

# A SAS macro for general power of GEE analysis of multi-period CRTs

Ying Zhang, John S. Preisser



UNC

GILLINGS SCHOOL OF  
GLOBAL PUBLIC HEALTH

# Table of contents

- 1 Development of the General GEE Power Methods
- 2 Introduction of the SAS macro
- 3 General GEE Power Method in the SAS macro
- 4 Inputs and Outputs in the SAS macro
- 5 Examples of the SAS macro
- 6 Download the CRTFASTGEEPWR SAS macro
- 7 References



# Development of General GEE Power Methods

- 1 The general power procedure for the GEE analysis was first proposed by Rochon, [1] which has been extended to sample size and power formulae for GEE analysis of parallel-groups CRTs, including cross-sectional and cohort CRTs, [2, 3] and multi-level CRTs. [4, 5]
- 2 Prior works [6, 7] have shown that the general GEE power for marginal mean (e.g, intervention) parameters in complete SW-CRTs agrees well with simulated power based on GEE with finite-sample sandwich variance estimators for as few as eight clusters.
- 3 However, there is a **gap** of statistical software tools for the general GEE power for multi-period CRTs that can accommodate both complete and incomplete designs



# Overview of the SAS macro

- 1 We develop a SAS macro, **CRTFASTGEEPWR**, which calculates the general GEE power for both categorical and continuous responses of multi-period CRTs with complete and incomplete designs.
- 2 The important features in the SAS macro are summarized below:
  - Binary, count and continuous outcomes with choice of link function
  - Two types of intervention effects models
  - Specification of categorical or linear period effects
  - Four correlation structures
  - Specification of the dispersion parameter for count or continuous data
  - Design pattern matrix for incomplete multi-period CRTs
  - Specification of varying cluster period sizes for cross-sectional multi-period CRTs
  - Consistency checks for the SAS macro inputs



# Marginal Mean Model Specification

In a multi-period CRT, there are  $J$  periods,  $S$  sequences,  $I$  clusters and  $m_s$  clusters in sequence  $s$ , such that  $I = \sum_{s=1}^S m_s$ . Let  $y_{ijk}$  with marginal mean  $\mu_{ijk}$  denote the response of the individual  $k$  from cluster  $i$  during time point  $j$ ,  $i = 1, \dots, I; j = 1, \dots, J; k = 1, \dots, N_{ij}$ .  $\mu_{ijk}$  is related to the intervention effect  $\delta$  and  $j$ th **categorical period effects**  $\beta_j$  with link function  $g(\cdot)$  via the marginal mean model:

$$g(\mu_{ijk}) = \beta_j + u_{ij}\delta \quad (1)$$

For multi-period CRTs with fewer number of clusters than the number of periods ( $I < J$ ), the marginal mean model with **linear period effects** is:

$$g(\mu_{ijk}) = \beta_0 + \beta_1 (t_{ij} - 1) + u_{ij}\delta \quad (2)$$

where  $\beta_0$  is the intercept, and  $\beta_1$  is the increment in the mean response on the scale of the link function for a unit increase in calendar period  $t_{ij}$



# Choice of Treatment Effects

In the marginal mean model,  $u_{ij}$  is the treatment status in cluster  $i$  at the time point  $j$  (equation (1) or (2)). We include two types of intervention effects model in the SAS macro.

- 1 Average intervention effect model:**  $u_{ij}$  is the period-specific treatment indicator (1=intervention; 0=control) and  $\delta$  is the intervention effect on the link function scale.
- 2 Incremental intervention effect model,**  $u_{ij} = (j - b_i)/m$ ,  $j = b_i + 1, \dots, b_i + m_i$  where the  $i_{th}$  cluster has  $b_i$  baseline and  $m_i$  intervention condition periods of measurement, respectively,  $u_{ij} = 0$  for control periods  $j = 1, \dots, b_i$ . And  $\delta$  is the maximal intervention effect by the end of the intervention period for the cluster(s) with the longest intervention period,  $m = \max\{m_i\}$ .



