HHS Public Access

Author manuscript

Clin Trials. Author manuscript; available in PMC 2022 August 12.

Published in final edited form as:

Clin Trials. 2022 August ; 19(4): 353–362. doi:10.1177/17407745211063423.

Influential Methods Reports for Group-Randomized Trials and Related Designs

David M Murray[1]

[1]Office of Disease Prevention, National Institutes of Health (North Bethesda, MD, USA)

Abstract

This paper identifies the most influential methods reports for group-randomized trials and related designs published through 2020. Many interventions are delivered to participants in real or virtual groups or in groups defined by a shared interventionist so that there is an expectation for positive correlation among observations taken on participants in the same group. These interventions are typically evaluated using a group- or cluster-randomized trial, an individually-randomized group treatment trial, or a stepped wedge group- or cluster-randomized trial. These trials face methodological issues beyond those encountered in the more familiar individually-randomized controlled trial. PubMed was searched to identify candidate methods reports; that search was supplemented by reports known to the author. Candidate reports were reviewed by the author to include only those focused on the designs of interest. Citation counts and the relative citation ratio, a new bibliometric tool developed at the National Institutes of Health, were used to identify influential reports. The relative citation ratio measures influence at the article level by comparing the citation rate of the reference article to the citation rates of the articles cited by other articles that also cite the reference article. 1043 reports were identified that were published through 2020. Fifty-five were deemed to be most influential based on their relative citation ratio or their citation count using criteria specific to each of the three designs, with 32 group-randomized trial reports, 7 individually-randomized group treatment trial reports, and 16 stepped wedge group-randomized trial reports. Many of the influential reports were early publications that drew attention to the issues that distinguish these designs from the more familiar individually-randomized controlled trial. Others were textbooks that covered a wide range of issues for these designs. Others were “first reports” on analytic methods appropriate for a specific type of data (e.g., binary data, ordinal data), for features commonly encountered in these studies (e.g., unequal cluster size, attrition), or for important variations in study design (e.g., repeated measures, cohort vs cross-section). Many presented methods for sample size calculations. Others described how these designs could be applied to a new area (e.g., dissemination and implementation research). Among the reports with the highest relative citation ratios were the CONSORT statements for each design. Collectively, the influential reports address topics of great interest to investigators who might consider using one of these designs and need guidance on selecting the most appropriate design for their research question and on the best methods for design, analysis, and sample size.

Corresponding Author: Dr. David M. Murray, Ph.D., 6705 Rockledge Drive, Room 733, MSC 7990, Bethesda, Maryland 20892, david.murray2@nih.gov, Phone: 301-827-5561, Fax: 301-480-7660.

Keywords

Group-randomized trial; cluster-randomized trial; individually-randomized group treatment trial; stepped wedge group-randomized trial; stepped wedge cluster-randomized trial

Introduction

Murray et al.[1] recently reviewed the development of methods for the design and analysis of group- or cluster-randomized trials, individually randomized group treatment trials and stepped wedge group- or cluster-randomized trials based on a review of reports published through 2018. This paper provides an update to that article, reviewing reports published through 2020. After describing the key features for these three types of randomized trials, the most influential reports for the development of these methods are identified and described.

Key features

Parallel group- or cluster-randomized trials

A group-randomized trial randomizes groups rather than individuals to study conditions and outcomes are measured for participants from each group.[1–7] In the parallel group-randomized trial considered here, there is no crossover of groups to a different study condition during the trial. The parallel group-randomized trial is the best comparative design available when there is a good reason for randomization of groups rather than individuals. The usual reasons are 1) concern for contamination across conditions if delivered within the same group or 2) the use of a group-based intervention.

The key feature of the group-randomized trial is the randomization of groups to study conditions. Outcomes on participants from the same group are expected to be positively correlated as a result of common exposures, shared experience, or participant interaction.[8] This correlation violates the assumption of independence of errors that underlies the familiar analytic methods for individually-randomized controlled trials.[1–7] This correlation is often measured by the intraclass correlation.

Murray et al.[9] recently characterized the design features for group-randomized trials involving cancer or cancer-related outcomes. Most group-randomized trials compared two study conditions using a pretest-posttest design. Some used a posttest-only design while others included multiple pretest and/or posttest measures. Most employed a cohort design observing the same participants from each group at each measurement occasion. Others employed a cross-sectional design observing different participants from each group at each measurement occasion. Still others included both a cohort design and a cross-sectional design in the same study. Most employed some form of restricted randomization, such as stratification, matching, or constrained randomization.

Despite the availability of many books and hundreds of papers on the design and analytic methods for group-randomized trials, reviews have regularly shown that a large proportion of published group-randomized trials fail to account for the intraclass correlation in either

sample size calculations or in data analysis.[9–20] This is problematic because undersized studies will have insufficient power for a valid analysis and because invalid analyses will have an inflated type 1 error rate.[9, 12–19, 21–37]

Stepped wedge group-randomized trials

Stepped wedge group-randomized trials have increased in popularity over the last 15 years and are now commonly used to evaluate interventions that test new approaches to delivery of care.[50] Stepped wedge group-randomized trials are more complex than either parallel group-randomized trials or individually-randomized group treatment trials and they also face greater risks for bias.[51, 52]

The key feature of the stepped wedge group-randomized trial is the crossover of each group from the control to the intervention condition in a random order and on a staggered schedule.[53] Observations from participants from the same group will be positively correlated as in parallel group-randomized trials; however, the impact of the intraclass correlation is reduced in the stepped wedge group-randomized trial because the groups are crossed with study conditions rather than being nested within study conditions. Unlike parallel group-randomized trials, the intervention effect is confounded with calendar time.[53, 54] Moreover, the effect of the intervention may vary depending on how much time has passed since the intervention was introduced;[54, 55] that is important in a stepped wedge group-randomized trial because they are often longer than a parallel group-randomized trial. Finally, the pattern of correlation over time can be complex because stepped wedge group-randomized trials involve repeated measurements on the same groups and sometimes on the same participants, often for a prolonged period.[56–58]

The most common design for a stepped wedge group-randomized trial is a complete design; here, data are collected when all groups are in the control condition, again in each group when one or more groups crosses over to the intervention condition, and usually again after all groups are in the intervention condition. The incomplete design is less common; here, data are not collected from all groups at all steps.[59] Stepped wedge group-randomized trials vary in the number of groups that cross over in each step, the number of steps, and in the time between steps. They may employ restricted randomization, as described above, to balance groups in the sets to be randomized to the steps. In the continuous recruitment short exposure design, participants are recruited continuously and exposed for only a short period; participants may be measured only once or repeatedly. In the closed cohort design, participants are identified at the beginning of the study, participate throughout the study, and are measured repeatedly. In the open cohort design many participants are identified at the beginning of the study, but some may leave while other participants are recruited over time; participants may be measured only once or at multiple occasions.

