The Joanna Briggs Institute Reviewers’ Manual 2014

The Systematic Review of Prevalence and Incidence Data
Foreword

Every year the Joanna Briggs Institute publishes a Reviewers’ Manual, which is designed to support individuals who are undertaking systematic reviews following JBI methodologies and methods. This chapter represents the latest work and methodological development of the Institute that was not ready for inclusion in the 2014 edition of the Reviewers’ Manual that was published in January.

As with the Reviewers’ Manual we recommend that this chapter be utilized in conjunction with the JBI SUMARI User Guide. Please note that this chapter makes reference to forthcoming analytical modules that do not currently exist in the JBI SUMARI software suite, but should be available in 2015. For advice on how to best apply the current software to accommodate this new methodology please contact the Synthesis Science Unit of the Institute at jbisyntehsis@adelaide.edu.au.

We hope that the information contained herewith provides further insight into how to analyse and synthesise different types of evidence to inform clinical and policy decisions to improve global health outcomes.

Associate Professor Edoardo Aromataris
Director, Synthesis Science
## Contents

1: Background ................................................................. 6  
   Introduction. .............................................................. 6  
   1.1 Use of observational studies in health care .............................. 8  
   1.2 Observational study designs ....................................... 8  
   1.3 The systematic review of prevalence and incidence data ........... 10  

2: Protocol and title development ....................................... 11  
   2.1 Title ................................................................. 11  
   2.2 Review question/objective ....................................... 11  
   2.3 Inclusion criteria .................................................. 12  
   2.4 Types of studies .................................................. 12  
   2.5 Search strategy .................................................... 13  
   2.6 Assessment of methodological quality ................................ 13  
   2.7 Data collection .................................................... 14  
   2.8 Data synthesis ..................................................... 14  
   2.9 Conflicts of interest .............................................. 14  
   2.10 Acknowledgements ............................................... 14  
   2.11 References ....................................................... 15  
   2.12 Appendices .......................................................... 15  

3: The systematic review and synthesis of prevalence and incidence data . 16  
   3.1 Title ................................................................. 16  
   3.2 Reviewers ............................................................ 16  
   3.3 Executive summary ................................................ 16  
   3.4 Background ........................................................... 18  
   3.5 Review objectives/review questions ................................ 18  
   3.6 Inclusion criteria .................................................. 18  
   3.7 Search strategy ..................................................... 18  
   3.8 Methods of the review ............................................. 19  
   3.9 Data synthesis ..................................................... 20  
   3.10 Considerations for conducting a meta-analysis ...................... 21  
   3.11 Subgroup analysis (analysis of subgroups or subsets): .......... 25
3.12 Meta-regression: .......................................................... 25
3.13 Publication bias: ............................................................ 25
3.14 Results ................................................................. 26
3.15 Description of studies .................................................. 26
3.16 Methodological quality .................................................. 28
3.17 Findings of the review ...................................................... 28
3.18 Synthesis of research findings ........................................... 29
3.19 Discussion ................................................................. 29
3.20 Conclusions ............................................................... 30
3.21 Implications for practice .................................................. 30
3.22 Implications for research .................................................. 30
3.23 Conflicts of interest ......................................................... 30
3.24 Acknowledgements .......................................................... 30
3.25 References ................................................................. 30
3.26 Appendices ................................................................. 31

References ................................................................. 32

4: Appendices ................................................................. 33
  Appendix 1 ................................................................. 33
  Appendix 2 ................................................................. 37
1: Background

Introduction

The accurate measurement of disease (whether at a local, national or global level) is of critical importance for governments, policy makers, health professionals and the general population. It enables them to inform the development and delivery of health services. Accurate information regarding measures of disease can assist in planning management of disease services (by ensuring resources are available to cope with the burden of disease), setting of priorities regarding public health initiatives, and evaluation of changes and trends in diseases over time.

There are a number of measurements of disease, including the proportion of a population who have a certain disease (the prevalence) and how often a disease occurs (the incidence). Questions of disease prevalence and incidence are often asked by researchers and the findings of their research provide crucial data for policy makers. Both of these measures enable health researchers to quantify disease amongst populations.¹

Prevalence

The prevalence of a disease indicates the number of people in a population that have the disease at a given point in time. This is often presented as a proportion or percentage, and sometimes as the number of cases out of a certain population. For example, the prevalence of HIV/AIDS in Australia in 2011 was 0.155%, or 115 out of every 100 000 people.² This is also known as ‘point’ prevalence as it provides a picture of what is occurring at a certain point of time.
The point prevalence is calculated using the following formula:

\[
\text{Prevalence} = \frac{\text{Number of people with disease at a given point in time}}{\text{Total number of people in the population}}
\]

‘Period’ prevalence is similar to “point” prevalence, except that it assesses the proportion of a population that has a disease at any time within a specified period of time. A third measure of disease prevalence is “lifetime” prevalence, and this measures the proportion of individuals in a population that will experience a disease at some time over the course of their lifetimes.

It is important to note that prevalence here does not exclusively refer to disease; prevalence of other variables such as a symptom, event, process or a risk factor may also be of interest to researchers and policymakers.

**Incidence**

The incidence of a disease indicates how many new cases of a disease occur in a defined period of time. Incidence data is obtained through tracking of an initially disease-free cohort and determining the number of those who developed the disease over a period of time.

Unlike prevalence, incidence is not affected by disease duration; it simply indicates the number of new occurrences of a disease in a certain period of time. The incidence of a disease in a one-year period, for example, would be found using the equation:

\[
\text{Incidence} = \frac{\text{Number of people who develop disease in one year}}{\text{Average number of people in the population in the same year}}
\]

For example, the number of new HIV diagnoses, or the incidence of HIV, in Australia in 2010 was 1,031 total cases. This means that during 2010, 0.0055% of the population was newly diagnosed with HIV, or in other words, approximately 5.5/100,000 people were newly diagnosed.

Cumulative incidence is similar to incidence, except that it considers only the portion of the population at risk of developing a disease during a specified time period; it gives the number of people who were at risk of developing a certain disease and who did develop the disease, in a certain time period. Cumulative incidence can be presented as either a proportion or percentage, or the number of cases out of the population. This can be calculated using the following equation:

\[
\text{Cumulative incidence} = \frac{\text{Number of people who develop disease in specified time period}}{\text{Number of people at risk of developing disease at the start of the time period}}
\]

Incidence can also be presented per person-years, which provides information on how quickly people develop a disease.

\[
\text{Incidence rate (or incidence density)} = \frac{\text{Number of people who develop disease}}{\text{Number of person-years when people were at risk of developing a disease}}
\]

For example, if measuring the incidence of prostate cancer in 1000 men over 10 years, and 80 men developed prostate cancer over that 10 year period, this can be presented as 80 per 10000 (1000x10) person years.

A relationship exists between prevalence and incidence. For example, two diseases can have a similar incidence, but if the duration of disease is longer for one of them (either because it is a chronic or longer lasting condition perhaps compared to an acute condition) than the prevalence will be much higher in the condition with the longer duration.
1.1 Use of observational studies in health care

Whilst randomized controlled trials (RCTs) are the best study design for answering questions of the effectiveness of interventions due to their ability to determine causality, they are not ideally suited to provide data of rates and patterns of disease occurrence. To address issues regarding prevalence and incidence, epidemiological studies, such as those classified under the term “observational and descriptive” studies, are required. These designs address questions such as: How many people have a disease? Who is getting the disease? Where is the disease occurring? This kind of information is particularly valuable for governments when making decisions regarding health policy. Furthermore, observational studies can often be used to infer correlations between two variables, for example, between a variable and a disease outcome. Data from observational studies can therefore be useful in formulating hypotheses regarding risk or preventive factors in disease development and progression. It is important to note that these studies are not able to determine causality; rather they are able only to infer correlations or relationships between variables.

