**STUDY PROTOCOL**

**Methods used to meta-analyse results from interrupted time series studies: A methodological systematic review protocol**

[version 1; peer review: awaiting peer review]

Elizabeth Korevaar

Amalia Karahalios

Andrew B. Forbes

Simon L. Turner

Steve McDonald

Monica Taljaard

Jeremy M. Grimshaw

Allen C. Cheng

Lisa Bero

Joanne E. McKenzie

1 School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, 3004, Australia

2 Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, K1Y 4E9, Canada

3 School of Epidemiology, Public Health and Preventive Medicine, University of Ottawa, Ottawa, Ontario, K1N 6N5, Canada

4 Department of Medicine, University of Ottawa, Ottawa, Ontario, K1H 8M5, Canada

5 Infection Prevention and Healthcare Epidemiology Unit, Alfred Health, Melbourne, Victoria, 3004, Australia

6 Faculty of Medicine and Health and Charles Perkins Centre, University of Sydney, Sydney, New South Wales, 2006, Australia

**Abstract**

**Background:** Systematic reviews are used to inform healthcare decision making. In reviews that aim to examine the effects of organisational, policy change or public health interventions, or exposures, evidence from interrupted time series (ITS) studies may be included. A core component of many systematic reviews is meta-analysis, which is the statistical synthesis of results across studies. There is currently a lack of guidance informing the choice of meta-analysis methods for combining results from ITS studies, and there have been no studies examining the meta-analysis methods used in practice. This study therefore aims to describe current meta-analysis methods used in a cohort of reviews of ITS studies.

**Methods:** We will identify 100 reviews that include meta-analyses of ITS studies from a search of eight electronic databases covering several disciplines (public health, psychology, education, economics). Study selection will be undertaken independently by two authors. Data extraction will be undertaken by one author, and for a random sample of the reviews, two authors. From eligible reviews we will extract details at the review level including discipline and type of interruption; at the meta-analytic level we will extract type of outcome, effect measure(s), meta-analytic methods, and any methods used to re-analyse the individual ITS studies. Descriptive statistics will be used to summarise the data.

**Conclusions:** This review will describe the methods used to meta-analyse results from ITS studies. Results from this review will inform future methods research examining how different meta-analysis methods perform, and ultimately, the development of guidance.

**Keywords**

Interrupted time series, meta-analysis, systematic review
Corresponding author: Amalia Karahalios (emily.karahalios@monash.edu)

Author roles: Korevaar E: Funding Acquisition, Methodology, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing; Karahalios A: Conceptualization, Funding Acquisition, Methodology, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; Forbes AB: Conceptualization, Funding Acquisition, Supervision, Writing – Review & Editing; Turner SL: Methodology, Writing – Review & Editing; McDonald S: Methodology, Writing – Review & Editing; Taljaard M: Funding Acquisition, Methodology, Writing – Review & Editing; Grimshaw JM: Funding Acquisition, Writing – Review & Editing; Cheng AC: Funding Acquisition, Writing – Review & Editing; Bero L: Funding Acquisition, Writing – Review & Editing; McKenzie JE: Conceptualization, Funding Acquisition, Methodology, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: EK and SLT are supported through an Australian Government Research Training Program (RTP) Scholarship administered by Monash University, Australia. JEM is supported by an NHMRC Career Development Fellowship (1143429). The project is funded by the Australian National Health and Medical Research Council (NHMRC) project grant GNT1145273, "How should we analyse, synthesize, and interpret evidence from interrupted time series studies? Making the best use of available evidence", McKenzie JE, Forbes A, Taljaard M, Cheng A, Grimshaw J, Bero L, Karahalios A.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2020 Korevaar E et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Korevaar E, Karahalios A, Forbes AB et al. Methods used to meta-analyse results from interrupted time series studies: A methodological systematic review protocol [version 1; peer review: awaiting peer review] F1000Research 2020, 9:110 (https://doi.org/10.12688/f1000research.22226.1)

Introduction

Systematic reviews aim to collate and synthesise all available evidence on a particular topic. They are used to inform healthcare decision making, either directly, or through their inclusion in knowledge tools such as clinical practice guidelines. Many reviews examining the effects of clinical interventions are appropriately limited in scope to inclusion of randomised trials. However, in reviews where the aim is to examine the effects of organisational, policy change or public health interventions or exposures (e.g. chemical exposures), evidence from non-randomised studies may offer the only evidence, or provide important additional evidence to that gained from randomised trials.