The literature on the design and analytic methods for stepped wedge group-randomized trials has developed rapidly over the last 15 years. Reviews have noted deficiencies in reporting of study design, sample size, analytic methods, and ethical conduct.[60–69]

Individually randomized group treatment trials

An individually-randomized group treatment trial differs from an individually-randomized controlled trial in that the method by which the intervention is delivered creates a level of intraclass correlation that otherwise would not exist.[34] Correlated observations can result if participants receive at least some of their intervention in a group format (e.g., attend the same weight-loss class), if participants share the same interventionist (e.g., have the same instructor, therapist, or surgeon), or if participants interact with one another in some other way that is related to the method in which the intervention is delivered (e.g., through a virtual chat room created for participants in the same study condition). The individually-randomized group treatment trial is the best comparative design available if randomization of individuals is possible but it is necessary or more efficient to deliver at least some of the intervention in a group format or through a shared interventionist.

The key feature of the individually-randomized group treatment trial is that the method by which the intervention is delivered generates some level of correlation among outcomes taken on groups of participants within the same study condition, creating the same type of intraclass correlation seen in parallel group-randomized trials. Investigators must account for the intraclass correlation in the sample size to avoid low power and in the analysis to avoid type I errors.[34, 38–42] If the method of intervention delivery creates multiple overlapping groups or if the group structure changes over time, the situation is even more complicated, further increasing the risk of an inflated type I error rate if the investigators do not account for the complex pattern of correlation.[43–46]

The methods literature for individually-randomized group treatment trials is much more limited than for the other designs considered here and the issues are not widely recognized. Reviews suggest that that most investigators who employ the individually-randomized group treatment trial design are not aware of it and do not use appropriate methods for sample size or analysis.[34, 46–49]

Influential methods reports for these designs

Methods

The primary objective of this paper was to identify the most influential reports in the development of methods for group-randomized trials and related designs through 2020. In the development of an earlier paper,[1] PubMed was searched to identify all methods papers related to these designs through 2018. Search terms used in the initial PubMed search were: (“Randomized Controlled Trials as Topic”[MeSH Terms] AND “Cluster analysis”[MeSH Terms] AND (“sample size”[MeSH Terms] OR “computer simulation”[MeSH Terms] OR “research design”[MeSH Terms] OR “Analysis of Variance”[MeSH Terms] OR “Epidemiologic Research Design”[MeSH Terms] OR “Random Allocation”[MeSH Terms]) AND (“1975/01/01”[PDAT] : “2018/12/31”[PDAT])) OR ((cluster*[Title] OR group*[Title] OR community[Title]) AND (random*[Title] OR RCT[Title]) AND (analysis*[Title] OR design[Title] OR method[Title] OR sampl* [Title]) AND (“1975/01/01”[PDAT] : “2018/12/31”[PDAT])). Additional searches were conducted for all papers published by each of the first authors identified in the initial search.

The results were augmented by articles, books, chapters, and reports known to the authors of the earlier report. PubMed was then searched for other papers by any of the authors of the identified reports, yielding a total of 4514 candidates. The author reviewed each report and any that were not focused on design or analytic methods for the designs of interest were excluded, leaving 924 reports; 797 focused on parallel group-randomized trials, 49 on individually-randomized group treatment trials, 74 on stepped wedge group-randomized trials, and 4 addressed all three designs.

Using the same methods, another 119 reports were identified through 2020, bringing the total to 1043 reports through 2020. That total included 873 reports focused on group-randomized trials, 55 focused on individually-randomized group treatment trials, 105 focused on stepped wedge group-randomized trials, and 10 that addressed two or more of these designs.

Citation counts and the relative citation ratio were used to assess the influence of each report, accessing data on April 27, 2021. The relative citation ratio uses relative citation rates to measure influence at the article level, standardized across areas of science defined by the relative citation network (the articles cited by other articles that also cite the reference article)[70]. In brief, the citation rate of the reference article is compared to the citation rates of the relative citation network. Relative citation ratios change over time as the relative citation network grows, particularly in the first few years after a report is published; as a result, the relative citation ratio for some of the reports included in the earlier paper had changed by the time the data were accessed for this paper two years later.

The relative citation ratio was available for 898 reports; for the remaining 145 reports, the Web of Science, Scopus, and Google Scholar were used to obtain citation counts. group-randomized trial reports (N=32) and reports addressing two or more of the study designs (N=0) with an RCR≥7.98 (99th percentile for all PubMed entries) or without an RCR but with ≥200 citations were retained. Individually-randomized group treatment trial reports (N=7) with an RCR≥3.45 (95th percentile) or without an RCR but with ≥100 citations were retained. Stepped wedge group-randomized trial reports (N=16) with an RCR≥4.91 (97.5th percentile) or without an RCR but with >150 citations were retained. The citation count thresholds generally discriminated between the reports that fell above and below the relative citation ratio thresholds; the sliding scale reflected the number of reports for each design with many more for group-randomized trials and many fewer for individually-randomized group treatment trials. This provided 55 methods reports related to these designs that were deemed most influential (Tables 1–3). These reports do not necessarily represent the current state of the science; recent summaries of the state of the science are available elsewhere.[1, 71–73]

The next three sections identify the influential reports for the three designs. The earliest reports are mentioned first in each section and the remaining reports are grouped by theme.

Parallel group-randomized trials

In 1978, Cornfield published the first methods paper for trials involving group randomization in the biomedical literature. He identified two penalties associated with group

randomization: extra variation attributable to the group and limited degrees of freedom for the test of the intervention effect.[74] Both penalties must be addressed in studies that randomize groups rather than individuals.

Several influential papers addressed general issues for parallel group-randomized trials. A 1997 commentary drew attention to the methodological issues inherent in these trials.[75] A 1997 paper presented features for the optimal design of a parallel group-randomized trial.[76] A 2003 paper reported on the use of parallel group-randomized trials for evaluating the effectiveness of change and improvement strategies.[77] A 2013 paper reported on methods for process evaluation.[78] A 2018 paper reported on the use of parallel group-randomized trials in dissemination and implementation research.[79]

Others presented methods for sample size calculations based either on the intraclass correlation[80–84] or on methods based on the coefficient of variation.[85] Eldridge et al. and Rutterford et al. addressed the question of unequal sample size at the group or cluster level.[86, 87] Rutterford et al, also addressed the effect of attrition, non-compliance, adjustment for baseline covariates, and repeated measures on sample size estimation.[87]

Others focused on specific analytic methods. Those included methods for random coefficient or growth curve models,[88] binary data,[89] ordinal data,[90, 91] Poisson regression,[92] meta-analysis,[93, 94] mediation analysis,[95, 96] and survival analysis.[97]

The first textbook on the design and analysis of parallel group-randomized trials was published in 1998[2] followed in 2000 by the second.[3] Three subsequent textbooks also met the criteria to be judged an influential report.[4, 7, 98]

The first CONSORT statement on parallel group-randomized trials was published in 2004[99] and provided a checklist to identify the methodological information to include in trial reports. An update was published in 2012.[100]

Murray published a review of methodological issues in parallel group-randomized trials in 2004,[101] summarizing work on both design and analytic methods.

