
Research

Precisely equal group size and allocation bias in nursing randomized controlled trials: a scientiometric study

Richard Gray ^{1,*}, Daniel Bressington ², Martin Jones ³, David Thompson ⁴

1. School of Nursing and Midwifery, La Trobe University, Melbourne; r.gray@latrobe.edu.au

2. College of Nursing and Midwifery, Charles Darwin University, Darwin, Australia; Daniel.bressington@cdu.edu.au

3. Academic Department of Rural Health, University of South Australia, Adelaide; martin.jones@unisa.edu.au

4. School of Nursing and Midwifery, Queens University, Belfast; david.thompson@qub.ac.uk

* Correspondence: r.gray@latrobe.edu.au; Tel.: +61 (0) 481248182

Abstract: The manipulation of participant allocation in randomized controlled trials to achieve equal groups sizes may introduce allocation bias potentially leading to larger treatment effect estimates. This study aimed to estimate the proportion of nursing trials that have precisely equal group sizes and examine if there was an association with trial outcome. Data were extracted from a sample of 148 randomized controlled trials published in nursing science journals in 2017. One hundred trials (68%) had precisely equal group sizes. Respectively, a positive outcome was reported in 70% and 58% of trials with equal/unequal groups. Trials from Asia were more likely to have equal group sizes than those from the rest of the world. Most trials reported a sample size calculation (n=105, 71%). In a third of trials (n=36, 34%), the number of participants recruited precisely matched the requirement of the sample size calculation; this was significantly more common in studies with equal group sizes. The high number of nursing trials with equal groups may suggest nurses conducting clinical trials are manipulating participant allocation to ensure equal group size increasing the risk of bias.

Keywords: Randomized Controlled Trial; Equal Group Size; Nursing; Allocation Bias; Effect Size

1. Introduction

Allocation bias is a particular form of bias that only relates to randomized controlled trials. Sedgwick (1) defines allocation bias as a systematic difference in how participants are assigned to treatment groups. If researchers know or can guess which group the participant will be allocated to, this information will influence how they approach potential subjects. For example, suppose the researcher knows that the next participant will be assigned to the experimental group. In that case, they may be more inclined to select patients they think will better respond to treatment (2). The impact of allocation bias on estimates of treatment effect can be considerable. Authors of meta-epidemiological studies report a 30-40% increase in effect sizes in trials with inadequately concealed allocation (3,4).

One way allocation bias is commonly introduced in randomized trials is by manipulating group allocation to achieve equal group sizes. Schulz and Grimes (3) have written about the "cringe worthiness" of the commonly held belief that a clinical trial should have groups of equal sizes. To achieve equal group sizes, researchers will often restrict their randomization sequence to a pre-specified number. Consequently, the more participants are randomized in a trial, the easier it becomes to guess group allocation. For example, in a trial of adherence therapy for patients with schizophrenia, an independent statistician prepared a computer-generated allocation sequence. One hundred and fourteen

(determined by the sample size calculation) sealed envelopes are prepared, each determining participants allocation to the experimental or control conditions. Two-thirds of the way through the trial, 75 patients have been successfully randomized, 47 to the control and 28 to the experimental groups. Consequently, the remaining patients will be allocated to the experimental condition. Forearmed with this information, the researcher may focus on selecting patients that they believe more likely to respond to the adherence intervention.

The unpredictability of group allocation is one of the central tenants of the randomized controlled trial. In simple randomization (akin to flipping a coin), equal group size can, of course, happen by chance but is unlikely. We note that block randomization procedures (5) – where participants are randomized within random sized blocks – are deliberately designed to ensure balanced groups, although the procedure is criticized for being predictable.

Previous studies have demonstrated that equal group sizes are common in RCTs reported in some specialist medical journals. For example, Schulz et al. (4) reported that of 206 trials published in obstetrics and gynecology journals, equal group sizes were observed in 54% of the studies that used simple randomization. In a review of 73 dermatological trials, 38 used simple randomization techniques, of which 22 (58%) were identical (6). That equal group sizes are seemingly common may provide evidence that researchers are manipulating group allocation introducing allocation bias into their study designs.

Authors have not examined geographical variations in rates of equal group sizes in previous studies. We are aware of geographical differences in other aspects of trial design and conduct (7), including in our own work on trial registration (8). It would be informative to explore differences in the prevalence of precisely equal group sizes by geographical region. This may provide insights into cultural expectations as to how clinical trials are conducted and reported.

