Published in final edited form as:
Contemp Clin Trials. 2015 November ; 45(Pt B): 338–339. doi:10.1016/j.cct.2015.09.011.

A note on "Design and analysis of stepped wedge cluster randomized trials"

Xiaomei Liao, **Xin Zhou**, and **Donna Spiegelman**
Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, United States

Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, United States

Keywords

Step wedge design; Time effect; Variance

Editor

Hussey and Hughes [1] seminal paper on the design and analysis of stepped wedge cluster randomized trials (SWD) is an important contribution to the literature. In this paper, they derived the variance formula for the intervention effect in a SWD assuming a time effect, and also compared the relative efficiency for some alternative estimates assuming no time effect. In the course of our own research on SWDs, we found an error in Eq. (9) of Section 3.5 in the paper, in which a factor of T was missing in the denominator. Appendix A gives the correct derivation of the variance of the generalized least squares estimator for a SWD in a model that assumes no time effect, using the notation of Hussey and Hughes' paper. We believe that SWDs are most useful for the study of relatively short-term outcomes in which the intervention is also short-term, such as in a study we are currently working on that investigates the effect of post-partum IUD insertion on the prevention of unintended pregnancy 1.5 years later. In such instances, it may be reasonable to assume that there is no time effect in the primary analysis; hence, the correction to the variance that we provide is potentially important for planners of SWDs.

Eq. (9) was given by Hussey and Hughes to illustrate the efficiency advantage of an analysis that includes both between- and within-cluster comparisons ($\hat{\theta}$), relative to that that only considers within-cluster comparisons ($\tilde{\theta}$). In Appendix B, we prove that the efficiency advantage they discuss still holds despite the error we discovered, although the relative advantage is less than previously reported and will lessen as T, the number of randomization steps in the SWD, increases, something not readily apparent in the original paper.

Appendix A

$Var(\hat{\theta})$ assuming no time effect

If there are no time effects, the $IT \times 2$ design matrix $\boldsymbol{Z}$ corresponding to the parameter vector $\eta = (\mu, \theta)$ for a stepped wedge design is given by $\boldsymbol{Z} = [\boldsymbol{Z}_1, \boldsymbol{Z}_2, \cdots, \boldsymbol{Z}_I]'$, where

$\boldsymbol{Z}_i=\begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_{i1} & X_{i2} & \cdots & X_{iT} \end{bmatrix}$, $i = 1,2, \cdots, I$. Here X_{ij} is an indicator of the treatment mode in cluster i at time j (1 = intervention; 0 = control).

Then the covariance matrix of the weighted least square (WLS) estimates $\hat{\eta}$ is $(\boldsymbol{Z}'\boldsymbol{V}^{-1}\boldsymbol{Z})^{-1}$, where $\boldsymbol{V}$ is an $IT\times IT$ block diagonal matrix, i.e. $\boldsymbol{V} = Diag(\boldsymbol{V}_1, \boldsymbol{V}_2, \cdots, \boldsymbol{V}_I)$, and each $T\times T$ block $\boldsymbol{V}_i$ within $\boldsymbol{V}$ has the structure

$$\boldsymbol{V}_i=\begin{bmatrix} \sigma^2+\tau^2 & \tau^2 & \cdots & \tau^2 \\ \tau^2 & \ddots & \tau^2 & \vdots \\ \vdots & \tau^2 & \ddots & \tau^2 \\ \tau^2 & \cdots & \tau^2 & \sigma^2+\tau^2 \end{bmatrix}.$$

Then we can obtain

$$\boldsymbol{V}_i^{-1}=\frac{1}{\sigma^4+T\sigma^2\tau^2}\cdot\begin{bmatrix} \sigma^2+(T-1)\tau^2 & -\tau^2 & \cdots & -\tau^2 \\ -\tau^2 & \ddots & -\tau^2 & \vdots \\ \vdots & -\tau^2 & \ddots & -\tau^2 \\ -\tau^2 & \cdots & -\tau^2 & \sigma^2+(T-1)\tau^2 \end{bmatrix}. \quad (1)$$

