

# Current practices in missing data handling for interrupted time series: a scoping review

Protocol Version: 1  
Protocol Date: 14<sup>th</sup> Feb 2020

PhD Student

Juan Carlos Bazo Alvarez



Signature

14<sup>th</sup> Feb 2020

Date

Supervisor

Irene Petersen



Signature

14<sup>th</sup> Feb 2020

Date

Supervisor

Tim Morris



Signature

14<sup>th</sup> Feb 2020

Date

Supervisor

James Carpenter



Signature

14<sup>th</sup> Feb 2020

Date

## Version History

Amendment no	Protocol version no	Description of changes	Date of protocol
	Version 1	New document	

1. **Title:** Current practices in missing data handling for interrupted time series: a scoping review
2. **General Language:** English
3. **Start:** Feb 05, 2020
4. **Early Termination Date:** Mar 05, 2020
5. **Review Stage:** In progress
6. **Named Contact**  
Juan Carlos Bazo-Alvarez
7. **Named Contact Email**  
[juan.alvarez.16@ucl.ac.uk](mailto:juan.alvarez.16@ucl.ac.uk)  
[juan.bazo.a@gmail.com](mailto:juan.bazo.a@gmail.com)
8. **Address**  
Research Department of Primary Care and Population Health  
Upper Third Floor  
UCL Medical School (Royal Free Campus)  
Rowland Hill Street  
London  
NW3 2PF
9. **Contact Telephone Number**  
+44 07376076260
10. **Organizational Affiliation of the Review**
  1. Department of Primary Care and Population Health, University College London (UCL), London, United Kingdom.
  2. Instituto de Investigación, Universidad Católica Los Ángeles de Chimbote, Chimbote, Peru.
  3. MRC Clinical Trials Unit at UCL, London, United Kingdom.
  4. Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London, United Kingdom.
  5. Department of Clinical Epidemiology, Aarhus University, Denmark.
11. **Members**  
Juan Carlos Bazo-Alvarez<sup>1,2</sup>, Tim P Morris<sup>3</sup>, James R Carpenter<sup>3,4</sup>, Irene Petersen<sup>1,5</sup>
12. **Financing**  
JCB is sponsored by FONDECYT-CONCYTEC (grant contract number 231-2015-FONDECYT). TPM and JRC are supported by the Medical Research Council (grant numbers MC\_UU\_12023/21 and MC\_UU\_12023/29). The study sponsors only had a funding role in this research. Thus, researchers will work with total independence from their sponsors.
13. **Conflicts of Interest**  
None
14. **Collaborators**  
Frank Peralta
15. **Review Questions**  
In health care studies using the interrupted time series (ITS) approach:
  1. How are the data management and statistical analysis?
  2. How researchers report and handle the main methodological issues of ITS statistical analysis?
  3. If so, how is missing data reported and handled?
16. **Search**  
We will use the MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations and Daily database (Ovid version) in February 2020 for finding ITS studies published in 2019. Search strategies will use a combination of free text terms and subject headings, and authors will be consulted for inclusion of appropriate terminology. Where appropriate, the validated filters will be used for limiting searches to interrupted time series designs. The search strategy will be reviewed by an information specialist using the Peer Review checklist of electronic search strategies. Studies whose full-text is not available will not be included. Current version of the search strategy is in the appendix.
17. **URL to Search for the Strategy**  
See section 27 and Appendix 4 For details on the search strategy.

**18. Study Condition or Domain**

Evidence of reporting missing data handling procedures in healthcare studies using the interrupted time series approach.

**19. Participants**

We will include all original ITS studies with a minimum of two data points before and after the initiation of a healthcare intervention (e.g. programs, policies, or educational interventions). Systematic reviews, meta-analysis, RCTs or studies that did not use and ITS-type analysis will be excluded. Studies with no access to individual-level data will be excluded as well. There will be no restrictions on participants, language of study, or the type of outcome.

**20. Intervention**

This is not applicable to the present systematic review as it focuses on missing data handling in interrupted time series studies that evaluate healthcare interventions.

**21. Comparator**

This is not applicable to the present systematic review as it focuses on missing data handling in interrupted time series studies that evaluate healthcare interventions.

**22. Types of Studies to be Included**

See section 19.

