ Does not apply □
CONFOUNDING			
1.13	The main potential confounders are identified and taken into account in the design and analysis.	Yes □ Can't say □	No □
STATISTICAL ANALYSIS			
1.14	Have confidence intervals been provided?	Yes □	No □
Section 2: OVERALL ASSESSMENT OF THE STUDY			
2.1	How well was the study done to minimise the risk of bias or confounding?	High quality (++) □ Acceptable (+) □ Unacceptable – reject 0	
2.2	Taking into account clinical considerations, your evaluation of the methodology used, and the statistical power of the study, do you think there is clear evidence of an association between exposure and outcome?	Yes Can't say	No
2.3	Are the results of this study directly applicable to the patient group targeted in this guideline?	Yes □	No □
2.4	**Notes.** Summarise the authors conclusions. Add any comments on your own assessment of the study, and the extent to which it answers your question and mention any areas of uncertainty raised above.		

SIGN

Methodology Checklist 1: Systematic Reviews and Meta-analyses

SIGN gratefully acknowledges the permission received from the authors of the AMSTAR tool to base this checklist on their work: *Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C,. et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Medical Research Methodology 2007, **7**:10 doi:10.1186/1471-2288-7-10. Available from http://www.biomedcentral.com/1471-2288/7/10 [cited 10 Sep 2012]*

Study identification (*Include author, title, year of publication, journal title, pages*)

Guideline topic:	Key Question No:

Before completing this checklist, consider:

1. Is the paper a systematic review or meta-analysis? IF NO reject. IF YES continue.
2. Is the paper relevant to key question? Analyse using PICO (Patient or Population Intervention Comparison Outcome). IF NO reject. IF YES complete the checklist.

Checklist completed by:

Section 1: Internal validity

In a well conducted systematic review:		***Does this study do it?***
1.1	The study addresses a clearly defined research question.[i]	Yes ☐ No ☐ Can't say ☐
1.2	At least two people should select studies and extract data.[ii]	Yes ☐ No ☐ Can't say ☐
1.3	A comprehensive literature search is carried out.[iii]	Yes ☐ No ☐ Can't say ☐ Does not apply ☐
1.4	The authors clearly state if or how they limited their review by publication type.[iv]	Yes ☐ No ☐
1.5	The included and excluded studies are listed.[v]	Yes ☐ No ☐
1.6	The characteristics of the included studies are provided.[vi]	Yes ☐ No ☐
1.7	The scientific quality of the included studies is assessed and documented.[vii]	Yes ☐ No ☐

1.8	The scientific quality of the included studies was assessed appropriately.[viii]	Yes ☐ No ☐ Can't say ☐
1.9	Appropriate methods are used to combine the individual study findings.[ix]	Yes ☐ No ☐ Can't say ☐
1.10	The likelihood of publication bias is assessed.[x]	Yes ☐ No ☐ Can't say ☐
1.11	Conflicts of interest are declared.[xi]	Yes ☐ No ☐
SECTION 2: OVERALL ASSESSMENT OF THE STUDY		
2.1	What is your overall assessment of the methodological quality of this review? [xii]	High quality (++) ☐ Acceptable (+) ☐ Unacceptable – reject 0 ☐
2.2	Are the results of this study directly applicable to the patient group targeted by this guideline?	Yes ☐ No ☐
2.3	**Notes:**[xiii]	

[i] The research question and inclusion criteria should be established before the review is conducted. To score a 'yes' for this factor there must be reference to a protocol, ethics approval, or pre-determined/a priori published research objectives.

[ii] At least two people should select papers and extract data. There should be a consensus procedure to resolve any differences.

[iii] At least two major electronic databases should be searched. The report must include years and databases searched (e.g., Central, EMBASE, MEDLINE, OpenGrey, 1999-2009). Key words and/or MESH terms must be stated and where feasible the search strategy should be provided. All searches should be supplemented by consulting current contents, reviews, textbooks, specialized registers, or experts in the particular field of study, and by reviewing the references in the studies found. In rare cases this may not apply where authors have carried out a meta analysis focusing on a specified range of major trials in their field.

[iv] The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status or language. If searching sources that contains both grey and non-grey literature, must specify that they were searching for both.