Hence, $(\boldsymbol{Z}'\boldsymbol{V}^{-1}\boldsymbol{Z})^{-1}=(\sum_{i=1}^{I}\boldsymbol{Z}_i\boldsymbol{V}_i^{-1}\boldsymbol{Z}_i')^{-1}$, where each $\boldsymbol{Z}_i\boldsymbol{V}_i^{-1}\boldsymbol{Z}_i'$, $i = 1,2, \cdots, I$, has the form

$$\begin{bmatrix} \sum_{k=1}^{T}\sum_{j=1}^{T}V_{jk} & \sum_{k=1}^{T}\left(\sum_{j=1}^{T}V_{jk}\right)X_{ik} \\ \sum_{j=1}^{T}\left(\sum_{k=1}^{T}V_{jk}\right)X_{ij} & \sum_{k=1}^{T}\left(\sum_{j=1}^{T}X_{ij}V_{jk}\right)X_{ik} \end{bmatrix} \quad (2)$$

with V_{jk} as the (j,k) entry of the matrix $\boldsymbol{V}_i^{-1}$ in Eq. (1).

From Eq. (1), we have $\sum_{k=1}^{T}\sum_{j=1}^{T}V_{jk}=\frac{T}{\sigma^2+T\tau^2}$, $\sum_{j=1}^{T}V_{\text{jk}}=\sum_{k=1}^{T}V_{jk}=\frac{1}{\sigma^2+T\tau^2}$, and

$$\begin{aligned}\sum_{k=1}^{T}\left(\sum_{j=1}^{T}X_{\text{ij}}V_{\text{jk}}\right)X_{\text{ik}}&=\sum_{k=1}^{T}\left(\sum_{j=1}^{T}X_{\text{ij}}X_{\text{ik}}\right)V_{\text{jk}}\\ &=\sum_{j=1}^{T}X_{\text{ij}}^2V_{\text{jj}}+\sum_{k\neq j}X_{\text{ij}}X_{\text{ik}}V_{\text{jk}}\\ &=\frac{\sigma^2+(T-1)\tau^2}{\sigma^4+T\sigma\tau^2}\sum_{j=1}^{T}X_{\text{ij}}^2-\frac{\tau^2}{\sigma^4+T\sigma^2\tau^2}\sum_{k\neq j}X_{\text{ij}}X_{\text{ik}}=\frac{1}{\sigma^2}\sum_{j=1}^{T}X_{\text{ij}}^2-\frac{\tau^2}{\sigma^4+T\sigma^2\tau^2}\left(\sum_{j=1}^{T}X_{\text{ij}}\right)^2.\end{aligned}$$

Hence,

$$\boldsymbol{Z}'\boldsymbol{V}^{-1}\boldsymbol{Z}=\sum_{i=1}^{I}\boldsymbol{Z}_i\boldsymbol{V}_i^{-1}\boldsymbol{Z}_i'=\begin{bmatrix}\frac{\mathrm{IT}}{\sigma^2+T\tau^2} & \frac{\sum_{i=1}^{I}\sum_{j=1}^{T}X_{\mathrm{ij}}}{\sigma^2+T\tau^2}\\ \frac{\sum_{i=1}^{I}\sum_{j=1}^{T}X_{\mathrm{ij}}}{\sigma^2+T\tau^2} & \frac{1}{\sigma^2}\sum_{i=1}^{I}\sum_{j=1}^{T}X_{\mathrm{ij}}^2-\frac{\tau^2}{\sigma^4+T\sigma^2\tau^2}\sum_{i=1}^{I}\left(\sum_{j=1}^{T}X_{\mathrm{ij}}\right)^2\end{bmatrix}.$$

