pmc painmanagement collaboratory

Supporting Research in Pain Management for Veterans and Military Service Members

A working document from the Pain Management Collaboratory Biostatistics/Design Workgroup. This work is supported by the National Institutes of Health (NIH), the Department of Defense (DoD), and the Department of Veterans Affairs (VA). The views presented here are solely the responsibility of the authors and do not necessarily represent the official views of the Funding Agencies

Intraclass Correlation Coefficients for Cluster Randomized Trials with Pain Outcomes

Background

Cluster randomized trial (CRT) designs are commonly used for testing interventions in real world settings particularly when contamination of the intervention across individual participants or patients is a concern.

Typical sample size calculations for individual randomized parallel arm trials require input on Type I error rate (α), Type II error rate (β), clinically meaningful difference, and a variance estimate (or proportion of participants expected to have outcome if the outcome is categorical). Sample size calculations for cluster randomized trials require an additional input, the intraclass correlation coefficient (ICC). The ICC estimates the correlation among individuals' responses within the same cluster and can have a large impact on the sample sizes needed to appropriately power cluster randomized trials.

Often in planning sample sizes, biostatisticians use best guesses of ICC's based on very limited information. ICC's vary widely depending on the clustering unit and outcomes being measured [1]. For this white paper, our goals were to provide an overview of the ICC, assess the impact on sample size calculations, describe a range of ICC's either assumed or observed in cluster randomized trials where pain was the primary outcome measure, and provide recommendations for trialists planning cluster randomized trials with pain outcomes.

Methodological Considerations

A commonly used mixed effects model for a continuous outcome in a cluster randomized trial with i=1,…..n clusters and j=1,…….m_i individuals per cluster is expressed as:

$$Y_{ij} = \beta_0 + \beta_1 X_i + \gamma_i + \varepsilon_{ij},$$

where X_i is an indicator for the treatment assigned to the i^{th} cluster, $\gamma_i \sim N(0, \sigma_b^2)$ and $\varepsilon_{ij} \sim N(0, \sigma_w^2)$.

Prepared by:	*Charity G. Patterson, PhD, MSPH (University of Pittsburgh), Scott S. Coggeshall, PhD (VA Puget Sound), Tom Greene, PhD (University of Utah)*
Acknowledgement:	*Clair Smith, MS (University of Pittsburgh)*
Contributions and Review by:	*NIH-VA-DoD Pain Management Collaboratory Biostatistics/Design Workgroup*
Version:	*Version #1.0, last updated December 10, 2019*

The ICC is defined as

$$\rho = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_w^2},$$

and represents the similarity in the outcome of individuals within the same cluster as the proportion of the total variation in the outcome attributed to the differences between clusters. Note the ICC will range between 0 and 1.

An analogous mixed effects model for a binary outcome in a cluster randomized trial is expressed as

$$logit(p_{ij}) = \beta_0 + \beta X_i + \alpha_i$$

where p_{ij} denotes the probability of the binary response, X_i is again an indicator for the treatment (0,1), and $\alpha_i \sim N(0, \sigma^2)$. This model can be seen as a latent-response model for Y_{ij}:

$$Y_{ij}^* = \beta_0 + \beta X_i + \alpha_i + \varepsilon_{ij}$$

where $Y_{ij}=1$ if $Y_{ij}^* > 0$ and 0 otherwise and ε_{ij} is assumed to have a logistic normal distribution with mean 0 and variance $\pi^2/3$. In this formulation, the ICC is given by

$$\rho = \frac{\sigma^2}{\sigma^2 + \pi^2/3}$$

The random effect/latent variable derivation of the ICC introduced here is one way of characterizing correlation within a cluster. Donner and Klar [2] provide a general overview of this and other approaches to defining and estimating an ICC in cluster randomized trials. Consideration of how to define the ICC is particularly important for binary outcomes, as we discuss below. Wu and colleagues [3] provides a comparison of 5 different methods for estimating the ICC for binary outcomes.

Standard sample size calculation

Sample size calculations for cluster randomized trials must take into account the design effect induced by the cluster randomized design. The design effect (DE) is$1+(m-1)\rho$ where m=the average number of individuals per cluster. The ICC (ρ) assumed in power analysis has a substantial impact on the number of individuals and clusters required for a trial. For example, suppose we plan a study to detect an effect size of 0.5 on a pain scale between intervention and control groups. The sample size needed for 80% power without adjustment for clustering is total N=128. If the trial was designed as a cluster randomized trial and we assume the ICC=0.05 and fixed number of individuals per cluster of 8, we would need to inflate the sample size by 35% [DE=1+(8-1)*0.05=1.35 (N*=176)].

