ABSTRACT

Introduction An interrupted time series (ITS) design is an important observational design used to examine the effects of an intervention or exposure. This design has particular utility in public health where it may be impracticable or infeasible to use a randomised trial to evaluate health system-wide policies, or examine the impact of exposures (such as earthquakes). There have been relatively few studies examining the design characteristics and statistical methods used to analyse ITS designs. Further, there is a lack of guidance to inform the design and analysis of ITS studies. This is the first study in a larger project that aims to provide tools and guidance for researchers in the design and analysis of ITS studies. The objectives of this study are to (1) examine and report the design characteristics and statistical methods used in a random sample of contemporary ITS studies examining public health interventions or exposures that impact on health-related outcomes, and (2) create a repository of time series data extracted from ITS studies. Results from this study will inform the remainder of the project which will investigate the performance of a range of commonly used statistical methods, and create a repository of input parameters required for sample size calculation.

Methods and analysis We will collate 200 ITS studies evaluating public health interventions or the impact of exposures. ITS studies will be identified from a search of the bibliometric database PubMed between the years 2013 and 2017, combined with stratified random sampling. From eligible studies, we will extract study characteristics, details of the statistical models and estimation methods, effect metrics and parameter estimates. Further, we will extract the time series data when available. We will use systematic review methods in the screening, application of inclusion and exclusion criteria, and extraction of data. Descriptive statistics will be used to summarise the data.

Ethics and dissemination Ethics approval is not required since information will only be extracted from published studies. Dissemination of the results will be through peer-reviewed publications and presentations at conferences. A repository of data extracted from the published ITS studies will be made publicly available.
they can be used to examine the impact of exposures such as earthquakes (eg,5 6) or nuclear power station leaks (eg,7). An important benefit of an ITS design is that it can account for the preintervention trend in estimating the effect of the intervention.8 Figure 1 provides a graphical representation of a simple ITS design, with data collected in preintervention and postintervention phases.

A feature of data collected over time is that the data points tend to be correlated. This is known as autocorrelation or serial correlation.9 10 This association could be positive (whereby data points close together in time are more similar than data points further apart) or negative (whereby data points close together are more dissimilar than data points further apart). A specific type of autocorrelation that may be observed in ITS designs with data collected over a long period of time is seasonality, which refers to periodic, repetitive and predictable patterns in the levels of the time series. Influenza rates, for example, may demonstrate patterns of higher levels in winter and lower levels in summer months, recurring every calendar year. Some methods that are used to analyse ITS studies, such as classical linear regression, assume that the observations are independent. If positive autocorrelation is present and not accounted for, this may lead to standard errors that are too small, with resulting confidence intervals that are too narrow and p values that are too small.10 Autocorrelation is also one of the key parameters required for sample size estimation.11

In addition to accounting for autocorrelation, other important aspects of analysing ITS data include specification of the structure of the model (ie, number of segments and their shape12), statistical estimation methods and choice of effect metrics. Regarding the latter, even in the circumstance where a simple segmented linear model is fitted (figure 1), a variety of effect metrics may be calculated. The effect of the intervention on the outcome can be calculated by estimating the change in level, or the change in slope of the preintervention and postintervention trends, or both (see figure 1). A combination of the two allows estimation of the intervention effect at a specified time post intervention based on the predicted counterfactual (ie, using the trend in the preintervention period to predict what would have occurred in the postintervention period, in the absence of the intervention). Further, these effect metrics can be expressed in absolute or relative terms with associated confidence intervals. For example, a level change corresponding to a drop in mortality could be expressed as an absolute effect such as number of deaths or as a relative effect such as percentage change in the number of deaths.13 Choosing which metric is most helpful for understanding and communicating the impact of an intervention is not straightforward.14

There have been relatively few studies examining the design characteristics and statistical methods used to analyse ITS designs. Ramsay et al15 examined ITS studies included in two systematic reviews and demonstrated that when reanalysed using more appropriate statistical methods, approximately half of the studies that had found statistically significant intervention effects, had their results overturned (ie, statistically significant results became non-significant). Jandoc et al16 examined a cohort of 220 ITS studies, published from 1984 to 2013, in drug

![Figure 1](image-url)  
*Figure 1* The rate of *Clostridium difficile* infections (per 1000 patient days) prebleach and postbleach disinfection per month.28 Various effect estimates can be constructed from the preintervention and postintervention slopes, such as the change in level and change in slopes.
utilisation research and found that important elements of an ITS design were not being accounted for in the analyses, such as taking into account autocorrelation and seasonality. Ewusie et al. are undertaking a scoping review of ITS methods used to analyse health research and the way in which the results are reported, focusing on the strengths and limitations of each method. None of these reviews has focused on ITS studies evaluating public health interventions, or exposures, that impact on health-related outcomes.

