HHS Public Access

Author manuscript

J Clin Epidemiol. Author manuscript; available in PMC 2018 June 01.

Published in final edited form as:

J Clin Epidemiol. 2017 June ; 86: 160–167. doi:10.1016/j.jclinepi.2017.04.004.

Proposed variations of the stepped-wedge design can be used to accommodate multiple interventions

Vivian H Lyons, MPH[a,b,*], **Lingyu Li, MS**[c], **James P Hughes, PhD**[c], and **Ali Rowhani-Rahbar, MD, PhD, MPH**[a,b]

[a]Department of Epidemiology, School of Public Health, University of Washington, Box 357236, Seattle, WA 98195-7236

[b]Harborview Injury Prevention & Research Center, 401 Broadway Ave, 4th floor, Seattle, WA 98122

[c]Department of Biostatistics, School of Public Health, University of Washington, Box 357232, Seattle, WA 98195-7232

Abstract

Objective—Stepped wedge design (SWD) cluster randomized trials have traditionally been used for evaluating a single intervention. We aimed to explore design variants suitable for evaluating multiple interventions in a SWD trial.

Study Design and Setting—We identified four specific variants of the traditional SWD that would allow two interventions to be conducted within a single cluster randomized trial: Concurrent, Replacement, Supplementation and Factorial SWDs. These variants were chosen to flexibly accommodate study characteristics that limit a one-size-fits-all approach for multiple interventions.

Results—In the Concurrent SWD, each cluster receives only one intervention, unlike the other variants. The Replacement SWD supports two interventions that will not or cannot be employed at the same time. The Supplementation SWD is appropriate when the second intervention requires the presence of the first intervention, and the Factorial SWD supports the evaluation of intervention interactions. The precision for estimating intervention effects varies across the four variants.

Conclusion—Selection of the appropriate design variant should be driven by the research question while considering the trade-off between the number of steps, number of clusters, restrictions for concurrent implementation of the interventions, lingering effects of each intervention, and precision of the intervention effect estimates.

Keywords

Multiple interventions; Pragmatic trials; Stepped wedge trial design

[*]**Corresponding Author:** Vivian H. Lyons – Department of Epidemiology, School of Public Health, University of Washington, 1959 NE Pacific Street, Health Sciences, Bldg, F-262, Box 357236, Seattle, WA 98195-7236. Phone: +1 206 412 5894. vlyons@uw.edu.

Conflicts of Interest: None.

1. Introduction

The use of stepped wedge designs (SWD) in public health and clinical research has gained popularity since the Gambia Hepatitis study [1]. The traditional SWD is a unidirectional crossover design in which time of the intervention implementation is randomized at the cluster level, with one or more clusters following the same randomization pattern within each cluster group (Figure 1). The SWD offers a pragmatic approach where clusters initially serve as a controls before receiving the intervention at a subsequent time step; eventually, all clusters receive the intervention [2]. This is a major departure from parallel cluster randomized trials where assignments for intervention and control groups are decided at the beginning of the study and not changed subsequently [3].

In a SWD trial, the intervention effect is estimated using both within-cluster and between-cluster information. Indeed, one of the strengths of SWD is that by including within-cluster information, time-invariant confounding on the cluster level can be avoided and precision may be gained [3]. The stepped rollout can also increase logistic feasibility when simultaneous intervention implementation in many clusters may be prohibitive. A SWD can be the preferred design choice when it is necessary or desirable for all clusters to receive the intervention by the end of the study (e.g. when randomizing to a control group is unethical), thus precluding the parallel or bidirectional crossover cluster randomized trial design. Another common motivation for using a SWD is having too few clusters for a parallel design. In these situations a SWD can provide higher quality evidence than purely observational studies or pre-post-assessments [4–6]. While some argue that two-arm parallel group or crossover cluster randomized trial design, when possible, is preferable to a traditional SWD, a SWD can be more efficient than a parallel group cluster randomized trial design under certain conditions [4,7,8]. Other scientific and logistical reasons for choosing a SWD are described in Mdege, et al. 2012 and elsewhere [7,5].

There are also inherent limitations to a SWD [4–11]. By design, the treatment effect in SWD is partially confounded by time, thus estimation of the intervention effect is often model-dependent. When using a SWD, consideration must be given to both the amount of time required to begin the intervention rollout for each cluster as well as the length and placement of any wash-in or wash-out periods [12]. Other limitations have been discussed elsewhere [2,3].