Interrupted time series (ITS) studies are a type of non-randomised design in which measurements on a group of individuals (e.g. a community) are taken repeatedly both before and after an ‘interruption’. The interruption may be intended (e.g. a government-implemented policy), although will not necessarily be initiated or designed by the ITS investigators (e.g. by researchers within a university), or may be unintended (e.g. an exposure such as a natural disaster). The key benefit of the ITS design is that the period before the interruption can be used to estimate the underlying time trend. If modelled correctly, this before-trend can be projected into the post-interruption period to provide a counterfactual for what would have occurred in the absence of the interruption. ITS studies with controls (e.g. an internal or external control series, control outcome) may provide more certainty in causally attributing any observed effects to the interruption. Several effect estimates can be obtained from an ITS study to characterise both short and long-terms effects of the interruption (e.g. level change and slope change).

Meta-analysis is the statistical synthesis of results across studies leading to combined effect estimates. Meta-analysis (and its extensions) is a core component of many systematic reviews. The benefits of meta-analysis have long been established, including the ability to more precisely estimate effects, examine and quantify inconsistency of the effects across studies, and identify factors that may potentially modify the size of the effects. Two approaches for meta-analysing results from ITS studies include the two-stage or one-stage approach. In the two-stage approach, effect estimates from each series are first computed, and these are then combined across series using a meta-analysis method (e.g. DerSimonian and Laird). In the one-stage approach, a single model including all series is fitted to simultaneously obtain the combined effect estimates. The one-stage approach requires the raw time series data to be available for all series, but has the proposed advantage of being more efficient since the data across all the series are used in estimating the effects.

For two-stage meta-analysis, a notable challenge is that many primary ITS studies are analysed incorrectly. For example, ITS studies may be analysed as though the study was a before-after design, or analysed as an ITS design, but without taking account of the correlation between observations over time (known as autocorrelation). The former is likely to result in estimates of the effect of the interruption that are biased, while the latter is likely to result in estimates of standard errors that are too small. Both have important implications for a two-stage meta-analysis in terms of bias, the weights that studies receive, and in turn, the precision of the combined estimate. A further challenge is that the effect measures chosen and reported by the primary study authors (e.g. level change) may not match those of interest to the systematic reviewer (e.g. slope change).

In some studies, the raw time series data may be available through extraction of data from graphs or their availability in tables. In this circumstance, it may be possible to overcome some of the above challenges through re-analysis of the ITS studies by appropriately accounting for the design and autocorrelation, or re-analysing the raw data to obtain the desired effect measure for the meta-analysis when it differs from that reported in the primary study. These computed effects may then be combined using two-stage meta-analysis. Alternatively, each study’s raw data may be analysed in a single model using a one-stage meta-analysis approach.

To our knowledge, there have been no reviews examining the approaches and methods used to meta-analyse effect estimates from ITS studies. In this review we therefore aim to: 1) investigate whether reviewers re-analyse primary ITS studies included in reviews, and if so, what re-analysis methods are used; 2) what meta-analysis methods are used; 3) what effect measures are used, and how completely the estimated combined effects are reported; and 4) what tools and domains are used to assess the risks of bias or methodological quality of the included ITS studies. Here, we report the planned design of our review, including the criteria that we will use to identify eligible studies, as well as the information we will extract and describe.