Prepared by:	*Charity G. Patterson, PhD, MSPH (University of Pittsburgh), Scott S. Coggeshall, PhD (VA Puget Sound), Tom Greene, PhD (University of Utah)*
Acknowledgement:	*Clair Smith, MS (University of Pittsburgh)*
Contributions and Review by:	*NIH-VA-DoD Pain Management Collaboratory Biostatistics/Design Workgroup*
Version:	*Version #1.0, last updated December 10, 2019*

Thus, we would need k=22 clusters (i.e. 176/8=22). If the ICC were incorrectly assumed to be 0.01, the design effect would be DE=1+(8-1)*0.01=1.07 (N*=144) resulting in fewer clusters (k=18) and loss of power if the ICC was truly higher.

Varying the cluster size for different ICCs demonstrates the impact on overall sample size and number of clusters for a fixed power. For example, the unadjusted sample size of N=128 participants would be inflated 15% for an ICC=0.01 and 75% for an ICC=0.05 with m=16 participants per clusters resulting in overall sample sizes 148-224 and 10-14 clusters. In contrast, with m=8 participants per cluster the unadjusted sample size is only inflated 7% to 35% for ICC=0.01 and ICC=0.05 resulting in smaller sample sizes (137-173) but markedly more clusters (k=18-22).

Assumed and Observed ICCs for Previous Cluster Randomized Trials with Pain Outcomes

In our search for cluster randomized trials with pain outcomes, we found a range of assumed ICCs from 0.01 to 0.2 and observed ICCs from 0 to 0.1 depending on the cluster unit and outcome (See Table 1).

All but one study used an ICC of equal or higher value to what was observed in the actual trial. Unfortunately, to date most trials have had to make assumptions about the ICC with limited prior evidence. We encourage investigators to report their observed ICCs once their trials are completed, thereby facilitating improved design of future cluster trials in pain research.

Recommendations for trialists designing trials involving clusters

- Conduct sample size analysis using conservative but reasonable values based on the study design, clustering unit, and the outcome of interest. Proceed with caution when using estimates of ICCs from studies with a small number of clusters particularly if the studies are pilot studies. Pilot studies should be used to assess feasibility and get preliminary estimates of nuisance parameters. Consider using the upper bound of the confidence intervals for the ICC[4] in addition to estimates from the literature.
- Inflate the sample size to take into account the variation in cluster sizes relative to the average cluster size[5, 6]
- If planning a trial with a binary outcome, consider the impact of different ICC values under the alternative hypothesis.
- When planning a trial, accurately account for all clustering within a cluster randomized design before finalizing sample size analysis. In addition to the standard cluster randomized trial setting in which clusters of individuals are randomized to different interventions, it is also common to see randomized trials of pain interventions in which randomization occurs at the

Prepared by:	*Charity G. Patterson, PhD, MSPH (University of Pittsburgh), Scott S. Coggeshall, PhD (VA Puget Sound), Tom Greene, PhD (University of Utah)*
Acknowledgement:	*Clair Smith, MS (University of Pittsburgh)*
Contributions and Review by:	*NIH-VA-DoD Pain Management Collaboratory Biostatistics/Design Workgroup*
Version:	*Version #1.0, last updated December 10, 2019*

individual level, but multiple subjects in one or more treatment groups are treated by the same provider (examples: physical therapists or cognitive behavioral counselors). In the past, such trials have usually been analyzed at the individual level, but proper statistical inference requires that this type of clustering be accounted for using random effects which are specific to the intervention groups in which the clustering occurs. See, for example, Bauer et al [7], Lee and Thompson [8], Hedges and Citkowicz [9] for additional details.

- Consider monitoring the ICC as a nuisance parameter during the course of the trial. If the observed ICC is larger than anticipated, sample size re-estimation may be necessary[10].
- After a study is completed, report both the ICC assumed in the sample size analysis and the observed with a confidence interval.
- If considering a cluster crossover trial, stepped-wedge trial, or other longitudinal design, be aware that additional assumptions must be made about correlation between participants within the same cluster at different time points and the correlation of cluster means across time periods (within cluster between period correlation)[11, 12].