Observational studies do not involve manipulation on the part of the researcher. These studies rely on the natural or “ecological” events of exposures and disease, where the researcher simply observes certain characteristics of the sample population as they occur “naturally”, and records the relevant data. In this way they can be distinguished from experimental or quasi-experimental studies (such as RCTs and controlled clinical trials) where there is researcher manipulation of the independent variable. Observational studies have a number of advantages over experimental study designs and are particularly valuable in instances where conduct of an RCT is unethical, such as investigating the effects of exposure to harmful substances, for example, the effects of asbestos exposure.

1.2 Observational study designs

Observational study designs include prospective and retrospective cohort studies, case-control studies, cross-sectional studies, case series and case reports, and can be broken down into the broad categories of analytical studies and descriptive studies. Generally, descriptive studies describe the occurrence/presence of an outcome only, whereas analytical studies describe the relationship between a variable and an outcome. Due to the nature of observational study designs, they are more at risk of confounding factors and different sources of bias, which will be discussed further below. Despite this, observational studies are essential in answering questions of prevalence and incidence.

Descriptive studies

Descriptive studies aim to collect information about a given individual or group and can be used to provide data on the distribution of disease. Examples of descriptive study designs are case reports and case series. In health care, these types of studies are typically used to describe the occurrence of disease or a risk factor. One powerful example of the value of descriptive studies is a case series detailing cases of Kaposi’s Sarcoma among homosexual men in California and New York which provided the first evidence of what would later be found to be HIV infection. Case reports and case series are often used to report unusual or unexpected occurrences of a disease or a unique finding, and they can be particularly informative for rare or emerging diseases.
Analytical studies
Analytical studies seek to make comparisons between different subgroups of the population and ask why a disease occurs in a subset of people or why a disease has a particular distribution. For example, a study may involve two groups; one group may have a disease and the other group may not, or one group may have experienced a certain event or been exposed to something, and another group have not. By assessing relationships between certain exposures, or indeed any variable such as behavior, risk factors, or characteristics, and a disease, researchers can form hypotheses regarding why disease occurs.

Cohort studies
Cohort studies are the most trusted observational study design and most closely resemble experimental studies. These longitudinal studies are typically used to analyze relationships between exposure and disease by comparing the outcomes between individuals who are exposed to the factor under study (risk factor, intervention, exposure, etc.) and others who are not. Sampling in cohort studies is based on the presence or absence of a certain exposure, and participants are followed over time to observe development of any disease outcomes. A prospective cohort study begins prior to or following an exposure, and participants are followed forward through time to observe any outcomes that may occur. A retrospective cohort study begins after outcomes are already present. In these studies, a researcher may sift through patient records and groups patients according to exposures, and identify any differences in outcomes.

Case-control studies
Case-control studies, on the other hand, select participants based on the presence of disease or a specific condition, and look for prior exposures or possible risk factors that may have led to the disease or condition developing. In this study design, those with the disease (cases) are matched with comparable individuals who do not have the disease (controls), and both groups are studied to determine if any differences in characteristics or past exposures exist. As the study population in case-control designs is based on individuals already having a disease, rates of prevalence or incidence of disease cannot be derived.

Cross sectional studies
Cross sectional studies are used to describe characteristics of a population at a given point in time, and as such provide a single “snapshot” of the prevalence of disease and other variables of the relevant population.

Cross sectional studies have been used to “examine the relationship between diseases (or other health-related characteristics) and other variables of interest as they exist in a defined population at one particular time.” Data can be used to infer relationships between a disease and other variables; however as the data is gathered simultaneously, chronological sequences of exposures and outcomes cannot be determined.

These types of studies collect information from sample populations that are selected regardless of exposure or disease status; they seek to describe the prevalence of exposure or disease in the population as a whole. Therefore, for this study design it is crucial that the sample obtained for the study represents the population as a whole, and random sampling methods should be used. Rare diseases can be problematic as they may require particularly large sample sizes to be captured.
Interpreting prevalence data from cross sectional studies must be done cautiously as the duration of a disease will impact its prevalence in a population. Diseases that are incurable, take a long time to cure, or have a long survival time will be more prevalent in a population at a given point in time than diseases that can be either cured quickly or quickly lead to death. For example, if two diseases have the same incidence, a disease with long survival will affect more people in a population at a single point in time compared with a disease that rapidly causes the death of affected individuals.

**Randomized controlled trials**

Randomized controlled trials can also be used to derive prevalence data to answer certain questions. For example, if one was interested in ascertaining the prevalence of a specific symptom among a population with a common condition, such as fatigue in type 2 diabetics, using baseline data from an RCT involving patients with type 2 diabetes, where the prevalence of fatigue is measured at baseline, would be acceptable. As the conduct of the subsequent trial would not be of interest to a question of prevalence, it is important to note that a critical appraisal tool designed for RCTs would not be appropriate to use.

**1.3 The systematic review of prevalence and incidence data**

The systematic review of prevalence and incidence data is important in the description of geographical distribution of a variable and the variation between subgroups (such as gender or socioeconomic status), and for informing health care planning and resource allocation. Pooling of such data is necessary to monitor trends in disease burden and emergence and to contribute to the design of further etiological studies. Systematic reviews are of particular relevance where individual studies are limited by small sample sizes.

There are many indications to conduct a review of these types of data, including the following:

- To assist in answering questions of global disease burden
- To help measure global disease burden (incidence data can be used to determine disability adjusted life years)
- Where it is not practical to do a large global survey
- For questions larger than a national scale
- Where cumulative meta-analysis can show changes and trends over time, and can highlight emerging or decreasing diseases
- To assist policy makers and funding models
- To elucidate geographical distributions and comparisons of subgroups
- To inform healthcare professionals of diseases and symptoms of disease
- To compare prevalence between groups
- To inform further research priorities.

The systematic review of studies to answer questions of prevalence and incidence data still follows the same basic principles of systematic reviews of other types of data. A protocol must be written for the conduct of the systematic review, comprehensive searching must be performed and critical appraisal of retrieved studies must be carried out. These steps will be further discussed in the following sections of this chapter.
2: Protocol and title development

This section outlines the components of a systematic review protocol of prevalence and incidence evidence and provides guidance on the information that each component should contain.

2.1 Title

The title should be clear, explicit and reflect the core elements of the protocol. Titles should not be phrased as questions or conclusions and there should be congruency between the title, review objectives/questions and inclusion criteria. The title needs to include the phrase “a systematic review protocol”.

The title should give an indication on the type of data that will be reported (descriptive, analytical or a combination of both) by including the epidemiological indicator or a term that reflects the analysis that will be used to measure the variables of interest. Generally, the measure of disease should appear in the title (prevalence, incidence).

The factors or events of interest (health condition or disease of interest) are defined by the time period, the place and the population at risk. Accordingly, the title should specify the defining characteristics of the population (i.e. gender, age) as well as the place and time of occurrence where relevant.

