The Role of Systematic Reviews and Meta-Analyses of Incidence and Prevalence Studies in Neuroepidemiology

K.M. Fiest a–c  T. Pringsheim a–c  S.B. Patten a, b  L.W. Svenson a, d, e  N. Jetté a, c

a Department of Community Health Sciences and Institute for Public Health, b Department of Psychiatry and Mathison Center for Mental Health Research and Education, and c Department of Clinical Neurosciences and Hotchkiss Brain Institute, University of Calgary, Calgary, Alta., Canada

Abstract

Background: Systematic reviews and meta-analyses on the incidence and prevalence of neurological conditions are important methods of quantifying the burden and risk of disease. Methods: The rigorous methodology required in order to minimize publication bias, account for study heterogeneity, and variation in study quality are described. When appropriate, a meta-analysis is a powerful statistical tool that can help synthesize a vast literature quantitatively, taking into account study heterogeneity. As the epidemiology of neurological conditions continue to be widely studied internationally, systematic reviews and meta-analyses have become essential. Results: If not conducted carefully, systematic reviews and meta-analyses in neuroepidemiology may lead to erroneous conclusions. It is important to consider various methodological, clinical and statistical factors at all stages of the review and analysis process. Detailed documentation should be kept to assist in the reporting process. Conclusions: Published reporting standards should be consulted when conducting systematic reviews and meta-analyses of the incidence and prevalence of neurological conditions, though reporting standards specific to neuroepidemiology are urgently needed.

Key Words
Epidemiology · Systematic review · Meta-analysis · Meta-regression

Introduction

Growing attention has been paid to the importance of neuroepidemiology – the study of the distribution and determinants of neurological conditions [1]. By describing the incidence, prevalence, comorbidities, mortality and burden of disease, neuroepidemiology has become increasingly relevant to the practice of clinical neurology, public health and health policy [1, 2]. After disease diagnosis and definition, the most commonly asked question in neuroepidemiology is: how common is a disease [2]? It is possible to quantify the frequency of neurological diseases using measures such as incidence and prevalence; in a defined population, incidence is the number (often expressed as a proportion or a rate) of new disease cases in an at-risk population and prevalence is the number of existing cases in that population [3]. Cumulative inci-
dence, or incidence proportion, is calculated by dividing the number of new cases in a population by the total number of at-risk subjects in a defined period of time [3]. The incidence rate is calculated differently, by dividing the number of new cases in the population by the total amount of person-time at risk (typically person-years), again during a defined period of time [3]. Incidence rates can accommodate different periods of follow-up caused by varying times of study withdrawals or entrances [3].

One method of compiling, reviewing and evaluating the wealth of information from these measures are through systematic reviews and meta-analyses [4]. A systematic review is a method of reviewing and synthesizing research findings, and is designed to minimize biases and errors in the review process [4]. A meta-analysis involves a quantitative synthesis of the results of systematically reviewed studies [4]. Meta-analyses may not always be appropriate, though a systematic and rigorous approach to reviewing the literature is always important.

Systematic reviews and meta-analyses are powerful and important tools in neuroepidemiology, as a means of integrating and evaluating vast amounts of primary research. One of the most important aspects of a systematic review is its transparency – reviews should be easily reproducible from the methodology reported by the authors [4].

This review will: (1) describe the clinical relevance of systematic reviews and meta-analyses of incidence and prevalence studies in neurological conditions; (2) describe the methodological importance of systematic reviews and meta-analyses in this context, and (3) make recommendations for conducting systematic reviews and meta-analyses of incidence and prevalence studies in neurological conditions.