Methods

Overview

This study aims to identify and describe reviews that include meta-analyses of ITS studies. The reviews will be identified by searching several electronic databases including MEDLINE (Ovid), EMBASE (Ovid), Campbell Systematic Reviews, EconLit (EBSCOhost), 3ie, PsycINFO (Ovid), ERIC (ProQuest) and the Cochrane Database of Systematic Reviews (CDSR). Study selection will be undertaken independently by two authors; data extraction will be undertaken by one author, and for a minimum of 20% of randomly selected reviews, two authors. We will extract details at the systematic review level, including: discipline (public health, psychology, education, economics), type of interruption, assessment of risk of bias and methodological quality; and at the meta-analytic level: type of outcome, effect measure(s), meta-analytic methods, and any methods used to re-analyse the individual ITS studies. These aspects will be analysed and described using summary statistics, tables and figures.

Eligibility criteria

Studies which meet our eligibility criteria (described below) will be included. We will not restrict inclusion of reviews based on discipline or any of the PICO elements (i.e. participants/
populations, interventions/interruptions, comparators, or outcomes).

Inclusion criteria. Studies meeting the following criteria will be included:

1. the study is a review that includes at least two ITS studies which meet the review authors’ definition of an ITS design; and

2. the review includes at least one meta-analysis of ITS studies.

Our definition of a ‘review’ is very broad. It includes systematic reviews, reviews of selected studies (i.e. between-study meta-analysis), and studies that combine multiple ITS across sites within the same study (i.e. within-study meta-analysis). We have opted for broad inclusion since our primary interest is in the meta-analysis methods, which apply regardless of the particular study design. We will not restrict the meta-analysis by approach, that is, we will include both one-stage and two-stage meta-analyses. We will only include meta-analyses that combine estimates of model parameters, or combinations of these (e.g. pre-intervention fitted trend, slope change, level change).

Exclusion criteria. Studies will be excluded if they meet one or more of the following criteria. The study is:

1. written in a language other than English;

2. a methodological review that describes or evaluates methods to synthesise results from ITS studies;

3. a review of ITS studies reported in a conference abstract, letter, book, or dissertation;

4. a protocol for a review of ITS studies; or

5. a stepped-wedge randomised trial.

Criterion 1 is included because we are not able to translate studies written in a language other than English due to resource constraints. Criterion 2 excludes methodological reviews that describe or evaluate methods to synthesise data from ITS studies, as our aim is to describe current statistical methods applied in practice.

Search methods

Several databases will be searched to capture the broad range of disciplines that use the ITS study design. To capture reviews in health, we will search MEDLINE (Ovid), EMBASE (Ovid), Campbell Systematic Reviews, the CDSR and 3ie. For CDSR, we will directly search the ‘Characteristics of included studies’ table included in each systematic review for ITS studies. This will allow more specific identification of eligible reviews. The search of MEDLINE (Ovid) will also capture systematic reviews from the Joanna Briggs Institute Database of Systematic Reviews and Implementation Reports. To capture reviews in economics, we will search EconLit (EBSCOhost), and for psychology and education disciplines, we will search PsycINFO (Ovid) and ERIC (ProQuest).

Our search strategy has been informed by previous publications that have reviewed ITS studies. Reviews of ITS studies will be identified using terms adapted from the search strategies of these publications and then combined with terms to identify meta-analyses and systematic reviews. As there is little consistency in the terminology used to describe ITS studies, our search terms are intentionally broad to achieve greater search sensitivity. Terms will be searched both as free text in the titles, abstracts and keywords fields, and as MeSH terms (or equivalent) where applicable. The MEDLINE (Ovid) strategy is presented in Table 1, and the search strategies for the remaining databases are presented in Appendix 1 (see Extended data). The search is limited to the period 1 Jan 2000 to 11 Oct 2019 for all databases except CDSR which is limited to the period 1 Jan 2000 to 9 Aug 2019.

