Methodological guidance for systematic reviews of observational epidemiological studies reporting prevalence and cumulative incidence data

Zachary Munn PhD, Sandeep Moola MClinSc, Karolina Lisy PhD, Dagmara Riitano BHSci (Hons) and Catalin Tufanaru MD

Faculty of Health Sciences, The Joanna Briggs Institute, The University of Adelaide, Adelaide, South Australia, Australia

ABSTRACT

Aim: There currently does not exist guidance for authors aiming to undertake systematic reviews of observational epidemiological studies, such as those reporting prevalence and incidence information. These reviews are particularly useful to measure global disease burden and changes in disease over time. The aim of this article is to provide guidance for conducting these types of reviews.

Methods: A methodological working group of the Joanna Briggs Institute, Adelaide, South Australia, Australia, was formed to create guidance for conducting systematic reviews of studies reporting prevalence and cumulative incidence information. All methodological output of the group was subject to peer review and feedback by members of the international evidence synthesis community.

Results: Systematic reviews of prevalence and incidence data should follow the same structured steps as systematic reviews of effectiveness. However, many of these steps need to be tailored for this type of evidence, particularly surrounding the stages of critical appraisal and synthesis.

Conclusion: Prevalence and incidence systematic review and meta-analysis is an emerging methodology in the field of evidence synthesis. These reviews can provide useful information for healthcare professionals and policymakers on the burden of disease, show changes and trends over time in disease, and inform geographical distributions of disease and conditions.

Key words: epidemiology, incidence, observational, prevalence, systematic review


Introduction

The accurate measurement of disease among populations, whether at a local, national, or global level, is of critical importance for governments, policy makers, health professionals, and the general population to inform the development, delivery, and use of health services. There are a number of measurements of disease, including determining the proportion of a population who have a certain disease at a specific point in time (the point prevalence), or within a given time frame (the period prevalence), and how often new cases of a disease occur within a given time frame (the incidence).1 This information can assist in planning management of disease services (by ensuring resources are available to cope with the burden of disease), set priorities regarding public health initiatives, and evaluate changes and trends in diseases over time. As such, systematic reviews of prevalence and incidence data are becoming increasingly important as policy makers acknowledge the usefulness of syntheses of this type of information. These syntheses have the potential to better inform social and health professionals, policy makers, and consumers in making decisions related to a range of healthcare issues, but particularly regarding the burden of disease both now and into the future.
When faced with a question to answer, one of the first decisions a researcher needs to consider is the most appropriate methodology and study design to answer that question. For questions of prevalence, cross-sectional studies using probabilistic random samples conducted in the relevant geographical setting or location are the most appropriate study design, whereas questions of incidence are best answered in cohort studies. However, when the conduct of such primary studies is not feasible or possible, a systematic review of the existing data already collected may be the most appropriate approach in order to provide at least some indication of the likely magnitude of the problem in absence of primary data. A synthesis may also be useful when primary studies have already been conducted in local settings and the aim is to provide a national or global overview of prevalence of a specific variable. Systematic reviews of prevalence or incidence data can therefore assist in answering questions of national or global disease burden; help to measure global disease burden (i.e., incidence data can be used to determine disability adjusted life years); show changes and trends over time (highlighting emerging or decreasing diseases); inform geographical distributions of disease and conditions; inform healthcare professionals of prevalence and incidence of diseases and symptoms of disease; and inform further research priorities. These types of reviews may not only address the prevalence or incidence of disease, but also include many types of variables such as symptoms, conditions, events, practices, behaviors, or factors. The following selected review topics show the diverse focus of reviews assessing prevalence and incidence, for example, prevalence and incidence of medication errors, worldwide prevalence of attention-deficit hyperactivity disorder, clausrophobic reactions in MRI, barriers to adherence with treatment, incidence of schizophrenia, and electronic health record adoption.

Recently, there has been a substantial increase in the number of published systematic reviews of prevalence and incidence data. Across these reviews, there is considerable variation in terms of the approaches to inclusion criteria, critical appraisal, data extraction, and data synthesis. This is likely due to the current lack of formal guidance to inform the conduct of systematic reviews of these types of data. In 2012, a working group was formed within the Joanna Briggs Institute, Adelaide, South Australia, Australia, to evaluate systematic reviews of prevalence and incidence data. The aim of this article is to describe the development of the Joanna Briggs Institute approach to these types of systematic reviews.