Stepped wedge group-randomized trials

Though the concept of the stepped wedge design was introduced in 1987,[106] Hussey and Hughes published the first analytic methods for that design in 2007.[53] Copas et al.[59] later delineated three types of stepped wedge group-randomized trials and discussed the number and length of the steps, incomplete and complete designs, and randomization methods, including restricted randomization methods. Hemming et al. provided guidance on the rationale, design, analysis, and reporting for this design.[54] They noted that the stepped wedge group-randomized trial is particularly well-suited for evaluations of health service delivery interventions. Barker et al. reviewed the statistical methods used in stepped wedge group-randomized trials.[65] Hemming et al. summarized methods appropriate for stepped wedge group-randomized trials with repeated cross-sectional samples.[107]

Several papers addressed sample size methods. Woertman et al. noted that the sample size will depend on group size, the intraclass correlation, the number of steps, the number of

baseline measurements, and the number of measurements between steps;[108] a subsequent letter[109] corrected an important error in that paper which was accepted by the authors.[110] Baio et al. presented simulation methods for sample size estimation.[111] Hemming et al. considered power for several different stepped wedge group-randomized trial designs, including incomplete cross-sectional designs, designs with multiple levels of nesting, and complete designs.[112] Girling and Hemming proposed an algorithm to optimize the design of stepped wedge group-randomized trials and showed that for large studies, the best design may be a hybrid of parallel group-randomized trial and stepped wedge group-randomized trial design components.[113] Hemming and Taljaard compared power for parallel group-randomized trials and stepped wedge group-randomized trials when the number of groups is fixed and reported that the parallel group-randomized trial tends to be more efficient when the intraclass correlation is small and that the stepped wedge group-randomized trial tends to be more efficient when the intraclass correlation is large, dependent on group size.[114] Hooper et al. presented formulaic methods for sample size estimation based on the intraclass correlation and the cluster and individual autocorrelations.[115] Kasza et al. reported on the effect of decaying correlation over time on sample size and power.[56]

Several state-of-the-practice reviews of stepped wedge group-randomized trials have been published,[60–62] reporting wide variation in data analytic methods and reporting standards. The recent CONSORT statement for this design[116] was written in part to try to reduce that variation.

Individually randomized group treatment trials

Several early reports addressed the risk of type 1 error in studies in which individually-randomized participants receive their intervention in a group format or from a shared interventionist. These papers appeared in the biomedical,[102] psychological,[48, 103] and educational literature.[104]

Roberts and Roberts presented the first report to address the analytic challenges specific to individually-randomized group treatment trials in 2005.[38] Consistent with guidance for parallel group-randomized trials, they noted that mixed models specifying the groups as levels of a random effect provided an appropriate analysis.

In 2008, Boutron et al. extended the CONSORT statement to non-pharmacologic interventions which include many individually-randomized group treatment trials.[40] They pointed to the need to provide details on how the intervention was delivered (e.g., individually, in groups, via a common interventionist), and to address the implications for analysis of having correlated observations within one or more study conditions. Boutron et al. published an update to that CONSORT statement in 2017.[105]

Summary

This paper identifies the 55 most influential reports contributing to the methods for the evaluation of group- or cluster-randomized trials, individually randomized group treatment trials, and stepped wedge group- or cluster-randomized trials through 2020, adding two years of data to an earlier paper.[1] There was substantial overlap with the 50 reports listed

in the earlier paper, but there was also turnover in the list of influential reports, based on changes to the relative citation ratios and citation counts as often happens with the passage of time. There were 6 new parallel group-randomized trial reports, 1 new individually-randomized group treatment trial report, and 5 new stepped wedge group-randomized trial reports included here. Of the original 50 reports, 3 individually-randomized group treatment trial reports, 2 stepped wedge group-randomized trial reports, and 2 reports that addressed more than two designs no longer met the inclusion criteria and so were not included in this report. Reports that had been included previously but failed to meet the inclusion criteria for this update generally had low but qualifying relative citation ratios for the original report and those values declined with the additional follow-up time. Relative citation ratios are generally stable after a few years and most of the reports for which the ratio changed were relatively new in the original report.

The relative citation ratio and citation counts were used to gauge the influence of a report on the work published later. Influence does not equate to quality, so there is no claim that the most influential reports are also the highest quality reports.

Many of the influential reports were early publications that drew attention to the issues that distinguish these designs from the more familiar individually-randomized controlled trial. Others were textbook treatments that covered a wide range of issues for these designs. Others were “first reports” on analytic methods appropriate for a specific type of data (e.g., binary data, ordinal data), for features commonly encountered in these studies (e.g., unequal cluster size, attrition), or for important variations in study design (e.g., repeated measures, cohort vs cross-sectional). Many presented methods for sample size calculations. Others described how these designs could be applied to a new area (e.g., dissemination and implementation research). Among the most influential reports were CONSORT statements which provide guidance for how to present the methods and results from a study based on its design. Collectively, they address topics of great interest to investigators who might consider conducting a group- or cluster-randomized trial, an individually randomized group-treatment trial, or a stepped-wedge group or cluster-randomized trial and need information to guide their planning for design, analysis, and sample size. As a set, they make an excellent reading list for anyone interested in learning about the methods used in these designs and their appropriate applications in public health and medicine.

References

1. Murray DM, Taljaard M, Turner EL, et al. Essential Ingredients and Innovations in the Design and Analysis of Group-Randomized Trials. Annu Rev Public Health 2020; 41: 1–19. [PubMed: 31869281]
2. Murray DM. Design and Analysis of Group-Randomized Trials. New York, NY: Oxford University Press, 1998, p.467.
3. Donner A and Klar N. Design and Analysis of Cluster Randomization Trials in Health Research. London: Arnold, 2000, p.178.
4. Hayes RJ and Moulton LH. Cluster Randomised Trials. Boca Raton, FL: CRC Press, 2009.
5. Eldridge S and Kerry S. A Practical Guide to Cluster Randomised Trials in Health Services Research. London: Arnold, 2012.
6. Campbell MJ and Walters SJ. How to Design, Analyse and Report Cluster Randomised Trials in Medicine and Health Related Research. Chichester: John Wiley & Sons Ltd., 2014.

