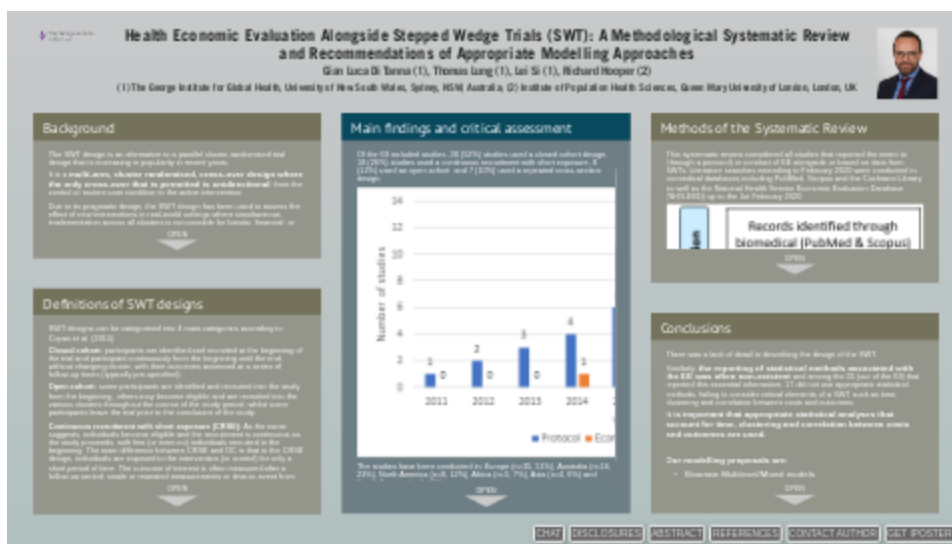


Health Economic Evaluation Alongside Stepped Wedge Trials (SWT): A Methodological Systematic Review and Recommendations of Appropriate Modelling Approaches



Gian Luca Di Tanna (1), Thomas Lung (1), Lei Si (1), Richard Hooper (2)

(1) The George Institute for Global Health, University of New South Wales, Sydney, NSW, Australia; (2) Institute of Population Health Sciences, Queen Mary University of London, London, UK



PRESENTED AT:



BACKGROUND

The SWT design is an alternative to a parallel cluster-randomised trial design that is increasing in popularity in recent years.

It is a **multi-arm, cluster randomised, cross-over design where the only cross-over that is permitted is unidirectional**: from the control or routine care condition to the active intervention.

Due to its pragmatic design, the SWT design has been used to assess the effect of new interventions in real-world settings where simultaneous implementation across all clusters is not possible for logistic, financial or ethical reasons

A SWT may be chosen in preference to a parallel group cluster randomised design for reasons of practicality and statistical efficiency

Statistical methods for economic evaluations (EE) alongside cluster randomised controlled trials have recommended addressing clustering in both costs and outcomes as well as correlation between individual- and cluster-level costs and outcomes

DEFINITIONS OF SWT DESIGNS

SWT designs can be categorised into 4 main categories according to Copas et al. (2015).

Closed cohort: participants are identified and recruited at the beginning of the trial and participate continuously from the beginning until the end, without changing cluster, with their outcomes assessed at a series of follow-up times (typically pre-specified).