Methods
A methodology working group was formed to investigate and develop guidance for conducting systematic reviews of prevalence and incidence data. The working group, composed of researchers from the Joanna Briggs Institute, was experienced in conducting systematic reviews of various methods. The working group met monthly to discuss, define, and develop methods for reviews of these types of data. Consensus was reached through discussion and testing. In August 2013, the newly proposed methodology was presented to the international members of the Scientific Committee of the Joanna Briggs Institute for further consideration, discussion, and, ultimately, approval. Following this, it was ratified at an Institute board meeting. In October 2013, the methodology was presented in a workshop during the Joanna Briggs Institute Convention, allowing international colleagues an opportunity to provide critique and feedback on the process that had been devised. In addition, the methodology was presented to the Joanna Briggs Institute Committee of Directors, comprising over 90 international experts in research synthesis from over 20 countries for further discussion and feedback. A cyclic process of feedback and review was used at all stages of the development process.

Results
The group was successful in developing guidance for conducting prevalence and incidence systematic reviews. The systematic review is an accepted type of research design that provides a reliable summary of the literature to assist health professionals to keep up to date. Key features of a systematic review include the creation of an a-priori protocol, clear selection criteria to guide inclusion of studies in the review, a structured and systematic search process, critical appraisal of studies, and a formal process of data extraction followed by methods to synthesize or combine these data. In this way, systematic reviews extend beyond the subjective, narrative reporting characteristic of a traditional literature review to provide a comprehensive, rigorous, and transparent synthesis of the literature on a certain topic. Systematic reviews of prevalence and incidence data follow the same structured process as systematic reviews of effects, risks or harms.

Inclusion criteria
The traditional approach to inclusion criteria for systematic reviews of effects is to follow the population, intervention, comparator, and outcome structure. However, this approach clearly does not align with questions relating to prevalence and incidence. As such, when
determining the inclusion criteria for these types of reviews, we recommend the CoCoPop mnemonic (condition, context, and population) for reviews assessing prevalence and incidence data.

**Population**
The population should be appropriate in order to satisfy the review objectives. Sufficient explanation of reasons for the inclusion or exclusion of participants should be given in the background such that it is transparent to the reader. It is important that the population or study participants are clearly defined and described in detail. This includes outlining the specific or defining characteristics of the population, such as age, sex, ethnicity, educational status, employment status, individual behavior, and sociodemographic factors. An example statement follows: ‘we will include studies involving adult pregnant women aged 18–45 years at any trimester up to delivery and for 6 months post-birth’. Exclusion criteria should also be outlined wherever relevant. For example, studies examining pregnancies with neural tube defects, intrauterine 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, behavior, practice, or factor. It is important that the variable of interest is clearly stated and defined. For example, malaria could alternatively be stated as *Plasmodium falciparum* infection, *P. vivax* infection, or disease due to malarial infection. Description of the variable of interest may include information regarding how it will be measured, diagnosed, or confirmed.

**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 of interest, a specific community or setting (e.g., inpatient vs. outpatient, or urban vs. rural), or the time period to be considered, given that some conditions may peak at a particular season (e.g., the incidence of influenza in different seasons and years).

**Critical appraisal**
Critical appraisal (or assessment of methodological quality/risk of bias) is a process conducted in systematic reviews to establish the validity and risk of bias of studies identified during the systematic search. A proportion of reviews assessing prevalence and incidence do not include a process of critical appraisal, but instead include ‘minimum criteria’ for inclusion. However, it is the approach of the Joanna Briggs Institute to conduct critical appraisal in all systematic reviews, regardless of the type of data under investigation. The Joanna Briggs Institute have a number of tools already developed for assessing the quality of various quantitative study designs, and these tools were deemed appropriate to use when assessing questions of incidence. However, these tools were not considered appropriate for assessing questions of prevalence, wherein the nature of the data to be extracted demands a different set of quality criteria, and as such, a new tool was developed specifically for studies reporting prevalence data. In 2013, the critical appraisal checklist developed for systematic reviews of prevalence was pilot tested. Based on this pilot test and further refinements, a revised tool was created (Box 1).