7. Hayes RJ and Moulton LH. Cluster Randomised Trials. 2nd ed. Boca Raton, FL: CRC Press, 2017.
8. Kish L Survey Sampling. New York, NY: John Wiley & Sons, 1965, p.643.
9. Murray DM, Pals SL, George SM, et al. Design and analysis of group-randomized trials in cancer: A review of current practices. Prev Med 2018; 111: 241–247. [PubMed: 29551717]
10. Brown AW, Li P, Bohan Brown MM, et al. Best (but oft-forgotten) practices: designing, analyzing, and reporting cluster randomized controlled trials. Am J Clin Nutr 2015; 102: 241–248. [PubMed: 26016864]
11. Crespi CM, Maxwell AE and Wu S. Cluster randomized trials of cancer screening interventions: are appropriate statistical methods being used? Contemp Clin Trials 2011; 32: 477–484. [PubMed: 21382513]
12. Diaz-Ordaz K, Froud R, Sheehan B, et al. A systematic review of cluster randomised trials in residential facilities for older people suggests how to improve quality. BMC Med Res Methodol 2013; 13: 127. [PubMed: 24148859]
13. Donner A, Brown KS and Brasher P. A methodological review of non-therapeutic intervention trials employing cluster randomization, 1979–1989. Int J Epidemiol 1990; 19: 795–800. [PubMed: 2084005]
14. Eldridge S, Ashby D, Bennett C, et al. Internal and external validity of cluster randomised trials: systematic review of recent trials. BMJ 2008; 336: 876–880. [PubMed: 18364360]
15. Ivers NM, Taljaard M, Dixon S, et al. Impact of CONSORT extension for cluster randomised trials on quality of reporting and study methodology: review of random sample of 300 trials, 2000–8. BMJ 2011; 343: d5886. [PubMed: 21948873]
16. Murray DM, Pals SL, Blitstein JL, et al. Design and analysis of group-randomized trials in cancer: a review of current practices. J Natl Cancer Inst 2008; 100: 483–491. [PubMed: 18364501]
17. Rutterford C, Taljaard M, Dixon S, et al. Reporting and methodological quality of sample size calculations in cluster randomized trials could be improved: a review. J Clin Epidemiol 2015; 68: 716–723. [PubMed: 25523375]
18. Simpson JM, Klar N and Donner A. Accounting for cluster randomization: a review of primary prevention trials, 1990 through 1993. Am J Public Health 1995; 85: 1378–1383. [PubMed: 7573621]
19. Varnell SP, Murray DM, Janega JB, et al. Design and analysis of group-randomized trials: a review of recent practices. Am J Public Health 2004; 94: 393–399. [PubMed: 14998802]
20. Cook DJ, Rutherford WB, Scales DC, et al. Rationale, Methodological Quality, and Reporting of Cluster-Randomized Controlled Trials in Critical Care Medicine: A Systematic Review. Crit Care Med 2021; 49: 977–987. [PubMed: 33591020]
21. Puffer S, Torgerson D and Watson J. Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medical journals. BMJ 2003; 327: 785–789. [PubMed: 14525877]
22. Eldridge SM, Ashby D, Feder GS, et al. Lessons for cluster randomized trials in the twenty-first century: a systematic review of trials in primary care. Clin Trials 2004; 1: 80–90. [PubMed: 16281464]
23. Bowater RJ, Abdelmalik SM and Lilford RJ. The methodological quality of cluster randomised controlled trials for managing tropical parasitic disease: a review of trials published from 1998 to 2007. Transactions of the Royal Society of Tropical Medicine and Hygiene 2009; 103: 429–436. [PubMed: 19232658]
24. Froud R, Eldridge S, Diaz Ordaz K, et al. Quality of cluster randomized controlled trials in oral health: a systematic review of reports published between 2005 and 2009. Community dentistry and oral epidemiology 2012; 40 Suppl 1: 3–14. [PubMed: 22369703]
25. Ivers NM, Halperin IJ, Barnsley J, et al. Allocation techniques for balance at baseline in cluster randomized trials: a methodological review. Trials 2012; 13: 120. [PubMed: 22853820]
26. Diaz-Ordaz K, Kenward MG, Cohen A, et al. Are missing data adequately handled in cluster randomised trials? A systematic review and guidelines. Clin Trials 2014; 11: 590–600. [PubMed: 24902924]

27. Mdege ND, Brabyn S, Hewitt C, et al. The 2 × 2 cluster randomized controlled factorial trial design is mainly used for efficiency and to explore intervention interactions: a systematic review. J Clin Epidemiol 2014; 67: 1083–1092. [PubMed: 25063555]
28. Wright N, Ivers N, Eldridge S, et al. A review of the use of covariates in cluster randomized trials uncovers marked discrepancies between guidance and practice. J Clin Epidemiol 2015; 68: 603–609. [PubMed: 25648791]
29. Fiero MH, Huang S, Oren E, et al. Statistical analysis and handling of missing data in cluster randomized trials: a systematic review. Trials 2016; 17: 72. [PubMed: 26862034]
30. Richardson M, Garner P and Donegan S. Cluster Randomised Trials in Cochrane Reviews: Evaluation of Methodological and Reporting Practice. PLoS One 2016; 11: e0151818. [PubMed: 26982697]
31. Huang S, Fiero MH and Bell ML. Generalized estimating equations in cluster randomized trials with a small number of clusters: Review of practice and simulation study. Clin Trials 2016; 13: 445–449. [PubMed: 27094487]
32. Agbla SC and DiazOrdaz K. Reporting non-adherence in cluster randomised trials: A systematic review. Clin Trials 2018; 15: 294–304. [PubMed: 29608096]
33. Heo M, Nair SR, Wylie-Rosett J, et al. Trial Characteristics and Appropriateness of Statistical Methods Applied for Design and Analysis of Randomized School-Based Studies Addressing Weight-Related Issues: A Literature Review. Journal of obesity 2018; 2018: 8767315. [PubMed: 30046468]
34. Pals SL, Murray DM, Alfano CM, et al. Individually randomized group treatment trials: a critical appraisal of frequently used design and analytic approaches. Am J Public Health 2008; 98: 1418–1424. [PubMed: 18556603]
35. El Alili M, van Dongen JM, Goldfeld KS, et al. Taking the Analysis of Trial-Based Economic Evaluations to the Next Level: The Importance of Accounting for Clustering. Pharmacoeconomics 2020; 38: 1247–1261. [PubMed: 32729091]
36. Golzarri-Arroyo L, Oakes JM, Brown AW, et al. Incorrect Analyses of Cluster-Randomized Trials that Do Not Take Clustering and Nesting into Account Likely Lead to p-Values that Are Too Small. Child Obes 2020; 16: 65–66. [PubMed: 32078372]
37. O'Hara LM, Blanco N, Leekha S, et al. Design, implementation, and analysis considerations for cluster-randomized trials in infection control and hospital epidemiology: A systematic review. Infect Control Hosp Epidemiol 2019; 40: 686–692. [PubMed: 31043183]
38. Roberts C and Roberts SA. Design and analysis of clinical trials with clustering effects due to treatment. Clin Trials 2005; 2: 152–162. [PubMed: 16279137]
39. Baldwin SA, Bauer DJ, Stice E, et al. Evaluating models for partially clustered designs. Psychol Methods 2011; 16: 149–165. [PubMed: 21517179]
40. Boutron I, Moher D, Altman DG, et al. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. Ann Intern Med 2008; 148: 295–309. [PubMed: 18283207]
41. Walwyn R and Roberts C. Therapist variation within randomised trials of psychotherapy: implications for precision, internal and external validity. Stat Methods Med Res 2010; 19: 291–315. [PubMed: 19608603]
42. Kahan BC and Morris TP. Assessing potential sources of clustering in individually randomised trials. BMC Med Res Methodol 2013; 13: 58. [PubMed: 23590245]
43. Roberts C and Walwyn R. Design and analysis of non-pharmacological treatment trials with multiple therapists per patient. Stat Med 2013; 32: 81–98. [PubMed: 22865729]
44. Andridge RR, Shoben AB, Muller KE, et al. Analytic methods for individually randomized group treatment trials and group-randomized trials when subjects belong to multiple groups. Stat Med 2014; 33: 2178–2190. [PubMed: 24399701]
45. Bauer DJ, Gottfredson NC, Dean D, et al. Analyzing repeated measures data on individuals nested within groups: accounting for dynamic group effects. Psychol Methods 2013; 18: 1–14. [PubMed: 23148474]