SWD have traditionally been used to evaluate a single intervention. However, the need and desire to evaluate multiple interventions in a specific setting is not uncommon [13–15]. For example, to combat childhood obesity a school might implement both a policy change to school lunch requirements and a new nutrition curriculum [16]. Evaluating multiple interventions in a single trial using a SWD has the potential to decrease the required time until each cluster receives an intervention, improve participant engagement, decrease total funding required, enhance efficiency, and allow for assessing potential interaction between the interventions as well as decrease total clusters needed compared to conducting two separate trials. Here, we propose new SWD variants that can be used for rolling out and evaluating multiple interventions and outline their features and efficiency. Note that these variants apply only in the context of multiple interventions, as without the second

intervention, the models would essentially reduce to a single intervention SWD. Examples from the existing literature where the proposed design variants could have been considered are provided throughout. The overarching goal is to provide a framework for researchers who plan to examine the effectiveness of multiple interventions using SWD in their work.

2. Methods

This paper was motivated by the need to evaluate the effects of a new healthcare protocol in conjunction with a technological support piece. The technological support piece could only be applied following the implementation of the new healthcare protocol (described in section 3.1.3. below), and we were interested not only in evaluating their joint effect, but also the effect of the healthcare protocol by itself. Due to the cost of implementation per cluster, limited number of available clusters and each cluster conditioning their participation on receiving the intervention, a SWD was chosen as the best design for this study, assuming it could be modified to accommodate the two interventions. In response to this identified gap in methodology, we developed four variants of SWD to support the evaluation of two interventions in one trial. We denote these variants as Concurrent, Replacement, Supplementation, and Factorial SWD. Each design utilizes unidirectional crossover, a hallmark of SWD. The designs are described in more detail in section 3.

To understand the relative efficiency of the different design options, and as an aid to power calculation, we also developed methods to calculate the variance of each of the two interventions. We used the following equation which extends the original model proposed by Hussey and Hughes (i.e., Intervention 1 and Intervention 2) [12],

$$\mu_{\mathrm{ij}}=\mu+\alpha_i+\beta_j+\sum_{m=1}^{M}X_{\mathrm{mij}}\eta_m \quad (1)$$

where α_i is a random effect for cluster i such that $\alpha_i \sim N(0,\tau^2)$ (i in 1, …, I), β_j is a fixed effect corresponding to time interval j (j in 1, …, $T-1$, $\beta_T=0$ for identifiability), X_{mij} is an indicator of the treatment mode for the m^{th} intervention in cluster i at time j (1 = intervention; 0 = control), η_m is the treatment effect for the m^{th} intervention, I is the number of clusters, T is the number of time points, and M is the number of interventions. Note that this equation assumes no interaction between the interventions, although it could be readily extended to include such a term. An equation for the observed cluster means is:

$$Y_{\mathrm{ij}}=\mu_{\mathrm{ij}}+e_{\mathrm{ij}} \quad (2)$$

where $e_{ij} \sim N(0,\sigma^2)$. Therefore, we have $Y_{ij} \sim N(\mu_{ij}, \tau^2+\sigma^2)$. The dependence between observations in the same cluster is often parameterized in terms of the intracluster correlation coefficient (ICC) defined as $\mathrm{ICC}=\tau^2/(\tau^2+\sigma_e^2)$ where $\sigma^2=\sigma_e^2/N$. N is the number of individuals sampled per cluster per time interval.

Let $\theta=(\mu, \beta_1,\cdots, \beta_{T-1}, \eta_1, \cdots, \eta_M)$. For each design variant, we can derive a corresponding design matrix [16], Z, and the variance of $\hat{\theta}$ can be calculated by:

$$\mathrm{cov}(\hat{\theta}) = (Z'V^{-1}Z)^{-1} \quad (3)$$

where V is an $IT\ x\ IT$ block diagonal matrix. Each $T\ x\ T$ block is of the form:

$$\begin{pmatrix} \sigma^2+\tau^2 & \tau^2 & \cdots & \tau^2 \\ \tau^2 & \sigma^2+\tau^2 & \cdots & \tau^2 \\ \vdots & \vdots & \ddots & \vdots \\ \tau^2 & \tau^2 & \cdots & \sigma^2+\tau^2 \end{pmatrix}$$

We used this approach to evaluate the variance of $\hat{\eta}_1$ and $\hat{\eta}_2$ for each of the design variants (see the appendix of Hughes, Granston and Haegerty [3] for more details and an approach for dealing with more complex models). Unless otherwise specified, all analyses compare each intervention to the control.

3. Results

We identified four major design variants of SWD for multiple interventions – Concurrent, Replacement, Supplementation, and Factorial SWDs. Each is suitable for addressing certain scientific questions of interest regarding the effect of two interventions in a SWD.