For example: “Prevalence and incidence of depression amongst adolescents globally: a systematic review protocol.”

2.2 Review question/objective

The review objective and any specific review question(s) must be clearly stated.

The overarching objective of reviews of prevalence and incidence data is to report on the frequency and distribution of specific factors, health states or conditions in a defined population.

Reviews of this type are broadly classified as having two primary objectives.

Reviews that aim to describe the distribution of existing variables or seek to answer the question “how common is a particular disease or condition in a specific group of individuals?” are often classified as descriptive and will utilize measures of prevalence and incidence to answer such lines of enquiry.

The objective of these reviews is to describe the health issue (what), those affected by it (who) as well as the location (where) and the time period (when) in which the health issue occurred.

Accordingly, the review question should outline the factor, disease, symptom or health condition of interest, the epidemiological indicator used to measure its frequency (prevalence, incidence), the population or groups at risk, as well as the context/location (e.g. limited to specific geographic areas) and time period (e.g. peaks at a particular season) where relevant.

For example: The objective of this review is to assess the prevalence and incidence of peri-natal depression among women in Australia.

Reviews focusing on how and why are predominantly analytic in nature. The objective of reviews of explanatory or analytic studies is to contribute to and improve our understanding of the causes of health-related events or outcomes by isolating the association between specific factors. This element is non-existent or lacking in studies that are purely descriptive.
While studies that report prevalence and incidence only are broadly classified as descriptive, and those that examine associations between exposures and outcomes are broadly classified as analytical, a clear-cut distinction between analytical and descriptive study designs is not possible. Data generated from these studies can be measured and reported in different ways and the review question will indicate whether the review seeks to report data that is descriptive, analytical or a combination of both.

2.3 Inclusion criteria

This section of the protocol details the basis on which studies will be considered for inclusion into the systematic review and should be as clear and unambiguous as possible.

When determining the inclusion criteria, the CoCoPop mnemonic (Condition, Context and Population) can be used for reviews assessing prevalence and incidence data.

Types of participants (population)

The types of participants should be appropriate for the review objectives. The reasons for the inclusion or exclusion of participants should be explained in the background.

It is important that the population or study subjects are clearly defined and described in detail. This includes outlining the specific or defining characteristics of the population such as age, sex, race gender, educational status, individual behavior, socio-demographic factors, etc.

For example: we will include studies involving adult pregnant women aged 18 – 45 years at any trimester and up to delivery.

Exclusion criteria should also be outlined where relevant. For example: studies examining pregnancies with neural tube defects, intra-uterine growth retardation and early pregnancy loss, and those involving adolescent pregnancies and anemic mothers will be excluded.

Condition

This refers to the variable of interest and may refer to a health condition, disease, symptom, event or factor. It is important that the variable of interest is clearly stated and defined. For example, malaria could be P. falciparum infection, P. vivax infection, or disease due to malarial infection.

This may include providing information on how the condition will be measured, diagnosed, or confirmed (i.e. by a blood test, observation of symptoms, a validated checklist, etc.).

Context

Environmental factors can have a substantial impact on the prevalence or incidence of a condition. Accordingly, it is important that authors define the context or specific setting relevant to their review question. For example, this may include defining the geographic area or country, specific community or setting (inpatient vs outpatient) and the time period given that some conditions may peak at a particular season (e.g. the incidence of influenza in different seasons and years).

2.4 Types of studies

Reviews of prevalence and incidence are predominantly derived from observational studies.
However many study designs may provide prevalence and incidence information.

- A cross-sectional study is the appropriate study design to determine the prevalence of a particular health problem. Cross-sectional surveys are typically used to estimate the point prevalence of common conditions of long duration and are generally not appropriate for rare or temporary diseases.

- As incidence is the number of new cases of a particular illness within a population over time, study participants need to be followed up. Therefore, cohort studies that have a prospective or longitudinal design and a follow-up of each subject over a suitable period of time are the best way to establish the incidence of a disease or the natural history of a condition.

### 2.5 Search strategy

This section details how the reviewers plan to search for relevant papers. A JBI review should consider papers published by both commercial publishers and literature that is not commercially published (grey literature). The timeframe chosen for the search should be justified and any language restrictions stated (e.g., only studies published in English will be considered for inclusion). The databases to be searched must be listed along with the initial keywords to be used for the search. Appropriate databases to search should be included, including specification from the outset of the platform used to search a particular database.

Prevalence and incidence data are reported within the published, peer-reviewed literature and accordingly the standard JBI three-step search strategy can be applied to locating this type of evidence.

There are also many and various sources of epidemiological data, within the grey literature, particularly for estimates of prevalence and incidence.

Some examples include:

- Administrative sources (clinical records, insurance data)
- Vital statistics data, government surveillance data and reports, Centers for Disease Control and Prevention data, population censuses and surveys (i.e., national or state health survey data), health care utilization records and disease registries (population-based disease registries established to record cases of certain serious diseases)
- Disease associations (e.g., American Diabetes Association)
- Medical books, grey literature and reports from experts.

What sources are chosen will obviously depend on the specific research question and its scope. For example, estimating the worldwide prevalence of a common condition (chronic disease) will need to include many more sources than a review examining the prevalence of a condition within a specific regional setting.

### 2.6 Assessment of methodological quality

The protocol should detail the criteria considered when determining methodological quality of papers to include in the review. JBI tools should be used preferentially; if not, clear reasoning should be provided. Critical appraisal tools must be appended to the protocol. For questions assessing incidence of a condition or disease, the critical appraisal tool should be selected based on the type of study design retrieved from the search process.
However, as prevalence data may be sourced from a number of study designs (including RCTs), a critical appraisal checklist specifically for prevalence studies has been developed (Appendix 1). Critical appraisal must be conducted by two reviewers independently of each other. The reviewers should then meet to discuss the results of their critical appraisal for their final appraisal. If the two reviewers disagree on the final critical appraisal and this cannot be resolved through discussion, a third reviewer may be required.

2.7 Data collection
Standardized data extraction tools allow the extraction of the same types of data across the included studies and are required for JBI systematic reviews. The protocol should detail what data the reviewers plan to extract from the included studies and the data extraction tool should be appended to the protocol. The data extraction sheet should suit the collection and stratification of the variables of interest from the included studies. It is important to extract data that reflects points of difference/heterogeneous characteristics between studies that affect the interpretation of the findings and synthesis of data. Whether data synthesis can be performed will depend on the heterogeneity of the variables of interest across included studies. To facilitate such a comparison, it is critical that data extraction details the variables that will be extracted and compared.

The description of disease patterns often includes analysis of demographic, geographical, social, seasonal and other risk factors. It is also likely to include the setting/location, dates of survey, definitions of conditions and populations, inclusion and exclusion criteria, mean age, sex, sample size, estimates of prevalence, incidence, etc.

For example, it may be important to consider gender categorization for diseases with a large gender gap (sexually transmitted diseases) and geographical distribution to describe diseases linked to environmental conditions.

Data extraction tools are outlined in Appendix 2.

2.8 Data synthesis
The protocol should detail how the reviewers plan to synthesize data extracted from included studies. The types of data it is anticipated will be synthesized should be consistent with the methods used for data collection and the included study designs. Refer to the next section for more detail on synthesis of prevalence and incidence data.