[v] A list of included and excluded studies should be provided. Limiting the excluded studies to references is acceptable.

[vi] In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions and outcomes. The ranges of characteristics in all the studies analyzed e.g., age, race, sex, relevant socioeconomic data, disease status, duration, severity, or other diseases should be reported. (**Note** that a format other than a table is acceptable, as long as the information noted here is provided).

[vii] This relates to the scientific quality of the studies included in the review. I can include use of a quality scoring tool or checklist, e.g., Jadad scale, risk of bias, sensitivity analysis, or a description of quality items, with some kind of result for EACH study ("low" or "high" is fine, as long as it is clear which studies scored "low" and which scored "high"; a summary score/range for all studies is not acceptable).

[viii] The methodological rigor and scientific quality should be considered in the analysis and the conclusions of the review, and explicitly stated in formulating recommendations. (**Note**: The review might say something like "the results should be interpreted with caution due to poor quality of included studies." Cannot score "yes" for this question if scored "no" for question 7).

[ix] For the pooled results, a test should be done to ensure the studies were combinable, to assess their homogeneity (i.e., Chi-squared test for homogeneity, *I*2). If heterogeneity exists a random effects model should be used and/or the clinical appropriateness of combining should be taken into consideration (i.e., is it sensible to combine?). Indicate "yes" where the authors mention or describe heterogeneity or variability between results and discuss the consequences (eg where authors declare they cannot pool results because of heterogeneity).

[x] An assessment of publication bias should include a combination of graphical aids (e.g., funnel plot, other available tests) and/or statistical tests (e.g., Egger regression test, Hedges-Olken). (**Note**: Score "Can't say" if there were fewer than 10 included studies).

[xi] Potential sources of support should be clearly acknowledged in both the systematic review and the included studies.

[xii] Rate the overall methodological quality of the study, using the following as a guide: **High quality** (++): Majority of criteria met. Little or no risk of bias. Results unlikely to be changed by further research. **Acceptable** (+): Most criteria met. Some flaws in the study with an associated risk of bias, Conclusions may change in the light of further studies. **Low quality** (0): Either most criteria not met, or significant flaws relating to key aspects of study design. Conclusions likely to change in the light of further studies.

[xiii] Add any comments on your own assessment of the study, and the extent to which it answers your question and mention any areas of uncertainty raised above. This is a very important part of the evaluation and will feature in the evidence table. PLEASE FILL IN.

Methodology Checklist 5: Studies of Diagnostic Accuracy

This checklist is based on the work of the QUADAS2 team at Bristol Univeristy (http://www.bris.ac.uk/quadas/).

Study identification (*Include author, title, reference, year of publication*)

Guideline topic:	Key Question No:

Before completing this checklist, consider:

1. Is the paper really a study of diagnostic accuracy? It should be comparing a specific diagnostic test against another, and **not** a general paper or comment on diagnosis.
2. Is the paper relevant to key question? Analyse using PICO (Patient or Population Intervention Comparison Outcome). IF NO REJECT (give reason below). IF YES complete the checklist..

Reason for rejection: Reason for rejection: 1. Paper not relevant to key question □ 2. Other reason □ (please specify):

Checklist completed by:

All the questions in the following sections have associated footnotes providing short explanations behind each of the questions. Users who want more detailed explanations should consult the QUADAS-2: Background Document.

DOMAIN 1 – PATIENT SELECTION

Risk of bias

In a well conducted diagnostic study…		***Is that true in this study?***	
1.1	A consecutive sequence or random selection of patients is enrolled.[i]	Yes ☐ No ☐	Can't say ☐
1.2	Case – control methods are not used.[ii]	Yes ☐ No ☐	Can't say ☐
1.3	Inappropriate exclusions are avoided.[iii]	Yes ☐ No ☐	Can't say ☐

Applicability

1.4	The included patients and settings match the key question.[iv]	Yes ☐ No ☐	Can't say ☐

DOMAIN 2 – INDEX TEST

Risk of bias

In a well conducted diagnostic study…		***Is that true in this study?***	
2.1	The index test results interpreted without knowledge of the results of the reference standard.[v]	Yes ☐ No ☐	Can't say ☐
2.2	If a threshold is used, it is pre-specified.[vi]	Yes ☐ No ☐	Can't say ☐