3.1. Classification and features of designs

3.1.1. Variant 1: Concurrent—This design is conceptually similar to the traditional SWD, with two separate interventions evaluated *at the same time* and in the same setting, but in different clusters (Figure 2A). Each cluster receives only Intervention 1 or Intervention 2. Both interventions are rolled out in a traditional SWD fashion and the study ends when all clusters receive either Intervention 1 or 2. The difference between this design and the traditional SWD is that the same time periods can serve as a control periods for both interventions, thereby increasing efficiency in estimating intervention effects.

A recent protocol published by Reuther, et al. in 2014 proposed a form of the Concurrent design to study the impact of two types of case conferences [WELCOME-IdA or WELCOME-NEO] (Web Figure 1) on quality of life in dementia patients in nursing homes [17]. With 12 nursing homes, they plan to conduct two separate SWD studies with 6 nursing homes in each, where each nursing home receives only one of the case conference methodologies. By using a Concurrent design, the investigators could achieve more precise estimates of each intervention's effect as well as a direct comparison of the two interventions (if desired).

3.1.2. Variant 2: Replacement—In the Replacement design, introduction of Intervention 1 follows the same pattern as the introduction of the intervention in the traditional SWD. Subsequently, however, Intervention 2 *replaces* Intervention 1 in each group of clusters (Figure 2B). The study ends when all groups of clusters have received Intervention 2. This variation is appropriate for assessing two interventions that will not or cannot be employed

at the same time and when the number of clusters available for the study is limited. This design could easily be modified to include a washout period to minimize the potential for any carryover effects from Intervention 1 to 2. Since it is not possible for a cluster to cross over directly from the control period to Intervention 2, we hypothesized that this design would underperform compared with other variants in evaluating the effect of Intervention 2.

Haines, et al. published a protocol for two nested SWD trials taking place in the same hospitals within one larger study but without any time periods that contain both SWD trials; the second SWD trial begins in the time period after the first SWD trial (Web Figure 2) [18]. For investigators interested in rolling out two interventions back to back, using the Replacement design may decrease funding required as well as total study duration needed to carry out the project as compared to conducting two separate SWD studies in series.

3.1.3. Variant 3: Supplementation—The Supplementation design introduces Intervention 1 and then *adds* Intervention 2 (Figure 2C). This is an especially suitable design for two interventions where the second intervention requires the presence of the first intervention. The study ends once all groups of clusters have received the combination of Intervention 1 + 2. While it would be possible to use this design when the two interventions could be offered both individually and simultaneously, the Factorial design (discussed below) would be the most appropriate choice in that case. The Supplementation design allows investigators to directly analyze the effect of Intervention 1 and combined Intervention 1 + 2. The Supplementation design may also be used to estimate the independent effect of Intervention 2 (which is never observed on its own) under the assumption of no interaction between the two interventions. Note that the Supplementation design can be considered as a reparameterization of the Replacement design since the combined effect of Intervention 1+2 in the Supplementation design is equivalent to the Intervention 2 effect in the Replacement design. Equivalently, the Intervention 2 effect in the Supplementation design is equivalent to the difference between Interventions 2 and 1 in the Replacement design. Intervention 1 is equivalent in both designs (Figure 3).

The Supplementation design could be used to determine the incremental benefit of, for example, a technological support piece (e.g., a near-real time support system as an add-on) during the rollout of a new healthcare protocol. Training the staff and implementing the new healthcare protocol would be Intervention 1 with the addition of the technological support piece employed as Intervention 2. In addition to reducing the complexity of introducing the entire intervention package at once, the Supplementation design provides an estimate of the contribution of each component to the overall intervention effect.

3.1.4. Design 4: Factorial—The Factorial design is characterized by concurrent rollout of Interventions 1 and 2, thereby allowing estimation of Intervention 1 and 2 separately as well as together (Figure 2D). The study ends once all clusters have received both Intervention 1 and 2. This design is appropriate for two interventions that can be employed concurrently or separately, allowing investigators to estimate both individual effects and the interaction between the two interventions. While this design could be chosen specifically for its ability to assess interaction between the two Interventions, it can also be chosen to reduce the financial and logistical burden of conducting two separate traditional SWD trials when the

two interventions are independent. The Factorial design may improve recruitment as it provides an intervention at an earlier time step for clusters or individuals that would otherwise receive interventions last in a traditional SWD.

