RESEARCH ARTICLE

WILEY Statistics in Medicine

Modeling clustering and treatment effect heterogeneity in parallel and stepped-wedge cluster randomized trials

Karla Hemming[1] | Monica Taljaard[2] | Andrew Forbes[3]

[1]Institute of Applied Health Research, University of Birmingham, Birmingham, B15 2TT, UK

[2]Ottawa Health Research Institute, Ottawa, Ontario K1Y 4E9, Canada

[3]Monash University, Clayton, Victoria 3800, Australia

Correspondence
Karla Hemming, Institute of Applied Health Research, University of Birmingham, Birmingham B15 2TT, UK.
Email: k.hemming@bham.ac.uk

Funding information
NIHR CLAHRC West Midlands+ initiative; National Health and Medical Research Council of Australia, Grant/Award Number: 1108283

Cluster randomized trials are frequently used in health service evaluation. It is common practice to use an analysis model with a random effect to allow for clustering at the analysis stage. In designs where clusters are exposed to both control and treatment conditions, it may be of interest to examine treatment effect heterogeneity across clusters. In designs where clusters are not exposed to both control and treatment conditions, it can also be of interest to allow heterogeneity in the degree of clustering between arms. These two types of heterogeneity are related. It has been proposed in both parallel cluster trials, stepped-wedge, and other cross-over designs that this heterogeneity can be allowed for by incorporating additional random effect(s) into the model. Here, we show that the choice of model parameterization needs careful consideration as some parameterizations for additional heterogeneity induce unnecessary or implausible assumptions. We suggest more appropriate parameterizations, discuss their relative advantages, and demonstrate the implications of these model choices using a real example of a parallel cluster trial and a simulated stepped-wedge trial.

KEYWORDS
cluster randomized trial, ICC, stepped-wedge, treatment effect heterogeneity

1 | INTRODUCTION

Cluster randomized trials (CRTs) randomize entire clusters of individuals to treatment or control conditions.[1] Another related design, which again randomizes entire clusters, is one in which all clusters are randomized to a treatment sequence that is either treatment followed by control, or control followed by treatment.[2] Stepped-wedge cluster randomized trials (SW-CRTs) also randomize entire clusters but randomize clusters to a sequence of time periods spent in the control condition followed by time periods spent in the treatment condition.[3] All cluster trials need to allow for the non-independence of observations within the same cluster. This is typically done by using a linear or generalized linear mixed model with a random effect for cluster.[1]

The effect of the treatment might vary across clusters, and it can be of interest to explore any treatment effect heterogeneity at the analysis stage. In stepped-wedge designs or other cluster randomized designs in which clusters are exposed to both treatment and control, this treatment effect heterogeneity can be identified.[4-7] In the conventional parallel CRT, there may also be treatment effect heterogeneity. However, in trials such as the parallel CRT in which each cluster is exclusively exposed to either the treatment or control condition, this heterogeneity cannot be separated out from the

cluster effect. But a related issue is where the variability between clusters may differ according to treatment arm in parallel cluster trials.[8] While these two sources of heterogeneity are inherently different, they have similar implications for design and analysis when this heterogeneity is modeled using random effects. There is a confusing array of different parameterizations that have been proposed in the literature to model these different but related sources of heterogeneity, which all make different assumptions. This paper aims to review these models and make some recommendations for which models make fewer assumptions.

In this paper, we demonstrate that in cross-sectional CRTs, the choice of model parameterization to incorporate treatment effect or cluster heterogeneity has important underlying assumptions. We demonstrate the implications of these model choices using practical examples of a parallel cluster trial and a simulated stepped-wedge trial.

2 | THEORETICAL MODELS

2.1 | Differential clustering in parallel arm CRTs

In parallel arm CRTs, clusters are either fully exposed or unexposed to the treatment. Therefore, any treatment effect heterogeneity cannot be disentangled from the cluster effect. However, the correlation between observations within clusters might vary across treatment arms. The potential for such differential correlation is intuitive because different treatments might be expected to induce homogeneity—or even heterogeneity. This is particularly relevant in trials with differential clustering between arms—perhaps where a group therapy is compared with an individual therapy. Here, we review the parameterizations that have been proposed for these models in the context of parallel cluster trials and outline their respective assumptions. In Appendix A, we show derivations of the correlations for model 2b (one of the more complex models). In the next section, we show how these models for differential clustering are related to treatment effect heterogeneity models.

2.1.1 | Basic model: single random effect

Let us consider a 2-arm parallel CRT. We assume that there is a continuous outcome y_{ij}, where $i = 1, \ldots, m$ represents the individual and $j = 1, \ldots, k$ represents the clusters. We assume that there are k clusters each of equal size m. One analysis model for this simple setup is[1]

$$y_{ij} = \mu + x_{ij}\theta + \alpha_j + e_{ij} \quad \alpha_j \sim N[0, \tau^2] \quad e_{ij} \sim N[0, \sigma_w^2], \tag{1}$$

where x_{ij} represents the treatment indicator. We code this treatment indicator as 1 for the treatment and 0 for the control (coding of contrasts is important, a point to which we return in the discussion). Then, θ is the treatment effect, and α_j is a random effect for cluster j. Under this model, the correlation between two observations in the same cluster, the intra-cluster correlation (ICC), will be

$$\rho = \frac{\tau^2}{\tau^2 + \sigma_w^2}. \tag{2}$$

The model thus assumes that for any cluster, observations within that cluster share a common correlation ρ and this is the same for both treatment and control clusters.

2.1.2 | Model extension 1: two separate random effects

To allow for different within- and between-cluster variability in treatment and control clusters, two separate random effects, one for treatment and one for control, are incorporated:

$$y_{ij} = \mu + x_{ij}\theta + x_{ij}\alpha(T)_j + (1 - x_{ij})\alpha(C)_j + e_{ij} \quad \alpha(T)_j \sim N[0, \tau_T^2] \quad \alpha(C)_j \sim N[0, \tau_C^2] \quad e_{ij} \sim N[0, \sigma_w^2], \tag{3}$$

where $\alpha(T)$ and $\alpha(C)$ represent random cluster effects for the treatment and control clusters. We use the sub-scripts capital T and capital C to denote the treatment and control clusters. This induces differing ICCs in control and treatment clusters. The ICC in the control clusters will be

$$\rho_C = \frac{\tau_C^2}{\tau_C^2 + \sigma_w^2}, \tag{4}$$

and the ICC in the treatment clusters will be

$$\rho_T = \frac{\tau_T^2}{\tau_T^2 + \sigma_w^2}. \tag{5}$$

Using this parameterization, there is no restriction on whether ρ_C is bigger or smaller than ρ_T. The model allows for random variation between the control clusters (variance τ_C^2), and a different random variation between treatment clusters, τ_T^2.