Additional considerations for binary outcomes

Binary outcomes present additional challenges for defining and modeling the ICC compared to continuous outcomes. It is well-known that binary random variables exhibit a mean-variance relationship, in which the variance depends on the mean through the relationship

Variance = mean*(1 – mean)

There is empirical evidence for mean-ICC relationship [3, 13] which makes sense from a theoretical standpoint when the ICC is defined on the probability scale, given the mean-variance relationship. In a two-arm cluster randomized study with a binary outcome, we would therefore expect the ICC among the treatment clusters to differ from the ICC among the control clusters on the probability scale in the presence of a true underlying effect of the intervention. For models such as logistic regression with random intercepts that model the ICC on the logit scale, it is theoretically possible to have a single logit-scale ICC shared by all clusters, even when an intervention effect causes means to differ between treatment and control clusters. At the same time, the aforementioned issues suggest that this assumption may often be violated in real data. Researchers should be aware of a potential mean-ICC relationship when choosing and assessing the analysis model and consider allowing for varying ICCs across treatment groups when appropriate. Allowance for varying ICCs can be achieved in mixed effects models, for instance, by including additional variance terms[14].

Prepared by:	*Charity G. Patterson, PhD, MSPH (University of Pittsburgh), Scott S. Coggeshall, PhD (VA Puget Sound), Tom Greene, PhD (University of Utah)*
Acknowledgement:	*Clair Smith, MS (University of Pittsburgh)*
Contributions and Review by:	*NIH-VA-DoD Pain Management Collaboratory Biostatistics/Design Workgroup*
Version:	*Version #1.0, last updated December 10, 2019*

Table 1

Citation	Design or Results	Assumed level of clustering	Unit of measurement	Primary outcome (Pain) and time point	Assumed ICC in power analysis	Observed ICC
Adams G, Gulliford MC, Ukoumunne OC, Eldridge S, Chinn S, Campbell MJ. Patterns of intra-cluster correlation from primary care research to inform study design and analysis. *Journal of clinical epidemiology*. 2004;57(8):785-794 [1]	Results	Clinic	Patient	SF-36 pain	Not applicable	0, 0.012, 0.039, 0.053
Albaladejo C, Kovacs FM, Royuela A, del Pino R, Zamora J, Network SBPR. The efficacy of a short education program and a short physiotherapy program for treating low back pain in primary care: a cluster randomized trial. *Spine*. 2010;35(5):483-496.[15]	Results	PCP	Patient	R&M disability baseline, 90, 180 days (VAS as well)	0.2	Not reported
Becker A, Leonhardt C, Kochen MM, et al. Effects of two guideline implementation strategies on patient outcomes in primary care: a cluster randomized controlled trial. *Spine*. 2008;33(5):473-480.[16]	Results	PCP	Patient	Days in pain baseline, 6, 12 months	0.03	Not reported
Delitto, Patterson, ….Saper Study protocol for targeted interventions to prevent chronic low back pain in high-risk patients: A multi-site pragmatic cluster randomized controlled trial (TARGET Trial). *Contemporary Clinical Trials* 2019: Jul;82:66-76. [17]	Design and Unpublished	Clinic	Patient	Transition to chronic low back pain at 6 months	0.05	0.03 (unofficialunpublished)
Lonsdale C, Hall AM, Williams GC, et al. Communication style and exercise compliance in physiotherapy (CONNECT). A cluster randomized controlled trial to test a theory-based intervention to increase chronic low back pain patients' adherence to physiotherapists' recommendations: study rationale, design, and methods. *BMC musculoskeletal disorders*. 2012;13(1):104....[18]	Design	Clinic	Patient	NPRS baseline, 1, 4, 12, 24 weeks	0.03	Not reported
Morone NE, Greco CM, Moore CG, Rollman BL, Lane B, Morrow LA, Glynn NW, Weiner DK. A mind body program for older adults with chronic low back pain: A randomized controlled trial. *JAMA Intern Med* 2016: Mar;176(3):329-3. [19]	Results	Participant but interventions delivered in class cohorts (IRGT)	Participant	R&M disability	0.01-0.02	0.016/0.021
Schmidt CO, Chenot J-F, Pfingsten M, et al. Assessing a risk tailored intervention to prevent disabling low back pain-protocol of a cluster randomized controlled trial. *BMC Musculoskeletal Disorders*. 2010;11(1):5. (results published in German *Der Schmerz* June 2019) [20]	Design	Clinic	Patient	Graded chronic pain scale baseline, 6, 12 months	0.025	Not in the results paper.