2.9 Conflicts of interest
A statement which either declares the absence of any conflicts of interest or which describes a specified or potential conflict of interest should be made by the reviewers in this section.

2.10 Acknowledgements
Any acknowledgements should be made in this section, e.g. sources of external funding or the contribution of colleagues or institutions. If the systematic review is to count towards a degree award, it should be noted.
2.11 References
All references should be listed in full using the Vancouver referencing style, in the order in which they appear in the protocol.

2.12 Appendices
The critical appraisal and data extraction tools should be appended in this section. They must match the study designs from the Inclusion Criteria section. Appendices should be numbered using Roman numerals.
3: The systematic review and synthesis of prevalence and incidence data

This section provides information on how to synthesize evidence relating to prevalence and incidence data. It provides guidance on the components that should make up a JBI systematic review of prevalence and incidence data, and the information that each component should contain. This section also provides a brief outline of how the systematic review should be formatted and the stylistic conventions that should be used to ensure the review meets the criteria for publication in the JBI Database of Systematic Reviews and Implementation Reports (JBISRIR). Specifically, guidance is provided on the following components: layout of the report, inclusion criteria, search strategy, critical appraisal, data extraction, data synthesis, results and conclusions.

3.1 Title
The title should be clear, explicit and reflect the core elements of the review. Titles should not be phrased as questions or conclusions and there should be congruency between the title, review objectives/questions and inclusion criteria. The title should include the phrase “a systematic review”. Refer to the guidance above for title development.

3.2 Reviewers
Each reviewer should have their names and post-nomial qualifications listed. Affiliations for each author need to be stated, including the affiliation with a JBI collaborating centre if there is one. An email address for the corresponding author only needs to be provided.

3.3 Executive summary
This section forms a structured abstract of the main features of the systematic review. It must be no longer than 500 words and should contain no abbreviations or references. The executive summary must accurately reflect and summarize the systematic review. The executive summary should include the following headings:

Background
This section should briefly describe the issue under review including the population, condition and context that are documented in the literature. The background should be an overview of the main issues. It should provide sufficient detail to justify why the review was conducted and the choice of the various elements such as the condition and context.

Objectives
The review objectives should be stated in full, as detailed in the protocol section.

Inclusion criteria:
Types of participants:
The report should provide details about the types of participants included in the review. Useful details include: age range, gender, profession, etc. Information supporting decisions about the types of participants should be explained in the background.
Condition: This section should present all the conditions/diseases/factors examined, as detailed in the protocol.

Context: This section should present all the contexts examined, as detailed in the protocol.

Types of studies:
As per the protocol section, the types of studies that were considered for the review should be included. There should be a statement about the target study type and indication if none of this type was found. The types of study identified by the search and those included should be detailed in the report.

Search strategy
A brief description of the search strategy should be included. This section should detail search activity (e.g. relevant databases searched, initial search terms or keywords, and any limitations) for the review, as predetermined in the protocol.

Methodological quality
Reviewers should mention how the studies included in the review were appraised and using which appraisal instrument.

Data extraction
This section should include a brief description of the types of data collected and the instrument (as specified in the protocol) used to extract data.

Data synthesis
This section should include a brief description of how the data was synthesized – as a meta-analysis or as a narrative summary with graphs and tables.

Results
This section should include a brief description of the findings of the review. Where possible, actual values of the overall prevalence or incidence should be provided for the reader.

Conclusions
This section should include a brief description of the conclusions of the review.

Implications for practice
This section should include a brief description of how the findings and conclusions of the review may be applied in practice, as well as any implications that the findings may have on current practice.

Implications for research
This section should include a brief description of how the findings of the review may lead to further research in the area, such as gaps identified in the body of knowledge.
3.4 Background
The background section should be comprehensive and cover all the main elements of the topic under review. Many reviewers will find that the background provided with the protocol needs modification or extension following the conduct of the review proper. The background should detail any definitions important to the review. The information in the background section must be sufficient to put the inclusion criteria into context. The background section should conclude with a statement that a preliminary search for previous systematic reviews on the topic was conducted (state the databases searched, e.g. JBISRIR, CINAHL and PubMed). If there is a previous systematic review on the topic, it should be specified how the proposed review differs. The Vancouver style referencing should be used throughout the review with superscript numbers without brackets used for in-text citations. JBI places significant emphasis on a comprehensive, clear and meaningful background section to every systematic review, particularly given the international circulation of systematic reviews, variation in local understandings of clinical practice, health service management and client or patient experiences. It is recommended that all JBI systematic reviews should contain a sentence clearly indicating:

“The objectives, inclusion criteria and methods of analysis for this review were specified in advance and documented in a protocol. Ref” (The reference should be to the appropriate citation in JBISRIR).

This sentence should appear as the final line of the background/introduction section of the review report and complies with the recommendations for reporting of systematic reviews detailed in the PRISMA guidelines.

3.5 Review objectives/review questions
As discussed previously in the protocol section, the objective(s) of the review should be clearly stated. This should be followed by specific question(s) that have been addressed in the review.

3.6 Inclusion criteria
As detailed in the protocol, this section of the review should detail the basis on which studies were considered for inclusion in the systematic review and should be as clear and unambiguous as possible. For a systematic review of prevalence and incidence studies, aspects should include: Population, Condition and Context.

3.7 Search strategy
This section should detail how the reviewers searched for relevant papers. The databases that were searched must be listed along with the search dates. A detailed search strategy for at least one of the major databases searched should be appended to the review. The documentation of search strategies is a key element of the scientific validity of a systematic review. It enables readers to look at and evaluate the steps taken and decisions made to consider the comprehensiveness and exhaustiveness of the search strategy for each included database. A JBI review should consider both commercially published studies and unpublished or grey literature.

Each electronic database is likely to use a different system for indexing keywords within their search engines. Hence the search strategy will be tailored to each particular database. These variations are important and need to be captured and included in the systematic review report.
Additionally, if a comprehensive systematic review is being conducted through JBI CReMS, the search strategies for each database for each approach are recorded and reported via CReMS. Commonly, these are added as appendices. The timeframe chosen for the search should be justified and any language restrictions stated (e.g. only studies published in English were considered for inclusion).

### 3.8 Methods of the review

#### Assessment of methodological quality

This section should detail the approach to critical appraisal, not the assessment results, and should be consistent with the protocol. Any deviations from the protocol must be reported and explained. The report should detail the criteria that were considered when determining the methodological quality of papers considered for inclusion in the review. The JBI tools should be used preferentially, as detailed in the protocol section. Critical appraisal tools must be appended to the review.

The primary and secondary reviewer should discuss each item of appraisal for each study design included in their review. The discussions should focus on what is considered acceptable to the needs of the review in terms of the specific study characteristics. The reviewers should be clear on what constitutes acceptable levels of information to allocate positive appraisal compared with a negative, or response of “unclear”. This discussion should take place before independently conducting the appraisal. The critical appraisal tool should be appended to the review.

#### Data collection

This section of the review should include details of the types of data extracted from the included studies. If no data was available for particular outcomes, that should also be discussed. Standardized data extraction tools allow the extraction of the same types of data across the included studies and are recommended for JBI systematic reviews. The included studies may include several outcomes; however, the review should focus on extracting information related to the research questions and outcomes of interest. Information that may impact upon the generalizability of the review findings such as study methods, setting and population characteristics should also be extracted and reported. Population characteristics include factors such as age, past medical history, co-morbidities, complications or other potential confounders. Reviewers should aim to reduce errors in data extraction by using two independent reviewers and a standardized data extraction instrument. The data extraction tool used must be appended to the review.