# Three types of outcomes

Table 1: The outcome types in the SAS macro

| Outcome types | Distribution | Range                     | Link*    | Variance                   |
|---------------|--------------|---------------------------|----------|----------------------------|
| Binary        | Binomial     | $0 \leq \mu_{ijk} \leq 1$ | Logit    | $\mu_{ijk}(1 - \mu_{ijk})$ |
| Count         | Poisson      | $0 < \mu_{ijk}$           | Log      | $\varphi \mu_{ijk}$        |
| Continuous    | Normal       | -                         | Identity | $\varphi$                  |


- The canonical link function is the default choice of link in the SAS macro. Different link functions could be specified as long as the marginal mean is within range for the outcomes type, otherwise, computational issues may arise.
- $\varphi$  is dispersion parameter for count outcome and the constant variance for continuous outcome under normal distribution.



# Four within-cluster correlation structures

Table 2: The Intra-cluster correlation (ICC):  $\text{corr}(y_{ijk}, y_{ij'k'})$ , ( $i$ =cluster,  $j$ =period,  $k$ =individual) under the four correlation structures

| Design          | Correlation Structure          | $j = j'$   | $j \neq j', k \neq k'$              | $j \neq j', k =$           |
|-----------------|--------------------------------|------------|-------------------------------------|----------------------------|
| Cross-Sectional | Nested Exchangeable            | $\alpha_1$ | $\alpha_2$                          | —                          |
| Cohort          | Exponential Decay              | $\alpha_0$ | $\alpha_0 r_0^{ t_{ij} - t_{ij}' }$ | —                          |
|                 | Block Exchangeable             | $\alpha_1$ | $\alpha_2$                          | $\alpha_3$                 |
|                 | General proportional Decay [8] | $\alpha_0$ | $\alpha_0 r_0^{ t_{ij} - t_{ij}' }$ | $r_1^{ t_{ij} - t_{ij}' }$ |

- $\alpha_0$  or  $\alpha_1$  is the usual ICCs, which measures the correlation between responses from different individuals during the same period.
- $\alpha_2$  is the ICC between responses from different individuals during the different periods. For block exchangeable correlation under cohort design.
- $\alpha_3$  is the ICC between responses from the same individuals during the different periods.
- $r_0$  is the auto-regressive decay rate between-individual correlation and  $r_1$  is the auto-regressive decay rate for intra-individual correlation. When  $r_0 = r_1$ , it reduces to the proportional decay correlation structure [7]. 



# General model-based covariance matrix

By the general GEE power methodology in Rochon [1], the model-based covariance matrix of  $\theta$  in GEE is:

$$\text{cov}_{\text{MB}}(\hat{\theta}) = \left[ \sum_i^I \mathbf{D}_i' \mathbf{V}_i^{-1} \mathbf{D}_i \right]^{-1} \quad (3)$$

Under the GEE model,  $\mathbf{D}_i = (\partial \mu_i / \partial \theta')$ , with  $\theta = (\beta, \delta)$ . The covariance matrix within the  $i_{th}$  cluster across periods is  $\mathbf{V}_i = \varphi \mathbf{A}_i^{1/2} \mathbf{R}_i \mathbf{A}_i^{1/2}$ , in which  $\mathbf{A}_i = \text{diag}\{\text{Var}(y_{i11}), \dots, \text{Var}(y_{iJN_{iJ}})\}$ , and  $\mathbf{R}_i$  is the working correlation matrix in the  $i_{th}$  cluster.

The model-based variance of the intervention effect  $\text{Var}(\hat{\delta})$  could be obtained by the  $(p, p)_{th}$  element in the  $\text{cov}_{\text{MB}}(\hat{\theta})$ , in which  $p = J + 1$  for categorical period effects and  $p = 3$  for linear period effects in the marginal mean model (equation (1), (2)).