**23. Context**

None

**24. Outcomes**

We will extract data from the studies consisted of:

- **General Information:**
  - Author
  - Year
  - DOI
  - Definition of study design (e.g. ITS, before-and-after) [study label]
  - *Country of study*
  - Study objectives (population, intervention and the outcomes of interest)
  - *Type of intervention (e.g. policy, program, treatment)*
  - Level of intervention (e.g. individual, hospital, district, country)
  - *Participants (e.g. patients, doctors)*
- **Data handling and statistical analysis (research question #1):**
  - Data source (e.g. prospective, routinely collected)
  - Type of outcome (e.g. continuous, count, binary)
  - The number of data points collected pre-and post-intervention and the unit (e.g. week, month, year)
  - Averaging-step at each time point (for individual level data only: yes/no)
  - Main statistical model/tool (e.g. segmented regression, mixed models, other)
- **Methodological issues -reporting and handling- (research question #2):**
  - Autocorrelation
  - Seasonality
  - Time-varying confounders
  - Others
- **Missing data handling reported (yes/no) (research question #3):** If yes, we will extract:
  - Missing data proportion reported (yes/no; if yes, % reported)
  - Missing data mechanism reported (yes/no; if yes, which one was declared)
  - Missing data handling method applied (yes/no; if yes, which method was applied)
  - Sensitivity analysis for missing data assumption (yes/no; if yes, which analysis was performed)

We will base data extraction on the primary outcome and if no defined primary outcome is reported, we will use the first reported outcome.

**25. Timing and Effect Measures**

This is not applicable to the present systematic review as it focuses on missing data handling in interrupted time series studies that evaluate healthcare interventions.

**26. Additional Result**

None.

**27. Data Extraction (Selection and Coding)**

After the search and selection of articles, all titles will be treated by the Rayyan program; a list will be created and duplicates removed. Review process will be carried out by two reviewers A and B, with the support of a third person C for disagreements (names to be defined). Reviewer A will screen titles and abstracts identified by the search for inclusion. Reviewer B will assess 10% of the titles and abstracts and, if there are no disagreements, then reviewer A would proceed to single screening. Full-text copies for all the potential studies will be obtained and assessed for inclusion by A, with B double assessing 10% of them. Full-text review and data extraction will be done using an Excel template. In this systematic review, data on outcomes (section 24) will be extracted using a data collection sheet. Studies will be collected in any language, excluding those studies that do not have the full-text. All the excluded studies will be listed and enumerated indicating the reason for their exclusion.

**28. Assessment of Risk of Bias**

As a methodological study, risk of bias assessment was not performed on individual studies.

**29. Strategy for Data Synthesis**

We will summarize data using descriptive statistics (numbers and percentages or median, 25th, and 75th centile). Some graphs could be included to facilitate the communication of specific results.

**30. Subgroup or Set Analysis**

A sub-group analysis will be performed between studies with or without access to individual level data. In particular, we are interested in i) missing data methods applied (if so); ii) statistical models applied; iii) whether an averaging-step was performed before statistical models fitted (when individual level data were available).

**31. Type or Method of Revision**

Scoping review

**32. Language**

English

**33. Country**

UK

**34. Other Registration Details**

None

**35. Protocol URL**

(to be defined)

**36. Dissemination Plans**

The results will be incorporated in a PhD thesis and later published in a scientific journal. Presentations will be made at an epidemiology congress.

**37. Keywords**

Interrupted Time Series Analysis; Segmented Regression; Missing Data; Multiple Imputation.

**38. Details of any other Existing Revisions**

No other existing reviews were reported in PROSPERO with similar goals. Hudson et al [7] published a similar study, but using 2015 data and with different objectives. We are using a similar research strategy.

**39. Status of the Current Revision**

Ongoing

**40. Additional Information**

None

**41. Details**

None

**42. References**

Hudson J, Fielding S, Ramsay CR. Methodology and reporting characteristics of studies using interrupted time series design in healthcare. BMC Med Res Methodol 2019;19:137. <https://doi.org/10.1186/s12874-019-0777-x>.

**Appendix: Search strategy**

1. Interrupted Time Series Analysis/
2. interrupted time series.tw,kw.
3. (segmented adj3 regression).tw,kw.
4. arima.tw,kw.
5. autoregressive integrated moving average.tw,kw.
6. 1 or 2 or 3 or 4 or 5
7. limit 6 to yr="2019"