The data collection should include the following items and a brief description be provided for each item (a data extraction form has been appended [Appendix 2] with the following items listed):

**Study details**

- Reviewer – ostly includes details or ID of the primary reviewer.
- Study ID/record number – a numeric code to identify the study from which the effect size estimate was obtained
- Date – the date when this data extraction form was filled
- Study title – the full title of the study
• Author – an alphabetic or character code which is usually the first few characters of the primary study author’s name. This serves as an easy way to identify the study in the bibliography.
• Year – the year of publication.
• Journal – the journal in which the article was published.

Study method
• Aims of the study – as stated in the report.
• Setting – may refer to hospital or community or aged care facility. May also refer to rural/urban, etc.
• Study design – briefly describing the type of study design, e.g. if it is a randomized controlled trial or quasi-randomized controlled trial.
• Follow-up or study duration – any details on the duration of the study or follow-up with the participants.
• Subject characteristics – includes age, sex, country/location, sample size, diagnosis and other relevant characteristics.
• Dependent variable – ????
• Outcomes – the primary outcome measured and where relevant includes associated secondary outcomes.
• Outcome measurements – the scales or tools used to measure the outcomes, e.g. a standardized pain scale to measure pain.
• Ethical approval – yes/no.
• Method of data analysis – ????

Results
• Prevalence n/N (%).
• Proportion and 95% Confidence Intervals.
• Incidence n/N (%).
• Proportion and 95% Confidence Intervals and duration of recruitment or the study.
• Authors’ comments.
• Reviewers’ comments.

3.9 Data synthesis

This section should detail the approach to data synthesis, not the results of the synthesis. The review should detail how the reviewers synthesized the data extracted from included studies and detail the approach and how it was applied consistently across all included studies. The types of data detailed in this section should be consistent with the methods used for data collection and the included study designs.

Different methods exist for presenting a synthesis and these are outlined below:
• Narrative.
• Tabular.
• Graphical.
• Meta-analysis.
If the data is heterogeneous and is presented as a narrative summary, sources of heterogeneity should be discussed (e.g. clinical, methodological or statistical) as well as the basis on which it was determined inappropriate to combine the data statistically (such as differences in populations, study designs or clinical or statistical heterogeneity). Where meta-analysis was used, the statistical methods and the software used should be described.

There are established methods for conducting meta-analyses of randomized controlled trials and some observational study designs. However, no clear guidance exists on synthesizing frequency data from incidence and prevalence estimates.

### 3.10 Considerations for conducting a meta-analysis

A meta-analysis is a statistical process that essentially calculates effect sizes for individual studies, converts them to a common metric, and then combines them to obtain an average effect size.\(^6\)

**Generic outcomes for meta-analyses:**

- **Binary data (event)**
  - Incidence, Odds Ratio, Relative Risk, and Risk Difference.
  - Mean Difference, and Standardized Mean Difference.
- **Continuous data**
  - Hazard Ratio, Correlation coefficient, etc.

A meta-analysis or pooling of data from observational and descriptive studies presents several issues that include heterogeneity, prospectively planning data collection and pooling data across surveys. Effect sizes from observational and descriptive studies are prone to sampling and non-sampling errors. Sampling errors are differences between the sample and population values and non-sampling errors are a result of various factors other than sampling, which include non-response, coverage error and measurement or response error.\(^7\)

**Effect size**

The effect size represents the study results and is represented by a square on a forest plot. In traditional effectiveness reviews, this could be the impact of a new therapy on mortality rates or the effect of a new teaching method on exam scores. The effect size could be a single number such as for a prevalence study or a ratio such as a risk ratio. The effect size has been described as being the “currency of the systematic review” as the aim of the meta-analysis is to summarize the effect size of each included study to obtain a summary effect.\(^8\) The summary effect is shown as a diamond on a forest plot. When effect sizes are statistically combined, the methods used make certain assumptions.

**Statistical combination of data**

In meta-analysis, the results of similar, individual studies are combined to determine the overall effect. In meta-analysis, the effect size and weight of each study are calculated. The effect size indicates the direction and magnitude of the results of a particular study (i.e. do the results favor the treatment or control, and if so, by how much), while the weight is indicative of how much information a study provides to the overall analysis when all studies are combined together.

It has been suggested that there are three important criteria for choosing a summary statistic for a meta-analysis: (i) consistency of effect across studies, (ii) mathematical properties, and (iii) ease of interpretation.\(^9\)
1. Consistency of effect is important because the aim of meta-analysis is to bring together the results of several studies into a single result.

2. The main mathematical property required by summary statistics is the availability of a reliable variance estimate. Consensus about the other two mathematical properties (reliance on which of the two outcome states [e.g. mortality/survival] is coded as the event and the odds ratio (OR) being the only statistic which is unbounded) has not yet been reached.

3. Ease of interpretation.

Essentially there are three popular approaches to conduct meta-analyses for all types of data: Hedge and Olkin technique, Rosenthal and Rubin technique and the Hunter and Schmidt technique. Hedge and Olkin developed both fixed- and random-effects models for pooling data, Rosenthal and Rubin developed a fixed-effects model only, and Hunter and Schmidt developed a random-effects model.

**Statistical assumptions in meta-analysis**

Meta-analyses can be based on either of two statistical assumptions – fixed or random effects. It is important to distinguish between fixed- and random-effects models when conducting a meta-analysis, as it can lead to false assumptions about statistical significance of the pooled estimate.

The main difference between fixed and random effects models is in the calculation of standard errors associated with the combined effect size. Fixed effects models use only within-study variability in their error term because all other ‘unknowns’ in the model are assumed not to affect the effect size. In contrast, in random effects models it is necessary to account for the errors associated with sampling from populations that themselves have been sampled from a superpopulation. As such the error term contains two components: within-study variability and variability arising from differences between studies.\(^6\)

The fixed effects model assumes that there is one true effect for the population underlying the studies in the analysis and that all the differences in the data are due to sampling error or chance within each study and that there is no heterogeneity between the studies. A fixed effect model is statistically stringent and should be used when there is little heterogeneity, as determined by Chi-square (or I²). This model therefore assumes that the overall sample consists of samples that all belong to the same underlying population.\(^10\) The between-study variability will be zero in this model as it assumes that the population effect size is identical for all studies. In an analysis based on a fixed effects model, inference is applicable or generalizable (“conditional”) based on statistical justification only on the studies actually done.\(^11\) The fixed effects model assumes that there is little interest or value in generalizing the results to other studies.\(^12, 13\)

A random effects model allows more flexibility, assuming that there may be other factors influencing the data than error or chance, within and between studies. As a result, in an analysis based on a random effects model, inference relies on the assumption that the studies used in the analysis are a random sample of some hypothetical population of studies.\(^11, 13\) For example, the effect size may be influenced in studies where the participants are more educated, older or healthier or if a more intense intervention is being used. The effect size is assumed to follow a normal distribution and consequently has a mean and variance. The random-effects model considers both between-study variability and within-study variability. The random-effects model enables generalizations beyond the population included in the studies.
There is no consensus about whether fixed or random effects models should be used in meta-analysis. In many cases when heterogeneity is absent, the two methods will give similar overall results. When heterogeneity is present, the random effects estimate provides a more conservative estimate of the overall effect size, and is less likely to detect significant differences. For this reason, random effects models are sometimes employed when heterogeneity is not severe; however, the random effects model does not actually analyze the heterogeneity and should not be considered as a substitute for a thorough investigation into the reasons for heterogeneity. Additionally, random effects models give relatively more weight to the results of smaller studies – this may not be desirable because smaller studies are typically more prone to bias and are often lower quality than larger studies.