As a group of journal editors, we handle many manuscripts reporting trials in our respective journals and have observed that precisely equal group sizes are seemingly common. However, no studies have specifically examined equal group size as a source of bias in nursing randomized trials.

The objective of this scientiometric study was therefore to estimate the prevalence of precisely equal group sizes in randomized controlled trials published in nursing science journals and determine if precisely equal group sizes were associated with trial outcome, method of sequence generation, allocation concealment, the country where the research was conducted and method of randomization.

2. Materials and Methods

We used a dataset of 151 randomized controlled trials previously used to study trial registration status (8). The dataset included all randomized controlled trials published in JCR (Journal Citation Report) nursing science journals in 2017. Trials were identified by manually searching the table of contents of 110 English language journals. Only trials where participants were randomized to at least two trial arms were included.

2.1 Data extraction

We extracted the following data from included trials that were entered into a spreadsheet and checked by at least two researchers:

- Number of trial arms, the total number of participants randomized, number of subjects allocated to each trial arm.
- The geographic region where the corresponding authors were based (initially coded Asia, Oceania, Europe, the Americas).
- Trial outcome (coded positive, negative, feasibility/pilot, equivalence, economic). Trials were coded as having a positive outcome if the study authors reported a significant (using the alpha value in the manuscript) between-group

difference in the primary outcome. If a primary outcome was not listed, we used the first outcome listed in the paper.

- CONSORT diagram reported.
- Method of randomization (coded simple, block, stratified, external randomization service [e.g., pharmacy, clinical trials unit], could not determine).
- The allocation sequence was random (coded yes if the authors used computer-generated random numbers, referred to a random number table or coin tossing were used)
- Allocation concealed (coded yes if a remote or centrally administered method to allocate or envelopes were used appropriately)
- Journal citation report impact factor for the publishing journal
- Sample size calculation (or justification of sample size in pilot/feasibility trials) and required participant numbers.

2.2 Analysis and recoding

Trials were coded as having a precisely equal group size if groups were precisely equal or as equal as possible in a sample with an odd number (3). The region where fieldwork was conducted was recoded into two groups (1. Asia, 2. rest of the world) because we anticipated small numbers of trials from some geographical regions. Finally, we also checked if the sample size calculation and the number of participants recruited matched. We calculated an X^2 statistic to test for associations.

3. Results

Data were extracted from 148 clinical trials (we could not access three trials) and are summarized in table 1.

Table 1, characteristics of included studies (n, %) or mean (standard deviation) where appropriate.

	Equal group size (n=100)	Unequal group size (n=48)	Total (n=148)
Trial outcome			
Positive	70 (70%)	28 (58%)	98 (66%)
Negative	25 (25%)	16 (33%)	41 (28%)
Feasibility/pilot	3 (3%)	3 (6%)	6 (4%)
Equivalence	1 (1%)	1 (2%)	2 (1%)
Economic evaluation	1 (1%)	0 (0%)	1 (1%)
Geographical region			
Asia	57 (57%)	18 (38%)	75 (51%)
Africa	2 (2%)	0 (0%)	2 (1%)
Europe	19 (19%)	12 (25%)	31 (21%)
The Americas	15 (15%)	13 (27%)	28 (19%)
Oceania	7 (7%)	5 (10%)	12 (8%)
Method of randomisation			
Simple	35 (35%)	16 (33%)	51 (34%)
Blocked	25 (25%)	7 (15%)	32 (22%)
Stratified	4 (4%)	2 (4%)	6 (4%)
Could not determine	32 (32%)	14 (29%)	46 (31%)

External randomisation service	4 (4%)	8 (17%)	12 (8%)
Allocation ratio 1:1	100 (100%)	47 (98%)	147 (99%)
Allocation sequence was random	50 (50%)	40 (83%)	90 (61%)
Allocation sequence was concealed	27 (27%)	13 (27%)	40 (27%)
Reported a sample size calculation or justification (pilot or feasibility trials)	73 (73%)	32 (67%)	105 (71%)
Sample size calculation and number recruited identical*	34/73 (47%)	2/32 (6%)	36/105 (34%)
Number of trial arms			
Two	89 (89%)	44 (92%)	133 (90%)
Three	8 (8%)	4 (8%)	12 (8%)
Four	2 (2%)	0 (0%)	2 (1%)
Six	1 (1%)	0 (0%)	1 (1%)
JCR Impact factor (mean, SD)	1.81 (0.97)	1.84 (0.82)	1.82 (.92)
Sample size (mean, SD)	138.36 (240.17)	143.21 (125.52)	139.93 (209.50)
CONSORT diagram reported	82 (82%)	36 (75%)	118 (80%)