2.1.3 | Model extension 2: a random interaction

Seen perhaps as a more intuitive way to model heterogeneity in cluster variability across arms, an alternative to the above model is to use a parameterization that includes a random (simple multiplicative) interaction between the treatment covariate and cluster:

$$y_{ij} = \mu + x_{ij}\theta + \alpha(M)_j + x_{ij}\alpha(I)_j + e_{ij} \quad \alpha(M)_j \sim N[0, \tau_M^2] \quad \alpha(I)_j \sim N[0, \tau_I^2] \quad e_{ij} \sim N[0, \sigma_w^2], \tag{6}$$

where $\alpha(M)$ and $\alpha(I)$ are independent. We use the notation M to represent the main effect term, and I the interaction term. We have used different notation to that assumed in model extension 1 to make explicit that this is a different model. Under this assumption, the ICC in the control clusters will be

$$\rho_C = \frac{\tau_M^2}{\tau_M^2 + \sigma_w^2}, \tag{7}$$

and the ICC in the treatment clusters will be

$$\rho_T = \frac{\tau_M^2 + \tau_I^2}{\tau_I^2 + \tau_M^2 + \sigma_w^2}. \tag{8}$$

Using this parameterization, then $\rho_C <= \rho_T$. That is, there is an implicit assumption that the total variance in the treatment clusters is greater than or equal to the total variance in the control clusters. This of course is not a tenable assumption in all situations. This means that if the treatment comparison was switched around, so that the control clusters were coded as 1 and the treatment clusters as 0 model results would differ, an aspect of a model fit which is clearly undesirable.

2.2 | Treatment effect heterogeneity in cluster randomized designs

The previous section focused on differential clustering by arm (or heteroscedasticity), as appropriate in parallel designs as treatment is nested within clusters. We now move on to consider treatment effect heterogeneity in designs in which the treatment is crossed with cluster (cross-over and stepped-wedge trials).

We consider a conventional stepped-wedge study, with S sequences to which the k clusters are randomly allocated, and where at each of the $S + 1$ time points, a cross-section of observations is taken.[3,9] Because each cluster in a SW-CRT is exposed both to the treatment and control condition, treatment effects can be estimated within each cluster. While for simplicity, we focus on the SW-CRT, models that follow will be generalizable to other cluster trials in which treatment is crossed with cluster. We outline various different models along with the corresponding correlations and give more details on relationships between the various parameterizations in Appendix B.

2.2.1 | Basic model: single random effect

We extend the basic model (1) for parallel CRTs to the SW-CRT, by incorporating fixed effects for each period[9]:

$$y_{ijs} = \mu + x_{ijs}\theta + \alpha_j + \pi_s + e_{ijs} \quad \alpha_j \sim N[0, \tau^2] \quad e_{ijs} \sim N[0, \sigma_w^2], \tag{9}$$

where $s = 1, \ldots, S$ denotes the period and π_s is a fixed effect for each period, and m now represents the size of each cluster at each period, y_{ijs} the outcome for individual i in cluster j at time s, and x_{ijs} the corresponding treatment indicator.

2.2.2 | Model extension 1: two separate random effects

The first model we consider allows for different between-cluster variability in treatment and control clusters and follows from model extension 1 above for CRTs. Here, we incorporate two separate random effects, one for treatment and one for control. In contrast to when the clustering is nested within treatment arm, when the clustering is crossed with treatment, the model with a non-zero covariance between these two random effects becomes identifiable:

$$y_{ijs} = \mu + x_{ijs}\theta + \pi_s + x_{ijs}\alpha(T)_j + (1 - x_{ijs})\alpha(C)_j + e_{ijs} \quad e_{ijs} \sim N[0, \sigma_w^2], \tag{10}$$

again $\alpha(T)$ and $\alpha(C)$ represent random cluster effects for the treatment and control conditions, where

$$\begin{pmatrix} \alpha(T)_j \\ \alpha(C)_j \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_C^2 & \sigma_{CT} \\ \sigma_{CT} & \tau_T^2 \end{pmatrix} \right)$$

and where σ_{CT} is the possibly non-zero covariance between these two random effects.

2.2.3 | Model extension 1a: two separate independent random effects

In model extension 1a, we make the assumption that the correlation between these two separate random effects is zero (ie, $\sigma_{CT} = 0$). This induces differing ICCs in control and treated conditions. The correlation between two observations in the same cluster, both exposed to the control condition, will be

$$\rho_{CC} = \frac{\tau_C^2}{\tau_C^2 + \sigma_w^2}, \tag{11}$$

and the correlation between two observations in the same cluster, both exposed to the treatment condition, will be

$$\rho_{TT} = \frac{\tau_T^2}{\tau_T^2 + \sigma_w^2}. \tag{12}$$

Using this parameterization, there is no restriction on whether ρ_{CC} is bigger or smaller than ρ_{TT}. However, an additional correlation also exists: the correlation between two observations within the same cluster, but one in the control condition and one in the treatment condition, which we call ρ_{CT}. Under this model parameterization, this correlation, ρ_{CT}, is assumed to be zero (clearly undesirable).

2.2.4 | Model extension 1b: two separate nonindependent random effects

In model extension 1b, the correlation between these two separate random effects is allowed to be non-zero. This induces differing ICCs in control and treatment clusters, and importantly a non-zero correlation between observations within the same cluster but different treatment exposure. These correlations are

$$\rho_{CC} = \frac{\tau_C^2}{\tau_C^2 + \sigma_w^2}, \tag{13}$$

$$\rho_{TT} = \frac{\tau_T^2}{\tau_T^2 + \sigma_w^2}, \tag{14}$$

$$\rho_{CT} = \frac{\sigma_{CT}}{\sqrt{\tau_T^2 + \sigma_w^2}\sqrt{\tau_C^2 + \sigma_w^2}}. \tag{15}$$

As in model 1a above, there are no restrictions on the relative magnitude of ρ_{TT} and ρ_{CC}, but in contrast to model 1a, this model no longer includes the restriction that $\rho_{CT} = 0$.

2.2.5 | Model extension 2: a random interaction

We now consider extensions that clearly include an interaction between treatment and cluster. Again, because the clustering is crossed with treatment, the model with a non-zero covariance between these two random effects becomes identifiable:

$$y_{ijs} = \mu + x_{ijs}\theta + \pi_s + \alpha(M)_j + x_{ijs}\alpha(I)_j + e_{ijs} \quad e_{ijs} \sim N[0, \sigma_w^2], \tag{16}$$

where the two random effects $\alpha(M)$ and $\alpha(I)$ (where M represents the main effect term, and I the interaction term) have the following distribution:

$$\begin{pmatrix} \alpha(M)_j \\ \alpha(I)_j \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_M^2 & \sigma_{MI} \\ \sigma_{MI} & \tau_I^2 \end{pmatrix} \right).$$

2.2.6 | Model extension 2a: an independent random interaction

When the covariance term σ_{MI} between the two random effects is assumed to be zero, the respective correlations are

$$\rho_{CC} = \frac{\tau_M^2}{\tau_M^2 + \sigma_w^2}, \tag{17}$$

$$\rho_{TT} = \frac{\tau_M^2 + \tau_I^2}{\tau_I^2 + \tau_M^2 + \sigma_w^2}, \tag{18}$$

$$\rho_{CT} = \frac{\tau_M^2}{\sqrt{\tau_I^2 + \tau_M^2 + \sigma_w^2}\sqrt{\tau_M^2 + \sigma_w^2}}. \tag{19}$$

Using this parameterization, we see that there is again the restriction that ρ_{TT} is greater than or equal to ρ_{CC}, as was the case when this model was used to allow for differential clustering between arms; furthermore, there is also the restriction that ρ_{CT} is greater than or equal to ρ_{CC}.