We identified several studies that could have potentially used the Factorial design to assess the effects of multiple interventions in the same cluster group. One of these studies compared two behavioral intervention programs [Intervention 1: Stepping Stones Triple P and Intervention 2: Stepping Stones Triple P plus Acceptance and Commitment Therapy] to reduce behavioral problems in children with cerebral palsy [19]. Parents of children with cerebral palsy were assigned to one of three arms of the study, Intervention 1, Intervention 2 or waitlist control. Inclusion of an Acceptance and Commitment Therapy-only arm (as used in other studies) [20–23] and use of the SWD Factorial design could allow investigators to provide all patients with an intervention sooner than with the traditional SWD, or a Supplementation design. It would also allow investigators to assess the effect of the Acceptance and Commitment Therapy intervention alone, as well as the potential synergistic or antagonistic effects between the two interventions. For investigators interested in testing different interventions in the same population, the Factorial design could be used for the roll out of multiple interventions concurrently. This is conceptually similar to running two separate stepped wedge studies in the same study population at the same time. Durovni and colleagues recently completed two separate SWD studies in Brazil. Their first study examined the impact of tuberculosis screening training at HIV clinics on tuberculosis incidence [24], and their second study examined the impact of changing smear microscopy on the notification rate of lab-confirmed tuberculosis [25]. While these studies took place in different settings (HIV clinics vs. laboratories), they could have been run in concert with each other using a Factorial design if they had chosen HIV clinics with attached laboratories for their study, thereby decreasing total clusters and study cost needed to assess both interventions. A Factorial design could also measure the impact of dual interventions on a single outcome or, as in the Physician's Health Study (i.e., an individually-randomized trial with the traditional factorial design) [26], a SWD Factorial design can also estimate Intervention 1 specific outcomes and Intervention 2 specific outcomes separately.

3.2. Efficiency of design variants

Although the choice of design should be driven by the research question, in some cases it may be possible for investigators to choose between designs. In this section, we explore the statistical efficiency of the various designs for estimating the treatment effects. Each design was analyzed with 5 time intervals and 24 clusters equally divided between the design variant's cluster groups to allow comparability, although each design could be arranged with varying numbers of clusters or time periods. We assumed 100 subjects sampled cross-sectionally in each cluster at each time point, an outcome with variance $\sigma_e^2=1$, a nonparametric time effect (as in equation 1) and no interaction between Interventions 1 and 2. All variances were calculated using equation 3 (Figure 3).

In both the Concurrent and the Factorial designs, the variances of $\hat{\eta}_1$ and $\hat{\eta}_2$ are equal. In the Replacement design, the variance of $\hat{\eta}_1$ is generally lower than the variance of $\hat{\eta}_2$. To better understand this, consider the independent case (ICC=0). When the ICC=0 and the model

shown in equations (1) and (2) is used for analysis, only between-cluster comparisons contribute to the estimation of the treatment effects. Since there are generally more between cluster comparisons of Intervention 1 to control than Intervention 2 to control, Intervention 1 is estimated more precisely. In general, the Replacement design is a weak choice, in terms of precision, for evaluating Intervention 2. In the Supplementation design, the variances of the Intervention 1 and 2 effects are equal. Additionally, due to the equivalence of the Replacement and Supplementation designs noted previously, the variance of the combined effect of Interventions 1 and 2 in the Supplementation design is equal to the (relatively large) variance of Intervention 2 alone in the Replacement design.

The variance of the comparison between the interventions (Intervention 2 minus Intervention 1) is similar (in this example) for the Replacement and Factorial designs, highly ICC dependent for the Concurrent design, and relatively large for the Supplementation design. Indeed, in the Supplementation design, the variance of the Intervention 2 versus Intervention 1 contrast is much larger than the variance of either Intervention compared to the control, and not ICC dependent (suggesting that all the information for this contrast arises from within-cluster comparisons). For the Concurrent design, the variance of Intervention 2 versus Intervention 1 is similar to the variance of $\hat{\eta}_1$ and $\hat{\eta}_2$ (Figure 3).

4. Discussion

These design variants offer alternative approaches for studying more than one intervention using a stepped wedge cluster-randomized trial and can provide an opportunity for additional data collection and improved efficiency. The selection of the appropriate design variant should be primarily driven by the research question and nature of the interventions, with additional consideration given to ethical and logistical issues, minimum number of available clusters, restrictions for concurrent implementation based on intervention characteristics, plausibility of assumptions, statistical power and sample size, and the precision of the intervention effect estimates. Indeed, as noted by Hargreaves et al., ethical issues may preclude use of a stepped wedge design at all in certain scenarios [27]. Although we do not advocate for the use of a SWD in all cases when evaluating multiple interventions, there are certain instances where a multiple intervention SWD trial is a natural choice. For example, if the investigator were testing two separate vaccines, a bi-directional cluster crossover design would not be possible as the vaccine-acquired immunity is not removable. Although a parallel group cluster randomized controlled trial (let’s say with 4 arms) could be another option, if all clusters would like to receive both vaccines during the course of the study, a Factorial design SWD would be more appropriate.