# General GEE power method for CRTs

We use Wald statistics to test the intervention effect of size  $\delta$  under the hypothesis,  $H_0 : \delta = 0$  vs.  $\delta \neq 0$  under a nominal type I error rate  $\alpha$ . Following Li et al., [6] the general GEE power for CRTs is:

- 1  $\Phi \left( z_{\alpha/2} + |\delta| / \sqrt{\text{var}(\hat{\delta})} \right)$  using the z-test with standard normal cumulative distribution function  $\Phi(\cdot)$  and normal quantile  $z_{\alpha/2}$  such that  $\Phi(z_{\alpha/2}) = 1 - \alpha/2$ .
- 2  $\Phi_{t,df} \left( t_{\alpha/2,df} + |\delta| / \sqrt{\text{var}(\hat{\delta})} \right)$  using the t-test, which is recommended for small number of clusters in CRTs.
- 3 The degrees of freedom of the  $t$ -test in CRTs is typically set to  $df = I - p$ , and some authors have used the alternative formula  $df = I - 2$ . [7, 9]



# Design pattern for incomplete CRTs

For incomplete CRTs, we specify a Design Pattern (DP) matrix inspired by Hemming et al. [10] to represent the experimental design, where 0 means the control periods, 1 means intervention periods and 2 means missing periods without data collection.

To illustrate the specification of DP matrix, an example of incomplete SW-CRT [10] is used here with  $S = 2$  treatment sequences,  $T = 4$  periods, and an implementation period that occurs in period 2 for the first sequence and in period 3 for the second sequence. The design pattern matrix is specified for the incomplete SW-CRT.

$$\text{DesignPattern} = \begin{pmatrix} 0 & 2 & 1 & 1 \\ 0 & 0 & 2 & 1 \end{pmatrix}$$

The characteristics of multi-period CRTs with complete or incomplete designs can be specified by the design pattern matrix for parallel, crossover, and stepped-wedge designs.



# Option arguments in the SAS macro

| Macro Variable*          | Input                             | Description  |
|--------------------------|-----------------------------------|--|
| DESIGNPATTERN            | Variable                          | A matrix with dimension $S \times T$ of 0, 1 and 2, with 0 representing control periods, 1 standing for intervention periods, 2 presenting periods without data collection.  |
| CP_SIZE_MATRIX           | Variable                          | Number of cluster-period sizes, a matrix vector with dimension $S \times T$ . For cross-sectional design, the SAS macro allows varying cluster period size by and within sequences. While for cohort design, the cohort size could vary by sequences but keeps the same or decays within sequences |
| M                        | Variable                          | Number of clusters in each sequence, a vector with dimension $S \times 1$  |
| DIST                     | = BINARY<br>= POISSON<br>= NORMAL | The distribution for the outcomes  |
| LINK                     | = LOGIT<br>= LOG<br>= IDENTITY    | Link function for the outcomes, For binary, count, and continuous responses, the default link is logit, log and identity, respectively   |
| PHI                      | Variable                          | The dispersion parameter   |
| INTERVENTION_EFFECT_TYPE | = AVE<br>= INC<br>= INC_EX        | Average intervention effects model<br>Incremental intervention effects model<br>Extended incremental intervention effects model  |
| PERIOD_EFFECT_TYPE       | = CAT<br>= LIN                    | Categorical time period model<br>Linear period effects model   |
| DELTA                    | Variable                          | The intervention effects on scale of link functions  |