46. Conroy EJ, Rosala-Hallas A, Blazeby JM, et al. Randomized trials involving surgery did not routinely report considerations of learning and clustering effects. J Clin Epidemiol 2019; 107: 27–35. [PubMed: 30439544]
47. Pals SL, Wiegand RE and Murray DM. Ignoring the group in group-level HIV/AIDS intervention trials: a review of reported design and analytic methods. Aids 2011; 25: 989–996. [PubMed: 21487252]
48. Baldwin SA, Murray DM and Shadish WR. Empirically supported treatments or type I errors? Problems with the analysis of data from group-administered treatments. J Consult Clin Psychol 2005; 73: 924–935. [PubMed: 16287392]
49. Lee KJ and Thompson SG. Clustering by health professional in individually randomised trials. BMJ 2005; 330: 142–144. [PubMed: 15649931]
50. Hemming K, Taljaard M and Grimshaw J. Introducing the new CONSORT extension for stepped-wedge cluster randomised trials. Trials 2019; 20: 68. [PubMed: 30658677]
51. Thompson JA, Fielding KL, Davey C, et al. Bias and inference from misspecified mixed-effect models in stepped wedge trial analysis. Stat Med 2017; 36: 3670–3682. [PubMed: 28556355]
52. Hemming K and Taljaard M. Reflection on modern methods: when is a stepped-wedge cluster randomized trial a good study design choice? Int J Epidemiol 2020; 49: 1043–1052. [PubMed: 32386407]
53. Hussey MA and Hughes JP. Design and analysis of stepped wedge cluster randomized trials. Contemp Clin Trials 2007; 28: 182–191. [PubMed: 16829207]
54. Hemming K, Haines TP, Chilton PJ, et al. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. BMJ 2015; 350: h391. [PubMed: 25662947]
55. Hughes JP, Granston TS and Heagerty PJ. Current issues in the design and analysis of stepped wedge trials. Contemp Clin Trials 2015; 45: 55–60. [PubMed: 26247569]
56. Kasza J, Hemming K, Hooper R, et al. Impact of non-uniform correlation structure on sample size and power in multiple-period cluster randomised trials. Stat Methods Med Res 2019; 28: 703–716. [PubMed: 29027505]
57. Li F. Design and analysis considerations for cohort stepped wedge cluster randomized trials with a decay correlation structure. Stat Med 2020; 39: 438–455. [PubMed: 31797438]
58. Kasza J and Forbes AB. Inference for the treatment effect in multiple-period cluster randomised trials when random effect correlation structure is misspecified. Stat Methods Med Res 2019; 28: 3112–3122. [PubMed: 30189794]
59. Copas AJ, Lewis JJ, Thompson JA, et al. Designing a stepped wedge trial: three main designs, carry-over effects and randomisation approaches. Trials 2015; 16: 352. [PubMed: 26279154]
60. Brown CA and Lilford RJ. The stepped wedge trial design: a systematic review. BMC Med Res Methodol 2006; 6: 54. [PubMed: 17092344]
61. Mdege ND, Man MS, Taylor Nee Brown CA, et al. Systematic review of stepped wedge cluster randomized trials shows that design is particularly used to evaluate interventions during routine implementation. J Clin Epidemiol 2011; 64: 936–948. [PubMed: 21411284]
62. Beard E, Lewis JJ, Copas A, et al. Stepped wedge randomised controlled trials: systematic review of studies published between 2010 and 2014. Trials 2015; 16: 353. [PubMed: 26278881]
63. Davey C, Hargreaves J, Thompson JA, et al. Analysis and reporting of stepped wedge randomised controlled trials: synthesis and critical appraisal of published studies, 2010 to 2014. Trials 2015; 16: 358. [PubMed: 26278667]
64. Martin J, Taljaard M, Girling A, et al. Systematic review finds major deficiencies in sample size methodology and reporting for stepped-wedge cluster randomised trials. BMJ Open 2016; 6: e010166.
65. Barker D, McElduff P, D'Este C, et al. Stepped wedge cluster randomised trials: a review of the statistical methodology used and available. BMC Med Res Methodol 2016; 16: 69. [PubMed: 27267471]
66. Grayling MJ, Wason JM and Mander AP. Stepped wedge cluster randomized controlled trial designs: a review of reporting quality and design features. Trials 2017; 18: 33. [PubMed: 28109321]

67. Taljaard M, Hemming K, Shah L, et al. Inadequacy of ethical conduct and reporting of stepped wedge cluster randomized trials: Results from a systematic review. Clin Trials 2017; 14: 333–341. [PubMed: 28393537]
68. Kristunas C, Morris T and Gray L. Unequal cluster sizes in stepped-wedge cluster randomised trials: a systematic review. BMJ Open 2017; 7: e017151.
69. Eichner FA, Groenwold RHH, Grobbee DE, et al. Systematic review showed that stepped-wedge cluster randomized trials often did not reach their planned sample size. J Clin Epidemiol 2019; 107: 89–100. [PubMed: 30458261]
70. Hutchins BI, Yuan X, Anderson JM, et al. Relative Citation Ratio (RCR): A New Metric That Uses Citation Rates to Measure Influence at the Article Level. PLoS Biol 2016; 14: e1002541. [PubMed: 27599104]
71. Turner EL, Li F, Gallis JA, et al. Review of Recent Methodological Developments in Group-Randomized Trials: Part 1-Design. Am J Public Health 2017; 107: 907–915. [PubMed: 28426295]
72. Turner EL, Prague M, Gallis JA, et al. Review of Recent Methodological Developments in Group-Randomized Trials: Part 2-Analysis. Am J Public Health 2017; 107: 1078–1086. [PubMed: 28520480]
73. Li F, Hughes JP, Hemming K, et al. Mixed-effects models for the design and analysis of stepped wedge cluster randomized trials: An overview. Stat Methods Med Res 2021; 30: 612–639. [PubMed: 32631142]
74. Cornfield J Randomization by group: a formal analysis. Am J Epidemiol 1978; 108: 100–102. [PubMed: 707470]
75. Bland JM and Kerry SM. Statistics notes. Trials randomised in clusters. BMJ 1997; 315: 600. [PubMed: 9302962]
76. Raudenbush SW. Statistical analysis and optimal design for cluster randomized trials. Psychological Methods 1997; 2: 173–185.
77. Eccles M, Grimshaw J, Campbell M, et al. Research designs for studies evaluating the effectiveness of change and improvement strategies. Quality & safety in health care 2003; 12: 47–52. [PubMed: 12571345]
78. Grant A, Treweek S, Dreischulte T, et al. Process evaluations for cluster-randomised trials of complex interventions: a proposed framework for design and reporting. Trials 2013; 14: 15. [PubMed: 23311722]
79. Handley MA, Lyles CR, McCulloch C, et al. Selecting and Improving Quasi-Experimental Designs in Effectiveness and Implementation Research. Annu Rev Public Health 2018; 39: 5–25. [PubMed: 29328873]
80. Donner A, Birkett N and Buck C. Randomization by cluster: sample size requirements and analysis. Am J Epidemiol 1981; 114: 906–914. [PubMed: 7315838]
81. Murray DM and Hannan PJ. Planning for the appropriate analysis in school-based drug-use prevention studies. J Consult Clin Psychol 1990; 58: 458–468. [PubMed: 2212183]
82. Gulliford MC, Ukoumunne OC and Chinn S. Components of variance and intraclass correlations for the design of community-based surveys and intervention studies: data from the Health Survey for England 1994. Am J Epidemiol 1999; 149: 876–883. [PubMed: 10221325]
83. Adams G, Gulliford MC, Ukoumunne OC, et al. Patterns of intra-cluster correlation from primary care research to inform study design and analysis. J Clin Epidemiol 2004; 57: 785–794. [PubMed: 15485730]
84. Hedges LV and Hedberg EC. Intraclass correlation values for planning group-randomized trials in education. Educ Eval Policy An 2007; 29: 60–87.
85. Hayes RJ and Bennett S. Simple sample size calculation for cluster-randomized trials. Int J Epidemiol 1999; 28: 319–326. [PubMed: 10342698]
86. Eldridge SM, Ashby D and Kerry S. Sample size for cluster randomized trials: effect of coefficient of variation of cluster size and analysis method. Int J Epidemiol 2006; 35: 1292–1300. [PubMed: 16943232]
87. Rutterford C, Copas A and Eldridge S. Methods for sample size determination in cluster randomized trials. Int J Epidemiol 2015; 44: 1051–1067. [PubMed: 26174515]