**Clinical Importance/Relevance**

Systematic reviews and meta-analyses have great utility in neuroepidemiology, particularly for clinical research. Individually, studies may not produce reliable or robust estimates of incidence or prevalence; by pooling multiple studies, it may be possible to see otherwise unidentified associations and to estimate associations with greater precision. Bias should always be considered separately from precision – possible mechanisms of bias should be investigated and elucidated. The overall burden of disease may be evaluated globally, and estimates from different regions of the world with similar characteristics combined (e.g. developing vs. developed countries). Variations between regions may become more robust when pooled and detailed subgroup analyses may be possible when a larger number of studies are identified. The pooling of multiple studies of rare neurological conditions can create more precise estimates of incidence and prevalence, and possibly allow for otherwise unavailable subgroup analyses.

A systematic review on the incidence and prevalence of epilepsy pooled and compared heterogeneous effect estimates through meta-analysis [5]. The incidence of epilepsy differs across many variables, such as location – this meta-analysis found a higher incidence of epilepsy in low- and middle-income countries. Meta-regression, a statistical method used to assess trends across extraneous variables, was employed to explore differences across study populations (population vs. hospital based), study design (prospective vs. retrospective) and years of data collection. The meta-regression also found that several factors were associated with heterogeneity in the study estimates, including study region (high-income vs. low-income countries) and study design (retrospective vs. prospective), for example. These comparisons are powerful, and would be difficult to make without meta-analysis.

Oskoui et al. [6] examined the prevalence of cerebral palsy in a recent systematic review. Through meta-analytic methods, the worldwide prevalence of cerebral palsy was estimated and important between-group differences explored. An important finding, recognized through meta-regression, is that although survival of at-risk pre-term infants improved, the prevalence of cerebral palsy per live birth has remained constant over time. Comparisons were also made according to birth weight, which is an important clinical factor that may drive heterogeneity; differences were found across this variable that may have been too small to detect in a single study. This meta-analysis provided very precise estimates, and used recommended standards to ensure the highest quality reporting.

A systematic review investigating the epidemiology of Alzheimer’s disease and dementia in China was recently published [7]. The authors performed a meta-analysis of over 300,000 individuals and identified previously unseen trends in the prevalence and incidence of dementia, particularly that the risk of dementia is increasing faster than earlier, smaller studies indicated. The authors concluded that smaller, individual studies may have underestimated the prevalence and incidence of dementia.

Recent systematic reviews and meta-analyses have explored the incidence and/or prevalence of movement disorders, including Huntington’s disease [8], tic disorders.
Risk factors for neurological conditions may be established and included in further analyses. In a systematic review and meta-analysis of multiple sclerosis, Simpson et al. [12] reported latitude to be a statistically significant factor associated with the prevalence of multiple sclerosis. Kwok et al. [13], using meta-analysis, found that individuals with atrial fibrillation were at greater risk of developing dementia – this risk was greatest in people who have also experienced a stroke.

Results from systematic reviews may also provide data that are essential for resource planning and interventions. Chan et al. [7] demonstrated that the burden of dementia was increasing faster than was generally assumed by the international community, which will have a great impact on planning and resource allocation in all areas of the world. The authors suggested introducing specialized care to target, prevent, diagnose and treat dementia; suggestions made possible through the power of meta-analytic techniques. The proposals of this review are not limited to dementia, and can be investigated and enacted for all neurological studies using meta-analysis as a justification.

Systematic reviews (without meta-analysis) are also adding to the body of literature on the incidence and prevalence of neurologic conditions, expanding the understanding of risk factors, clinical correlates and disease outcomes. For example, systematic reviews focusing on specific world regions were recently published [14–17]. These studies explored factors of heterogeneity that are specific to these areas, and provide important recommendations that may be used when conducting larger meta-analytic reviews. The prevalence of multiple sclerosis is known to vary across geographic regions, though this factor alone cannot predict the risk of disease [15]. The authors also examined factors such as methods of case ascertainment and diagnosis, which may differ in areas with varying resources [15].