# SAS macro argument Cont'd

|                                |                          |   |
|--------------------------------|--------------------------|---|
| <b>BETA_PERIOD_EFFECTS</b>     | Variable                 | The period effects on scale of link functions:<br>A T*1 vector for categorical period effects<br>A 2*1 vector for continuous period effects       |
| <b>CORR_TYPE</b>               | =NE<br>=ED<br>=BE<br>=PD | Nested exchangeable<br>Exponential decay<br>Block exchangeable<br>General proportional decay  |
| <b>ALPHA0</b>                  | Variable                 | The within-period correlation in exponential decay and proportional decay correlation structure   |
| <b>R0</b>                      | Variable                 | Between-subject correlation decay rate over time in ED and PD correlation structures  |
| <b>R1</b>                      | variable                 | Within-subject correlation decay rate over time in the PD correlation structure   |
| <b>ALPHA1</b>                  | Variable                 | and proportional decay correlation structure<br>The within-period correlation in nested exchangeable and block exchangeable correlation structure |
| <b>ALPHA2</b>                  | Variable                 | The inter-period correlation in nested exchangeable and block exchangeable correlation structure  |
| <b>ALPHA3</b>                  | Variable                 | The within-subject correlation in block exchangeable correlation structure  |
| <b>MAX_INTERVENTION_PERIOD</b> | Variable                 | The number of maximum intervention periods in incremental intervention effect model   |
| <b>ALPHA</b>                   | Variable                 | Significance level, default at 0.05   |
| <b>DF_CHOICE</b>               | =1<br><br>=2             | Default degrees of freedom (df), df = 1-p, p is the number of parameters in marginal model<br>df = 1-2  |

\*All required arguments are in boldface, where the required correlation parameters depend upon correlation type.



# SAS macro Output

| Output Variable | Descriptions  |
|-----------------|---|
| T               | Number of periods, a scalar   |
| S               | Number of sequences (steps), a scalar   |
| clusters        | Number of clusters in the CRT   |
| df              | Degrees of freedom in the test of the intervention effects                              |
| theta           | Value of parameters in marginal model   |
| totaln          | Total number of the participants in the CRT   |
| Dist            | The distribution for the outcomes   |
| link            | The link function used in the GEE model   |
| stddel          | Estimated standardized intervention effects, $ \delta /\sqrt{\text{var}(\hat{\delta})}$ |
| zpower          | Power using the z-test  |
| tpower          | Power using the t-test  |



# Consistency Check

- 1 Print out the number of periods T and sequences S in the output.
- 2 The number of period effects parameters and selected period effects model should be consistent. For the continuous period effects model, there are two parameters of time effects. While for categorical period effects model, there are T time effects parameters, which is derived from the Design pattern matrix.
- 3 Check if the value of  $\mu_{ijk}$  is within the boundaries. For binary outcomes,  $\mu_{ijk}$  are within [0,1]. For count outcomes,  $\mu_{ijk}$  are non-negative.
- 4 Whether the assumed working correlation is out of Frechet bounds for binary outcomes and whether the correlation structure is invertible.
- 5 For the missing cluster periods in the DESIGNPATTERN matrix, the corresponding locations in the CP\_SIZE\_Matrix need to be 0.

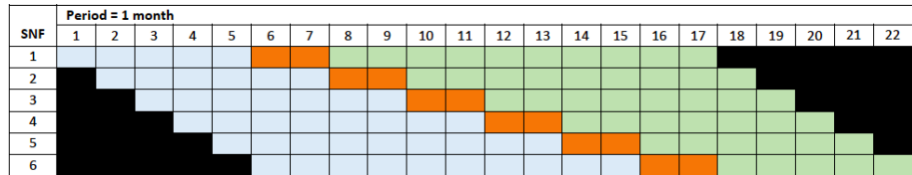


# Example 1: Connect-Home study

Connect-Home study is a cross-sectional SW-CRT and tests an intervention to improve outcomes for rehabilitation patients transitioning from skilled nursing facilities (SNFs) to home-based care. [11] The incomplete design was chosen based on considerations of internal validity and power under restrictions placed by available resources and logistical considerations.

There are 6 sequences with 22 periods in this design, having 7 periods without patient enrollment and 15 periods with patient enrollment in each cluster in Figure 1. The original design had six clusters or SNFs. [11]

Figure 1: The design pattern of the Connect-Home trial, black or orange boxes for cluster period without patient enrollment, blue boxes for control periods and green boxes for intervention periods