Prepared by:	*Charity G. Patterson, PhD, MSPH (University of Pittsburgh), Scott S. Coggeshall, PhD (VA Puget Sound), Tom Greene, PhD (University of Utah)*
Acknowledgement:	*Clair Smith, MS (University of Pittsburgh)*
Contributions and Review by:	*NIH-VA-DoD Pain Management Collaboratory Biostatistics/Design Workgroup*
Version:	*Version #1.0, last updated December 10, 2019*

Citation	Design or Results	Assumed level of clustering	Unit of measurement	Primary outcome (Pain) and time point	Assumed ICC in power analysis	Observed ICC
Tsai, P., Chang, J.Y., Beck, C., Kuo, Y., Keefe, F.J., A pilot cluster-randomized trial of a 20-week tai chi program in elders with cognitive impairment and osteoarthritic knee: effects on pain and other health outcomes. *Journal of Pain and Symptom Management*, 2013. 45(4): p. 660-669. [21]	Results	Site	Participant	WOMAC pain, week 1, week 5, week 9, week 13, week 17, week 21	0.01	0.005
Jakobsen, M. D., Sundstrup, E., Brandt, M., Jay, K., Aagaard, P., Andersen, L. L. Effect of workplace-versus home-based physical exercise on musculoskeletal pain among healthcare workers: A cluster randomized controlled trial. *Scandinavian Journal of Work, Environment & Health.* 2015; 41(2): p. 153-163. [22]	Results	Hospital department	Participant	Pain intensity on 0-10 scale baseline, 10 weeks	Not reported	0.0457
Rasmussen, C.D.N., et al. Prevention of low back pain and its consequences among nurses' aides in elderly care: a stepped-wedge multi-faceted cluster-randomized controlled trial. *BMC Public Health.* 2013. 13(1): p. 1088. [23] Results in *Pain* 2015;156(0): 1786-94.	Design	Work team	Participant	Pain intensity on 0-10 scale baseline, 3 months	0.05	0.007
Vitiello, M.V., et al. Cognitive-behavioral treatment for comorbid insomnia and osteoarthritis pain in primary care: The lifestyles randomized controlled trial. *Journal of the American Geriatrics Society.* 2013. 61(6): p. 947-956. [24]	Results	Group	Patient	Pain severity as measured by the Graded Chronic Pain Scale baseline, 9 months	0.022	0.10

References

1. Adams, G., et al., *Patterns of intra-cluster correlation from primary care research to inform study design and analysis.* J Clin Epidemiol, 2004. **57**(8): p. 785-94.
2. Donner, A. and N. Klar, *Design and Analysis of Cluster Randomization Trials in Health Research.* International Journal of Epidemiology, 2001. **30**(2): p. 407-408.
3. Wu, S., C.M. Crespi, and W.K. Wong, *Comparison of methods for estimating the intraclass correlation coefficient for binary responses in cancer prevention cluster randomized trials.* Contemp Clin Trials, 2012. **33**(5): p. 869-80.
4. Moore, C.G., et al., *Recommendations for planning pilot studies in clinical and translational research.* Clin Transl Sci, 2011. **4**(5): p. 332-7.

Prepared by:	*Charity G. Patterson, PhD, MSPH (University of Pittsburgh), Scott S. Coggeshall, PhD (VA Puget Sound), Tom Greene, PhD (University of Utah)*
Acknowledgement:	*Clair Smith, MS (University of Pittsburgh)*
Contributions and Review by:	*NIH-VA-DoD Pain Management Collaboratory Biostatistics/Design Workgroup*
Version:	*Version #1.0, last updated December 10, 2019*