There is a lack of available information and guidance to inform the design and analysis of ITS studies. First, there are no databases providing empirical estimates of key parameters (such as autocorrelation coefficients) required for sample size calculation. Second, there is an absence of clear guidance to inform the choice of statistical analysis. Third, the choice of intervention effect estimates and their standard errors that are most robust to misspecification of the model and estimation method is unclear.

This is the first study in a larger project that aims to address these issues by providing tools and guidance for researchers on the design and analysis of ITS studies. The project includes a series of studies that will examine the design features and statistical methods used in practice, investigate the performance of a range of commonly used statistical methods through numerical simulation and empirical evaluation, and create a repository of input parameters required for sample size calculation. Here, we report the planned design of the first study, the objectives of which are to (1) examine and report the design characteristics and statistical methods used in a random sample of contemporary ITS studies examining public health interventions or exposures that impact on health-related outcomes, and (2) create a repository of time series data extracted from ITS studies.

**METHODS AND ANALYSIS**

**Overview**

We will identify and describe ITS studies evaluating public health interventions or the impact of exposures. ITS studies will be identified from a search of the bibliometric database PubMed, combined with stratified random sampling. Study selection and data extraction will be undertaken by one author, and for a fraction of studies, two authors. From eligible studies, we will extract study characteristics, details of the statistical models and estimation methods, effect metrics and parameter estimates. Further, we will extract the time series data, when available.

**Eligibility criteria**

Following are the inclusion and exclusion criteria applied to determine which ITS studies will be included.

**Inclusion criteria**

Our definition of an ITS study is informed by the design features taxonomy for quasi-experimental designs proposed by the Cochrane Non-Randomized Studies Methods Group and further developed by Reeves et al. ITS studies meeting the following criteria will be included (with rationale and further explanations for each criterion provided below):

1. There are at least two segments separated by a clearly defined intervention or exposure with at least three points in each segment (eg, preintervention and postintervention time series, each with at least three points).
2. Observations are collected on a group of individuals (eg, community, hospital) at each time point.
3. The study is investigating the impact of a public health intervention or exposure that has public health implications (eg, patient health outcome, resource use).

Criterion 1 is that used by the Cochrane Effective Practice and Organisation of Care Group as minimum criteria for an ITS study to be considered eligible for inclusion in systematic reviews undertaken within this group. The rationale for this criterion is that three time points pre intervention and post intervention allow the possibility of using segmented time series regression (although using as few as three points would not be recommended).

Criterion 2 restricts the inclusion to ITS studies focused on examining the effects of public health interventions (or exposures) on populations, and importantly, excludes ITS studies that examine the effects of an intervention on individuals (also known as single-case designs or single-case experimental designs). The group of individuals studied at each time point may or may not include the same individuals. The study may also include multiple series (multiple intervention and control series measured in aggregate or groups (eg, hospitals, communities)).

Our definition of public health interventions (criterion 3) is informed by that used by the Cochrane Public Health Review Group, and includes interventions that aim to prevent or promote population health for communicable and non-communicable diseases (eg, vaccination and screening programmes; programmes aimed to reduce the use of tobacco or alcohol; public information/awareness campaigns for stroke recognition). The interventions may fall outside of the health service (such as education, work environment, housing and the built environment, natural environment interventions), but will be included if they aim to improve population health-related outcomes. Further, there will be no restriction on the level that the intervention is targeted at, which may be, for example, individuals, communities or health systems. Our definition of exposures are events that are not under investigators’ control (eg, earthquakes, financial crises, tsunamis, environmental chemicals).

**Exclusion criteria**

ITS studies meeting the following criteria will be excluded:

1. Studies written in a language other than English.
2. Single-case designs.
3. Methodological papers examining ITS studies.
Criterion 1 is included as we are not able to translate studies written in languages other than English due to resource constraints. Criterion 3 excludes papers that examine statistical methods for ITS studies. While these methodological papers often include motivating examples, which demonstrate the application of different methods to an ITS study selected from the literature, these examples may not be representative of published ITS studies.