There are a number of meta-analytical techniques available – the selection of a particular technique is governed by three things: the study type, the nature of the data extracted and the assumptions underlying the meta-analysis.

**Meta-analysis of prevalence and incidence data - Proportions**

Prevalence and incidence data is often reported as a proportion. When pooling proportions for a meta-analysis, a transformation of the data is required. There are two main ways to transform the data: the Freeman-Tukey transformation (arcsine square root transformation), and the Logit transformation. Both of these are used to calculate the weighted summary proportion under the fixed and random effects model. The resultant meta-analysis will give pooled proportion with 95% CI both for the fixed effects model and the random effects model and in addition, will list the proportions (expressed as a percentage), with their 95% CI, found in the individual studies included in the meta-analysis. The results are then presented graphically in a forest plot. For all meta-analyses, prevalence estimates are transformed to logits to improve their statistical properties. These are then back-transformed to prevalence.

There are different models for performing the meta-analysis as mentioned above. The reviewer has the option of both of these models. We recommend that the meta-analyses of the prevalence reported in the studies are grouped by a random-effects model and presented with 95% confidence intervals (95% CI). A random effects model is used when there is sufficient information on standard errors. However, bear in mind that the random-effects model gives a conservative estimate with a wider confidence interval. The fixed model can be chosen but the reviewer should be aware of its underlying principles, particularly in relation to its assumption that there is one true effect, which may not hold for prevalence and incidence data.

Heterogeneity of the results is tested by the I-squared, Tau-squared and Chi-squared (p > 0.05) tests. These tests of heterogeneity (Cochran’s Q test) evaluate whether the differences in prevalence estimates across studies are higher than expected by chance. To identify the sources of heterogeneity across studies, subgroup analysis or meta-regression can be used to assess the contribution of each variable (i.e. year of study, geographic location, characteristic of countries, study population, etc.) to the overall heterogeneity. Those variables significantly associated with the heterogeneity (p < 0.05) can be included in a multivariate hierarchical model. A p value of <0.05 is considered statistically significant in all the analyses.
How to interpret effect sizes

Once authors calculate effect sizes, they need to answer the question: What does the effect size mean?

An effect size is simply a number and its meaning and importance must be explained by the researcher. An effect size of any magnitude can mean different things depending on the research that produced it and the results of similar past studies. Therefore, it is the researcher’s responsibility to discuss the importance of his or her findings and this information requires comparing current effects to those obtained in previous work in the same research area.

Confidence Intervals (CIs) are an important way to evaluate the precision of a study’s findings by providing a range of likely values around the obtained effect size.

Heterogeneity

When used in relation to a meta-analysis, the term “heterogeneity” refers to the amount of variation in characteristics of included studies. For example, if three studies are to be included in a meta-analysis, does each of the included studies have similar demographics, and assess the same intervention? While some variation between studies will always occur due to chance alone, heterogeneity is said to occur if there are significant differences between studies, and under these circumstances meta-analysis is not valid and should not be undertaken.

There are three types of heterogeneity: clinical, methodological, and statistical heterogeneity. Differences in the characteristics of study populations and measurements represent clinical heterogeneity. Differences in study designs and methodological quality (risk of bias) represent methodological heterogeneity. Statistical heterogeneity is the variation of effects sizes between studies. Statistical heterogeneity may arise because of clinical heterogeneity, methodological heterogeneity, or simply by chance.

There is often heterogeneity amongst studies addressing prevalence and incidence. This is due to a number of reasons. Firstly, clinical heterogeneity may be present due to the measures used to determine the presence of a variable. For example, different scales exist to measure depression, and depending on the scale used, a person may be classified as suffering from depression whilst using one scale and not suffering based on a different scale. Additionally, prevalence and incidence studies often look at specific populations at a specific point of time, and the scope of the study may be limited by state or national borders. Another consideration with the population is whether those considered at risk or eligible for the disease have been included. For example, if you look at the prevalence or incidence of breast cancer, have these studies reported the proportion out of the whole population, all females, only adult females, and so on? These different populations may contribute to clinical heterogeneity.

Methodological heterogeneity is also important to consider. Prevalence and incidence data can arise from various study designs with differing levels of methodological quality. This can also result in differences amongst studies.

But how does one tell whether or not differences are significant?

Firstly, the studies should be assessed carefully to determine whether there is clinical or methodological heterogeneity present. If conducting a meta-analysis, then a visual inspection of the meta-analysis output (e.g. the forest plot) is the first stage of assessing heterogeneity. If the results are scattered across the forest plot and none of the confidence intervals overlap, this is a good indicator of heterogeneity.
A formal statistical test of the similarity of studies is provided by a test of homogeneity. The test calculates a probability (P value) from a Chi-square statistic calculated using estimates of the individual study weight, effect size and the overall effect size. Note, however, that this test suffers from lack of power – and will often fail to detect a significant difference when a difference actually exists – especially when there are relatively few studies included in the meta-analysis. Because of this low power, some review authors use a significance level of P < 0.1, rather than the conventional 0.05 value, in order to protect against the possibility of falsely stating that there is no heterogeneity present. Often when combining the results from a series of observational studies, this is the default significance level due to the increased heterogeneity associated inherent with the study design.

The I2 statistic is the percentage of observed total variation across, due to heterogeneity and not due to chance. A value of 0% indicates no observed heterogeneity and larger values show increasing heterogeneity.

If there is statistically significant heterogeneity, a narrative synthesis or graphical representation is recommended.

3.11 Subgroup analysis (analysis of subgroups or subsets):

Subgroup analysis is a means of investigating heterogeneous results and can be used to estimate the influence of various subsets including age group, gender, type of population and sampling strategy used to gather data (e.g. letter, phone, face-to-face). However, subgroups should be pre-specified a priori and there should not be too many subgroups. Subgroup analysis may be carried out.

Subgroup analysis could include:
- Subsets of studies
- Subsets of patient groups.

3.12 Meta-regression:

Meta-regression investigates whether particular covariates explain any of the heterogeneity of treatment effects between studies. A meta-regression is either a linear or logistic regression and could be a fixed-effect or random-effect model. The unit of analysis is a study and predictors in the regression are the study-level covariates.

3.13 Publication bias:

The research that appears in the published literature is systematically unrepresentative of the population of completed studies. File drawer problem or publication bias is a term coined by Rosenthal to mean the number of statistically non-significant studies (p > 0.05) that remain unpublished. A Funnel plot is used to detect publication bias.

Funnel plot:
- Scatter plot of effect estimate (x-axis) against inverse of its variance (y-axis).
- If there is no bias then the funnel will appear symmetric and inverted and if there is bias, the funnel will be asymmetric or skewed in shape.
3.14 Results
This section should allow the reader to clearly follow how the included studies were identified and selected for inclusion in the review. In addition, the number of papers excluded should also be stated. There should be a narrative description of the process accompanied by a flowchart of review process (from the PRISMA statement) detailing the flow from the search, through study selection, duplicates, full text retrieval, and any additions from third search, appraisal, extraction and synthesis (see Figure 1 for an example of a flowchart template).