2.2.7 | Model extension 2b: a nonindependent random interaction

Allowing the covariance term to be non-zero, the correlations become

$$\rho_{CC} = \frac{\tau_M^2}{\tau_M^2 + \sigma_w^2}, \tag{20}$$

$$\rho_{TT} = \frac{\tau_M^2 + \tau_I^2 + 2\sigma_{MI}}{\tau_I^2 + \tau_M^2 + \sigma_w^2 + 2\sigma_{MI}}, \tag{21}$$

$$\rho_{CT} = \frac{\tau_M^2 + \sigma_{MI}}{\sqrt{\tau_I^2 + \tau_M^2 + 2\sigma_{MI} + \sigma_w^2}\sqrt{\tau_M^2 + \sigma_w^2}}. \tag{22}$$

Inclusion of the non-zero covariance term means that the model does not induce assumptions on the variance in the treated periods being greater than the variance in the control periods, nor that ρ_{CT} is greater than or equal to ρ_{CC}.

2.2.8 | Model extension 3: an alternative random interaction

An alternative parameterization to allow for an interaction with treatment and cluster is one in which the random variation for the cluster effects is partitioned into a common part and an exposure specific part:

$$y_{ijs} = \mu + x_{ijs}\theta + \pi_s + \gamma_j + x_{ijs}\gamma(T)_j + (1 - x_{ijs})\gamma(C)_j + e_{ijs} \quad e_{ijs} \sim N[0, \sigma_w^2] \tag{23}$$

and that

$$\begin{pmatrix} \gamma_j \\ \gamma(C)_j \\ \gamma(T)_j \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_\gamma^2 & 0 & 0 \\ 0 & \tau_{\gamma T}^2 & 0 \\ 0 & 0 & \tau_{\gamma C}^2 \end{pmatrix} \right).$$

Note that we have used a change of notation from α to γ for the random effects to avoid confusion with models 1 and 2.

2.2.9 | Model extension 3a: an alternative random interaction with constrained total variance

When it is assumed that $\tau_{\gamma T}^2 = \tau_{\gamma C}^2 (= \tau_{\gamma TC}^2)$, the correlations between two observations in the same cluster and both exposed to the same treatment (whether that be treatment or control) will be

$$\rho_{CC} = \rho_{TT} = \frac{\tau_\gamma^2 + \tau_{\gamma TC}^2}{\tau_\gamma^2 + \tau_{\gamma TC}^2 + \sigma_w^2}, \tag{24}$$

and the correlations between two observations in the same cluster and both exposed to different treatments will be

$$\rho_{CT} = \frac{\tau_\gamma^2}{\tau_\gamma^2 + \tau_{\gamma TC}^2 + \sigma_w^2}. \tag{25}$$

This model assumes that the total variance under exposure to treatment is equivalent to the total variance under exposure to control. This parameterization therefore imposes a restriction on the correlations within clusters being identical for two observations that are both treated and two observations that are both not treated. Furthermore, this model also makes the assumption that ρ_{CT} is less than or equal to ρ_{TT} ($=\rho_{CC}$).

2.2.10 | Model extension 3b: an alternative random interaction without constrained total variance

An alternative parameterization, to model 3a, which relaxes the assumption that the total variance in the treatment arm is equivalent to the total variance in the treatment arm, is one in which the parameters $\tau_{\gamma T}^2$ and $\tau_{\gamma C}^2$ are not constrained to be the same. The correlations then become

$$\rho_{CC} = \frac{\tau_\gamma^2 + \tau_{\gamma C}^2}{\tau_\gamma^2 + \tau_{\gamma C}^2 + \sigma_w^2}, \tag{26}$$

$$\rho_{TT} = \frac{\tau_\gamma^2 + \tau_{\gamma T}^2}{\tau_\gamma^2 + \tau_{\gamma T}^2 + \sigma_w^2}, \tag{27}$$

$$\rho_{CT} = \frac{\tau_\gamma^2}{\sqrt{\tau_\gamma^2 + \tau_{\gamma T}^2 + \sigma_w^2}\sqrt{\tau_\gamma^2 + \tau_{\gamma C}^2 + \sigma_w^2}}, \tag{28}$$

While this model does not make the assumption that one of the ICCs is larger than the other, and neither makes the assumption that the correlation between two observations, one treated and one control, is zero, this model does make assumptions that the three random effects are uncorrelated. This does induce additional assumptions (Appendix B) that there is a positive correlation between the treatment and control random effects (in terms of model 1 notation as τ_C^2 increases so will τ_T^2 see Appendix B).

3 | MODEL FITTING

We have fitted these models in Stata 14 using the xtmixed function, REML methods, and default settings. Careful model specification and understanding of the language syntax is required to avoid model misspecification. We have included sample Stata code (Appendix C).

4 | EXAMPLE OF DIFFERENTIAL CLUSTERING IN A PARALLEL CRT

In this example, we illustrate results from fitting models 1 and 2 to a set of data from a parallel CRT. The trial we choose as an example is a parallel cluster trial conducted in 53 schools (clusters) of a behavioral treatment to prevent obesity in school-aged children.[10] The outcome we consider here is the child's body mass index measured at the end of the trial. There are a total of 689 observations in the treatment arm and 778 observations in the control arm, with an average cluster size of 24.

This dataset usefully highlights that one of the parameterizations (model 2) is not a good one. In this CRT, there is little variability between the treatment clusters, the estimated ICC from a simple analysis of variance on the treatment clusters only is very close to zero, whereas in the control clusters, it is higher at 0.057 (Table 1). Model 1 is able to capture this pattern in the data (estimating the ICCs in control and treatment arm as 0.050 and 0.000, respectively), whereas model 2 is unable to, estimating the ICCs in the control and treatment arm as 0.022 and 0.022 (Table 1). The estimated treatment effects differ slightly between the two approaches, but likelihoods do not indicate any preference for either model.

TABLE 1 Example of differential clustering in a parallel cluster randomized trial

	Log-likelihood	Treat effect, SE	First random effect, SE	Second random effect, SE	Residual variance, SE	ICC, 95% CI Control	Intervention
No differential clustering	−2100.47	0.111 (0.092)	0.195 (0.052)	1.288 (0.026)		0.022 (0.008, 0.062)	
Stratified model–control arm			0.304 (0.072)		1.236 (0.034)	0.057 (0.023, 0.134)	
Stratified model–treatment arm			0.000 (0.000)		1.337 (0.039)		0.000 (0.000, 0.000)
Differential clustering models							
Model 1—two separate random effects	−2097.09	0.098 (0.094)	0.000 (0.000)	0.296 (0.073)	1.285 (0.026)	0.050	0.000
Model 2—random interaction	−2100.47	0.111 (0.092)	0.000 (0.000)	0.195 (0.052)	1.288 (0.026)	0.022	0.022

Abbreviations: CI, confidence interval; ICC, intra-cluster correlation; SE, standard error.

5 | SIMULATION STUDY OF IMPACT OF MODEL CHOICE FOR TREATMENT EFFECT HETEROGENEITY IN CLUSTER RANDOMIZED DESIGNS

To investigate the actual consequences on estimation of model parameters when modeling treatment effect heterogeneity in cluster trials in which treatment is crossed with cluster, we undertook a small simulation study. In this simulation study, we have compared model performance statistics (ie, bias and coverage) across the 6 models (1a to 3b above) for the estimates of treatments effects (and standard errors) and the estimates of the variance of the random effects parameters (ie, the ICCs in both treatment and control arms).