* denominator is the number of trials where a sample size was reported

Most included studies were two-arm trials; the allocation ratio was 1:1 in all but one study. A CONSORT diagram was reported in 80% of included manuscripts. The mean sample size was 139.93, and most authors reported a sample size calculation (or justification in the case of pilot/feasibility trials). One hundred (68%) included studies met our criteria for precisely equal group sizes. Authors of a third of trials reported that they used simple randomization. We could not determine the method of randomization used – because there was insufficient information to make a judgement – in a further third of trials. External randomization services were rarely used. The allocation sequence was random in two-thirds of trials, and allocation concealed in a quarter of trials. A positive outcome was reported in two-thirds of studies. Half of the trials involved a corresponding author from a country in Asia.

The proportion of trials with a positive result was greater in studies with precisely equal compared to unequal groups, but this did not reach statistical significance ($X^2(1, n = 139) = 1.46, p = 0.23$). Trials from Asia were more likely to have precisely equal group sizes than the rest of the world ($X^2(1, n = 148) = 4.93, p = 0.03$). There was no association between the method of randomization and equal/unequal group size. The mean sample size did not significantly differ between groups.

There was a significant association between allocation concealment and equal/unequal group sizes ($X^2(1, n = 148) = 15.12, p < 0.01$), indicating that authors of trials where groups were unequal were more likely to report that the allocation sequence was concealed.

There was no association between reporting a sample size calculation and precisely equal/unequal group size. However, it was significantly more likely than trials with precisely equal group sizes also reported a sample size calculation that matched the number of patients recruited ($X^2(1, n = 105) = 16.06, p = 0.01$).

3.1 Post hoc observations

Post hoc, we noted that 48 (46%) of the 105 trials that reported a sample size calculation recruited more than the required number of participants, in some cases by a considerable margin.

3.2 Prevalence of round numbers

We noted, again post hoc, that of the 148 included trials 44 (30%) had sample sizes that were a round number (a number ending in zero digit). Of the 105 trials that reported sample size calculations, 48 (46%) were round numbers.

4. Discussion

This study aimed to estimate the prevalence of precisely equal group sizes in randomized controlled trials published in nursing science journals. In trials using simple randomization procedures, 35 of 51 studies (69%) had identical groups. The findings from our study suggest that precisely equal group sizes are somewhat more common in nursing trials than in obstetrics and gynaecology (54%) (4) or dermatology (58%) (6). It seems plausible that nurse trialists – wittingly or unwittingly – are manipulating group allocation to ensure precisely equal groups. The non-random manipulation of group allocation introduces bias to the study because researchers can predict which group participants have been allocated to and can use this information when allocating subjects or completing assessments. Likely, this means that effect sizes have been inflated. Compared to trials with unequal groups, studies with precisely equal group sizes were more likely to report positive outcomes; however, the association was not statistically significant. Appropriately powered studies are required to test if there is an association between precisely equal group size and a positive trial outcome.

The association between geographical region and precisely equal group size in trials using simple randomization is important. Issues with the integrity of research emerging from some Asian countries, notably China (Resnik & Zeng, 2010), have been reported previously; the homogeneity of groups sizes from Asian trials may provide evidence of deliberate manipulation of allocation to achieve a balanced number of participants in trial arms. We speculate that the high prevalence of precisely equal group sizes is a product of how researchers are trained in Asian countries. Further exploration – perhaps by interviewing nurse trialists – is necessary.

The current study shows that authors provided details of how the random sequence was generated and how allocation was concealed in 61% and 27% of trials, respectively, slightly lower than the proportions reported in a review of CONSORT compliance in 83 RCTs published in seven "leading" nursing journals (10). In leading medical journals, 80% and 48% of manuscripts – respectively – report sequence generation and allocation concealment (11). The poor reporting of crucial information about the randomization process means that trials will be considered to have a high risk of bias when included in systematic reviews. It may be that this poor reporting stems from authors' failure to check their reporting against CONSORT (12) guidelines and/or reviewers' and journal editors' failure to appraise submitted manuscripts appropriately.

We could not determine procedures used to randomize participants in almost a third of trials; this is concerning and requires comment. For example, Park et al. (13) stated, "A total of 120 participants were randomly assigned to the two experimental groups and the control group as 1:1:1 ratio." No other details as to how participants were allocated were provided. In another example, the authors provided a link to an external randomization service, but the link did not work, meaning we could not check how the authors randomized (14). Again, these observations suggest a lack of author, editor and reviewer diligence.