5. Eldridge, S.M., D. Ashby, and S. Kerry, *Sample size for cluster randomized trials: effect of coefficient of variation of cluster size and analysis method.* Int J Epidemiol, 2006. **35**(5): p. 1292-300.
6. Kong, S.-H., C.W. Ahn, and S.-H. Jung, *Sample Size Calculation for Dichotomous Outcomes in Cluster Randomization Trials with Varying Cluster Size.* Drug Information Journal, 2016. **37**(1): p. 109-114.
7. Bauer, D.J., S.K. Sterba, and D.D. Hallfors, *Evaluating Group-Based Interventions When Control Participants Are Ungrouped.* Multivariate Behav Res, 2008. **43**(2): p. 210-236.
8. Lee, K.J. and S.G. Thompson, *The use of random effects models to allow for clustering in individually randomized trials.* Clin Trials, 2005. **2**(2): p. 163-73.
9. Hedges, L.V. and M. Citkowicz, *Estimating effect size when there is clustering in one treatment group.* Behav Res Methods, 2015. **47**(4): p. 1295-1308.
10. Kada, A., et al., *Sample size estimation and re-estimation of cluster randomized controlled trials for real-time feedback, debriefing, and retraining system of cardiopulmonary resuscitation for out-of-hospital cardiac arrests.* Contemp Clin Trials Commun, 2019. **14**: p. 100316.
11. Arnup, S.J., et al., *Understanding the cluster randomised crossover design: a graphical illustraton of the components of variation and a sample size tutorial.* Trials, 2017. **18**(1): p. 381.
12. Hooper, R., et al., *Sample size calculation for stepped wedge and other longitudinal cluster randomised trials.* Stat Med, 2016. **35**(26): p. 4718-4728.
13. Gulliford, M.C., et al., *Intraclass correlation coefficient and outcome prevalence are associated in clustered binary data.* J Clin Epidemiol, 2005. **58**(3): p. 246-51.
14. Omar, R.Z. and S.G. Thompson, *Analysis of a cluster randomized trial with binary outcome data using a multi-level model.* Stat Med, 2000. **19**(19): p. 2675-88.
15. Albaladejo, C., et al., *The efficacy of a short education program and a short physiotherapy program for treating low back pain in primary care: a cluster randomized trial.* Spine (Phila Pa 1976), 2010. **35**(5): p. 483-96.
16. Becker, A., et al., *Effects of two guideline implementation strategies on patient outcomes in primary care: a cluster randomized controlled trial.* Spine (Phila Pa 1976), 2008. **33**(5): p. 473-80.
17. Delitto, A., et al., *Study protocol for targeted interventions to prevent chronic low back pain in high-risk patients: A multi-site pragmatic cluster randomized controlled trial (TARGET Trial).* Contemp Clin Trials, 2019. **82**: p. 66-76.
18. Lonsdale, C., et al., *Communication style and exercise compliance in physiotherapy (CONNECT): a cluster randomized controlled trial to test a theory-based intervention to increase chronic low back pain patients' adherence to physiotherapists' recommendations: study rationale, design, and methods.* BMC Musculoskelet Disord, 2012. **13**: p. 104.
19. Morone, N.E., et al., *A Mind-Body Program for Older Adults With Chronic Low Back Pain: A Randomized Clinical Trial.* JAMA Intern Med, 2016. **176**(3): p. 329-37.

Prepared by:	*Charity G. Patterson, PhD, MSPH (University of Pittsburgh), Scott S. Coggeshall, PhD (VA Puget Sound), Tom Greene, PhD (University of Utah)*
Acknowledgement:	*Clair Smith, MS (University of Pittsburgh)*
Contributions and Review by:	*NIH-VA-DoD Pain Management Collaboratory Biostatistics/Design Workgroup*
Version:	*Version #1.0, last updated December 10, 2019*

20. Schmidt, C.O., et al., *Assessing a risk tailored intervention to prevent disabling low back pain--protocol of a cluster randomized controlled trial.* BMC Musculoskelet Disord, 2010. **11**: p. 5.
21. Tsai, P.F., et al., *A pilot cluster-randomized trial of a 20-week Tai Chi program in elders with cognitive impairment and osteoarthritic knee: effects on pain and other health outcomes.* J Pain Symptom Manage, 2013. **45**(4): p. 660-9.
22. Jakobsen, M.D., et al., *Effect of Workplace- versus Home-Based Physical Exercise on Muscle Response to Sudden Trunk Perturbation among Healthcare Workers: A Cluster Randomized Controlled Trial.* Biomed Res Int, 2015. **2015**: p. 902896.
23. Rasmussen, C.D., et al., *Prevention of low back pain and its consequences among nurses' aides in elderly care: a stepped-wedge multi-faceted cluster-randomized controlled trial.* BMC Public Health, 2013. **13**: p. 1088.
24. Vitiello, M.V., et al., *Cognitive-behavioral treatment for comorbid insomnia and osteoarthritis pain in primary care: the lifestyles randomized controlled trial.* J Am Geriatr Soc, 2013. **61**(6): p. 947-56.

Prepared by:	*Charity G. Patterson, PhD, MSPH (University of Pittsburgh), Scott S. Coggeshall, PhD (VA Puget Sound), Tom Greene, PhD (University of Utah)*
Acknowledgement:	*Clair Smith, MS (University of Pittsburgh)*
Contributions and Review by:	*NIH-VA-DoD Pain Management Collaboratory Biostatistics/Design Workgroup*
Version:	*Version #1.0, last updated December 10, 2019*