<table>
<thead>
<tr>
<th>#</th>
<th>Searches</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Interrupted Time Series Analysis/</td>
</tr>
<tr>
<td>2</td>
<td>interrupted time series.mp.</td>
</tr>
<tr>
<td>3</td>
<td>(time series or time trend$ or trend analys?s?).mp.</td>
</tr>
<tr>
<td>4</td>
<td>(change point or repeated measures or phase design or multiple baseline$ or difference-in-difference$ or single case research or single case experimental).mp.</td>
</tr>
<tr>
<td>5</td>
<td>(ARIMA or autoregressive integrated moving average or integrated moving average or piecewise regression or segmented regression).mp.</td>
</tr>
<tr>
<td>6</td>
<td>or/1-5</td>
</tr>
<tr>
<td>7</td>
<td>(Systematic Review or Meta-Analysis).pt. or (meta-analys?s or pooled analysa?s).ti.ab.kw.</td>
</tr>
<tr>
<td>8</td>
<td>and/6-7</td>
</tr>
<tr>
<td>9</td>
<td>Limit 8 to (abstracts and English language and yr=&quot;2000-Current&quot;)</td>
</tr>
</tbody>
</table>
Study selection

Citations identified from the searches will be imported into EndNote X8 (Clarivate Analytics, Philadelphia) to remove duplicates. Titles and abstracts will be sorted by year in descending order and will be screened against the eligibility criteria, with each abstract assessed as: 1) ‘Yes/Maybe includes two or more ITS studies’ and ‘Yes/Maybe a meta-analysis of ITS studies has been undertaken’, 2) ‘Yes/Maybe includes two or more ITS studies’ and ‘No meta-analysis of ITS studies has been undertaken’, or 3) ‘No, does not include two or more ITS studies’. This process will be piloted on 20 studies by EK, SLT, AK and JEM. The remaining abstracts will be screened independently by at least two members of the review team (EK, and any of SLT, AK and JEM). The full-text articles of the titles and abstracts assessed as potentially meeting the eligibility criteria (i.e. group 1 above) will be retrieved, sorted by most recent first and screened against the eligibility criteria until all reviews (if less than 100), or the 100 most recently published reviews are identified. Conflicts in screening decisions at the abstract and full-text stages will be resolved via discussion between the screeners or through consultation with the broader team.

Sample size

Our sample size of 100 reviews was primarily selected for reasons of feasibility. A sample of this size will allow estimation of the percentage of reviews with a particular element (e.g. reviews that re-analyse the primary study data) to within a 10% margin of error (assuming a prevalence of 50%). This margin of error will decrease if the prevalence varies from 50%.

Selection of outcomes

Reviews may include several meta-analyses of ITS studies for different outcomes. We plan to examine the meta-analysis methods for only one outcome per review. The following set of rules will be applied hierarchically until a unique outcome is identified (for which there could be multiple meta-analyses of different effect estimates):

1) The outcome that has the largest number of effect measures (e.g. the outcome that has meta-analyses of level change and slope change estimates would be selected ahead of an outcome with only a meta-analysis of level change estimates);

2) The outcome with the largest number of ITS studies; or

3) The outcome that is first reported in the abstract, then the methods section, then the results section of the manuscript.

A single outcome is chosen as it is likely that the meta-analysis methods are consistent across outcomes within a review. Criterion 1 has been included so that we can capture the range of effect measures used. Uncertainty in the selection of the outcome will be resolved through discussion with the review team.

Data extraction and management

The data extraction form will be designed using the Research Electronic Data Capture (REDCap) online designer. The review team (EK, AK, SLT, ABF, and JEM) will pilot the data extraction form by independently extracting data from 10 reviews. This pilot testing will be used to revise the form if we uncover ambiguity or a lack of clarity in any items, identify missing items and test the logic of the form. Following piloting, data extraction will be undertaken by EK for all eligible studies and independently by at least two members of the review team (one of AK, SLT, ABF and JEM) for a further 20% of randomly selected reviews. Any inconsistencies in data extraction will be resolved via discussion between the data extractors or through consultation with the broader team. For any items where a large percentage of inconsistency is found, the percentage of studies with double data extraction will be increased.

A summary of the data extraction items is presented in Table 2. In brief, we will extract details of the review’s aims, meta-analysis methods (including the reviewer’s rationale) and methods used to assess the methodological quality and/or risk of bias of the included ITS studies. For the selected outcome, we will extract the type of effect measure(s), methods of synthesis, adjustment for autocorrelation and/or seasonality.