This simulation study is not exhaustive and does not cover the range of possible scenarios but serves to demonstrate only some of the possible implications of model choice. This simulation study is limited to a simple stepped-wedge study. We assume a large number of clusters (100) and a large cluster size (1000) to avoid any issue of small sample sizes and assume a stepped-wedge study with 4 steps (ie, 25 clusters each randomized to one of the 4 sequences). We assume a treatment effect of 0 and residual variance of 1 (fixed throughout), and no secular trend (ie, each $\pi_s = 0$).

Data were generated from a linear mixed model as in model extension 1, adding cluster specific variation derived from three specified ICCs, one for the control condition, one for the treatment condition, and one for both. From this, we then derived variance components and a non-zero covariance term. These ICCs (the scenarios) were chosen to represent a broad range of scenarios, some of which are concordant with the models presented here and some of which were discordant.

The first scenario (scenario 1) examines what happens when we fit a model that is more complex than warranted by the data. In the first scenario, there is no differential heterogeneity, and the ICC under both treatment conditions, and between treatment conditions, is set at a typical value of 0.01. In this scenario, the data are concordant with all model extensions presented in this paper, except model 1a (which assumes $\rho_{CT} = 0$).

The second scenario (scenario 2) illustrates what happens when we fit these models to data in which there is differential clustering—but differential clustering that is not too discordant with most models. In the second scenario, there

TABLE 2 Simulation study of impact of model choice for treatment effect heterogeneity in cluster randomized designs

	Model 1a	Model 1b	Model 2a	Model 2b	Model 3a	Model 3b
Scenario 1 (ICC = 0.01 in control clusters; ICC = 0.01 in intervention clusters; ICC = 0.01 in clusters crossed with treatment)						
Treatment effects						
Absolute bias	0.00	0.00	0.00	0.00	0.00	0.00
Coverage	100.0%	93.7%	95.4%	95.6%	95.6%	95.8%
ICC estimates (percentage bias)						
ICC in treatment arm	−0.03	0.46	1.51	0.48	0.42	0.28
ICC in control arm	0.08	0.56	−0.77	0.51	0.42	0.44
Scenario 2 (ICC = 0.01 in control clusters; ICC = 0.05 in intervention clusters; ICC = 0.005 in clusters crossed with treatment)						
Treatment effects						
Absolute bias	0.00	0.00	0.00	0.00	0.00	0.00
Coverage	97.2%	96.1%	95.6%	96.1%	96.2%	96.1%
ICC estimates (percentage bias)						
ICC in treatment arm	0.31	0.32	14.31	0.32	−40.4	0.24
ICC in control arm	−0.95	−0.95	−4.93	−0.96	210.8	−0.93
Scenario 3 (ICC = 0.05 in control clusters; ICC = 0.01 in intervention clusters; ICC = 0.001 in clusters crossed with treatment)						
Treatment effects						
Absolute bias	0.00	0.00	0.00	0.00	0.00	0.00
Coverage	96.0%	94.9%	93.0%	94.9%	94.6%	94.8%
ICC estimates (percentage bias)						
ICC in treatment arm	−0.29	−0.29	907.7	−0.29	208.8	−0.29
ICC in control arm	−0.63	−0.63	−8.81	−0.63	−40.7	−0.63

Abbreviation: ICC, intra-cluster correlation.

Model 1: two separate random effects one for cluster and one for treatment condition (a [b]: with a zero [non-zero] covariance term).

Model 2: random interaction between treatment and cluster (a [b]: with a zero [non-zero] covariance term).

Model 3: two separate random effects one for cluster and one for treatment condition, with a partition into a common part (a [b]: with same [different] variance in treatment and control arms).

is differential heterogeneity: the ICC in the control condition being 0.01, the ICC in the treatment condition being 0.05, and the ICC between control and treatment conditions 0.005. In this scenario, the data are discordant with model 2a (which assumes $\rho_{CT} > \rho_{CC}$); model extension 3a (which assumes $\rho_{CC} = \rho_{TT}$); and model 1a (which assumes $\rho_{CT} = 0$).

In the third scenario, there is again differential heterogeneity across both arms, but the direction is reversed: the ICC is 0.05 in the control condition and lower in the treatment condition (0.01), and between conditions 0.001. So in this situation, there is differential heterogeneity, and this is very discordant with some models. In this scenario, the data are discordant with model 2a (because $\rho_{CT} < \rho_{CC}$ and $\rho_{CC} > \rho_{TT}$); model extension 3a (because $\rho_{CC} \neq \rho_{TT}$); and model 1a (because $\rho_{CT} \neq 0$).

We ran 1000 simulations for each scenario, which meant nominal coverage would be estimated to between 93.6% and 96.4%. We observed most models converged (8/1000 did not converge under model 1a); treatment effects showed little evidence of bias (Table 2). For the treatment effects, we observed some small departures from nominal coverage in some models, and some small differences in coverage between different models. However, these differences were small and likely to be due to sampling variation (values of coverage outside of the range 93.6 and 96.4 are possibly due to sampling variation). Model 1a exhibited higher than nominal coverage (100%) under scenario 1, explained by the zero covariance assumption of this model (Appendix B).

Comparing the implications on the estimates of the ICCs, we observed that models 2a and 3a perform particularly poorly. Under scenario 2, model 3a shows a large percentage bias for the ICC in both the control and treatment conditions (as much as 200%); model 2a shows moderate degree of bias (up to 14%). For scenario 3, models 2a and 3a again perform particularly poorly (up to 900% bias for model 2a).

6 | DISCUSSION

We have demonstrated that when heterogeneity in between-cluster variability or treatment effect heterogeneity is included in models for the analysis of a CRT, the parameterization of the model and the resulting implicit assumptions are important for estimation of ICCs and in limited scenarios on the standard errors of treatment effects. We have identified and illustrated that some parameterizations make assumptions that are not only not obvious from the model itself but also will not always be tenable.

6.1 | What is already known

Individually randomized trials

Where individually randomized studies have included random effects to model treatment effect heterogeneity, random interaction models allowing for a correlation between center main effects and treatment by center interactions have been used.[11-13]

Of note, Lee and Thompson discuss treatment effect heterogeneity where there is clustering in individually randomized trials and allow for non-zero covariance terms (model 2b), and while they clearly express a preference for a model that makes fewer assumptions, the assumptions implicit in the parameterizations are not clear.[13] The model proposed by Brown makes an assumption of equal variances in both arms (ie, is our model 3a).[11] Others have not clearly outlined the model parameterization choice at all.[12] Thus, even these scholarly articles we identified in the individually randomized literature did not make it apparent what model assumptions were being made.