Reviews of neurological disease have specific challenges that distinguish this field from others in medicine. There are a substantial number of neurological conditions, and within those conditions heterogeneous subtypes, for example hundreds of epilepsy syndromes [18], many subtypes of multiple sclerosis (relapsing remitting, primary progressive, secondary progressive) [19] and a broad range of muscular dystrophy syndromes [20]. Unlike other medical conditions (e.g. lab tests for renal failure, diabetes) there are no gold standard tests for diagnosing most neurological disorders, and this must be taken into consideration before conducting meta-analyses. However, meta-analyses can also be very helpful in neuroepidemiological research. Most neurological conditions are very rare and thus pooling of many smaller studies may increase the precision of incidence or prevalence estimates. In addition, these meta-analyses may help uncover factors associated with each condition that may not be elucidated with smaller study sizes. Risk, geographical and environmental factors of neurological conditions can be described through meta-analyses; for example, there is a prominent geographic gradient in multiple sclerosis [12], a specific subtype of amyotrophic lateral sclerosis/parkinsonism dementia complex in Guam [21] and the negative association between smoking and Parkinson’s disease [22].

Methodological Importance/Relevance

It is necessary to discuss the methodological factors that must be considered when conducting systematic reviews and meta-analyses of incidence and prevalence studies in neuroepidemiology. Figure 1 outlines the steps involved in a systematic review and meta-analysis, which are discussed below. To reduce the possibility of bias, it is very important to pay close attention to numerous aspects of papers included in any systematic review of neuroepidemiological studies. For example, the study objectives and design, sample size, treatment and/or dosage used, control selection criteria, outcomes measured, length of follow-up and study quality should be critically evaluated for each included study.
**Search Strategy and Literature Search**

The first step in designing a high-quality neurological systematic review is developing a comprehensive literature search strategy. A clear and concise research question is the driver of the search strategy. A medical librarian should be consulted when developing the search, along with neurologists (e.g. multiple sclerosis), neurosurgeons (e.g. spinal cord injury), neuroepidemiologists and/or statisticians, and keywords should be generated and applied. The search approach should be guided by content knowledge about the condition and methodological understanding. The search should be carried out in several online databases to ensure all articles are found and to avoid selection bias, and the number of limits (year, language of publication) placed on the search should be minimized. For example, in a review of cerebral palsy [6] the authors used multiple databases and consulted a medical librarian, and the search was also made available in an online appendix. Previous systematic reviews for the neurological condition of interest (to identify key clinical terms) or in epidemiology (to identify search terms related to incidence and/or prevalence) are also an excellent resource when developing search strategies. Once the search is completed, abstracts should be indexed in bibliographic management software. Duplicates in general may be easily removed using such software, and the final list of articles can then be categorized. A large sample of articles may result from literature searches, which can be problematic if human resources are limited. Translation of foreign language articles may further delay progress, but the importance of reviewing articles in every language also limits selection bias. Some articles can be translated using free online software, but often additional collaborators who speak the languages of interest are required.
Abstract and Full-Text Review

The abstract and full-text review stages should be conducted thoroughly, with detailed records kept throughout. Criteria for abstract and full-text article inclusion should be established and clarified before reviewing begins. It is important to clearly define the population of interest (e.g., all ages, children only, etc.), any relevant study designs (case control, cross-sectional, etc.), settings (population based, registry based) and the exposures or outcomes of interest before abstract review begins. At least 2 people should independently review each title and abstract resulting from the search. Abstracts selected for inclusion by either reviewer will move onto the full-text review stage. Recent research has suggested that screening only the title of an abstract may be more efficient than screening both the title and abstract – this approach may also be considered [23]. It is possible that more stringent inclusion/exclusion criteria be applied in the full-text screening stage, to allow for the inclusion of articles using specific measures or tests. At this stage, consensus on articles moving on to data abstraction must be reached; if reviewers disagree on which articles to include, they should first discuss the discrepancies and attempt to achieve consensus. If consensus is not possible, a third reviewer should resolve any remaining discrepancies at the full-text stage. Agreement between reviewers during the review process (abstract review and full text review) can be assessed using the kappa statistic, which can be used to measure beyond-chance agreement on the inclusion of articles [24]. During the full-text review, reasons for article exclusion should be documented to assist in creating the study flow diagram [4].