**Data synthesis and meta-analysis**
As with all systematic reviews, there are various approaches to present the results, including a narrative, graphical, or tabular summary, or a meta-analysis. Cumulative graphs or forest plots (with or without a pooled estimate) may be particularly useful for displaying how prevalence and incidence estimates vary between studies, across populations, and over time.

Meta-analysis can be conducted to determine a single estimate of disease frequency and its variance across studies. This has led some to question whether meta-analysis of this type of data is appropriate, as it is likely that studies conducted in different populations will report significantly different prevalence and incidence estimates because of the characteristics of that population; as such providing an average, overall estimate may be of little use. As Saha et al. state, ‘we were not interested in a single, pooled value – an exercise in reductionism that sacrifices potentially informative variation’. This may well be true if the reason for conducting a systematic review (based on funding or policy reasons) is to inform the likely prevalence or incidence of a condition for a certain population, such as Australia, when there are no primary studies conducted within Australia. In this case, it may be more useful to conduct a systematic search to find studies that report prevalence and incidence that have similar population characteristics as the Australian population. The aim of such a review may then simply be to identify the most valid study (or studies) conducted within a setting relevant to the Australian context (or the population the review is
### Box 1 Joanna Briggs Institute’s critical appraisal checklist for studies reporting prevalence data

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
<th>Unclear</th>
<th>Not applicable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Was the sample frame appropriate to address the target population?</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Were study participants sampled in an appropriate way?</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>3.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Was the sample size adequate?</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>4.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Were the study subjects and the setting described in detail?</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>5.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Was the data analysis conducted with sufficient coverage of the identified sample?</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>6.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Were valid methods used for the identification of the condition?</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>7.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Was the condition measured in a standard, reliable way for all participants?</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>8.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Was there appropriate statistical analysis?</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>9.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Was the response rate adequate, and if not, was the low response rate managed appropriately?</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>
interested in), and meta-analysis may be of little use in this case. We label this the ‘best and brightest’ approach to systematic reviews of prevalence and incidence, and these may be performed in order to inform primary studies of prevalence and incidence.

However, for systematic reviews with different objectives, such as estimating the global burden of disease, or with the aim to summarize prevalence or incidence across a certain jurisdiction (i.e., when there are multiple studies within a country and the aim is to provide a single summary estimate for that country), meta-analysis may well be appropriate. In these cases, ‘meta-analysis involving a standardized recalculation of rates from many previous studies should generate a distribution allowing one to estimate with some probability the extent to which populations differ’. If risk adjustment is sought (and not meta-analysis to establish a ‘true estimate’), then the procedure for risk adjustment or standardization based on introducing subpopulation weights into a meta-analytic procedure, as described by Doi et al., can be considered.

Prevalence and cumulative incidence data are normally reported as a proportion, and the following section discusses meta-analysis of proportional data. When pooling proportions for meta-analysis, a transformation of the data is required. There are two main ways to transform the data, the Freeman–Tukey transformation (arc sine 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 a pooled proportion with a 95% confidence interval (CI) both for the fixed and the random-effects model, and will additionally list the proportions (expressed as a percentage), with their 95% CI, found in the individual studies included in the meta-analysis. The results may then be graphically presented 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. The logit transformation addresses the problem of CI estimates falling outside the 0–1 range. However, this approach does not succeed in stabilizing the variance, and as this is the case, Barendregt et al. recommend the double arc sine transformation as the preferred approach to meta-analysis. This approach solves the problem of variance instability in addition to addressing the problem of confidence limits falling outside the 0–1 range. For further information regarding meta-analysis and formulas, we suggest readers to refer to the article by Barendregt et al.

There are different models for performing the meta-analysis as mentioned earlier. We recommend that meta-analyses of the prevalence or cumulative incidence reported in studies is grouped by a random-effects model and presented with 95% CI. The random-effects model allows for between-study variation by assuming that individual study prevalence estimates follow a normal distribution. The fixed-effects 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 estimate, which may not hold for prevalence and incidence data.

