STEPPED-WEDGE CLUSTER RANDOMIZED TRIALS: PROTOCOL FOR DATASET COLLECTION AND ANALYSES

Team (alphabetical): Patrick Heagerty[1], Karla Hemming[2], James Hughes[1], Avi Kenny[3], Fan Li[4], Yongdong Ouyang[5], Julia Shaw[6], Monical Taljaard[6], Emily Voldal[7], Fan Xia[8]

[1] University of Washington, Seattle, WA, USA
[2] University of Birmingham, Birmingham, UK
[3] Duke University, Durham, NC, USA
[4] Yale University, New Haven, CN, USA
[5] The Hospital for Sick Children, Toronto, Ontario, Canada
[6] University of Ottawa, Ottawa, Ontario, CANADA
[7] Fred Hutchinson Cancer Center, Seattle, WA, USA
[8] University of California at San Francisco, San Francisco, CA, USA

Oct 24, 2023

Summary: Stepped wedge cluster randomized trials are often used in health care and medical settings to evaluate the impact of system-level interventions or policies. In recent years a large literature has developed on approaches to analyzing data from stepped wedge trials. The vast majority of these approaches rely on an assumption that the treatment effect in a stepped wedge trial is constant regardless of the time since introduction of the treatment. However, recent methodologic research has pointed out a potentially serious shortcoming in approaches that rely on this assumption. It is unclear, however, how often this assumption is violated in practice. Therefore, we propose to collect a large number of de-identified datasets from stepped wedge trials by downloading publicly available datasets and contacting individuals who have published results of stepped wedge cluster randomized trials. The data will be stored on a Shared Google Drive administered by Dr. Hughes and access will be limited to the investigators named in this protocol. We will analyze each dataset by methods that do and do not rely on the assumption of a constant treatment effect and compare the results.

Background:

Cluster randomized trials (CRTs) involve randomizing groups of individuals (e.g. classrooms, medical practices, communities) to a treatment or control condition and are often conducted when individual randomization is impractical or when the treatment, by its nature, must be delivered at the cluster level. Cluster randomized trials are often used to evaluate the impact of health care system level interventions through modification of the systematic processes used to treat patients. Importantly, although the treatment is delivered at the cluster-level, outcome measurements are collected on individuals within the clusters.

One type of CRT design is the stepped wedge. In a stepped wedge CRT, all clusters begin in the control condition and switch over to the treatment condition in a staggered manner – see figure 1. Each unique treatment pattern is called a sequence and clusters are randomly assigned to the different sequences. Data are typically collected from all clusters in each time period, most often cross-sectionally (ie different individuals contribute data in each period). Alternatively, a study may take repeated observations on the same individuals over time (cohort design). The strengths and limitations of the stepped wedge design have been extensively discussed in recent years (Hemming and Taljaard, 2020).

		Period				
		1	2	3	4	5
Cluster	1	C	T	T	T	T
	2	C	C	T	T	T
	3	C	C	C	T	T
	4	C	C	C	C	T

Figure 1 Stepped wedge design. "C" represents control; "T" represents a treatment

Most analyses of data from stepped wedge trials assume that the treatment effect is "instantaneous" (reaches full effect within the time period in which it is introduced) and constant (does not vary as a function of time since introduction – the "exposure time"). More generally, however, the treatment effect may vary as an arbitrary function of exposure time (time-varying treatment effect). Importantly, Kenny et al (2022) show that assuming the treatment effect is instantaneous and constant when it is not can lead to extremely misleading estimates of the treatment effect. Kenny et al (2022) (and, subsequently, others e.g. Maleyeff et al (2023)) have proposed alternative estimates of the treatment effect that do not rely on the instantaneous, constant treatment (IT) effect assumption. However, it remains unclear how often this assumption is violated in practice. We propose to collect a large number of datasets from stepped wedge CRTs and analyze them to determine the prevalence of time-varying treatment effects in real datasets. Such a database of trials could also be a valuable resource for addressing other methodological questions in the design and analysis of stepped wedge trials.

Objectives:

1) Collect and organize a large number of datasets from previously published stepped wedge CRTs.

2) Analyze these datasets to determine the prevalence of time-varying treatment effects and to demonstrate the statistical properties of various estimators in the presence of time-varying treatment effects.
3) Create a resource that can be used to address other questions about statistical methodology relevant to the design and analysis of stepped wedge CRT.

Data Collection:

We will aim to identify a comprehensive set of completed published stepped wedge CRTs. Any type of stepped wedge CRT (cross-sectional or cohort) will be eligible, provided there are at least two sequences and three periods, and a minimum of five independent randomized clusters. No restrictions will be placed on country of study conduct. The main source for identifying potential trials will be a recent systematic review (which included some of the investigators on this project as coauthors) that identified 160 stepped wedge trials published between 2016 – 2022 (Nevins et al, 2022). In addition, the investigators named on this protocol will identify additional stepped wedge studies where data are already publicly available or studies they were involved in as co-investigators

For trials where the data are already publicly available, the available data will be downloaded directly. For trials with a data sharing statement indicating that the trial data are available upon request, as well as trials without any statement about data sharing, we will obtain the email addresses of the contact person indicated in the data sharing statement and/or the corresponding author for the trial. Trials stating that data will not be available for sharing will be excluded. We will contact the corresponding author or investigator via email to request access to the dataset used in the published study or report. A draft of our proposed email is included in the appendix (note that there are slightly different versions of the email depending on whether the paper included a data sharing statement or not).

