9.3 Quality assessment of prevalence studies

What is a prevalence study?

Prevalence studies address the prevalence or incidence of conditions. They inform health policies and explain or predict healthcare utilization. It is important that health policy decisions are based on high quality research. Prevalence and risk factors for conditions such as chronic diseases are evaluated using observational studies. The methods that are used to assess the validity of prevalence studies therefore differ from those that might be used to assess studies of the effects of interventions.

The STROBE statement provides guidance to improve reporting of observational studies and facilitates critical appraisal and interpretation of studies by reviewers, journal editors and readers. It is not a quality assessment instrument, but a guide to good reporting.

Hoy and colleagues (2012) developed a checklist for assessing the quality of prevalence studies using strong research methods. It can be found using the DOI 10.1016/j.jclinepi.2011.11.014 (see instructions for accessing articles held in the Monash library using the DOI).

In the summary below, each risk of bias item in the checklist developed by Hoy and colleagues is presented, along with explanations about the target bias. In addition some notes have been added to support those who are new to critical appraisal and applying decision rules associated with high quality critical appraisal checklists.

External Validity

**Item 1 addresses who was targeted for assessment in the population of interest**

Was the study’s target population a close representation of the national population in relation to relevant variables, e.g. age, sex, occupation?

The target population refers to the group of people or entities to which the results of the study will be generalised.

The definition of terms used in describing populations and samples targeted for investigation is summarised here.

Response Options

**Yes (LOW RISK):** The study’s target population was a close representation of the national population. **No (HIGH RISK):** The study’s target population was clearly NOT representative of the national population.
Example responses

The study was a national health survey of people 15 years and over and the sample was drawn from a list that included all individuals in the population aged 15 years and over.

A suitable response would be Yes (LOW RISK).

The study was conducted in one province only, and it is not clear if this was representative of the national population.

A suitable response would be No (HIGH RISK)

The study was undertaken in one village only and it is clear this was not representative of the national population.

A suitable response would be No (HIGH RISK)

**Item 2 addresses how the subset of participants was sampled from the population**

Was the sampling frame a true or close representation of the target population?

The sampling frame is a list of the sampling units in the target population and the study sample is drawn from this list.

Response Options

**Yes (LOW RISK):** The sampling frame was a true or close representation of the target population.

**No (HIGH RISK):** The sampling frame was NOT a true or close representation of the target population.

Example responses

The sampling frame was a list of almost every individual within the target population.

A suitable response would be Yes (LOW RISK).

The cluster sampling method was used and the sample of clusters/villages was drawn from a list of all villages in the target population.

A suitable response would be Yes (LOW RISK).

The sampling frame was a list of just one particular ethnic group within the overall target population, which comprised many groups.
A suitable response would be No (HIGH RISK).

**Item 3 addresses how samples were selected from the samples that were available for consideration**

Was some form of random selection used to select the sample, OR, was a census undertaken?

A census collects information from every unit in the sampling frame. In a survey, only part of the sampling frame is sampled. In these instances, random selection of the sample helps minimise study bias.

**Response Options**

**Yes (LOW RISK):** A census was undertaken, OR, some form of random selection was used to select the sample (e.g. simple random sampling, stratified random sampling, cluster sampling, systematic sampling).

**No (HIGH RISK):** A census was NOT undertaken, AND some form of random selection was NOT used to select the sample.

**Example responses**

The sample was selected using simple random sampling.

A suitable response would be Yes (LOW RISK).

The target population was the village and every person in the village was sampled.

A suitable response would be Yes (LOW RISK).

The nearest villages to the capital city were selected in order to save on the cost of fuel.

A suitable response would be No (HIGH RISK).

**Item 4 addresses the response rate and provides some acceptable rules for classifying response rate as acceptable or unacceptable**

Was the likelihood of non-response bias minimal?

**Response Options**

**Yes (LOW RISK):** The response rate for the study was $\geq 75\%$, OR, an analysis was performed that showed no significant difference in relevant demographic characteristics between responders and non-responders.
**No (HIGH RISK):** The response rate was <75%, and if any analysis comparing responders and non-responders was done, it showed a significant difference in relevant demographic characteristics between responders and non-responders.