Reasons for employing a SWD variant with multiple interventions may differ from reasons for using a single intervention SWD. While a single intervention SWD is often used to assess an intervention’s effect during rollout of a public health program, use of the Concurrent, Replacement and Factorial designs may occur more frequently as a result of simultaneous rollout of two separate interventions (e.g. Concurrent and Factorial). In contrast, a Supplementation design is a natural choice when the first intervention must be present before the second can be implemented.

Unlike the single intervention SWD, where each cluster receives the same intervention, the Concurrent design would not provide both treatments to each cluster and the Replacement design would take away Intervention 1. Both of these features may preclude the use of the design for ethical or political reasons. However, similar to a single intervention SWD, the multiple intervention SWD variants would all ensure that each cluster received at least one of the interventions by the end of the study. Also like the single intervention SWD, use of one of the four SWD variants presented would likely decrease logistical challenges associated with simultaneous intervention implementation (compared to a parallel group cluster randomized trial) or the ethical/political challenge of maintaining a control group.

In this paper, we chose to use simplistic forms of these variants to demonstrate principals of evaluating multiple interventions with a SWD. Investigators could employ these four basic models flexibly with additional variations, including limiting the study to a single cluster per group (although this approach may not result in sufficiently powered SWD studies) and/or varying the number of time intervals spent in each intervention [28,29]. A recent study by Chinbuah, et al, showcases one of the many additional variations that can be employed – they used a form of the SWD Concurrent design to assess two drugs in the treatment of childhood fever substituting Intervention 1+2 for Intervention 2 [Intervention1: artesunate amodiaquine and Intervention 2: artesunate amodiaquine plus amoxicillin] (Web Figure 3) [30]. With any modifications to the design variants (e.g. adding time steps, adding interventions, etc.) the variances and precision would need to be re-calculated.

Our variance calculations were all done assuming “regular” stepping patterns and may not hold for irregular designs with unequal numbers of clusters in each step. Additionally, all calculations were done under the assumptions of equation 1 that includes only a random intercept term. Others have suggested adding random cluster*treatment or cluster*time terms to equation 1 [3]. Importantly, if random treatment effects (i.e. a cluster*treatment interaction) are added to equation 1, some findings (e.g., equal precision of the estimated intervention effects in the Concurrent, Supplementation and Factorial designs; Intervention 1 generally estimated more precisely than Intervention 2 in the Replacement design) may not hold, unless the two variances of the random treatment effects are equal. All of the calculations to compare the efficiency of models were also done assuming no interaction of the interventions although an interaction effect could be validly estimated in the Factorial design. Also, valid estimation of the Intervention 2 effect in the Replacement design assumed that there was no carry-over effect of Intervention 1. All variance calculations were computed using linear regression and normal errors. Finally, all our results assume cross sectional sampling; while it is possible that these SWD variants could be employed in a model that included correlation for repeated measures on each subject (cohort sampling), we did not investigate such a model and are not sure all our qualitative conclusions would hold. Similarly, our examination of the relative efficiency of designs was conducted only for one particular combination of clusters, subjects and time steps. It is unclear how changing those parameters would impact our findings. Further study is needed to identify the most efficient versions of these four SWD variants, or other variants for multiple interventions, much like the research done to optimize the traditional SWD [7,31].

5. Conclusion

These stepped wedge design variants offer alternative methods for studying multiple interventions using a cluster-randomized trial. The selection of the appropriate design should be driven by the research question. Other factors, such as the available number of clusters, time to conduct the study, interactions between the interventions, appropriate model structure and/or lingering effects of each intervention, should be considered. Appropriate use of these variants can provide investigators with potentially decreased cost and improved study feasibility, and may allow for addressing new scientific questions that were either unanswerable or challenging to address with the traditional stepped wedge design.

Acknowledgments

Special thanks to Dr. Monica Vavilala for her support and inspiration.

Funding/Support: Vivian Lyons was supported by the Maternal and Child Health Bureau, Health Resources and Services Administration (Grant #T76 MC00011); and James Hughes is the Principal Investigator of a grant from National Institute of Allergy and Infectious Diseases (Grant #AI29168), which contributes to his salary. Neither funding agency had a role in the study design, analysis, the writing of the report or the decision to submit the article for publication.