Data Abstraction

After all full-text articles have been reviewed, data should be abstracted, regardless of whether or not a meta-analysis is the final goal. Clinical input should be garnered to ensure important disease-specific variables are abstracted from the articles, such as disease severity indices or subtypes, if available. Other variables to abstract include study characteristics, such as sample size, years of data collection and study location. Items for assessing study quality (see next section) should also be abstracted at this stage. A pre-designed data abstraction form, whether in a spreadsheet or online tool, is essential to ensure consistency and methodological rigour. The form should be tested by both reviewers using a sample of articles (e.g., 2–5 articles) to ensure it is complete, easy to use and results in consistent abstraction between reviewers (e.g., all variables are clearly defined for the reviewers). The form should be modified if necessary before proceeding with abstraction of the rest of the articles. As with the full-text review, consensus must be reached during the data abstraction phase. If this is not possible, a third reviewer should resolve any disagreements.

Study Quality

The quality of individual studies may be assessed using published checklists and scales [25, 26]. There is great technical quality, data integrity and methodological rigour in properly conducted systematic reviews and meta-analyses, which are enhanced by the use of recommended study guidelines. The individual studies included in a review may vary in terms of their respective quality, which can be defined in many ways, but generally relates to the relative merits of each study. The most important aspects of study quality for systematic reviews of incidence and prevalence are the quality of reporting and the extent to which information is provided on the design and methodology of the study. Several methods for evaluating study quality exist, including the Jadad score for evaluating data from randomized controlled trials [27], guidelines by Boyle [25] on evaluating prevalence studies, and a rubric for determining the quality of health research literature [26]. Assessments of study quality can be used initially to assess study eligibility or directly included in meta-analyses through a sensitivity analysis, and can be used as a descriptive factor to explain heterogeneity. Study quality indices should be used with caution, however, as high overall scores may be present in spite of methodological flaws – several aspects of quality should be taken into account when interpreting scales and checklists. Evans et al. [15] reported the quality of all studies included in their review of the incidence and prevalence of multiple sclerosis; factors assessed included response rate, sample representativeness and the use of standardized data collection methods. A common concern with study quality is whether or not to include a lot of studies with a lower bar for quality, or a smaller number of studies judged to be of the highest quality – this choice will be guided by the research question and study objectives.

Heterogeneity and Meta-Analysis

Study estimates may differ clinically or statistically, based on methodological characteristics or based on their variance. It is important to assess both clinical and statistical heterogeneity when conducting systematic reviews and meta-analyses. Content-area experts should
be consulted when determining clinical heterogeneity – groups that differ according to important clinical characteristics, such as sex, age and disease severity should be examined separately, regardless of their statistical significance. If possible, standardizing the results of incidence or prevalence studies by age or sex should be considered as a first step, to enhance the comparability of results across studies. To assess significant between-study statistical heterogeneity, the Cochrane Q statistic should be calculated and I² used to quantify the magnitude of between-study heterogeneity [28]. The Q statistic provides a p value indicating the presence of heterogeneity and the I² statistic provides information on the percentage of heterogeneity that exists. Steeves et al. [10] used this method to quantify the amount of heterogeneity present in a recent review on the prevalence of primary dystonia.