Details of full-text articles retrieved for critical appraisal should be given. There should be separate appendices for details of included and excluded studies, and for excluded studies; reasons should be stated on why they were excluded.

3.15 Description of studies
This section of the results should also include an overall description of the included studies (with reference to the table in the appendices), with the main aim to provide some context to the results section and sufficient detail for the reader to determine if the included studies are similar enough to combine in a meta-analysis. Specific items/points of interest from individual studies may also be highlighted here. Additional details may include the assessment of methodological quality, characteristics of the participants and types of interventions and outcomes.

Where a systematic review has several foci, the results should be presented in a logical, structured way, relevant to the specific questions. The roles of tables and appendices should not be overlooked. Adding extensive detail on studies in the results section may crowd the findings, making them less accessible to readers, hence the use of tables and in text reference to specific appendices is encouraged.
Figure 1: Flowchart detailing identification and selection of studies for inclusion in the review


For more information, visit www.prisma-statement.org.
3.16 Methodological quality

This section should focus on methodological quality as determined by the relevant critical appraisal checklist. There should be a narrative summary of the overall methodological quality of the included studies, which can be supported (optional) by a table showing the results of the critical appraisal (see Table 1 for example). Where only few studies are identified, or there are specific items of interest from included studies, these should be addressed in the narrative also, particularly where studies were deficient, or particularly good, i.e. with clear narrative regarding risk of bias/rigor of included studies. Use of “N/A” should also be justified in the text.

Table 1: Critical appraisal results for included studies using the JBI Prevalence Critical Appraisal Checklist

<table>
<thead>
<tr>
<th>Study</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
<th>Q6</th>
<th>Q7</th>
<th>Q8</th>
<th>Q9</th>
<th>Q10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s) ref</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>U</td>
</tr>
</tbody>
</table>

Y - Yes, N - No, U - Unclear

3.17 Findings of the review

Although there is no defined structure for this section, the findings of the review should flow logically from the review objection/question, i.e. they must ultimately answer the question! Findings should be extracted and a narrative, tabular, graphical or meta-analysis should constitute part of this section.

With detail on the studies reported, the results section then focuses on providing a detailed description of the results of the review. For clarity and consistency of presentation, JBI recommends that the reviewer, in discussion with their review panel, give consideration to whether the specific review question has been used to structure the results section, or whether the findings can be reported under the conditions specified in the protocol. When a systematic review seeks to address multiple questions, the results may be structured in such a way that particular conditions are reported according to the specific questions.

Given there is no clear international standard or agreement on the structure or key components of this section of a review report, and the level of variation evidence in published systematic reviews, the advice here is general in nature. In general, findings are discussed textually and then supported with meta-graphs, tables and figures as appropriate. Graphs may be particularly useful for presenting prevalence and incidence data where a meta-analysis is not possible.

The focus should be on presenting information in a clear and concise manner. Any large or complex diagrams/tables/figures should be included as appendices so as not to break the flow of text. Meta-view graphs represent a specific item of analysis that can be incorporated into the results section of the review. However, the results are more than meta-view graphs, and whether this section is structured based on the intervention of interest, or some other structure, the content of this section needs to present the results with clarity.
3.18 Synthesis of research findings

It is important to combine the studies in an appropriate manner, otherwise the conclusions that are drawn will not be valid. Where possible, study results should be pooled in a statistical meta-analysis. All results must be double entered in order to avoid data entry errors. Where statistical pooling is not possible the findings can be presented in narrative summary or graphical form, as previously discussed.

The meta-analysis functions in SUMARI are made up of a number of dropdown menus that allow the user to specify the condition, the population, etc. These dropdown menus incorporate descriptions and data that have been previously entered. The SUMARI user guide is a recommended text for technical aspects of data synthesis.

This section of the report should describe the data type, the required effects model used (random/fixed), the statistical method of meta-analysis required and the size of confidence limits to be included in the calculations. The method used will depend on the data type. In terms of confidence intervals, the default setting is to calculate 95% confidence intervals; however this can be adjusted to either 90% or 99% as required. Once all the appropriate settings have been selected, the forest plot summarizing the results of individual studies and their combined meta-analysis can be generated.

3.19 Discussion

This section should discuss the results of the synthesis as well as any limitations of the primary studies included in the review and of the review itself (i.e. language, access, timeframe, study design, etc.). The results should be discussed in the context of current literature, practice and policy.

The aim of this section is to minimize and discuss the main findings – including the strength of the evidence, for each main outcome. It should address the issues arising from the conduct of the review including limitations and issues arising from the findings of the review (such as search limitations). The discussion does not bring in new literature or information that has not been reported in the results section. The discussion does seek to establish a line of argument based on the findings regarding the effectiveness of an intervention or its impact on the outcomes identified in the protocol. The application and relevance of the findings to relevant stakeholders (e.g. healthcare providers, patients and policy makers) should also be discussed in this section.

Points to consider this section include:

- Were any problems identified undertaking the search (perhaps there is little primary research on this topic or perhaps it is poorly indexed by the databases that were searched or perhaps the search was insufficient)?
- What limitations were found in the included primary research (e.g. were there inconsistencies or errors in reporting)?
- How do the review findings fit with what is currently known on the topic (from issues highlighted in the Background section)?
- Are the findings generalizable to other populations of participants/healthcare settings, etc?
3.20 Conclusions
This section should begin with an overall conclusion based on the results. The conclusions drawn should match with the review objective/question.

The conclusion section of a systematic review should provide a general interpretation of the findings in the context of other evidence and provide a detailed discussion of issues arising from the findings of the review and demonstrate the significance of the review findings to practice and research. Areas that may be addressed include:

- A summary of the major findings of the review
- Issues related to the quality of the research within the area of interest
- Other issues of relevance
- Potential limitations of the systematic review.

3.21 Implications for practice
It should be stated how the findings of the review impact on clinical practice or policy in the area of focus of the review. If there is sufficient evidence to make specific recommendations for practice, then the appropriate JBI Grades of Recommendation should be assigned to each recommendation based on the study design that led to the recommendation.

3.22 Implications for research
This section should include clear, specific recommendations for future research based on gaps in knowledge identified from the results of the review. Implications for research should avoid generalized statements calling for further research, but should be linked to specific issues.

3.23 Conflicts of interest
A statement which either declares the absence of any conflicts of interest or which describes a specified or potential conflict of interest should be made by the reviewers in this section.

3.24 Acknowledgements
Any acknowledgements should be made in this section, e.g. sources of external funding or the contribution of colleagues or institutions. If the systematic review is to count towards a degree award, it should be noted.

3.25 References
All references should be listed in full using the Vancouver referencing style, in the order in which they appear in the review. The references should be appropriate in content and volume and include background references and studies from the initial search.
3.26 Appendices

Appendices should be numbered using Roman numerals in the order in which they have been referred to in the body of the text. There are several required appendices for a JBI review:

Appendix I: Search strategy
A detailed search strategy for at least one of the major databases searched must be appended.

Appendix II: Critical appraisal instrument
The critical appraisal instrument used must be appended, i.e. JBI Critical Appraisal Checklist for prevalence studies.