References

1. Hall AJ, Inskip HM, Loik F, et al. for the gambia hepatitis study group. The gambia hepatitis intervention study. Cancer Res. 1987; 47(21):5782–5787. [PubMed: 2822233]
2. Mdege ND, Man M-S, Taylor Nee Brown CA, Torgerson DJ. Systematic review of stepped wedge cluster randomized trials shows that design is particularly used to evaluate interventions during routine implementation. J Clin Epidemiol. 2011; 64:936–48. DOI: 10.1016/j.jclinepi.2010.12.003 [PubMed: 21411284]
3. Hughes JP, Granston TS, Heagerty PJ. Current issues in the design and analysis of stepped wedge trials. Contemp Clin Trials. 2015; 45:55–60. DOI: 10.1016/j.cct.2015.07.006 [PubMed: 26247569]
4. Kotz D, Spigt M, Arts ICW, Crutzen R, Viechtbauer W. Use of the stepped wedge design cannot be recommended: a critical appraisal and comparison with the classic cluster randomized controlled trial design. J Clin Epidemiol. 2012; 65:1249–52. DOI: 10.1016/j.jclinepi.2012.06.004 [PubMed: 22964070]
5. Mdege ND, Man M-S, Taylor nee Brown CA, Torgerson DJ. There are some circumstances where the stepped-wedge cluster randomized trial is preferable to the alternative: no randomized trial at all. Response to the commentary by Kotz and colleagues. J Clin Epidemiol. 2012; 65:1253–4. DOI: 10.1016/j.jclinepi.2012.06.003 [PubMed: 22968178]
6. Kotz D, Spigt M, Arts ICW, Crutzen R, Viechtbauer W. The stepped wedge design does not inherently have more power than a cluster randomized controlled trial. J Clin Epidemiol. 2013; 66:1059–60. DOI: 10.1016/j.jclinepi.2013.05.004 [PubMed: 23953488]
7. Girling AJ, Hemming K. Statistical efficiency and optimal design for stepped cluster studies under linear mixed effects models. Stat Med. 2016; 35doi: 10.1002/sim.6850
8. Hemming K, Taljaard M. Sample size calculations for stepped wedge and cluster randomised trials: a unified approach. J Clin Epidemiol. 2016; 69:137–46. DOI: 10.1016/j.jclinepi.2015.08.015 [PubMed: 26344808]
9. Kotz D, Spigt M, Arts ICW, Crutzen R, Viechtbauer W. Researchers should convince policy makers to perform a classic cluster randomized controlled trial instead of a stepped wedge design when an intervention is rolled out. J Clin Epidemiol. 2012; 65:1255–6. DOI: 10.1016/j.jclinepi.2012.06.016 [PubMed: 22964069]