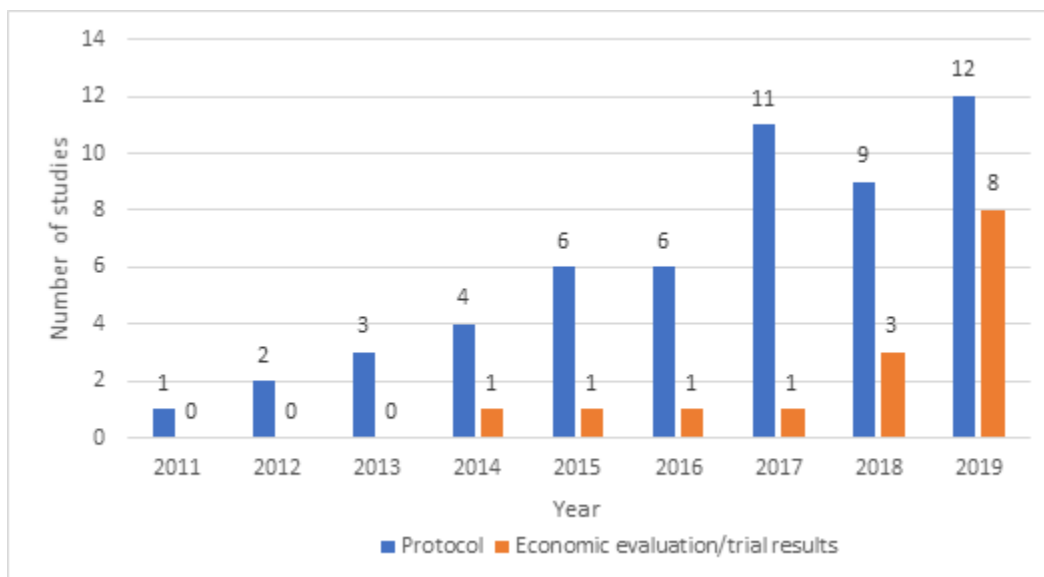
Open cohort: some participants are identified and recruited into the study from the beginning, others may become eligible and are recruited into the various clusters throughout the course of the study period, whilst some participants leave the trial prior to the conclusion of the study.

Continuous recruitment with short exposure (CRSE): As the name suggests, individuals become eligible and the recruitment is continuous as the study proceeds, with few (or even no) individuals recruited in the beginning. The main difference between CRSE and OC is that in the CRSE design, individuals are exposed to the intervention (or control) for only a short period of time. The outcome of interest is often measured after a follow-up period; single or repeated measurements or time-to-event from the beginning of their exposure period.

Repeated cross-section: A variation to the CRSE, whereby instead of continuous recruitment, participants are recruited at discrete time points, otherwise known as (vertical) cross-sections, with each cross-section acting as discrete snapshots at given times.

MAIN FINDINGS AND CRITICAL ASSESSMENT

Of the 69 included studies, 36 (52%) studies used a closed cohort design, 18 (26%) studies used a continuous recruitment with short exposure, 8 (12%) used an open cohort and 7 (10%) used a repeated cross-section design.



The studies have been conducted in: Europe (n=35, 51%), Australia (n=16, 23%), North America (n=8, 12%), Africa (n=5, 7%), Asia (n=4, 6%) and South America (n=1, 1%).

Of the 69 included studies, 54 (78%) were SWT protocols, 7 (10%) reported the findings from both the SWT and EE and 8 (12%) detailed the EE alone.

Cost-effectiveness analysis (n=46) was the most commonly used type of EE, followed by cost-utility analysis (n=43) and cost-benefit analysis (n=4). 24 (35%) studies used both cost-effectiveness and cost-utility analysis and 1 conducted both cost-utility and cost-benefit analysis. 2 protocol studies did not report the type of EE that they planned to conduct.

Statistical methods of economic evaluations

34/69 (49%) studies did not report any statistical methods, of which 31/34 (91%) studies were SWT protocols, 2/34 (6%) full EE [34, 80] and 1/34 (3%) trial results study

A total of 16 studies (9/23 (39%) protocols and 7/12 (58%) EE and trial results) reported methods that accounted for time, clustering and correlation between costs and outcomes by means of **mixed/multilevel models**.