A number of factors may determine whether a meta-analysis is conducted after a systematic review, including study heterogeneity. Two different modeling approaches are available for the pooling of study estimates through meta-analysis: fixed and random effects. A fixed-effect model assumes that one effect size is shared by all of the included studies, whereas a random-effects model allows for variance and distribution in the true effect size of studies. The weights assigned to each study differ in the two models. In a random-effects model, large studies will not overshadow smaller study effects to the same extent as in a fixed-effects model. If significant heterogeneity is present, a random-effects model should be chosen over a fixed-effects model – it should be noted that fixed-effects models may not be appropriate even if there is low heterogeneity. Random-effects models give more weight to smaller studies, which can be influenced more strongly by biases inherent in small studies. Exploring and detailing potential sources of heterogeneity is one the most important aspects of a meta-analysis – the use of random-effects models without explaining the heterogeneity is discouraged. Regardless of whether a meta-analysis is conducted, the study estimates should be plotted visually using a forest plot, which is a graphical depiction of all study estimates and their corresponding confidence intervals. For examples of forest plots for systematic reviews of incidence and prevalence of neurological conditions, see Pringsheim et al. [8], Steeves et al. [10] and Knight et al. [9].

**Potential Sources of Heterogeneity**

**Sources of Case Ascertaintment.** Methods of case ascertainment vary greatly in neuroepidemiology, with studies using methods such as direct clinical assessment in the community, mail or telephone surveys, registry data, administrative data or medical charts to determine incidence or prevalence [6, 15]. This may present difficulties when pooling studies, as they have differing abilities to capture cases. Most studies use multiple sources of data to ascertain cases, making pooling of these studies more difficult. It is important for studies to accurately report how they identified cases to facilitate categorization and comparison.

**Diagnostic Criteria.** Methods of disease diagnosis are an important factor to consider when combining study data, particularly in neurological conditions where many sets of diagnostic criteria exist. For example, in dementia many different criteria exist for diagnosis: the DSM (Diagnostic and Statistical Manual)-III, DSM-III-R, DSM-IV, DSM-V, ICD (International Classification of Disease)-9 and ICD-10. Erkinjuntti et al. [29] examined the effect of differing diagnostic criteria on the prevalence of dementia; prevalence estimates ranged from 3.1 to 29.1%. Some neurological conditions may lack diagnostic criteria, as the presentation is very heterogeneous, such as in epilepsy, which can present with multiple seizure types and/or with varied epilepsy syndromes [30]. Other conditions such as multiple sclerosis have well-established diagnostic criteria, although even these criteria may be modified over time [31]. In addition to evolving diagnostic criteria, changes in technology (e.g. the introduction of MRI in the 1980s) may make it difficult to compare studies over time; newer technologies may bias incidence estimates upward, when in fact the increase is merely an artifact of improved disease identification. Varying diagnostic criteria may make the pooling of results in a meta-analysis difficult and it is often advisable to avoid combining studies using different criteria. In those cases, it can be helpful to plot estimates over time to see the potential impact of new diagnostic criteria on disease estimates.

**Stratification of Results**

Calculating results based on levels of an extraneous variable is a powerful tool available in the meta-analysis of incidence and prevalence studies. The frequency of many neurological conditions differs based on demographic factors, and clinical judgment should be used when determining variables for stratification. For example, it is well known that the incidence of Parkinson’s disease [32] and dementia [33] increase with age, making age an important factor for stratification. By stratifying prevalence or incidence estimates on another factor,
such as age or sex, it is possible to see trends that would be obscured by pooling overall estimates [9, 11, 16]. An exploration of stratified and overall pooled results should be considered for meta-analyses of neurological conditions.