88. Feldman HA. Families of lines: random effects in linear regression analysis. J Appl Physiol 1988; 64: 1721–1732. [PubMed: 3379003]

89. Rao JN and Scott AJ. A simple method for the analysis of clustered binary data. Biometrics 1992; 48: 577–585. [PubMed: 1637980]

90. Hedeker D and Gibbons RD. A random-effects ordinal regression model for multilevel analysis. Biometrics 1994; 50: 933–944. [PubMed: 7787006]

91. Hedeker D and Gibbons RD. MIXOR: a computer program for mixed-effects ordinal regression analysis. Comput Methods Programs Biomed 1996; 49: 157–176. [PubMed: 8735023]

92. Zou GY and Donner A. Extension of the modified Poisson regression model to prospective studies with correlated binary data. Stat Methods Med Res 2013; 22: 661–670. [PubMed: 22072596]

93. Donner A and Klar N. Issues in the meta-analysis of cluster randomized trials. Stat Med 2002; 21: 2971–2980. [PubMed: 12325113]

94. Hedges LV. Effect sizes in cluster-randomized designs. J Educ Behav Stat 2007; 32: 341–370.

95. Krull JL and MacKinnon DP. Multilevel mediation modeling in group-based intervention studies. Eval Rev 1999; 23: 418–444. [PubMed: 10558394]

96. Emsley R, Dunn G and White IR. Mediation and moderation of treatment effects in randomised controlled trials of complex interventions. Stat Methods Med Res 2010; 19: 237–270. [PubMed: 19608601]

97. Austin PC. A Tutorial on Multilevel Survival Analysis: Methods, Models and Applications. Int Stat Rev 2017; 85: 185–203. [PubMed: 29307954]

98. Eldridge S and Kerry S. A practical guide to cluster randomized trials in health research. London: Arnold, 2012.

99. Campbell MK, Elbourne DR, Altman DG, et al. CONSORT statement: extension to cluster randomised trials. BMJ 2004; 328: 702–708. [PubMed: 15031246]

100. Campbell MK, Piaggio G, Elbourne DR, et al. CONSORT 2010 statement: extension to cluster randomised trials. BMJ 2012; 345: e5661. [PubMed: 22951546]

101. Murray DM, Varnell SP and Blitstein JL. Design and analysis of group-randomized trials: a review of recent methodological developments. Am J Public Health 2004; 94: 423–432. [PubMed: 14998806]

102. Whiting-O'Keefe QE, Henke C and Simborg DW. Choosing the correct unit of analysis in Medical Care experiments. Med Care 1984; 22: 1101–1114. [PubMed: 6513619]

103. Crits-Christoph P and Mintz J. Implications of therapist effects for the design and analysis of comparative studies of psychotherapies. J Consult Clin Psychol 1991; 59: 20–26. [PubMed: 2002139]

104. Nye B, Konstantopoulos S and Hedges L. How large are teacher effects? Educ Eval Policy An 2004; 26: 237–257.

105. Boutron I, Altman DG, Moher D, et al. CONSORT Statement for Randomized Trials of Nonpharmacologic Treatments: A 2017 Update and a CONSORT Extension for Nonpharmacologic Trial Abstracts. Ann Intern Med 2017; 167: 40–47. [PubMed: 28630973]

106. The Gambia Hepatitis Study Group. The Gambia Hepatitis Intervention Study. The Gambia Hepatitis Study Group. Cancer Res 1987; 47: 5782–5787. [PubMed: 2822233]

107. Hemming K, Taljaard M and Forbes A. Analysis of cluster randomised stepped wedge trials with repeated cross-sectional samples. Trials 2017; 18: 101. [PubMed: 28259174]

108. Woertman W, de Hoop E, Moerbeek M, et al. Stepped wedge designs could reduce the required sample size in cluster randomized trials. J Clin Epidemiol 2013; 66: 752–758. [PubMed: 23523551]

109. Hemming K and Girling A. The efficiency of stepped wedge vs. cluster randomized trials: stepped wedge studies do not always require a smaller sample size. J Clin Epidemiol 2013; 66: 1427–1428. [PubMed: 24035495]

110. de Hoop E, Woertman W and Teerenstra S. The stepped wedge cluster randomized trial always requires fewer clusters but not always fewer measurements, that is, participants than a parallel cluster randomized trial in a cross-sectional design. In reply. J Clin Epidemiol 2013; 66: 1428.

111. Baio G, Copas A, Ambler G, et al. Sample size calculation for a stepped wedge trial. Trials 2015; 16: 354. [PubMed: 26282553]

112. Hemming K, Lilford R and Girling AJ. Stepped-wedge cluster randomised controlled trials: a generic framework including parallel and multiple-level designs. Stat Med 2015; 34: 181–196. [PubMed: 25346484]

113. Girling AJ and Hemming K. Statistical efficiency and optimal design for stepped cluster studies under linear mixed effects models. Stat Med 2016; 35: 2149–2166. [PubMed: 26748662]

114. Hemming K and Taljaard M. Sample size calculations for stepped wedge and cluster randomised trials: a unified approach. J Clin Epidemiol 2016; 69: 137–146. [PubMed: 26344808]

115. Hooper R, Teerenstra S, de Hoop E, et al. Sample size calculation for stepped wedge and other longitudinal cluster randomised trials. Stat Med 2016; 35: 4718–4728. [PubMed: 27350420]

116. Hemming K, Taljaard M, McKenzie JE, et al. Reporting of stepped wedge cluster randomised trials: extension of the CONSORT 2010 statement with explanation and elaboration. BMJ 2018; 363: k1614. [PubMed: 30413417]

Table 1.