Since X_{ij} is an indicator variable, $\sum_{i=1}^{I}\sum_{j=1}^{T}X_{ij}=\sum_{i=1}^{I}\sum_{j=1}^{T}X_{ij}^2$. Denote $U=\sum_{i=1}^{I}\sum_{j=1}^{T}X_{ij}, V=\sum_{i=1}^{I}(\sum_{j=1}^{T}X_{ij})^2$, then

$$\boldsymbol{Z}'\boldsymbol{V}^{-1}\boldsymbol{Z}=\begin{bmatrix}\frac{IT}{\sigma^2+T\tau^2} & \frac{U}{\sigma^2+T\tau^2}\\ \frac{U}{\sigma^2+T\tau^2} & \frac{1}{\sigma^2}U-\frac{\tau^2}{\sigma^4+T\sigma^2\tau^2}V\end{bmatrix},$$

and it follows that, if there are no time effects, $Var(\hat{\theta})$, i.e. the (2,2) component of $(\boldsymbol{Z}'\boldsymbol{V}^{-1}\boldsymbol{Z})^{-1}$, is given by

$$\mathrm{Var}\left(\hat{\theta}\right)=\frac{IT(\sigma^2+T\tau^2)\sigma^2}{(ITU-U^2)\sigma^2+IT(TU-V)\tau^2}.$$

Thus, Eq. (9) of Section 3.5 in Hussey and Hughes [1] should be

$$\mathrm{effic}\left(\hat{\theta},\tilde{\theta}\right)=\frac{\mathrm{Var}(\tilde{\theta})}{\mathrm{Var}(\hat{\theta})}=\frac{\sum_i\left(\frac{1}{w_i}+\frac{1}{T-w_i}\right)\left[(ITU-U^2)\sigma^2+IT(TU-V)\tau^2\right]}{I^3T(\sigma^2+T\tau^2)},$$

which has an extra factor T on the denominator comparing to the original Eq. (9) in Hussey and Hughes [1].

Appendix B

The variance of the WLS estimator, *Var*(θ̂), is less than or equal to the variance of the within-cluster estimator, *Var*(θ̃), even with the extra factor, *T*, included in *Var*(θ̂)

To prove that $Var(\hat{\theta})$ is smaller than or equal to $Var(\tilde{\theta})$, even with the extra factor, T, included in $Var(\hat{\theta})$, we first write

$U=\sum_{i=1}^{I}(T-w_i)=IT-\sum_{i=1}^{I}w_i, V=\sum_{i=1}^{I}(T-w_i)^2=IT^2-2T\sum_{i=1}^{I}w_i+\sum_{i=1}^{I}w_i^2$. Hence,

$$\text{effic}\left(\hat{\theta},\tilde{\theta}\right)=T\left(\sum_{i=1}^{I}\frac{1}{w_i(T-w_i)}\right)\frac{\left[\left(U-\frac{U^2}{\text{TI}}\right)\sigma^2+\left(U-\frac{V}{T}\right)T\tau^2\right]}{I^2\sigma^2+T\tau^2}$$
$$=T\left(\sum_{i=1}^{I}\frac{1}{w_i(T-w_i)}\right)\frac{\left[\left(\sum_{i=1}^{I}w_i-\frac{\left(\sum_{i=1}^{I}w_i\right)^2}{\text{TI}}\right)\sigma^2+\left(\sum_{i=1}^{I}w_i-\frac{\sum_{i=1}^{I}w_i^2}{T}\right)T\tau^2\right]}{I^2(\sigma^2+T\tau^2)},$$

and using the fact $\sum_{i=1}^{I}w_i^2 \geq \frac{\left(\sum_{i=1}^{I}w_i\right)^2}{I}$ from the Cauchy-Schwarz inequality, we obtain

$$\text{effic}\left(\hat{\theta},\tilde{\theta}\right)\geq T\left(\sum_{i=1}^{I}\frac{1}{w_i(T-w_i)}\right)\frac{\left(\sum_{i=1}^{I}w_i-\frac{\sum_{i=1}^{I}w_i^2}{T}\right)}{I^2}=\frac{1}{I^2}\left(\sum_{i=1}^{I}\frac{1}{w_i(T-w_i)}\right)\left(\sum_{i=1}^{I}w_i(T-w_i)\right)\geq 1,$$

where the last inequality holds from the Cauchy-Schwarz inequality as well, completing the proof.

Reference

1. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. Contemp. Clin. Trials. 2007; 28:182–191. [PubMed: 16829207]