Parallel CRTs

One situation where between-cluster heterogeneity is likely to be important is in designs with differential clustering between arms.[14,15] However, to date, these models have concentrated on situations in which it is natural to expect that the correlation between the observations within the treatment clusters is greater than that in the control clusters. Situations like this arise, for example, when the treatment induces clustering, for example, group therapy.[8] Model choice for such data very naturally fits in line with model 1a, and this has been the model of choice.[8]

CRTs with treatment crossed with cluster

In the cross-over trial literature, Turner et al have used a parameterization very similar to our model 3b, and so while do not make assumptions about variances being larger in the treatment arm,[6] the model does make other "hidden" assumptions (Appendix B). Yet others have followed more restrictive forms of parameterizations in the stepped-wedge literature and have chosen to use model 2a without appreciating the assumptions implicit in this model.[4,5]

Individual patient data meta-analysis

In individual patient data meta-analysis models, which also use mixed models and also allow for treatment effect heterogeneity (across studies rather than clusters), it has become common practice to include study effects as fixed rather than random. This means the issue of how to parameterize the additional random study by treatment heterogeneity is non-problematic. Despite this, where random effects models have been used, model assumptions have been clearly articulated,[16] which means clear guidance does exist for those wishing to model treatment effect heterogeneity using random effects. Of some note, it has been observed that the fully flexible model (model 2b here) can sometimes not converge, and so an alternative contrast parameterization (+1/2, −1/2) has been advocated for the model that has zero covariance terms that has a different constraint, namely, that the variances are equal between arms.[17]

Limitations

There are alternative ways to circumvent this issues raised in this paper. For example, an appropriate choice of contrast coding or selection of the appropriate arm to be coded as the arm under which heterogeneity is assumed to be greater. However, we do not endorse the choice of models that depend on choice of contrast coding, as this approach offers more potential for error and model misspecification on behalf of the user.

Our simulation study was small insofar, as it considered only a very limited number of scenarios. However, it illustrates that even in the case of large studies, where estimates of the ICCs are needed, model choice can have large implications on bias of these parameters and could induce unnecessarily large standard errors on treatment effects. It is likely that in other scenarios, where, for example, the cluster sizes are small or the number of clusters are small, the implications on the ICCs would be greater. We also limited our consideration to mixed models, whereas population average models offer an alternative approach.

While considering an interaction with cluster and treatment, we did not consider interactions between time and treatment or time and cluster.[7,18] Such interactions are likely to be of importance in trials that are longitudinal in nature as a decay in the strength of the correlations within a cluster will be expected.

Recommendations

In parallel CRTs, clusters are nested within treatment arms, whereas in the SW-CRT and cross-over designs, clusters are crossed with treatment arms. So in parallel CRTs, it does not make sense to consider clusters crossed with treatment, but rather clusters nested within treatment arms. This leads directly to model 1 as the natural approach for parallel cluster trials to model differential clustering between arms. It is also apparent from this that there are only two parameters here, and only two parameters estimable because each cluster gets only one treatment. Therefore, immediately any model with three parameters for the random effects structure is over-parameterized.

In a SW-CRT and cross-over trial, however, clusters are now crossed with treatments, so interaction terms as well as potential for covariance between the two random effects are now estimable. Model 1a no longer becomes a good choice, as it is unlikely to be a reasonable assumption that the correlation between two observations in the same cluster but of different treatment status are independent. Model 2a is unlikely to be a good choice, as it assumes the correlation between two observations, one treated and the other not, is greater than two not treated. Model 3a is equally unlikely to be a good choice, as it assumes the total variance is equivalent across treatment conditions. Finally, model 3b, while not performing poorly in our simulation study, does induce additional assumptions. So, in the analysis of stepped-wedge and cross-over trials, it may be more natural to follow approaches that make minimal assumptions, that is, model 1b or 2b.[12] Testing for simpler models (1a or 2a) will involve a different set of assumptions, which are clearly evident when placing the models into a common parameterization (Appendix B).

ORCID

Karla Hemming http://orcid.org/0000-0002-2226-6550

REFERENCES

1. Donner A, Klar N. Design and analysis of cluster randomization trials in health research; 2000.

2. Parienti J-J, Kuss O. Cluster-cross-over design: a method for limiting clusters level effect in community-intervention studies. *Contemp Clin Trials*. 2007;28(3):316-323.

3. Hemming K, Haines TP, Chilton PJ, Girling AJ, Lilford RJ. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *BMJ*. 2015;350:h391.

4. Hughes JP, Granston TS, Heagerty PJ. Current issues in the design and analysis of stepped wedge trials. *Contemp Clin Trials*. 2015;55-60.

5. Baio G, Copas A, Ambler G, Hargreaves J, Beard E, Omar RZ. Sample size calculation for a stepped wedge trial. *Trials*. 2015;16(1):354.

6. Turner RM, White IR, Croudace T. Analysis of cluster randomized cross-over trial data: a comparison of methods. *Stat Med*. 2007;26(2):274-289.

7. Ukoumunne OC, Thompson SG. Analysis of cluster randomized trials with repeated cross-sectional binary measurements. *Stat Med*. 2001;20(3):417-433.

8. Roberts C, Roberts SA. Design and analysis of clinical trials with clustering effects due to treatment. *Clin Trials*. 2005;2(2):152-162.

9. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials*. 2007;28(2):182-191.

10. Adab P, Pallan MJ, Lancashire ER, et al. A cluster-randomised controlled trial to assess the effectiveness and cost-effectiveness of a childhood obesity prevention programme delivered through schools, targeting 6–7 year old children: the waves study protocol. *BMC Public Health*. 2015;15(1):1.

11. Brown H, Prescott R. *Applied Mixed Models in Medicine*, Vol. 2. New York: John Wiley & Sons; 2006.

12. Feaster DJ, Mikulich-Gilbertson S, Brincks AM. Modeling site effects in the design and analysis of multi-site trials. *Am J Drug Alcohol Abuse*. 2011;37(5):383-391.

13. Lee KJ, Thompson SG. The use of random effects models to allow for clustering in individually randomized trials. *Clin Trials*. 2005;2(2):163-173.

14. Murray DM, Hannan PJ, Wolfinger RD, Baker WL, Dwyer JH. Analysis of data from group-randomized trials with repeat observations on the same groups. *Stat Med*. 1998;17(14):1581-1600.

15. Austin PC. A comparison of the statistical power of different methods for the analysis of repeated cross-sectional cluster randomization trials with binary outcomes. *Int J Biostat*. 2010;6(1).

16. Higgins JPT, Whitehead A, Turner RM, Omar RZ, Thompson SG. Meta-analysis of continuous outcome data from individual patients. *Stat Med*. 2001;20(15):2219-2241.

17. Turner RM, Omar RZ, Yang M, Goldstein H, Thompson SG, et al. A multilevel model framework for meta-analysis of clinical trials with binary outcomes. *Stat Med*. 2000;19(24):3417-3432.

18. Thompson JA, Fielding KL, Davey C, Aiken AM, Hargreaves JR, Hayes RJ. Bias and inference from misspecified mixed-effect models in stepped wedge trial analysis. *Stat Med*. 2017;36(23):3670-3682.

APPENDIX A

We outline the derivation of the correlations for mode1 2b, one of the more complex models. Correlations from other models can be derived very similarly (not shown). The ICC is the correlation between two different observations (i is individual; j is cluster) y_{ij} and $y_{i'j}$. In the SW-CRT, this is irrespective of which period (s) the observation is from. For simplicity were therefore suppress the s notation from the following derivation.