Sample size
We plan to include 200 ITS studies, which will allow estimation of the percentage of ITS studies with a particular element (eg, studies taking autocorrelation into account) to within a maximum margin of error of 7% (assuming a prevalence of 50%); for a prevalence less (or greater) than 50%, the margin of error will be smaller. We will include 200 studies from 2013 to 2017. The studies will be stratified by year, and within each year, will be randomly sampled until 40 are identified that meet the eligibility criteria. If fewer than 40 studies are eligible for inclusion in any given year, we will randomly sample studies from earlier years until we meet our target sample size. We are sampling (using standard survey methodology) since this provides a valid approach for estimating the prevalence of characteristics of ITS studies. This is unlike systematic reviews that aim to estimate a combined treatment effect, for which it is imperative to identify all studies.

Search methods for the identification of studies
To identify potentially eligible ITS studies, we will search PubMed. PubMed has been chosen since it has the broadest coverage of public health and health services research journals, and as such, provides a sufficient sampling frame. We will search using free-text terms informed from previous search strategies developed to identify ITS studies, terms used to describe the ITS design in the methods section of previously published papers (eg, Cheng et al.2 Baker and Alonso,3 Milojevic et al4), and controlled vocabulary (Medical Subject Headings (MESH) terms) (table 1). In developing the strategy, we examined how well it captured a subset (10% random sample, 31 studies) of ITS studies included in two systematic reviews and refined accordingly. Studies not captured by our preliminary search strategy were investigated to identify additional search terms. After adding the new search terms, the process was repeated until all studies in the sample were captured. The strategy was not restricted by public health terms since we anticipated that there may be large variation in the terminology used.

Study selection
Titles and abstracts will be extracted into a screening and data collection programme built using Microsoft Access. Abstracts will be grouped by year of publication, and randomly sorted. In the piloting phase, four authors (SLT, JEM, ABF and AK) will independently assess 50 abstracts to ensure consistency of the application of the inclusion criteria.

Following piloting, two authors (SLT and one of JEM, ABF or AK) will use a two-phase screening process to identify ITS studies. One author (SLT) will screen all abstracts, and a second author will screen a 50% sample of the identified abstracts or until 20 ITS studies (per year) have been identified for inclusion (AK, JEM, ABF). In the first phase, abstracts which are independently assessed by both reviewers as not meeting at least one of the following criteria, will be excluded.

Following piloting, two authors (SLT and one of JEM, ABF or AK) will use a two-phase screening process to identify ITS studies. One author (SLT) will screen all abstracts, and a second author will screen a 50% sample of the identified abstracts or until 20 ITS studies (per year) have been identified for inclusion (AK, JEM, ABF). In the first phase, abstracts which are independently assessed by both reviewers as not meeting at least one of the following criteria, will be excluded.

Table 1 PubMed search strategy

<table>
<thead>
<tr>
<th>Search (#)</th>
<th>Search terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Interrupted time series analysis (MeSH term)</td>
</tr>
<tr>
<td>2</td>
<td>“Interrupted time series” (title/abstract)</td>
</tr>
<tr>
<td>3</td>
<td>“Change point” (title/abstract)</td>
</tr>
<tr>
<td>4</td>
<td>“Segmented regression” (title/abstract)</td>
</tr>
<tr>
<td>5</td>
<td>“Segmented linear regression” (title/abstract)</td>
</tr>
<tr>
<td>6</td>
<td>“Repeated measures study” (title/abstract)</td>
</tr>
<tr>
<td>7</td>
<td>“Piecewise regression” (title/abstract)</td>
</tr>
<tr>
<td>8</td>
<td>“Time-series intervention” (title/abstract)</td>
</tr>
<tr>
<td>9</td>
<td>“Phase design” (title/abstract)</td>
</tr>
<tr>
<td>10</td>
<td>“Multiple baseline” (title/abstract)</td>
</tr>
<tr>
<td>11</td>
<td>“ARIMA” (title/abstract)</td>
</tr>
<tr>
<td>12</td>
<td>“Integrated moving average” (title/abstract)</td>
</tr>
<tr>
<td>13</td>
<td>1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12</td>
</tr>
</tbody>
</table>

ARIMA, autoregressive integrated moving average.

Studies that are found to meet these criteria will be included.

1. Does the study appear to use an ITS design?
2. Were observations collected from a group of individuals?
3. Is this study written in English?
4. Does the study appear to evaluate the impact of a public health intervention or natural interruption?

In the second phase, the full text of each study will be retrieved and assessed against the criteria listed below. Studies that are found to meet these criteria will be included.

1. Does the study use an ITS design?
2. Were observations collected from a group of individuals?
3. Are there are at least two segments separated by a clearly defined intervention with at least three points in each segment?