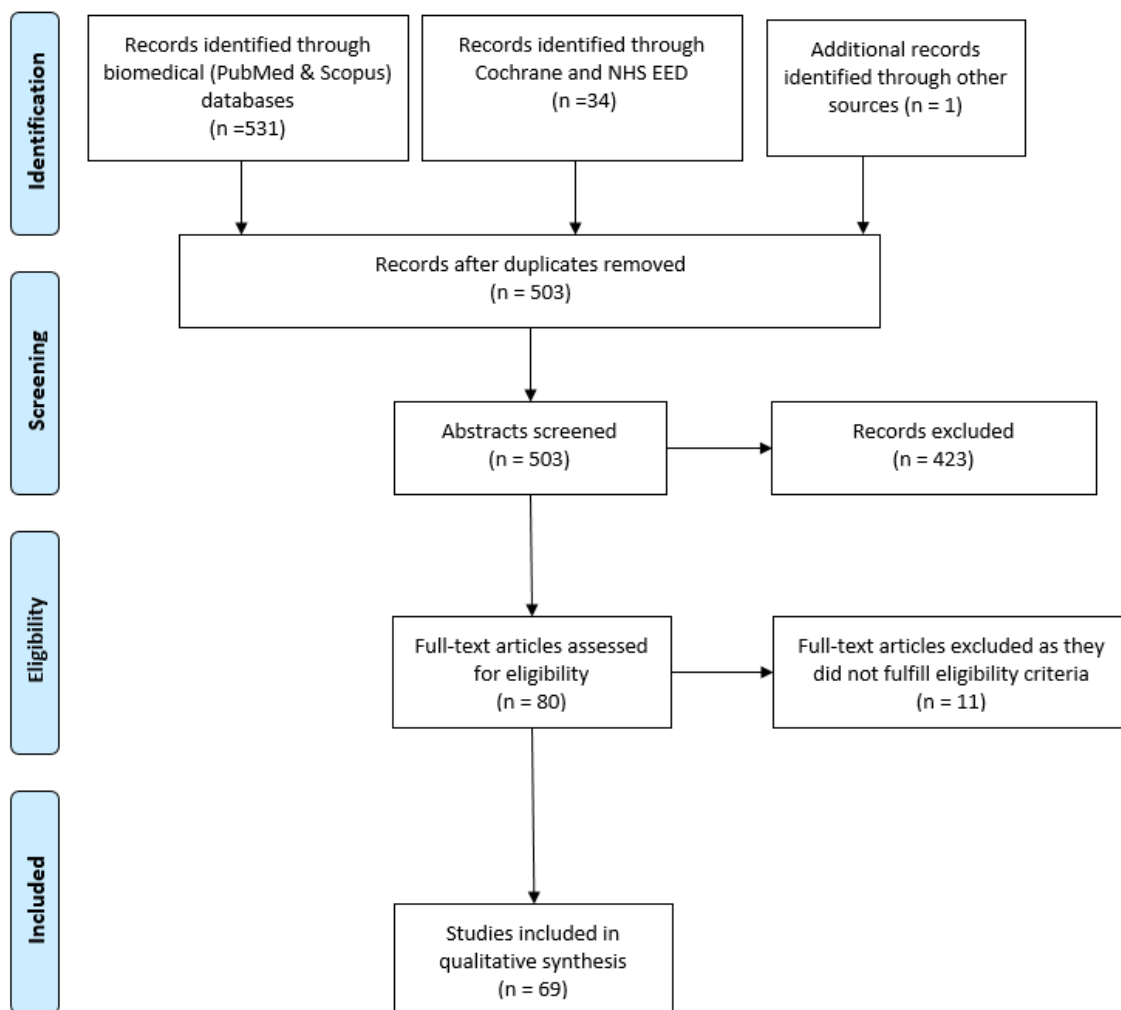
2 trial studies reported the use of **Seemingly Unrelated Regression**, one was a CC the other a CRSE. 8 papers (seven protocol, one trial results) discussed or used **generic bootstrap**, four of which was based on a CC type of SWT the other four on a CRSE design. 4 studies described the use of modelling techniques such as **Bayesian techniques** with Monte Carlo simulation (one protocol), Markov models (one protocol), decision trees (one EE) or generic probabilistic sensitivity analysis (one protocol). 4 protocol studies discussed generic regression methods (likely failing to take into account clustering, time trends and/or correlation between costs and outcomes) and one EE reported the use of empirical estimates of costs and effectiveness results.