Firstly, we derive the correlation between two observations both treated:

$$\rho_{TT} = corr(y_{ij}, y_{i'j}) = \frac{cov(y_{ij}, y_{i'j})}{\sqrt{var(y_{ij}) * var(y_{i'j})}}, \tag{A1}$$

where

$$
\begin{aligned}
cov(y_{ij}, y_{i'j}) &= cov(\alpha_{Mi} + \alpha_{Ii} + e_{ij}, \alpha_{Mi'} + \alpha_{Ii'} + e_{i'j}) \\
&= cov(\alpha_{Mi}, \alpha_{Mi'}) + cov(\alpha_{Mi}, \alpha_{Ii'}) + cov(\alpha_{Mi}, e_{i'j}) \\
&\quad + cov(\alpha_{Ii}, \alpha_{Mi'}) + cov(\alpha_{Ii}, \alpha_{Ii'}) + cov(\alpha_{Ii}, e_{i'j}) \\
&\quad + cov(e_{ij}, \alpha_{Mi'}) + cov(e_{ij}, \alpha_{Ii'}) + cov(e_{ij}, e_{i'j}) \\
&= var(\alpha_M) + var(\alpha_I) + 2cov(\alpha_M, \alpha_I)
\end{aligned} \tag{A2}
$$

$$\rho_{TT} = corr(y_{ij}, y_{i'j}) = \frac{\tau_M^2 + \tau_I^2 + 2\sigma_{MI}}{\tau_M^2 + \tau_I^2 + \sigma_w^2 + 2\sigma_{MI}}. \tag{A3}$$

For the control arm, the ICC is the correlation between two different observations, y_{ij} and $y_{i'j}$, which are again in the same cluster, but this time now not treated:

$$\rho_{CC} = corr(y_{ij}, y_{i'j}) = \frac{cov(y_{ij}, y_{i'j})}{\sqrt{var(y_{ij}) * var(y_{i'j})}}, \tag{A4}$$

where

$$
\begin{aligned}
cov(y_{ij}, y_{i'j}) &= cov(\alpha_{Mi} + e_{ij}, \alpha_{Mi'} + e_{i'j}) \\
&= cov(\alpha_{Mi}, \alpha_{Mi'}) + cov(\alpha_{Mi}, e_{ji'}) \\
&\quad + cov(e_{ij}, \alpha_{Mi'}) + cov(e_{ij}, e_{ji'}) \\
&= var(\alpha_M))
\end{aligned}
\tag{A5}
$$

so that

$$\rho_{CC} = corr(y_{ij}, y_{i'j}) = \frac{\tau_M^2}{\tau_M^2 + \sigma_w^2}. \tag{A6}$$

For the SW-CRT, there is also a correlation between observations in the same cluster, any period, but different treatment exposures:

$$\rho_{CT} = corr(y_{ij(T=1)}, y_{i'j(T=0)}) = \frac{cov(y_{ij(T=1)}, y_{i'j(T=0)})}{\sqrt{var(y_{ij(T=1)}) * var(y_{i'j(T=0)})}}, \tag{A7}$$

where

$$
\begin{aligned}
cov(y_{ij(T=1)}, y_{i'j(T=0)}) &= cov(\alpha_{Mi(T=1)} + \alpha_{Ii(T=1)} + e_{ij(T=1)}, \alpha_{Mi'(T=0)} + \alpha_{Ii'(T=0)} + e_{i'j(T=0)}) \\
&= cov(\alpha_{Mi(T=1)}, \alpha_{Mi'(T=0)}) + cov(\alpha_{Mi(T=1)}, \alpha_{Ii'(T=0)}) + cov(\alpha_{Mi(T=1)}, e_{ji'(T=0)}) \\
&\quad + cov(\alpha_{Ii(T=1)}, \alpha_{Mi'(T=0)}) + cov(\alpha_{Ii(T=1)}, \alpha_{Ii'(T=0)}) + cov(\alpha_{Ii(T=1)}, e_{ji'(T=0)}) \\
&\quad + cov(e_{ij(T=1)}, \alpha_{Mi'(T=0)}) + cov(e_{ij(T=1)}, \alpha_{Ii'(T=0)}) + cov(e_{ij(T=1)}, e_{ji'(T=0)}) \\
&= var(\alpha_M) + cov(\alpha_M, \alpha_I)
\end{aligned}
\tag{A8}
$$

Furthermore,

$$
\begin{aligned}
var(y_{ij(T=1)}) &= var(\alpha_{Mi(T=1)} + \alpha_{Ii(T=1)} + e_{ij(T=1)}) \\
&= \tau_M^2 + \sigma_I^2 + \sigma_w^2,
\end{aligned}
\tag{A9}
$$

and

$$
\begin{aligned}
var(y_{ij(T=0)}) &= var(\alpha_{Mi(T=0)} + \alpha_{Ii(T=0)} + e_{ij(T=0)}) \\
&= \tau_M^2 + \sigma_w^2,
\end{aligned}
\tag{A10}
$$

$$\rho_{CT} = corr(y_{ij(T=1)}, y_{i'j(T=0)}) = \frac{\tau_M^2 + \sigma_{IM}}{\sqrt{(\tau_M^2 + \tau_I^2 + 2\sigma_{MI} + \sigma_w^2)(\tau_M^2 + \sigma_W^2)}}. \tag{A11}$$

APPENDIX B

B.1 | General formulation of treatment heterogeneity models

Any of the 3 models can be the "general formulation" since can always obtain one from the other. Here, we use model extension 1, which has a random effect for cluster in treated conditions, and for cluster in control conditions, and allows a covariance between them.

Notation

j: clusters
s: periods

i: person

Let x_{ijs} be the treatment in the jth cluster and sth period

θ is the average treatment effect

π_s is the period effect at period s

y_{ijs} is the outcome for the ith person in period s in the jth cluster

B.1.1 | Model extension 1

$$y_{ijs} = \mu + \pi_s + x_{ijs}\theta + x_{ijs}\alpha(T)_j + \left(1 - x_{ij}\right)\alpha(C)_j + \epsilon_{ijs}, \tag{1}$$

where $\pi_1 = 0$ for identifiability.

$\epsilon_{ijs} \sim N(0, \sigma_w^2)$ is individual random errors, and the two random effects have a multivariate normal distribution: $(\alpha(T)_j\alpha(C)_j) \sim N((0, 0), (\tau_T^2, \sigma_{CT}\tau_C^2, \sigma_{CT}))$.

Interpretation:

μ is the expected average/population response in the control condition in the absence of period effects averaged over all clusters and individuals.

$\mu + \theta$ is the expected average/population response in the treated condition in the absence of period effects, averaged over all clusters and individuals.

$\alpha(C)_j$ is the deviation of the average response of cluster j in the control condition from the population average μ. It can be thought of as the totality of unmeasured covariates at cluster level at the time of implementation of the control condition.

$\alpha(T)_j$ is the deviation of the average response of cluster j in the treated condition from the population average $\mu + \theta$. It can be thought of as the totality of unmeasured covariates at cluster level at the time of implementation of the treated condition.

The ICCs for this model are

$$\rho_{CC} = \frac{\tau_C^2}{\tau_C^2 + \sigma_w^2}, \quad \rho_{TT} = \frac{\tau_T^2}{\tau_T^2 + \sigma_w^2}, \quad \rho_{CT} = \frac{\sigma_{CT}}{\sqrt{\tau_C^2 + \sigma_w^2}\sqrt{\tau_T^2 + \sigma_w^2}}, \tag{2}$$

And Cov(control random effect, random treatment difference) $= Cov\left(\alpha(C)_i, \theta + \alpha(T)_i - \alpha(C)_i\right) = \sigma_{CT} - \tau_C^2$. This can be positive or negative depending on the relative sizes of σ_{CT} and τ_C^2.