**Publication Bias**

An assessment of publication bias is necessary when conducting systematic reviews and meta-analyses to determine if any unpublished studies are missing from the initial systematic review search [4]. Large, statistically significant studies are more likely to be published, more likely to be cited and more likely to produce multiple publications, making them more likely to be captured in a systematic review [4]. The most common analyses of publication bias visually assess funnel plot asymmetry; funnel plots are a scatter plot of treatment effects and study size [4]. An asymmetric funnel plot would indicate the absence of studies of a certain size and magnitude [4]. Funnel plots may also be employed to assess possible biases in the identification and study selection phases, and should be produced whenever possible. Statistical tests for quantifying funnel plot asymmetry exist, including Begg’s and Egger’s tests [4]. Funnel plot asymmetry does not necessarily indicate the presence of publication bias; selection bias, true heterogeneity, data irregularities and chance may also explain an asymmetric funnel plot [4]. Methods such as the trim and fill should be considered as a sensitivity analysis if publication bias is present – the trim and fill method identifies funnel plot asymmetry by imputing the effect estimates of potentially missing studies and assessing the influence of these studies on the pooled estimate [34]. Simpler methods, such as ordering by study size, can also be informative; regardless of the method, sensitivity analyses are extremely important when dealing with publication bias.

**Meta-Regression**

Meta-regression is a powerful tool to assess for trends across an extraneous variable that is a study-level characteristic. For example, the incidence of a disease may be changing over time; in order to assess the effects of time, a meta-regression can be conducted across years. Meta-regression analyses produce a p value, the significance of which determines the difference in groups across that variable. Meta-regression allows for the inclusion of measured or continuous variables, where categorization is unnecessary and important between-group differences may be uncovered.

**Recommendations**

Reporting of results in primary research and systematic reviews and meta-analyses should be standardized to allow for comparisons between studies. When reporting on primary research data, authors should provide detailed information on methodology, to facilitate the evaluation and pooling of results. Use of the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement should be considered when conducting and reporting on systematic review data of randomized controlled trials [35]. The PRISMA statement includes a 27-item and four-phase flow diagram containing items considered essential for the reporting of systematic reviews, including identification, screening, eligibility and inclusion phases. The Meta-Analysis of Observational Studies in Epidemiology (MOOSE) criteria are an excellent guide for conducting systematic reviews and meta-analyses in any field of epidemiology [36]. The MOOSE criteria also include a checklist for conducting meta-analyses of observational studies; this checklist includes suggestions for reporting on the background, search strategy, methodology, results, discussion and conclusions.

**Conclusions**

Although in the past meta-analyses were mostly targeted at comparing effect sizes from outcomes research, they are likely to become increasingly common for incidence and prevalence studies as tools (e.g. software) are now widely available. The use of systematic review and meta-analytical methods advances our understanding of neuroepidemiology, offers a solid methodological foundation for understanding a vast and diverse literature, and can allow us to see patterns that may not be detectable in any single study. The approach can help us prioritize future research and have a direct influence on decisions from the patient level to the health system level.

It is essential that the validity of meta-analyses as a whole is considered, not only the constituent parts. The objectives of meta-analyses should be well defined, a thorough evaluation of bias should be conducted, sources of heterogeneity should be described and evaluated and data analysis methods should be justified. A comprehensive assessment of bias is very important in all phases of a systematic review and meta-analysis, from study selection to the analysis phase. By carefully exploring potential sources of bias, systematic reviews and meta-analyses of
incidence and prevalence studies in neuroepidemiology can be very powerful tools.

Minimum reporting standards for neuroepidemiological research should be established and widely instituted. More widely accepted and validated diagnostic criteria for neurologic conditions (which are lacking for most) are required – this is necessary until biomarkers or gold-standard tests are established for all conditions. Neurological and neurosurgical societies should make these standards and criterion a priority in their recommendations for research related to all neurological conditions. Ultimately, however, widely accepted criteria on the reporting of original neuroepidemiological studies are urgently needed to improve study quality and comparability between regions and sub-groups.

References


Acknowledgements

Kirsten Fiest holds an Alberta Innovates Health Solutions (AIHS) Mental Health Graduate Studentship. Dr. Scott Patten is supported by a Senior Health Scholar Award from AIHS. Dr. Nathalie Jette holds a Population Health salary award from AIHS and a Canada Research Chair Tier 2 in Neurological Health Services Research.

Disclosure Statement

K.M.F., T.P., S.B.P., L.W.S. and N.J. report no disclosures or conflicts of interest.