**Example responses**

The response rate was 68%; however, the researchers did an analysis and found no significant difference between responders and non-responders in terms of age, sex, occupation and socioeconomic status.

A suitable response would be **Yes (LOW RISK).**

The response rate was 65% and the researchers did NOT carry out an analysis to compare relevant demographic characteristics between responders and non-responders.

A suitable response would be **No (HIGH RISK).**

The response rate was 69% and the researchers did an analysis and found a significant difference in age, sex and socio-economic status between responders and non-responders.

A suitable response would be **No (HIGH RISK).**

**Internal Validity**

**Item 5 addresses how the data were collected**

Were data collected **directly from the subjects** (as opposed to a proxy)?

**Response Options**

**Yes (LOW RISK):** All data were collected directly from the subjects.

**No (HIGH RISK):** In some instances, data were collected from a proxy. A proxy is a representative of the subject.

Point of decision rule clarification 1: The term 'in some instances' may need to be defined before this scale is applied. For example, would you accept an occasional proxy response and, if so, would you define the proportion of proxy responses that might be acceptable.

Point of decision rule clarification 2: In some instances a proxy response may be the only way to get the data. For example, when seeking information about instances of hospitalisation in young children, it might be appropriate to seek this information from parents, and you might need to consider the age at which it would be more appropriate to seek this information directly from the child.
Example responses

All eligible subjects in the household were interviewed separately.

A suitable response would be **Yes (LOW RISK)**.

A representative of the household was interviewed and questioned about the presence of low back pain in each household member.

A suitable response would be **No (HIGH RISK)**.

**Item 6 addresses the issue of how the target for assessment was defined**

Was an acceptable case definition used in the study?

**Response Options**

**Yes (LOW RISK):** An acceptable case definition was used.

**No (HIGH RISK):** An acceptable case definition was NOT used.

**Example responses**

For a study on low back pain, the following case definition was used: "Low back pain is defined as activity-limiting pain lasting more than one day in the area on the posterior aspect of the body from the bottom of the 12th rib to the lower gluteal folds."

A suitable response would be **Yes (LOW RISK)**.

For a study on back pain, there was no description of the specific anatomical location ‘back’ referred to.

A suitable response would be **No (HIGH RISK)**.

For a study on osteoarthritis, the following case definition was used: “Symptomatic osteoarthritis of the hip or knee, radiologically confirmed as Kellgren-Lawrence grade 2-4”.

A suitable response would be **Yes (LOW RISK)**.

**Item 7 addresses the method used to measure the target condition**

Was the study instrument that measured the parameter of interest (e.g. prevalence of low back pain) shown to have **reliability and validity (if necessary)?**

**Response Options**
Yes (LOW RISK): The study instrument had been shown to have reliability and validity (if this was necessary), e.g. test-retest, piloting, validation in a previous study, etc.

No (HIGH RISK): The study instrument had NOT been shown to have reliability or validity (if this was necessary).

The authors used the COPCORD questionnaire, which had previously been validated. They also tested the inter-rater reliability of the questionnaire.

A suitable response would be Yes (LOW RISK).

The authors developed their own questionnaire and did not test this for validity or reliability.

A suitable response would be No (HIGH RISK).

Point of decision rule clarification 1: There is a possible point of decision rule confusion here for a number of reasons. Neither reliability nor validity have been defined. There is some error in all measurements so what would be ‘reliable enough’? It also requires quite a lot of work to assess what is known about the reliability of all the instruments that might be used in papers included in review (essentially you need to review reliability reports for each method unless summary documents are published).

Point of decision rule clarification 2: If the measurements do identify a relationship between factors of interest, then the method used to gather the data is clearly reliable enough (or the effect would have been obscured by error). So in this case, even if no reliability data on the measurement system are available, it is likely to be adequately reliable.

Point of decision rule clarification 3: Just because the authors claim reliability (or tested reliability) does not mean that the instrument is reliable enough for the intended purpose as it is the measurement (in context) not the instrument for which reliability is relevant.