Influential reports for Group-Randomized Trial methods sorted by Relative Citation Ratio (RCR) and citation count with explanation for selecting items; RCR and count data were retrieved April 27, 2021.

Citation	RCR	Count	Explanation
Campbell MK, Piaggio G, Elbourne DR, et al. CONSORT 2010 statement: extension to cluster randomised trials. *BMJ* 2012; 345: e5661. 2012/09/07. DOI: 10.1136/bmj.e5661.	47.84	849	High RCR and citation count
Campbell MK, Elbourne DR, Altman DG, et al. CONSORT statement: extension to cluster randomised trials. *BMJ* 2004; 328: 702–708. 2004/03/20. DOI: 10.1136/bmj.328.7441.702.	37.90	1042	High RCR and citation count
Donner A, Birkett N and Buck C. Randomization by cluster. Sample size requirements and analysis. *Am J Epidemiol* 1981; 114: 906–914. 1981/12/01. DOI: 10.1093/oxfordjournals.aje.a113261.	23.15	411	High RCR and citation count
Bland JM and Kerry SM. Statistics notes. Trials randomised in clusters. *BMJ* 1997; 315: 600. 1997/09/26. DOI: 10.1136/bmj.315.7108.600.	20.98	397	High RCR and citation count
Gulliford MC, Ukoumunne OC and Chinn S. Components of variance and intraclass correlations for the design of community-based surveys and intervention studies: data from the Health Survey for England 1994. *Am J Epidemiol* 1999; 149: 876–883. 1999/04/30. DOI: 10.1093/oxfordjournals.aje.a009904.	19.89	427	High RCR and citation count
Hayes RJ and Bennett S. Simple sample size calculation for cluster-randomized trials. *Int J Epidemiol* 1999; 28: 319–326. 1999/05/26. DOI: 10.1093/ije/28.2.319.	18.16	512	High RCR and citation count
Rao JN and Scott AJ. A simple method for the analysis of clustered binary data. *Biometrics* 1992; 48: 577–585. 1992/06/01.	15.64	316	High RCR and citation count
Grant A, Treweek S, Dreischulte T, et al. Process evaluations for cluster-randomised trials of complex interventions: a proposed framework for design and reporting. *Trials* 2013; 14: 15. 2013/01/15. DOI: 10.1186/1745-6215-14-15.	14.51	219	High RCR and citation count
Donner A and Klar N. Issues in the meta-analysis of cluster randomized trials. *Stat Med* 2002; 21: 2971–2980. 2002/09/27. DOI: 10.1002/sim.1301.	14.22	352	High RCR and citation count
Zou GY and Donner A. Extension of the modified Poisson regression model to prospective studies with correlated binary data. *Stat Methods Med Res* 2013; 22: 661–670. 2011/11/11. DOI: 10.1177/0962280211427759.	13.73	253	High RCR and citation count
Krull JL and MacKinnon DP. Multilevel Modeling of Individual and Group Level Mediated Effects. *Multivariate Behav Res* 2001; 36: 249–277. 2001/04/01. DOI: 10.1207/S15327906MBR3602_06.	13.49	359	High RCR and citation count
Hedeker D and Gibbons RD. A random-effects ordinal regression model for multilevel analysis. *Biometrics* 1994; 50: 933–944. 1994/12/01.	12.99	221	High RCR and citation count
Murray DM, Varnell SP and Blitstein JL. Design and analysis of group-randomized trials: a review of recent methodological developments. *Am J Public Health* 2004; 94: 423–432. 2004/03/05. DOI: 10.2105/ajph.94.3.423.	12.45	341	High RCR and citation count
Eccles M, Grimshaw J, Campbell M, et al. Research designs for studies evaluating the effectiveness of change and improvement strategies. *Quality & safety in health care* 2003; 12: 47–52. 2003/02/07. DOI: 10.1136/qhc.12.1.47.	11.40	321	High RCR and citation count
Murray DM and Hannan PJ. Planning for the appropriate analysis in school-based drug-use prevention studies. *J Consult Clin Psychol* 1990; 58: 458–468. 1990/08/01.	11.33	133	High RCR
Adams G, Gulliford MC, Ukoumunne OC, et al. Patterns of intra-cluster correlation from primary care research to inform study design and analysis. *J Clin Epidemiol* 2004; 57: 785–794. 2004/10/16. DOI: 10.1016/j.jclinepi.2003.12.013.	11.05	303	High RCR and citation count
Eldridge SM, Ashby D and Kerry S. Sample size for cluster randomized trials: effect of coefficient of variation of cluster size and analysis method. *Int J Epidemiol* 2006; 35: 1292–1300. 2006/09/01. DOI: 10.1093/ije/dyl129.	**9.56**	**260**	**High RCR and citation count**
Hedeker D and Gibbons RD. MIXOR: a computer program for mixed-effects ordinal regression analysis. *Comput Methods Programs Biomed* 1996; 49: 157–176. 1996/03/01.	9.33	180	High RCR

Citation	RCR	Count	Explanation
Feldman HA. Families of lines: random effects in linear regression analysis. *J Appl Physiol* 1988; 64: 1721–1732. 1988/04/01. DOI: 10.1152/jappl.1988.64.4.1721.	9.09	138	High RCR
Handley MA, Lyles CR, McCulloch C, et al. Selecting and Improving Quasi-Experimental Designs in Effectiveness and Implementation Research. *Annu Rev Public Health* 2018; 39: 5–25. 2018/01/13. DOI: 10.1146/annurev-publhealth-040617-014128.	**9.01**	**49**	**High RCR**
Emsley R, Dunn G and White IR. Mediation and moderation of treatment effects in randomised controlled trials of complex interventions. *Stat Methods Med Res* 2010; 19: 237–270. 2009/07/18. DOI: 10.1177/0962280209105014.	8.46	190	High RCR
Rutterford C, Copas A and Eldridge S. Methods for sample size determination in cluster randomized trials. *Int J Epidemiol* 2015; 44: 1051–1067. 2015/07/16. DOI: 10.1093/ije/dyv113.	**8.41**	**105**	**High RCR**
Austin PC. A Tutorial on Multilevel Survival Analysis: Methods, Models and Applications. *Int Stat Rev* 2017; 85: 185–203. 2018/01/09. DOI: 10.1111/insr.12214.	**8.11**	**82**	**High RCR**
Donner A and Klar N. *Design and Analysis of Cluster Randomization Trials in Health Research*. London: Arnold, 2000, p. 178.		2120	High Citation Count
Murray DM. *Design and Analysis of Group-Randomized Trials*. New York, NY: Oxford University Press, 1998, p.467.		1832	High Citation Count
Hayes RJ and Moulton LH. *Cluster Randomised Trials*. Boca Raton, FL: CRC Press, 2009.		**1033**	**High Citation Count**
Hayes RJ and Moulton LH. *Cluster Randomised Trials*. 2nd ed. Boca Raton, FL: CRC Press, 2017.		1033	High Citation Count
Hedges LV and Hedberg EC. Intraclass correlation values for planning group-randomized trials in education. *Educ Eval Policy An* 2007; 29: 60–87. DOI: 10.3102/0162373707299706.		414	High Citation Count
Raudenbush SW. Statistical analysis and optimal design for cluster randomized trials. *Psychological Methods* 1997; 2: 173–185. DOI: 10.1037/1082–989x.2.2.173.		383	High Citation Count
Eldridge S and Kerry S. *A Practical Guide to Cluster Randomised Trials in Health Services Research*. London: Arnold, 2012.		295	High Citation Count
Cornfield J. 1978. Randomization by group: a formal analysis. *Am J Epidemiol* 108: 100–2		289	High citation count
Hedges LV. Effect sizes in cluster-randomized designs. *J Educ Behav Stat* 2007; 32: 341–370. DOI: 10.3102/1076998606298043.		**230**	**High Citation Count**

Reports in bold type were not included as influential reports in the earlier paper (1).