Applicability			
2.3	The index test, its conduct, and its interpretation is similar to that used in practice with the target population of the guideline.[vii]	Yes ☐ No ☐	Can't say ☐
DOMAIN 3 – REFERENCE STANDARD			
Risk of bias			
In a well conducted diagnostic study…		***Is that true in this study?***	
3.1	The reference standard is likely to correctly identify the target condition.[viii]	Yes ☐ No ☐	Can't say ☐
3.2	Reference standard results are interpreted without knowledge of the results of the index test.[ix]	Yes ☐ No ☐	Can't say ☐
Applicability			
3.3	The target condition as defined by the reference standard matches that found in the target population of the guideline.[x]	Yes ☐ No ☐	Can't say ☐
DOMAIN 4 – FLOW AND TIMING			
Risk of bias			
In a well conducted diagnostic study…		***Is that true in this study?***	
4.1	There is an appropriate interval between the index test and reference standard.[xi]	Yes ☐ No ☐	Can't say ☐
4.2	All patients receive the same reference standard.[xii]	Yes ☐ No ☐	Can't say ☐
4.3	All patients recruited into the study are included in the analysis.[xiii]	Yes ☐ No ☐	Can't say ☐
SECTION 5: OVERALL ASSESSMENT OF THE STUDY			
5.1	How well was the study done to minimise bias? *Code as follows:*[xiv]	High quality (++)☐ Acceptable (+)☐ Unacceptable – reject 0 ☐	
5.2	What is your assessment of the **applicability** of this study to our target population?	Directly applicable ☐ Some indirectness ☐ (Please explain in the following section for **Notes**)	
5.2	**Notes.** Summarise the authors conclusions. Add any comments on your own assessment of the study, and the extent to which it answers your question.		

[i] Studies should enrol either all eligible patients suspected of having the target condition during a specified period, or a random sample of those patients. The essential point is that investigators should have no freedom of choice as to which individual patients are or are not included.

[ii] There is evidence that studies comparing patients with known disease with a control group without the condition tend to exaggerate diagnostic accuracy.
[iii] Inappropriate exclusions may result in either overestimates (eg by excluding 'difficult to diagnose' patients) or underestimates (eg by excluding patients with 'red flags' suggesting presence of disease) of the degree of diagnostic accuracy.
[iv] Patients included in the study should match the target population of the guideline in terms of severity of the target condition, demographic features, presence of differential diagnosis or co-morbidity, setting of the study and previous testing protocols.
[v] This is similar to the question of 'blinding' in intervention studies. The index test should always been done first, or by a separate investigator with no knowledge of the outcome of the reference test.
[vi] Bias can be introduced if a threshold level is set after data has been collected. Any minimum threshold should be specified at the start of the trial.
[vii] Variations in test technology, execution, or interpretation (eg use of a higher ultrasound transducer frequency) may affect estimates of diagnostic accuracy.
[viii] Estimates of test accuracy are based on the assumption that the reference standard is 100% sensitive (=accurately diagnoses the target condition).
[ix] This is the similar to question 2.1, but in this case relates to making sure the reference standard is applied without any prior knowledge of the outcome of previous tests.
[x] The definition of the target condition used when testing the reference standard may differ from that used by the NHS in Scotland. eg threshold levels used in laboratory cultures may differ.
[xi] The index test and reference standard should be performed as close together in time as possible, otherwise changes in the patients condition is likely to invalidate the results.
[xii] In some cases the choice of reference standard may be influenced by the outcome of the index test or the urgency of the need for diagnosis. Use of different reference standards is likely to lead to overestimates of both sensitivity and specificity.
[xiii] Not including all patients in the analysis may lead to bias as there may be some systematic difference between those lost to follow-up and those analysed.
[xiv] Rate the overall methodological quality of the study, using the following as a guide: **High quality** (++): Majority of criteria met. Little or no risk of bias. Results unlikely to be changed by further research. **Acceptable** (+): Most criteria met. Some flaws in the study with an associated risk of bias, Conclusions may change in the light of further studies. **Low quality** (0): Either most criteria not met, or significant flaws relating to key aspects of study design. Conclusions likely to change in the light of further studies.