Modified checklist to assess statistical methods of 15 EE & Trial results papers

Item	Yes	No	NA/Unclear
Was clustering accounted for in the univariate analysis of outcomes?	11 (73%)	0	4 (27%)
Was clustering accounted for in the univariate analysis of costs?	8 (53%)	0	7 (47%)
Did the statistical analysis account for the correlation between costs and outcomes?	9 (60%)	0	6 (40%)
Did the study account for clustering and correlation in the estimation of incremental cost-effectiveness?	9 (60%)	0	6 (40%)
Did the study explicitly make appropriate assumptions about the distribution of outcomes?	9 (60%)	2 (13%)	4 (27%)
Did the study explicitly make appropriate assumptions about the distribution of costs?	7 (47%)	2 (13%)	6 (40%)
Did the study adjust for time in their analysis?	5 (33%)	3 (20%)	7 (47%)

METHODS OF THE SYSTEMATIC REVIEW

This systematic review considered all studies that reported the intent to (through a protocol) or conduct of EE alongside or based on data from SWTs. Literature searches extending to February 2020 were conducted in biomedical databases including PubMed, Scopus and the Cochrane Library as well as the National Health Service Economic Evaluation Database (NHS-EED) up to the 1st February 2020



CONCLUSIONS

There was a lack of detail in describing the design of the SWT.

Similarly, **the reporting of statistical methods associated with the EE was often non-existent** and among the 35 (out of the 69) that reported this essential information, 17 did not use appropriate statistical methods, failing to consider critical elements of a SWT such as time, clustering and correlation between costs and outcomes.

It is important that appropriate statistical analyses that account for time, clustering and correlation between costs and outcomes are used.

Our modelling proposals are:

- Bivariate Multilevel/Mixed models
- Seemingly Unrelated Regressions
- Two-Stage Bootstrap method
- Bayesian Hierarchical models

From preliminary simulations conducted on the first three methods we have found MLM and SUR performed well under a variety of conditions, with a low mean bias and high confidence interval coverage.

DISCLOSURES

No disclosures, nor conflict of interests

ABSTRACT

OBJECTIVES : Recently, there has been an increase in use of the stepped wedge trial (SWT) design in the context of health services research, due to its pragmatic and methodological advantages over the parallel group design. Our objectives were twofold: review the statistical methods used when conducting economic evaluations alongside SWTs and make a proposal for appropriate methods to use.

METHODS : A methodological systematic literature search was conducted (up to February 2020) in the PubMed, Scopus, Cochrane and NHS-EED databases to find and evaluate studies where there was an intention to conduct an economic evaluation alongside a SWT.

RESULTS : A total of 69 studies were identified and included in this review. 54 studies were published protocols, 8 economic evaluations and 7 studies reporting full trial results. Included studies varied in terms of their reporting of statistical methods, in both detail and methodology. There were 34 studies that did not report any statistical methods for the economic evaluation and only 16 studies reported appropriate methods, mainly using some form of mixed/multilevel models, seemingly unrelated regression and Bayesian techniques with Monte Carlo simulation. The remaining studies failed to appropriate account for clustering, correlation or adjusted for time.

CONCLUSIONS : The use of appropriate statistical methods that account for time, clustering, and correlation between costs and outcomes is an important part of SWT health economics analysis and in particular we recommend the use of one of the following methods: hierarchical/multilevel models, seemingly unrelated regressions, two-stage bootstrap methods and bivariate Bayesian models. We also propose the extension of the CHEERS checklist in order to encourage transparent reporting of economic evaluations alongside SWTs.

REFERENCES

Lung T, Si L, Hooper R, Di Tanna GL. Health Economic Evaluation Alongside Stepped Wedge Trials: A Methodological Systematic Review. Pharmacoeconomics. 2021 Jan;39(1):63-80. doi: 10.1007/s40273-020-00963-x. Epub 2020 Oct 5

Copas AJ, Lewis JJ, Thompson JA, Davey C, Baio G, Hargreaves JR. Designing a stepped wedge trial: three main designs, carry-over effects and randomisation approaches. Trials. 2015;16:352. doi:10.1186/s13063-015-0842-7.