Table 2.

Influential reports for Individually Randomized Group Treatment trial methods sorted by Relative Citation Ratio (RCR) and citation count with explanation for selecting items; RCR and count data were retrieved April 27, 2021.

Citation	RCR	Count	Explanation
Boutron I, Moher D, Altman DG, et al. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann Intern Med* 2008; 148: 295–309. 2008/02/20. DOI: 10.7326/0003-4819-148-4-200802190-00008.	59.92	1360	High RCR and citation count
Boutron I, Altman DG, Moher D, et al. CONSORT Statement for Randomized Trials of Nonpharmacologic Treatments: A 2017 Update and a CONSORT Extension for Nonpharmacologic Trial Abstracts. *Ann Intern Med* 2017; 167: 40–47. 2017/06/21. DOI: 10.7326/M17-0046.	**41.80**	**321**	**High RCR and citation count**
Crits-Christoph P, Mintz J. 1991. Implications of therapist effects for the design and analysis of comparative studies of psychotherapies. *Journal of Consulting and Clinical Psychology* 59: 20–26	8.35	126	High RCR and citation count
Whiting-O'Keefe QE, Henke C, Simborg DW. 1984. Choosing the correct unit of analysis in medical care experiments. *Medical Care* 22: 1101–14	5.85	90	High RCR and citation count
Roberts C, Roberts SA. 2005. Design and analysis of clinical trials with clustering effects due to treatment. *Clin Trials* 2: 152–62	4.82	128	High RCR
Baldwin SA, Murray DM, Shadish WR. 2005. Empirically supported treatments or type I errors? Problems with the analysis of data from group-administered treatments. *J Consult Clin Psychol* 73: 924–35	3.57	88	High RCR
Nye B, Konstantopoulos S, Hedges L. 2004. How large are teacher effects? *Educational Evaluation and Policy Analysis* 26: 237–57		2465	High citation count

Reports in bold type were not included as influential reports in the earlier paper (1).

Table 3.

Influential reports for Stepped Wedge Group-Randomized Trial methods sorted by Relative Citation Ratio (RCR) and citation count with explanation for selecting items; RCR and count data were retrieved April 27, 2021.

Citation	RCR	Count	Explanation
Hemming K, Haines TP, Chilton PJ, et al. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *BMJ* 2015; 350: h391. 2015/02/11. DOI: 10.1136/bmj.h391.	31.13	416	High RCR and citation count
Hussey MA and Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials* 2007; 28: 182–191. 2006/07/11. DOI: 10.1016/j.cct.2006.05.007.	19.48	557	High RCR and citation count
Brown CA and Lilford RJ. The stepped wedge trial design: a systematic review. *BMC Med Res Methodol* 2006; 6: 54. 2006/11/10. DOI: 10.1186/1471-2288-6-54.	17.69	497	High RCR and citation count
Mdege ND, Man MS, Taylor Nee Brown CA, Torgerson DJ. 2011. Systematic review of stepped wedge cluster randomized trials shows that design is particularly used to evaluate interventions during routine implementation. *J Clin Epidemiol* 64: 936–48	12.51	252	High RCR and citation count
Kasza J, Hemming K, Hooper R, et al. Impact of non-uniform correlation structure on sample size and power in multiple-period cluster randomised trials. *Stat Methods Med Res* 2019; 28: 703–716. 2017/10/14. DOI: 10.1177/0962280217734981.	**11.42**	**35**	**High RCR**
Hemming K, Taljaard M, McKenzie JE, et al. Reporting of stepped wedge cluster randomised trials: extension of the CONSORT 2010 statement with explanation and elaboration. *Bmj* 2018; 363: k1614. 2018/11/11. DOI: 10.1136/bmj.k1614.	**10.90**	**67**	**High RCR**
Woertman W, de Hoop E, Moerbeek M, et al. Stepped wedge designs could reduce the required sample size in cluster randomized trials. *J Clin Epidemiol* 2013; 66: 752–758. 2013/03/26. DOI: 10.1016/j.jclinepi.2013.01.009.	9.22	148	High RCR
Hemming K, Lilford R, Girling AJ. 2015. Stepped-wedge cluster randomised controlled trials: a generic framework including parallel and multiple-level designs. *Stat Med* 34: 181–96	7.85	98	High RCR
Copas AJ, Lewis JJ, Thompson JA, Davey C, Baio G, Hargreaves JR. 2015. Designing a stepped wedge trial: three main designs, carry-over effects and randomisation approaches. *Trials* 16: 352	7.35	87	High RCR
Hemming K, Taljaard M. 2016. Sample size calculations for stepped wedge and cluster randomised trials: a unified approach. *J Clin Epidemiol* 69: 137–46	6.67	70	High RCR
Hooper R, Teerenstra S, de Hoop E, Eldridge S. 2016. Sample size calculation for stepped wedge and other longitudinal cluster randomised trials. *Stat Med* 35: 4718–28	**6.29**	**71**	**High RCR**
Beard E, Lewis JJ, Copas A, Davey C, Osrin D, et al. 2015. Stepped wedge randomised controlled trials: systematic review of studies published between 2010 and 2014. *Trials* 16: 353	5.86	68	High RCR
Baio G, Copas A, Ambler G, Hargreaves J, Beard E, Omar RZ. 2015. Sample size calculation for a stepped wedge trial. *Trials* 16: 354	5.70	71	High RCR
Hemming K, Taljaard M, Forbes A. 2017. Analysis of cluster randomised stepped wedge trials with repeated cross-sectional samples. *Trials* 18: 101	**5.37**	**45**	**High RCR**
Barker D, McElduff P, D'Este C, Campbell MJ. 2016. Stepped wedge cluster randomised trials: a review of the statistical methodology used and available. *BMC Med Res Methodol* 16: 69	**5.33**	**54**	**High RCR**
Girling AJ, Hemming K. 2016. Statistical efficiency and optimal design for stepped cluster studies under linear mixed effects models. *Stat Med* 35: 2149–66	4.93	51	High RCR

Reports in bold type were not included as influential reports in the earlier paper (1).