## Example 1.a: Connect-Home study cont'd

- In our illustration, there are 3 clusters in each sequence and a total of 18 clusters with 540 subjects in the trial, in which 2 patients enrolled in each cluster period.
- For the binary outcome, any acute care use for patients, we assume gently decreasing linear period effects such that  $\beta_0 = 0.85$  and  $\beta_1 = -0.01$  for  $t_{ij} = 1, \dots, 22$ , calculated from assumptions that baseline prevalence is  $\exp(\beta_0) / [1 + \exp(\beta_0)] = 0.7$  and negative time trend such that prevalence at the last period (22 months) is 0.65.
- The effect size at 10 months on treatment condition ( $m = 10$ ) for the incremental intervention effects model is  $\delta = \log(0.25) = -1.386$ , in which the odds ratio is assumed to be 0.25.
- ICCs are specified with  $(\alpha_0, r_0) = (0.03, 0.8)$  under exponential decay correlation to indicate a moderate decaying of within-cluster correlation for the cross-sectional design.



# Example 1.a: SAS macro specification

```
%CRTFASTGEEPWR(alpha=0.05, m=%str(J(6,1,3)), corr_type = ED,alpha0 = 0.03,r0 = 0.8 ,
intervention_effect_type=INC, delta = -1.386, period_effect_type=LIN,
beta_period_effects=%str({0.847,-0.01}) ,
dist = BINARY,phi=1,max_intervention_period=10,
  CP_size_matrix =
      %str({2 2 2 2 2 0 0 2 2 2 2 2 2 2 2 2 2 0 0 0 0 0,
            0 2 2 2 2 2 2 0 0 2 2 2 2 2 2 2 2 2 0 0 0 0,
            0 0 2 2 2 2 2 2 2 0 0 2 2 2 2 2 2 2 2 0 0 0,
            0 0 0 2 2 2 2 2 2 2 2 0 0 2 2 2 2 2 2 2 0 0,
            0 0 0 0 2 2 2 2 2 2 2 2 2 2 0 0 2 2 2 2 2 2,
            0 0 0 0 0 2 2 2 2 2 2 2 2 2 2 2 0 0 2 2 2 2}),
  DesignPattern =
      %str({0 0 0 0 0 2 2 1 1 1 1 1 1 1 1 1 2 2 2 2 2,
            2 0 0 0 0 0 0 2 2 1 1 1 1 1 1 1 1 2 2 2 2,
            2 2 0 0 0 0 0 0 0 2 2 1 1 1 1 1 1 1 2 2 2,
            2 2 2 0 0 0 0 0 0 0 0 2 2 1 1 1 1 1 1 2 2,
            2 2 2 2 0 0 0 0 0 0 0 0 0 2 2 1 1 1 1 1 2,
            2 2 2 2 2 0 0 0 0 0 0 0 0 0 0 2 2 1 1 1 1});
```



## Example 1.a: SAS macro outputs

From the results, the power using the z-test is greater than the *t*-test. Simulation studies have shown that the z-test is too optimistic and tends to have an inflated test size in SW-CRTs with a small number of clusters.

(4) Thus, we recommend calculating power with the *t*-test for the SW-CRTs.

```
The fast GEE power of BINARY outcomes with exponential decay correlation  
structure and (alpha0,r0):(0.03, 0.8)
```

```
Under incremental intervention effects model and delta = -1.386
```

```
T S clusters df theta totaln Dist Link stddel zpower tpower  
22 6 18 15 0.847 540 BINARY LOGIT 2.7585 0.7877 0.73  
-0.01  
-1.386
```



## Example 1.b: Closed-cohort design with decaying cluster sizes

- In this example, we assume the Connect-Home study used a closed-cohort study with decaying cluster size. There are 4 clusters in each sequence and a total of 24 clusters.
- For the cluster size, we assume there are 2 subjects followed up in each cluster and 1 patient tends to drop out after 15 periods of observation. Thus, there are with 28 observations in each cluster from 2 subjects by 15 periods.
- Parameters in the marginal mean model keep the same as in the former example.
- ICCs are specified with  $(\alpha_0, r_0, r_1) = (0.03, 0.8, 0.1)$  under general proportional decay correlation to indicate a moderate decaying of between-patient correlation and a small decaying of intra-patient correlation for the closed-cohort design.