Analysis

We will summarise the characteristics of included systematic reviews with descriptive statistics. For categorical data (e.g. the meta-analysis approach used, the risk of bias tool used) we will present frequency and percentage, and for numerical data (e.g. the number of meta-analysed ITS studies, the number of pooled estimates) we will present means (with standard deviations) or medians (with interquartile range). Statistical analyses will be undertaken using Stata version 15.0.

Discussion

To our knowledge, this will be the first review to examine methods for meta-analysis of ITS studies that are used in practice. Specifically, the choice of meta-analysis approach, effect measures, completeness of reporting, and tools for assessing quality or risk of bias (if undertaken). The results of this review will inform our broader research program which aims to examine how different meta-analysis methods of ITS studies perform, using statistical simulation and empirical evaluation, and provide guidance on the methods.

Strengths and limitations

There are several strengths to this study. The search, screening and data extraction methods have been prespecified, and the study has been registered with PROSPERO (submitted 4 Oct 2019). Further, we will search a broad range of databases, encompassing the areas of health, economics, psychology and education. This will allow us to identify a broader range of meta-analysis methods in use, not restricted to a particular discipline.
While the study will be limited by our ability to identify all potentially eligible reviews and meta-analyses of ITS studies, our search strategy attempts to capture the various ways these studies are described. However, given ITS studies are often not identified as such\(^\text{16}\), it is likely that we will not capture all reviews and meta-analyses that include ITS studies. Conversely, we may end up including reviews where no information regarding the definition of the included ITS studies is provided, or where an inappropriate label of ITS has been applied to included studies. While we will not exclude these reviews, we will record the reviewers’ definition of an ITS study.

**Conclusions**

The ITS design is often used to examine the effects of organisational, policy change or public health interventions or exposures. Meta-analysis of results from these studies provides the opportunity to estimate the interruption’s impact more precisely, and investigate factors that may modify the size of the impact. However, there is a paucity of guidance available for meta-analysing results from ITS studies. Results from this review will provide the first examination of meta-analysis methods used in practice to combine results from ITS studies. This will be used to inform future research that investigates how different methods perform, from which guidance will be developed.

**Data availability**

**Underlying data**

No underlying data are associated with this article.

**Extended data**

Figshare: Methods used to meta-analyse results from interrupted time series studies: A methodological systematic review protocol - Appendix 1 Search strategy. [https://doi.org/10.26180/5e3b5cc4acf30].

**Reporting guidelines**

Figshare: PRISMA-P checklist for ‘Methods used to meta-analyse results from interrupted time series studies: A methodological systematic review protocol’. [https://doi.org/10.26180/5e3b5d75c00a].

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

**Table 2. Summary of data extraction items for the selected meta-analysis.**

<table>
<thead>
<tr>
<th>Domain</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review characteristics</td>
<td>Author name; publication year; journal; target population or content area; number of ITS studies included in the review</td>
</tr>
<tr>
<td>Outcome and studies included</td>
<td>Description and type of outcome (e.g. continuous, count, rate); number of ITS studies measuring the chosen outcome</td>
</tr>
<tr>
<td>Methods for synthesising ITS results</td>
<td>Number of ITS studies meta-analysed; use of primary study data (i.e. re-analysis); pairwise or network meta-analysis; fixed/random effects model; methods to quantify between-study variation</td>
</tr>
<tr>
<td>Results/Estimates</td>
<td>Description and type of effect measures (e.g. change in level, change in slope, combination of change in level and slope (i.e., counterfactual)); completeness of reporting estimates (e.g. combined effect estimate, confidence interval, measure of heterogeneity)</td>
</tr>
<tr>
<td>Risk of bias and/or assessment of study quality</td>
<td>Description of assessment (if performed) of primary study risk of bias / methodological quality; tool or domains used for assessment</td>
</tr>
</tbody>
</table>

Abbreviations: ITS, interrupted time series

**References**


The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com