As seen in the draft letters, we will request a de-identified[1] dataset with only the few variables needed for our proposed analyses. The requested variables are

- primary outcome (and any secondary outcomes available under the data sharing statement);
- period indicator (indicating the interval or time period for each data point);
- treatment indicator (intervention or control);
- cluster indicator;
- any variables used in the randomization (e.g., stratification variables);
- the coded individual subject indicator (only needed if the design is a cohort design).

The preferred format is individual-level data but if it is more convenient or preferred, the investigator may choose to send us a dataset that is summarized to the cluster-period level (such summaries would only be appropriate for cross-sectional designs with binary outcomes; cohort designs and non-binary outcomes require individual-level data for analysis).

If the investigator indicates that the data cannot be shared, we may – depending on how many times this occurs – ask if he/she is willing to run data analysis code (that we would provide) and send us the output.

[1] Individual participants will always be deidentified. In some cases, the cluster (e.g. hospital) may be identifiable (to us) based on information given in the study publication.

If we receive no response to our initial email within 10 working days, we will send a followup email. If there is no response to that second email, we may opt to ask if the investigator is willing to run data analysis code (that we would provide) and send us the output. The maximum number of contacts will be three.

Data Storage:

De-identified datasets received from investigators will be stored on a secure and encrypted Shared Google Drive administered by Dr. Hughes. A Shared Google Drive is a cloud-based data storage and file sharing platform that allows a team of investigators to easily share documents and files. Google Drives have built-in protection against malware and ransomware, and are FERPA compliant. Files are automatically backed up by Google and any file that is accidentally deleted may be recovered up to 30 days following deletion.

Storing files on a Shared Google Drive eliminates the need for each investigator to maintain a local copy of the datafiles and other documents. To access the drive an investigator must login to their Google-affiliated account and the drive administrator must explicitly provide access to that account. For those given access, the shared drive shows up as a virtual network drive (e.g. G: drive) on their local computer and files may be accessed as if they were stored locally.

Access to the shared drive will be limited to the project investigators named on this protocol and their trainees . All investigators/trainees will be given read-access ("Viewer" permission) to the data. A subset of investigators who need to upload files to the drive will be given read/write/edit access ("Content Manager" permission). All investigators and trainees given access to the shared Google drive will provide written confirmation that they

i. will not attempt to identify individual participants in any study (which we believe would not be possible anyway)
ii. will only access datafiles for analysis through the Shared Google Drive and will not download any of the datafiles to their hard drive.

Data Analysis:

All analyses will be run on files stored on the Shared Google Drive. Files will not be downloaded for analysis.

1) Primary Analysis
   Our initial analysis of the stepped wedge datasets will focus on the issue of time-varying treatment effects. We will analyze all available stepped wedge trial datasets using different models and report the estimated treatment effect and precision given by each model. Estimated treatment effects of interest include those from models that

   i. Assume an instantaneous and constant treatment effect ("standard model")
   ii. Use the exposure time indicator method described by Kenny et al (2022) ("ETI model")
   iii. Assume treatment effect heterogeneity over exposure time as described by Maleyeff et al (2023) ("TEH model")

   Information about estimated treatment effects will be presented both in aggregate and by trial. In no case will results on individual trial participants or individual clusters within a trial be presented. Also,

while individual trial results may be presented, the trials will not be explicitly identified in any publications (although in some cases it is possible that unique characteristics of a trial may make it identifiable).

We also plan to investigate the impact of different modelling approaches (e.g. mixed models versus generalized estimating equations), different correlation structure assumptions, and trial characteristics (e.g. field of study, type of outcome, number and size of clusters, number of time periods) on the results.

We will offer to share the results of this analysis with participating researchers and we will cite the original publication of the data by contributing authors to acknowledge their contribution to the project.

2) Future analyses

The datasets collected under this protocol will form a valuable resource for investigating various methodological issues relevant to the design and analysis of stepped wedge randomized trials. As such, the investigators may wish to use these data for future studies of statistical methodology for stepped wedge design trials. Only projects that focus on answering a question about statistical methodology for stepped wedge trials will use the data from the repository. We will not use the data for general meta-analyses or other large scale attempts to reinterpret the trials.

References

Hemming K, Taljaard M. Reflection on modern methods: when is a stepped-wedge cluster randomized trial a good study design choice? International Journal of Epidemiology. 49:1043-1052, 2020.

Kenny A, Voldal E, Xia F, Heagerty PJ, Hughes JP. Analysis of stepped wedge cluster randomized trials in the presence of a time-varying treatment effect. Statistics in Medicine 41:4311-4339, 2022.

Maleyeff L, Li F, Haneuse S. Wang R. Assessing exposure-time treatment effect heterogeneity in stepped-wedge cluster randomized trials. Biometrics 79:2551-2564, 2023.

Nevins P, Davis-Plourde K, Macedo JAP, Ouyang Y, Ryan M, Tong G, Wang X, Meng C, Ortiz-Reyes L, Li F, Caille A, Taljaard M. A scoping review described diversity in methods of randomization and reporting of baseline balance in stepped-wedge cluster randomized trials. Clin Epidemiol. 157: 134–145, 2023.