Selection of outcome(s) for inclusion
Multiple outcomes per ITS study are potentially eligible for inclusion. Within each outcome type category (binary, continuous, count), we will select one outcome using the following hierarchy:

1. ITS data availability—outcomes with data available to be extracted (either from tables or figures) will be selected ahead of those without data.
Table 2  Example data extraction items

<table>
<thead>
<tr>
<th>Examples of data extraction items</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study characteristics</strong></td>
</tr>
<tr>
<td><strong>Design</strong></td>
</tr>
<tr>
<td><strong>Outcome</strong></td>
</tr>
<tr>
<td><strong>Model</strong></td>
</tr>
<tr>
<td><strong>Statistical methods</strong></td>
</tr>
<tr>
<td><strong>Effect measures</strong></td>
</tr>
</tbody>
</table>

2. Stated primary outcome (or reported in the title or objectives).
3. First reported result outcome in the abstract.
4. First reported outcome in the results.

Uncertainty in the selection of the review outcomes will be discussed among the review team.

Data extraction and management

The data extraction process will initially be piloted by four reviewers (SLT, JEM, ABF and AK) on a sample of 10 ITS studies to ensure consistency of data extraction and will be adjusted as necessary. Following piloting, we will use double data extraction for all items (except extraction of the time series data) on a randomly selected 20% of studies. Discrepancies and uncertainty in coding will be discussed in meetings with three reviewers. For any items where we observe a high degree of inconsistency, we will undertake double data extraction for these items on a further randomly selected sample of studies.

We will extract data pertaining to the study characteristics, design, outcome, model, statistical methods and effect measures. Further details are provided in table 2, with the complete list of items available in online supplementary additional file 1.

The time series data for each included outcome will be extracted when possible. For studies that present their data in graphical form only, we will extract from graphs using WebPlotDigitizer software. This software has been shown to give reliable and valid results. The data will be kept in a Microsoft Access 2016 database on a secure server. Any free text responses will be input as prespecified options or categories wherever possible.

Analysis

We will calculate descriptive summary statistics. For categorical data, such as the type of model used, we will present percentages and frequencies. For counts, such as the number of time intervals, we will present means (with standard deviations) and medians (with interquartile range). Statistical analyses will be undertaken in Stata Release 15. The data sets arising from this study will be made available on figshare.

Patient and public involvement

No patients will be involved in this project and information will only be extracted from published studies.

DISCUSSION

This is a first study in a larger project that aims to provide tools and guidance for researchers in the design and analysis of ITS studies. This study will provide information on design characteristics of a contemporary sample of ITS studies in public health, statistical methods used in practice, and provide a repository of ITS data. Results from this study will underpin the remainder of the project by informing a numerical simulation study to investigate the performance of commonly used statistical methods; an empirical study to investigate the impact of using different statistical methods for analysing ITS on real data; and a study to create a repository of parameter values for sample size calculation, along with generalisable ‘rules of thumb’ on the selection of values.

Strengths and limitations

There are several strengths to our study. To our knowledge, this will be the first study specifically examining
the use of ITS studies in public health. Further, a priori specified systematic review methods will be used in the screening, application of inclusion and exclusion criteria, and data extraction.

However, there are some limitations. For some items, the sample size may not be large enough to precisely estimate the percentage of ITS studies with a particular element. Our search strategy is unlikely to locate all published ITS studies in public health, since studies will use terminology other than our search terms. Further, we will only search a single database (ie, PubMed); however, this database has the broadest coverage of public health and health services research journals.

**Implications of this research**

Previous reviews have described the details of ITS studies including whether autocorrelation, seasonality and/or stationarity were accounted for. However, these studies have focused on specific systematic reviews or drug utilisation studies. Our study will extend this research, with a focus on ITS studies in public health. Results of this study will inform the larger project, which aims to provide tools and guidance for researchers designing and analysing ITS studies. The repository of ITS data that we will curate as part of this project will also be of value for future methodological and statistical research.

**CONCLUSION**

ITS studies are commonly used designs in public health to examine whether an intervention or an exposure has had an effect on health outcomes. However, there is a lack of available information and guidance to inform the design and analysis of ITS studies. Results of this study will help to address this gap by providing information on current design and analysis practice of ITS studies in public health.

**REFERENCES**

20. EPOC. Cochrane Effective Practice and Organisation of Care (EPOC). EPOC Resources for review authors. 2017 epoc.cochrane.org/epoc-resources-review-authors.