# Example 1.b: SAS macro specification

```
%CRFASTGEEPWR(alpha=0.05, m =%str(J(6,1,4)), corr_type = PD,alpha0 = 0.03,r0 = 0.8 , r1 = 0.1 ,
  intervention_effect_type=INC, delta = -1.386, period_effect_type=LIN,
beta_period_effects =%str({0.847,-0.01}) ,dist = BINARY,phi=1,max_intervention_period=10,
  CP_size_matrix =
      %str({2 2 2 2 2 0 0 2 2 2 2 2 2 2 2 2 1 1 0 0 0 0 0,
            0 2 2 2 2 2 2 0 0 2 2 2 2 2 2 2 2 1 1 0 0 0 0,
            0 0 2 2 2 2 2 2 2 0 0 2 2 2 2 2 2 2 1 1 0 0 0,
            0 0 0 2 2 2 2 2 2 2 0 0 2 2 2 2 2 2 1 1 0 0 ,
            0 0 0 0 2 2 2 2 2 2 2 2 0 0 2 2 2 2 2 2 1 1 0,
            0 0 0 0 0 2 2 2 2 2 2 2 2 2 2 2 0 0 2 2 2 1 1}),
DesignPattern = %str({0 0 0 0 0 2 2 1 1 1 1 1 1 1 1 1 2 2 2 2 2,
                      2 0 0 0 0 0 0 2 2 1 1 1 1 1 1 1 1 2 2 2 2,
                      2 2 0 0 0 0 0 0 2 2 1 1 1 1 1 1 1 2 2 2,
                      2 2 2 0 0 0 0 0 0 0 2 2 1 1 1 1 1 1 2 2,
                      2 2 2 2 0 0 0 0 0 0 0 0 2 2 1 1 1 1 1 2,
                      2 2 2 2 2 0 0 0 0 0 0 0 0 0 2 2 1 1 1 1 1});
```



## Example 1.b: SAS macro outputs

Fixing all the marginal mean parameters, powers are higher than those in the former example due to a greater number of clusters. The z-test still gives higher power than t-test, but their values are more similar and all above 0.77.

```
The fast GEE power of BINARY outcomes with proportional decay correlation structure and (alpha0,r0,r1):(0.03, 0.8,0.1)
```

```
Under incremental intervention effects model and delta = -1.386
```

```
T S clusters df theta totaln Dist Link stddel zpower tpower
22 6 24      21 0.847 672  BINARY LOGIT 2.8408 0.8108 0.7725
      -0.01
      -1.386
```



## Example 2: A complete SW-CRT

The second example calculates power for a hypothetical complete SW-CRT with categorical period effects.

- 4 sequences and 5 periods, with the conventional specification of one less sequences than the number of periods.
- There are 24 clusters in total and 6 clusters in each sequence, which includes 100 participants in each cluster period and 12000 participants in total.
- We are interested in calculating power for binary responses with a logit link. The categorical period effects are assumed to be constant over time with  $-2.944$  at the logit link scale and the effect size is  $\log(0.55) = -0.598$  assumed with an odds ratio 0.55 under the average intervention effects model.
- ICCs under the nested exchangeable correlation is specified with  $(\alpha_1, \alpha_2) = (0.01, 0.005)$ .



## Example 2: SAS macro specification

```
%CRTFASTGEEPWR(alpha=0.05, m=%str(J(4,1,6)), corr_type = NE, alpha1 = 0.01,alpha2 = 0.005 ,  
  intervention_effect_type=AVE, delta = -0.598, period_effect_type=CAT,  
beta_period_effects=%str(J(5,1,-2.944)), dist = BINARY,phi=1,  
  CP_size_matrix = %str(J(4,5,100)),  
  
  DesignPattern = %str({0 1 1 1 1,  
                        0 0 1 1 1,  
                        0 0 0 1 1,  
                        0 0 0 0 1});
```