Point of decision rule clarification 4: If there is no relationship between the factors of interest, inadequate reliability MIGHT explain this outcome, and can be raised as an issue.

Point of decision rule clarification 5: If an instrument is valid (that is, it have been compared to a measurement we believe accurately measures the construct of interest and it produces comparable measurements) then it is clearly reliable enough to show this relationship, i.e. validity REQUIRES reliability. On the other hand an instrument can be reliable (low measurement error) but not valid (produces scores that are not related to the construct of interest).

Point of decision rule clarification 6: This is a ‘two target’ item and would be better redesigned as two items.
Given these barriers to interpretation of the item you could consider dropping or modifying this item.

**Item 8 addresses consistency in the methods used to collect data**

Was the **same mode of data collection** used for all subjects?

The mode of data collection is the method used for collecting information from the subjects. The most common modes are face-to-face interviews, telephone interviews and self-administered questionnaires.

Response Options

**Yes (LOW RISK):** The same mode of data collection was used for all subjects.

**No (HIGH RISK):** The same mode of data collection was NOT used for all subjects.

**Example responses**

All eligible subjects had a face-to-face interview.

A suitable response would be **Yes (LOW RISK).**

Some subjects were interviewed over the telephone and some filled in postal questionnaires.

A suitable response would be **No (HIGH RISK).**

Point of clarification: Researchers might undertake a study to determine whether comparable information is obtained under different data collection procedures. If this was done, and confidence that data collection options were comparable was conferred, you might consider amending the decision rule.

**Item 9 addresses the time period referred to in data collection**

Was the **length of the shortest prevalence period** for the parameter of interest appropriate?

The prevalence period is the period that the subject is asked about e.g. “Have you experienced low back pain over the previous year?” In this example, the prevalence period is one year. The longer the prevalence period, the greater the likelihood of the subject forgetting if they experienced the symptom of interest (e.g. low back pain).

Response Options

**Yes (LOW RISK):** The shortest prevalence period for the parameter of interest was appropriate (e.g. point prevalence, one-week prevalence, one-year prevalence).
**No (HIGH RISK):** The shortest prevalence period for the parameter of interest was not appropriate (e.g. lifetime prevalence)

**Example responses**

Subjects were asked about pain over the past week.

A suitable response would be **Yes (LOW RISK).**

Subjects were only asked about pain over the past three years.

A suitable response would be **No (HIGH RISK).**

Point of clarification: A decision (based on evidence) would need to be made regarding what was considered appropriate before quality assessment commenced.

**Item 10 addresses the accuracy of reported data**

*Were the numerator(s) and denominator(s) for the parameter of interest appropriate?*

There may be errors in the calculation and/or reporting of the numerator and/or denominator.

**Response Options**

**Yes (LOW RISK):** The paper presented appropriate numerator(s) AND denominator(s) for the parameter of interest (e.g. the prevalence of low back pain).

**No (HIGH RISK):** The paper did present numerator(s) AND denominator(s) for the parameter of interest but one or more of these were inappropriate.

**Example responses**

There were no errors in the reporting of the numerator(s) AND denominator(s) for the prevalence of low back pain.

A suitable response would be **Yes (LOW RISK).**

In reporting the overall prevalence of low back pain (in both men and women), the authors accidentally used the population of women as the denominator rather than the combined population.

A suitable response would be **No (HIGH RISK).**

Point of clarification 1: The item uses the term ‘appropriate’; an alternative term might be ‘accurate’.
Point of clarification 2: It is not clear what to do if no data are reported that enable cross checking of prevalence statistics; a decision rule might include a high risk classification if data for cross checking were not reported

**Item 11. Summary item on the overall risk of study bias**

Response Options

**LOW RISK OF BIAS:** Further research is very unlikely to change our confidence in the estimate.

**MODERATE RISK OF BIAS:** Further research is likely to have an important impact on our confidence in the estimate and may change the estimate.

**HIGH RISK OF BIAS:** Further research is very likely to have an important impact on our confidence in the estimate and is likely to change the estimate.

Point of clarification: It is not clear how to make a decision regarding low, moderate or high. This item requires a decision rule

**References**

Examples and help with decisions are found in the [full instrument](#):