It is important to note that there are emerging developments in approaches to meta-analysis. Stanley and Doucouliagos recommend the routine use of a simple unrestricted weighted least squares estimator that corrects for the poor coverage of the fixed-effect estimator and has been shown to be useful over random-effects meta-analysis estimates when there is either publication selection or small-sample bias. Doi et al. propose an inverse variance heterogeneity model of meta-analysis, which they recommend should replace the traditional fixed and random-effect models. As further methodological work is conducted in this area, recommendations regarding models and approaches to meta-analysis of prevalence and incidence data are likely to evolve.

Heterogeneity of the results can be explored by the $\hat{I}^2$, $i^2$, and Cochran’s $Q$ test ($X^2$) tests. 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, and study population) to the overall heterogeneity. Those variables significantly associated with the heterogeneity ($P < 0.05$) can be included in a multivariate hierarchical model. The common value of $P$ less than 0.05 is usually employed for the rejection of the null hypothesis (i.e., there is a statistically significant result).

It is suggested that $i^2$, rather than $\hat{I}^2$, be considered as the appropriate measure in assessing clinically relevant heterogeneity, particularly when deciding whether or not to pool estimates in a meta-analysis. Rücker et al. suggest that statistical test of heterogeneity should never be the basis when deciding between a fixed and a random-effects meta-analysis.

Funnel plots are often presented in systematic reviews with meta-analysis as a way of assessing publication bias. At present, there is a lack of published research or guidance regarding the use of funnel plots in meta-analyses of proportional data. However, one study found that conventional funnel plots appear inaccurate when there are either low or high proportion outcomes. As such, funnel plots ‘using sample size as
the measure of accuracy instead of the inverse of the standard error on the y-axis' were viewed favorably.22

An example meta-analysis is presented in Fig. 1. In this example, seven studies were included in the meta-analysis, with the overall effect size equaling 0.0141 (95% CI 0.0134–0.0149). Therefore, the prevalence of the condition can be reported as 1.4%.

Discussion
Historically, the focus of systematic reviews has largely been on the effectiveness of interventions or practices on social and health outcomes. However, decisions made in healthcare require more information than can be provided by the simple question ‘does this work?’ As such, there now exist methods and guidance for conducting reviews of various forms of evidence, including qualitative research, cost data, diagnostics, prognostics, harms, and risks.13,23 The official guidance summarized here and provided by the Joanna Briggs Institute appears to be the first of its kind for reviews of prevalence and incidence. It is hoped that by providing formal guidance for these types of reviews, the inconsistencies and variation in the conduct and rigor of reviews of these types will be improved.

Systematic reviews of prevalence and incidence may not always be the most appropriate approach for answering questions of prevalence and incidence. Systematic reviews of prevalence and incidence data may be plagued by issues related to differences between included studies, such as differences in study design, instruments or measures of disease, or geographical locations. For example, different measurement instruments or classifications or definitions of disease across both space and time (i.e., with the evolution of the International Classification of Disease) may result in it being inappropriate to combine studies statistically in a meta-analysis. Genuine differences in prevalence are very likely across different populations and settings, for example, in the case of one review assessing the prevalence of attention-deficit hyperactivity disorder globally, higher rates were found in North America than in Africa and the Middle East.3 However, it is our opinion that when a systematic review of these types of data is conducted rigorously and data analyzed appropriately, these types of reviews can be of great benefit to healthcare professionals, policy makers, and funders.

Conclusion
Prevalence and incidence systematic reviews are an emerging methodology in the field of evidence synthesis. These reviews can provide useful information for healthcare professionals and policy makers on the burden of disease, show changes and trends over time in disease, and inform geographical distributions of disease and conditions. The standardized approach outlined here offers a rigorous and transparent method to conduct these reviews.
Acknowledgements

Conflicts of interest

Z.M. lead the methodological group and drafted the article. S.M., D.R., and K.L. were members of the working group and provided substantial input regarding its development and testing. C.T. provided additional advice and support.

There are no conflicts of interest.

References