10. Hemming K, Girling A. The efficiency of stepped wedge vs. cluster randomized trials: stepped wedge studies do not always require a smaller sample size. J Clin Epidemiol. 2013; 66:1427–8. DOI: 10.1016/j.jclinepi.2013.07.007 [PubMed: 24035495]
11. Hemming K, Girling A, Martin J, Bond SJ. Stepped wedge cluster randomized trials are efficient and provide a method of evaluation without which some interventions would not be evaluated. J Clin Epidemiol. 2013; 66:1058–9. DOI: 10.1016/j.jclinepi.2012.12.020 [PubMed: 23849737]
12. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. Contemp Clin Trials. 2007; 28:182–91. DOI: 10.1016/j.cct.2006.05.007 [PubMed: 16829207]
13. Parker AG, Hetrick SE, Jorm AF, Mackinnon AJ, McGorry PD, Yung AR, et al. The effectiveness of simple psychological and physical activity interventions for high prevalence mental health problems in young people: A factorial randomised controlled trial. J Affect Disord. 2016; 196:200–9. DOI: 10.1016/j.jad.2016.02.043 [PubMed: 26926659]
14. Pears S, Bijker M, Morton K, Vasconcelos J, Parker RA, Westgate K, et al. A randomised controlled trial of three very brief interventions for physical activity in primary care. BMC Public Health. 2016; 16:1033.doi: 10.1186/s12889-016-3684-7 [PubMed: 27716297]
15. Salmon J, Arundell L, Hume C, Brown H, Hesketh K, Dunstan DW, et al. A cluster-randomized controlled trial to reduce sedentary behavior and promote physical activity and health of 8–9 year olds: the Transform-Us! study. BMC Public Health. 2011; 11:759.doi: 10.1186/1471-2458-11-759 [PubMed: 21970511]
16. Seber, GA., Lee, AJ. Linear regression analysis. Vol. 936. John Wiley & Sons; 2012.
17. Reuther S, Holle D, Buscher I, Dortmann O, Muller R, Bartholomeyczik S, et al. Effect evaluation of two types of dementia-specific case conferences in German nursing homes (FallDem) using a stepped-wedge design: study protocol for a randomized controlled trial. Trials. 2014; 15:319.doi: 10.1186/1745-6215-15-319 [PubMed: 25118091]
18. Haines TP, O'Brien L, Mitchell D, Bowles K-A, Haas R, Markham D, et al. Study protocol for two randomized controlled trials examining the effectiveness and safety of current weekend allied health services and a new stakeholder-driven model for acute medical/surgical patients versus no weekend allied health services. Trials. 2015; 16:133.doi: 10.1186/s13063-015-0619-z [PubMed: 25873250]
19. Whittingham K, Sanders M, McKinlay L, Boyd RN. Interventions to reduce behavioral problems in children with cerebral palsy: An RCT. Pediatrics. 2014; 133:e1249–57. DOI: 10.1542/peds.2013-3620 [PubMed: 24709926]
20. Johns LC, Oliver JE, Khondoker M, Byrne M, Jolley S, Wykes T, et al. The feasibility and acceptability of a brief Acceptance and Commitment Therapy (ACT) group intervention for people with psychosis: The "ACT for life" study. J Behav Ther Exp Psychiatry. 2016; 50:257–63. DOI: 10.1016/j.jbtep.2015.10.001 [PubMed: 26480469]
21. Kohle N, Drossaert CHC, Schreurs KMG, Hagedoorn M, Verdonck-de Leeuw IM, Bohlmeijer ET. A web-based self-help intervention for partners of cancer patients based on Acceptance and Commitment Therapy: a protocol of a randomized controlled trial. BMC Public Health. 2015; 15:303.doi: 10.1186/s12889-015-1656-y [PubMed: 25884187]
22. Mak YW, Loke AY. The acceptance and commitment therapy for smoking cessation in the primary health care setting: a study protocol. BMC Public Health. 2015; 15:105.doi: 10.1186/s12889-015-1485-z [PubMed: 25879419]
23. Weineland S, Arvidsson D, Kakoulidis TP, Dahl J. Acceptance and commitment therapy for bariatric surgery patients, a pilot RCT. Obes Res Clin Pract. 2012; 6:e1–90. DOI: 10.1016/j.orcp.2011.04.004
24. Durovni B, Saraceni V, Moulton LH, Pacheco AG, Cavalcante SC, King BS, et al. Effect of improved tuberculosis screening and isoniazid preventive therapy on incidence of tuberculosis and death in patients with HIV in clinics in Rio de Janeiro, Brazil: a stepped wedge, cluster-randomised trial. Lancet Infect Dis. 2013; 13:852–8. DOI: 10.1016/S1473-3099(13)70187-7 [PubMed: 23954450]
25. Durovni B, Saraceni V, van den Hof S, Trajman A, Cordeiro-Santos M, Cavalcante S, et al. Impact of replacing smear microscopy with Xpert MTB/RIF for diagnosing tuberculosis in Brazil: a stepped-wedge cluster-randomized trial. PLoS Med. 2014; 11:e1001766.doi: 10.1371/journal.pmed.1001766 [PubMed: 25490549]

26. Christen WG, Gaziano JM, Hennekens CH. Design of Physicians' Health Study II--a randomized trial of beta-carotene, vitamins E and C, and multivitamins, in prevention of cancer, cardiovascular disease, and eye disease, and review of results of completed trials. Ann Epidemiol. 2000; 10:125–34. [PubMed: 10691066]
27. Hargreaves JR, Copas AJ, Beard E, Osrin D, Lewis JJ, Davey C, et al. Five questions to consider before conducting a stepped wedge trial. Trials. 2015; 16:350.doi: 10.1186/s13063-015-0841-8 [PubMed: 26279013]
28. Lyons, VH., Li, L., Hughes, JP., Rowhani-Rahbar, A. University of Washington Department of Biostatistics Technical Report. Working Paper 411. Seattle, WA: 2016. Evaluation of multiple interventions using a stepped wedge design. Avaliable at: http://biostats.bepress.com/uwbiostat/paper411/
29. Taljaard M, Teerenstra S, Ivers NM, Fergusson DA. Substantial risks associated with few clusters in cluster randomized and stepped wedge designs. Clin Trials. 2016; 13:459–63. DOI: 10.1177/1740774516634316 [PubMed: 26940696]
30. Chinbuah MA, Kager PA, Abbey M, Gyapong M, Awini E, Nonvignon J, et al. Impact of community management of fever (using antimalarials with or without antibiotics) on childhood mortality: a cluster-randomized controlled trial in Ghana. Am J Trop Med Hyg. 2012; 87:11–20. DOI: 10.4269/ajtmh.2012.12-0078
31. Lawrie, J., Carlin, JB., Forbes, AB. Optimal stepped wedge designs; Stat Probab Lett. 2015. p. 210-4.doi:doi.org/10.1016/j.spl.2015.01.024

WHAT IS NEW?