Model extension 1A

This model has no correlation between $\alpha(C)_i$ and $\alpha(T)_i$, ie, $\sigma_{CT} = 0$.

This also implies Cov(control random effect, random treatment difference) $= -\tau_C^2 < 0$, meaning the larger the control response is, the smaller (or more negative) the treatment difference is for that cluster.

B.1.2 | Model extension 2

$$y_{ijs} = \mu + \pi_s + x_{ijs}\theta + \alpha(M)_j + x_{ijs}\alpha(I)_j + \epsilon_{ijs}$$

Expressing this in terms of model 1, we have

$$\alpha(M)_j = \alpha(C)_j; \quad \alpha(I)_j = \alpha(T)_j - \alpha(C)_j,$$
$$Var\left(\alpha(M)_j\right) = \tau_M^2; \quad Var\left(\alpha(I)_j\right) = \tau_I^2; Cov = \sigma_{MI}.$$

Then expressing model 2 parameters in terms of model 1 parameters, $\tau_M^2 = \tau_C^2, \tau_I^2 = \tau_C^2 + \tau_T^2 - 2\sigma_{CT}, \sigma_{MI} = \sigma_{CT} - \tau_C^2$, and $\sigma_{CT} =$Cov(control random effect, treated random effect) $= Cov\left(\alpha(M)_i, \alpha(M)_i + \alpha(I)_i\right) = \tau_M^2 + 2\sigma_{MI} > 0$, which may be positive or negative. Then also, Cov(control random effect, random treatment difference) $= Cov\left(\alpha(M)_i, \theta + \alpha(I)_i\right) = \sigma_{MI} = \sigma_{CT} - \tau_C^2$, which may be positive or negative.

And expressing model 1 parameters in terms of model 2,

$$\tau_C^2 = \tau_M^2; \quad \sigma_{CT} = \sigma_{MI} + \tau_M^2$$
$$\tau_T^2 = \tau_I^2 - \tau_C^2 + 2\sigma_{CT} = \tau_I^2 - \tau_M^2 + 2\left(\sigma_{MI} + \tau_M^2\right) = \tau_I^2 + \tau_C^2 + 2\sigma_{MI}.$$

Extension 2A

Model extension 2A has no correlation between $\alpha(M)_i$ and $(I)_i$ so that $\sigma_{MI} = 0$.

Hence, $\tau_T^2 = \tau_I^2 + \tau_C^2 > \tau_C^2$, hence treatment random effect variance is always larger than control. And $\sigma_{MI} = 0$ corresponds to assuming $\sigma_{CT} = \tau_M^2 = \tau_C^2 > 0$, implying that there is a positive correlation between the random effects under treatment $\alpha(T)_i$ and control $\alpha(C)_i$

And also by definition, it means that Cov(control random effect, random treatment difference)= 0 exactly. The size of the treatment difference in cluster j does not depend on the control mean.

In model 1 notation,

$$\rho_{CC} = \frac{\tau_C^2}{\tau_C^2 + \sigma_w^2}, \quad \rho_{TT} = \frac{\tau_T^2}{\tau_T^2 + \sigma_w^2}, \quad \rho_{CT} = \frac{\sigma_{CT}}{\sqrt{\tau_C^2 + \sigma_w^2}\sqrt{\tau_T^2 + \sigma_w^2}} = \frac{\tau_C^2}{\sqrt{\tau_C^2 + \sigma_w^2}\sqrt{\tau_T^2 + \sigma_w^2}}.$$

In model 2 notation, $\rho_{CC} = \frac{\tau_M^2}{\tau_M^2 + \sigma_w^2}$, $\rho_{TT} = \frac{\tau_M^2 + \tau_I^2}{\tau_M^2 + \tau_I^2 + \sigma_w^2}$, $\rho_{CT} = \frac{\tau_M^2}{\sqrt{\tau_M^2 + \sigma_w^2}\sqrt{\tau_M^2 + \tau_I^2 + \sigma_w^2}}$, and from this, it is clear that $\rho_{TT} > \rho_{CC}$ if $\tau_I^2 > 0$.

Extension 2B

Model extension 2B allows covariance σ_{MI}. So that $\sigma_{MI} = \sigma_{CT} - \tau_C^2$ (and this can be positive or negative). And $\tau_T^2 = \tau_I^2 + \tau_C^2 + 2\sigma_{MI}$ can be greater or less than τ_C^2.

Substituting back to model 2 notation,

$$\rho_{CC} = \frac{\tau_M^2}{\tau_M^2 + \sigma_w^2}, \quad \rho_{TT} = \frac{\tau_M^2 + \tau_I^2 + 2\sigma_{MI}}{\tau_M^2 + \tau_I^2 + 2\sigma_{MI} + \sigma_w^2}, \quad \rho_{CT} = \frac{\tau_M^2 + \sigma_{MI}}{\sqrt{\tau_M^2 + \sigma_w^2}\sqrt{\tau_M^2 + \tau_I^2 + 2\sigma_{MI} + \sigma_w^2}},$$

where $\tau_T^2 = \tau_I^2 + \tau_C^2 + 2\sigma_{MI}$.

B.1.3 | Model extension 3

$$Y_{ijs} = \mu + \pi_s + x_{ijs}\theta + \gamma_j + (1 - x_{ij})\gamma(C)_j + x_{ij}\gamma(T)_j + \epsilon_{ijs},$$
$$Var(\gamma_i) = \tau_\gamma^2; \; Var(\gamma(C)_i) = \tau_{\gamma C}^2; \quad Var(\gamma(T)_i) = \tau_{\gamma T}^2.$$

And $Cov(\gamma_i, \gamma(C)_i) = 0$; $Cov(\gamma(C)_i, \gamma(T)_i) = 0$.

Then,

$$\gamma_i + \gamma(C)_i = \alpha(C)_i; \quad \gamma_i + \gamma(T)_i = \alpha(T)_i.$$

This thus gives that $\sigma_{CT} =$ Cov(control random effect, treated random effect)$= Cov(\gamma_i + \gamma(C)_i, \gamma_i + \gamma(T)_i) = \tau_\gamma^2 > 0$. Hence, there is always a positive correlation between treated and control random effects.

Then, in terms of model 1 parameters, $\tau_\gamma^2 + \tau_{\gamma C}^2 = \tau_C^2$, $\tau_\gamma^2 + \tau_{\gamma T}^2 = \tau_T^2$, $\tau_\gamma^2 = \sigma_{CT}$

$$\tau_{\gamma C}^2 = \tau_C^2 - \sigma_{CT}; \; \tau_{\gamma T}^2 = \tau_T^2 - \sigma_{CT}.$$

This therefore means that Cov(control random effect, random treatment difference)$= Cov(\gamma_i + \gamma(C)_i, \theta + \gamma(T)_i - \gamma(C)_i) = -\tau_{\gamma C}^2 < 0$.

And expressing model 1 parameters in terms of model 3,

$$\tau_C^2 = \tau_\gamma^2 + \tau_{\gamma C}^2; \quad \tau_T^2 = \tau_\gamma^2 + \tau_{\gamma T}^2; \; \sigma_{CT} = \tau_\gamma^2.$$

B.1.4 | Extension 3A

Model extension 3A has equal variance for $\gamma(C)_i$ and $(T)_i$, ie, $\tau_{\gamma C}^2 = \tau_{\gamma T}^2$, hence corresponds to assuming equal variances of treated and control random effects $\tau_C^2 = \tau_T^2$.