Appendix IV: Table of included studies
A table of included studies is crucial to allow a snapshot of the studies included in the review.

Appendix V: List of excluded studies
At a minimum, a list of studies excluded at the critical appraisal stage must be appended and reasons for exclusion should be provided for each study (these reasons should relate to the methodological quality of the study, not study selection). Studies excluded following examination of the full-text may also be listed along with their reason for exclusion at that stage (i.e. a mismatch with the inclusion criteria). This may be as a separate appendix or itemized in some fashion within the one appendix.

Appropriate appendices (appraisal, extraction tools) as they appear from CReMS should be provided and referred to in the chapter.
References


## Appendix 1

### JBI Critical Appraisal Checklist for Studies Reporting Prevalence Data

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>Unclear</th>
<th>Not applicable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Was the sample representative of the target population?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Were study participants recruited in an appropriate way?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Was the sample size adequate?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Were the study subjects and the setting described in detail?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Was the data analysis conducted with sufficient coverage of the identified sample?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Were objective, standard criteria used for the measurement of the condition?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Was the condition measured reliably?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Was there appropriate statistical analysis?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Are all important confounding factors/ subgroups/ differences identified and accounted for?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Were subpopulations identified using objective criteria?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Overall appraisal: Include [ ] Exclude [ ] Seek further info [ ]
Prevalence Critical Appraisal Tool

Answers: Yes, No, Unclear or Not/Applicable

1. Was the sample representative of the target population?

This question relies upon knowledge of the broader characteristics of the population of interest. If the study is of women with breast cancer, knowledge of at least the characteristics, demographics and medical history is needed. The term “target population” should not be taken to infer every individual from everywhere or with similar disease or exposure characteristics. Instead, give consideration to specific population characteristics in the study, including age range, gender, morbidities, medications, and other potentially influential factors. For example, a sample may not be representative of the target population if a certain group has been used (such as those working for one organisation, or one profession) and the results then inferred to the target population (i.e. working adults).

2. Were study participants recruited in an appropriate way?

Recruitment is the calling or advertising strategy for gaining interest in the study, and is not the same as sampling. Studies may report random sampling from a population, and the methods section should report how sampling was performed. What source of data were study participants recruited from? Was the sampling frame appropriate? For example, census data is a good example of appropriate recruitment as a good census will identify everybody. Was everybody included who should have been included? Were any groups of persons excluded? Was the whole population of interest surveyed? If not, was random sampling from a defined subset of the population employed? Was stratified random sampling with eligibility criteria used to ensure the sample was representative of the population that the researchers were generalizing to?

3. Was the sample size adequate?

An adequate sample size is important to ensure good precision of the final estimate. Ideally we are looking for evidence that the authors conducted a sample size calculation to determine an adequate sample size. This will estimate how many subjects are needed to produce a reliable estimate of the measure(s) of interest. For conditions with a low prevalence, a larger sample size is needed. Also consider sample sizes for subgroup (or characteristics) analyses, and whether these are appropriate. Sometimes, the study will be large enough (as in large national surveys) whereby a sample size calculation is not required. In these cases, sample size can be considered adequate.
When there is no sample size calculation and it is not a large national survey, the reviewers may consider conducting their own sample size analysis using the following formula:\textsuperscript{15, 16}

\[ N = \frac{Z^2 P(1-P)}{d^2} \]

Where:
- \( N \) = sample size
- \( Z \) = \( Z \) statistic for a level of confidence
- \( P \) = Expected prevalence or proportion (in proportion of one; if 20%, \( P = 0.2 \))
- \( d \) = precision (in proportion of one; if 5%, \( d = 0.05 \))

4. **Were the study subjects and setting described in detail?**

Certain diseases or conditions vary in prevalence across different geographic regions and populations (e.g. women vs. men, sociodemographic variables between countries). Has the study sample been described in sufficient detail so that other researchers can determine if it is comparable to the population of interest to them?

5. **Is the data analysis conducted with sufficient coverage of the identified sample?**

A large number of dropouts, refusals or “not founds” amongst selected subjects may diminish a study’s validity, as can low response rates for survey studies.

- Did the authors describe the reasons for non-response and compare persons in the study to those not in the study, particularly with regards to their socio-demographic characteristics?

- Could the not-responders have led to an underestimate of prevalence of the disease or condition under investigation?

- If reasons for non-response appear to be unrelated to the outcome measured and the characteristics of non-responders are comparable to those in the study, the researchers may be able to justify a more modest response rate.

- Did the means of assessment or measurement negatively affect the response rate (measurement should be easily accessible, conveniently timed for participants, acceptable in length and suitable in content).

6. **Were objective, standard criteria used for measurement of the condition?**

Here we are looking for measurement or classification bias. Many health problems are not easily diagnosed or defined and some measures may not be capable of including or excluding appropriate levels or stages of the health problem. If the outcomes were assessed based on existing definitions or diagnostic criteria, then the answer to this question is likely to be yes. If the outcomes were assessed using observer reported, or self-reported scales, the risk of over- or under-reporting is increased, and objectivity is compromised. Importantly, determine if the measurement tools used were validated instruments as this has a significant impact on outcome assessment validity.
7. **Was the condition measured reliably?**

Considerable judgment is required to determine the presence of some health outcomes. Having established the objectivity of the outcome measurement instrument (see item 6 of this scale), it is important to establish how the measurement was conducted. Were those involved in collecting data trained or educated in the use of the instrument/s? If there was more than one data collector, were they similar in terms of level of education, clinical or research experience, or level of responsibility in the piece of research being appraised?

- Has the researcher justified the methods chosen?
- Has the researcher made the methods explicit? (For interview method, how were interviews conducted?)

8. **Was there appropriate statistical analysis?**

As with any consideration of statistical analysis, consideration should be given to whether there was a more appropriate alternate statistical method that could have been used. The methods section should be detailed enough for reviewers to identify the analytical technique used and how specific variables were measured. Additionally, it is also important to assess the appropriateness of the analytical strategy in terms of the assumptions associated with the approach as differing methods of analysis are based on differing assumptions about the data and how it will respond. Prevalence rates found in studies only provide estimates of the true prevalence of a problem in the larger population. Since some subgroups are very small, 95% confidence intervals are usually given.

9. **Are all important confounding factors/ subgroups/differences identified and accounted for?**

Incidence and prevalence studies often draw or report findings regarding the differences between groups. It is important that authors of these studies identify all important confounding factors, subgroups and differences and account for these.

10. **Were subpopulations identified using objective criteria?**

Objective criteria should also be used where possible to identify subgroups (refer to question 6).
Appendix 2
JBI Data Extraction Form

JBI Data Extraction Form for Prevalence and Incidence Studies

Study details
Reviewer –
Study ID/Record Number -
Date –
Study title –
Author –
Year –
Journal –
Aims of the study –

Study Method
Setting –
Study design –
Follow-up or study duration –
Subject characteristics –
Dependent variable -
Outcomes –
Outcome measurements –
Ethical approval –
Method of data analysis -

Results
Prevalence n/N (%)
Proportion and 95% Confidence Intervals

Incidence n/N (%)
Proportion and 95% Confidence Intervals and duration of recruitment or the study

Authors’ comments

Reviewer comments

Copyright © The Joanna Briggs Institute 2014