## Example 2: SAS macro outputs

From the following results, the power using z-test and *t*-test are all above 0.8 but not extremely high considering the large cluster size in the trial. Hemming et al. [12] and others have shown that the cluster size often makes no material contribution to the power of CRTs beyond a specific value given a fixed number of clusters so the number of clusters and cluster size should be determined concurrently.

```
The fast GEE power of BINARY outcomes with nested exchangeable  
correlation structure and (alpha1,alpha2):(0.01, 0.005)  
Under average intervention effects model and delta = -0.598
```

```
T S clusters df theta totaln Dist Link stddel zpower tpower  
5 4 24      18 -2.944 12000 BINARY LOGIT 3.0663 0.8657 0.8264  
      -2.944  
      -2.944  
      -2.944  
      -2.944  
      -0.598
```



# Download the CRTFASTGEEPWR SAS macro

- The CRTFASTGEEPWR SAS macro is available at John Preisser's personal software website:  
<http://www.bios.unc.edu/~preisser/personal/software.html> ▶ Link
- Any questions related to the SAS macro should be emailed to [zyingunc@gmail.com](mailto:zyingunc@gmail.com)



# References I

- [1] James Rochon. Application of gee procedures for sample size calculations in repeated measures experiments. *Statistics in Medicine*, 17(14):1643–1658, 1998.
- [2] John S. Preisser, Mary L. Young, Daniel J. Zaccaro, and Mark Wolfson. An integrated population-averaged approach to the design, analysis and sample size determination of cluster-unit trials. *Statistics in Medicine*, 22:1235–54, Apr 2003.
- [3] John S. Preisser, Beth A. Reboussin, Eun-Young Song, and Mark Wolfson. The importance and role of intracluster correlations in planning cluster trials. *Epidemiology (Cambridge, Mass.)*, 18: 552–60, Sep 2007.



## References II

- [4] BA Reboussin, JS Preisser, E-Y Song, and M Wolfson. Sample size estimation for alternating logistic regression analysis of multilevel randomized community trials of underage drinking. *Journal of the Royal Statistical Society, Series A*, 175:691–712, 2012.
- [5] Steven Teerenstra, Bing Lu, John S. Preisser, Theo van Achterberg, and George F. Borm. Sample size considerations for gee analyses of three-level cluster randomized trials. *Biometrics*, 66:1230–7, Dec 2010.
- [6] Fan Li, Elizabeth L. Turner, and John S. Preisser. Sample size determination for gee analyses of stepped wedge cluster randomized trials. *Biometrics*, 74(4):1450–1458, 2018.
- [7] Fan Li. Design and analysis considerations for cohort stepped wedge cluster randomized trials with a decay correlation structure. *Statistics in Medicine*, 39(4):438–455, 2020.



## References III

- [8] John A. Gallis, Xueqi Wang, Paul J. Rathouz, John S. Preisser, Fan Li, and Elizabeth L. Turner. Power swgee: Gee-based power calculations in stepped wedge cluster randomized trials. *Stata J*, 22(4):811–841, Dec 2022. doi: 10.1177/1536867x221140953.
- [9] Whitney P. Ford and Philip M. Westgate. Maintaining the validity of inference in small-sample stepped wedge cluster randomized trials with binary outcomes when using generalized estimating equations. *Statistics in Medicine*, 39(21):2779–2792, 2020. doi: <https://doi.org/10.1002/sim.8575>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.8575>.
- [10] Karla Hemming, Richard Lilford, and Alan J. Girling. Stepped-wedge cluster randomised controlled trials: a generic framework including parallel and multiple-level designs. *Statistics in Medicine*, 34:181–96, 2015.



## References IV

- [11] M. Toles, C. Colón-Emeric, LC. Hanson, M. Naylor, M. Weinberger, J. Covington, and JS. Preisser. Transitional care from skilled nursing facilities to home: study protocol for a stepped wedge cluster randomized trial *Trials*. pages 22–120, 2021.
- [12] K Hemming, S Eldridge, G Forbes, C Weijer, and M Taljaard. How to design efficient cluster randomised trials. *BMJ*, 358, 2017. ISSN 0959-8138. doi: 10.1136/bmj.j3064. URL <https://www.bmj.com/content/358/bmj.j3064>.