Key findings

- Evaluating multiple interventions using a stepped wedge design is feasible.
- Investigators could employ these four basic stepped wedge designs (i.e. Concurrent, Replacement, Supplementation, and Factorial stepped wedge designs) when evaluating multiple interventions with additional variations.

What this adds to what was known?

- Previous stepped wedge studies have been limited to evaluation of a single intervention; investigators may evaluate multiple interventions using these four designs.

What is the implication and what should change now?

- Use of the Replacement, Supplementation and Factorial stepped wedge designs allow investigators to assess the effect of multiple interventions within the same trial.
- Use of the Factorial stepped wedge design allows investigators to assess potential interactions between interventions.

	Time 0	Time 1	Time 2	Time 3
Group 1	0	1	1	1
Group 2	0	0	1	1
Group 3	0	0	0	1

Figure 1.
The traditional stepped wedge study
0=Control period; 1=Intervention period

A: Design 1 – Concurrent

	Time 0	Time 1	Time 2	Time 3	Time 4
Group 1	0	1	1	1	1
Group 2	0	0	1	1	1
Group 3	0	0	0	1	1
Group 4	0	0	0	0	1
Group 5	0	2	2	2	2
Group 6	0	0	2	2	2
Group 7	0	0	0	2	2
Group 8	0	0	0	0	2

B: Design 2 - Replacement

	Time 0	Time 1	Time 2	Time 3	Time 4
Group 1	0	1	2	2	2
Group 2	0	0	1	2	2
Group 3	0	0	0	1	2

C: Design 3 – Supplementation

	Time 0	Time 1	Time 2	Time 3	Time 4
Group 1	0	1	1+2	1+2	1+2
Group 2	0	0	1	1+2	1+2
Group 3	0	0	0	1	1+2

D: Design 4 – Factorial

	Time 0	Time 1	Time 2	Time 3	Time 4
Group 1	0	1	1	1+2	1+2
Group 2	0	0	1	1	1+2
Group 3	0	0	2	2	1+2
Group 4	0	2	2	1+2	1+2

Figure 2.

Stepped Wedge Design Variations*

0=Control period; 1=Intervention 1 period; 2=Intervention 2 period; 1+2=Period with both Intervention 1 and Intervention 2.

*To aid in comparisons, each design is constructed to have 5 time intervals and 24 clusters (i.e. each group contains 3, 8, 8, and 6 clusters for the Concurrent, Replacement, Supplementation and Factorial design, respectively). Variants of each design are possible.

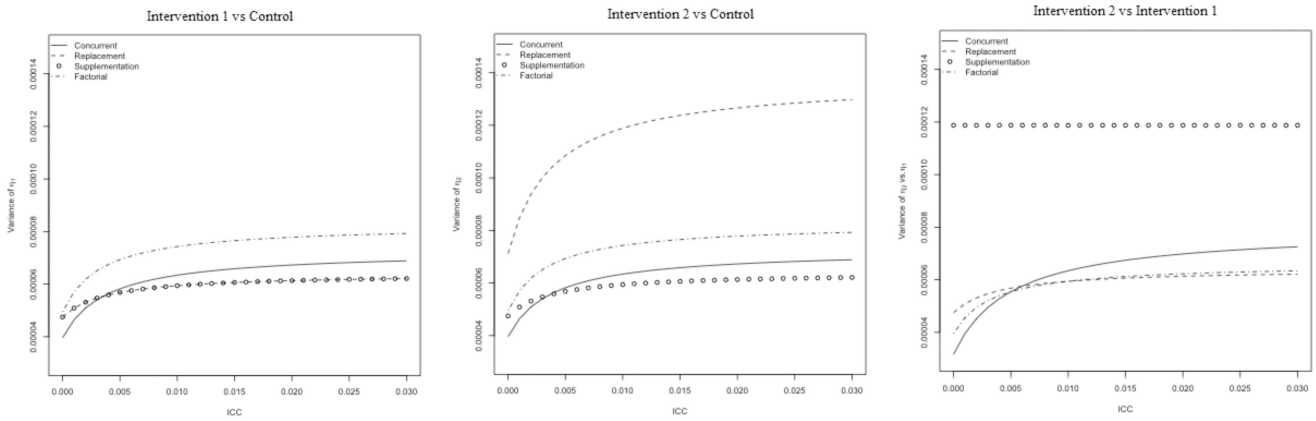

Figure 3.

Variances of the estimates of intervention effects by design*

* Each design was analyzed with 24 clusters, 5 time steps and 100 individuals sampled cross-sectionally per time step per cluster. Note that the variance of the Intervention 1 estimate is identical for the Supplementation and Replacement designs.