In model 3 notation,

$$\rho_{CC} = \frac{\tau_\gamma^2 + \tau_{\gamma C}^2}{\tau_\gamma^2 + \tau_{\gamma C}^2 + \sigma_w^2}, \quad \rho_{TT} = \frac{\tau_\gamma^2 + \tau_{\gamma C}^2}{\tau_\gamma^2 + \tau_{\gamma C}^2 + \sigma_w^2}, \quad \rho_{CT} = \frac{\sigma_{CT}}{\sqrt{\tau_C^2 + \sigma_w^2}\sqrt{\tau_T^2 + \sigma_w^2}} = \frac{\tau_\gamma^2}{\tau_\gamma^2 + \tau_{\gamma C}^2 + \sigma_w^2}.$$

Substituting back to model 3 notation, it is therefore clear that $\rho_{CC} = \rho_{TT} = \frac{\tau_C^2}{\tau_C^2 + \sigma_w^2}$, $\rho_{TC} = \frac{\sigma_{CT}}{\tau_C^2 + \sigma_w^2}$ (or equiv with τ_T^2).

B.1.5 | Extension 3B

Model extension 3B allows unequal variances for $\gamma(C)_i$ and $(T)_i$, ie, $\tau_{\gamma C}^2 \neq \tau_{\gamma T}^2$. Using model 1 notation,

$$\rho_{CC} = \frac{\tau_\gamma^2 + \tau_{\gamma C}^2}{\tau_\gamma^2 + \tau_{\gamma C}^2 + \sigma_w^2}, \quad \rho_{TT} = \frac{\tau_\gamma^2 + \tau_{\gamma T}^2}{\tau_\gamma^2 + \tau_{\gamma T}^2 + \sigma_w^2},$$

$$\rho_{TC} = \frac{\sigma_{CT}}{\sqrt{\tau_C^2 + \sigma_w^2}\sqrt{\tau_T^2 + \sigma_w^2}} = \frac{\tau_\gamma^2}{\sqrt{\tau_\gamma^2 + \tau_{\gamma C}^2 + \sigma_w^2}\sqrt{\tau_\gamma^2 + \tau_{\gamma T}^2 + \sigma_w^2}}.$$

This means that since $\sigma_{CT} = \text{Cov}(\text{control random effect, treated random effect}) = \tau_\gamma^2 > 0$). And furthermore that the larger the control response, the smaller the treatment difference is for that cluster, since Cov(control random effect, random treatment difference) $= -\tau_{\gamma C}^2 < 0$.

APPENDIX C

C.1 | Stata code to fit models 1a to 3b

```
*y represents the outcome
*trt represents the treatment indicator (1 treated 0 control)
*cluster represents the grouping by cluster
*clustertrt represents a grouping by cluster and treatment
egen clustertrt = group(cluster trt)
*model 1a
xtmixed y trt || cluster: trt notreat, nocons var reml
*model 1b
xtmixed y trt || cluster: trt notreat, nocons cov(uns) var reml
*model 2a
xtmixed y trt || cluster: trt, var reml
*model 2b
xtmixed y trt || cluster: trt, cov(uns) var reml
*model 3a
xtmixed y trt || cluster: || clustertrt: ,variance reml
*model 3b
xtmixed y trt || cluster: || clustertrt: trt notreat, nocons variance reml
*Code to extract variance terms
*Model 1a
local M1a_var_r_trt = exp(2*_b[lns1_1_1:_cons])
local M1a_var_r_control = exp(2*_b[lns1_1_2:_cons])
*Model 1b
local M1b_var_r_trt = exp(2*_b[lns1_1_1:_cons])
local M1b_var_r_control = exp(2*_b[lns1_1_2:_cons])
local M1b_corr_r_trt_control = tanh(_b[atr1_1_1_2:_cons])
local M1b_cov_r_trt_control = 'M2_corr_r_trt_control'*sqrt('M2_var_r_trt'*'M2_var_r_control')
*Model 2a
local M2a_var_r_control = exp(2*_b[lns1_1_2:_cons])
local M2a_var_r_trt = exp(2*_b[lns1_1_1:_cons])+exp(2*_b[lns1_1_2:_cons])
```

*Model 2b

local M2b_var_r_control = exp(2*_b[lns1_1_2:_cons])

local M2b_var_r_trt = exp(2*_b[lns1_1_1:_cons])

local M2b_corr_r_trt_control = tanh(_b[atr1_1_1_2:_cons])

local M2b_cov_r_trt_control = 'M4_corr_r_trt_control'*sqrt('M4_var_r_trt'*'M4_var_r_control'

local M2b_var_r_trt = 'M4_var_r_trt' + 'M4_var_r_control' + 2*'M4_cov_r_trt_control'

*Model 3a

local M3a_var_r_control = exp(2*_b[lns1_1_1:_cons]) + exp(2*_b[lns2_1_1:_cons])

local M3b_var_r_trt = 'M5_var_r_control'

*Model 3b

local M3b_pvalue=2*(1-abs(normal('M6_test_valu')))

local M3b_var_r_control = exp(2*_b[lns1_1_1:_cons]) + exp(2*_b[lns2_1_2:_cons])

local M3b_var_r_trt = exp(2*_b[lns1_1_1:_cons]) + exp(2*_b[lns2_1_1:_cons])

APPENDIX D: SUMMARY

D.1 | Model extension 1:

Model extension 1A has the restriction that there is no correlation between cluster random effects of control and treatment conditions. This implies that

(a) The larger the control response is, the smaller the treatment difference is for that cluster, since the Cov(control random effect, random treatment difference) $= -\tau_C^2 < 0$.
(b) The correlation between two observations in the same cluster, one observed under the control and the other under the treatment condition, is zero ($\rho_{TT} = 0$).

Model extension 1B is unrestricted for the variances, correlations, and ICCs.

D.2 | Model extension 2:

Model extension 2A has the restriction of no correlation between main random effect and the interaction random effect This implies that

(a) The treatment random effect variance is always larger than the control random effect variance ($\tau_T^2 = \tau_I^2 + \tau_C^2 > \tau_C^2$).
(b) The ICC for treatment clusters is always larger than the ICC for control clusters ($\rho_{TT} > \rho_{CC}$).
(c) There is a positive correlation between the random effects under treatment $\alpha(T)_i$ and control $\alpha(C)_i$ (since $\sigma_{MI} = 0$ corresponds to assuming $\sigma_{CT} = \tau_C^2 > 0$).

Model extension 2B is unrestricted variances, correlations, and ICCs

D.3 | Model extension 3:

Model extension 3A is restricted to equal variances of all cluster-treatment random effects. This implies that

(a) The variances of treated and control random effects are equivalent ($\tau_C^2 = \tau_T^2$).
(b) The correlation between two observations both treated is equivalent to the correlation between two observations both observed under the control condition ($\rho_{TT} = \rho_{CC}$).

Model extension 3B has the restriction that the correlation between the 3 model random effects is zero. This implies that

(a) There is a positive correlation between treated and control random effects, since $\sigma_{CT} =$ Cov(control random effect, treated random effect) $= \tau_\gamma^2 > 0$.
(b) The larger the control response, the smaller the treatment difference is for that cluster, since Cov(control random effect, random treatment difference) $= -\tau_{\gamma C}^2 < 0$.