Several trials reported that they were randomized studies, but careful inspection of the procedures may suggest they were not. For example, Dönmez & Kapucu (15) stated: "The block randomization was computer-derived, allowing for orderly distribution into two groups and to reduce the risk of irregular distribution to accomplish randomization, the first patient was assigned to the intervention group and the next one to the control group." Similarly, Gürkan & Kömürçü (16) stated they were using the "systematic randomization method." The reported procedure involved "allocating names corresponding to uneven numbers were recruited into the intervention group and those corresponding to even numbers into the control group." In both examples, the authors state that randomization procedures were explicitly done to ensure equal group size. Describing trials that are not randomized as randomized may – at least in part – explain the large number of studies with precisely equal group sizes.

Sample size calculations are intended to ensure that researchers recruit the requisite number of participants to detect a difference between groups if a difference exists. We note the substantial number of studies that over recruited. For example, Zhang et al. (17) needed to recruit 96 participants, according to the sample size calculation, to show that the "hands and knee" position effectively reduced rates of episiotomy during labour. However, the authors recruited 1,400 participants to the trial because "... some hospitals were hypothesized to have enough cases to draw their own research conclusion..." [sic] (17). Over recruiting to a trial may be unethical because it places participants at unnecessary risk because they are given an experimental treatment unnecessarily. With this case, we decided that we needed to write – which we did on the 8th of October 2020 – to the editors of *Midwifery*, the journal that published this trial, expressing our concern

about this manuscript because, in our view, participants may have had unnecessary episiotomies (18). The authors of the trial and journal editors did publish a response to our letter (19).

Although not an aim of the study, it is necessary that we comment on the third of trials we included where the estimated sample size and number of participants recruited were identical. In our experience recruiting participants to clinical trials is complex, often the clinical team recruiting participants will ask multiple people at the same time if they are interested in taking part in the trial and often there will be variation as to when patients respond. Consequently, it is typical that the number of participants recruited to a trial approximates but doesn't match the sample size calculation. We also note that if the number of participants recruited and requirement of the sample size calculation was much more common in trials where groups were precisely equal. The most likely explanation is that researchers are manipulating multiple aspects of their study methodology (e.g., sample size calculation, participant allocation) to produce a trial that "looks" aesthetically pleasing. Again, this may be to do with how nurse researchers have been trained.

That 30% of sample sizes ended in a zero is probably an example of Benford's (or first digit) law (20), which states that the smallest number (in this case 0) appears as the leading digit about 30% of the time. If the numbers were uniformly distributed, one would expect each number to occur 10% of the time. Of the authors that reported a sample size calculation, half ended in a zero exceeding the first digit law, perhaps providing evidence of some researchers manipulation of sample size calculations.

4.1 Limitations

Deciding if a study was reporting a positive or negative outcome was problematic because authors – in our judgement – tended to put a positive spin on what were, in fact, negative trials. Presumably, this is because authors – quite correctly – perceive that trials with positive results are more likely to be published in better journals (21).

A further limitation is that our findings may not reflect current practice because we used a list of nursing trials published in 2017, the fieldwork for which was typically conducted in the mid-2010s. There may have been an improvement in the conduct of research over time – use of web-based randomization services, for example – that we have not captured.

We also note that trials conducted by nurses are not always published in nursing journals. Many nurses report what they perceive as their more important studies in medical journals considered more prestigious. Consequently, our observations cannot be generalized to all nursing trials rather trials published in nursing journals.

5. Conclusions

Unpredictability is a cornerstone of the randomized-controlled trial. Each time a patient is randomized, there should be an equal chance they are allocated to one of the trial arms. There is a preponderance of precisely equal group sizes in trials reported in nursing journals that suggest manipulation of group allocation by authors; inevitably, this

introduces a substantial risk of allocation bias and likely inflates the reported effect size. There is a pressing need to improve the conduct and reporting of randomization procedures in nursing trials, particularly as the results of these studies can directly influence care and treatment administered to patients.

Author Contributions: Conceptualization, RG, DB, MJ, DT.; methodology, RG, DB, MJ, DT; formal analysis, RG, DB; data curation, RG; writing—original draft preparation, RG; writing—review and editing, RG, DB, MJ, DT. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding

Institutional Review Board Statement: Not applicable

Data Availability Statement: A complete list of included studies can be accessed via this link <https://latrobe.figshare.com/s/70698023de6acf7a669b>.

Conflicts of Interest: The authors declare no conflict of interest

References

1. Sedgwick P. Selection bias versus allocation bias. *BMJ*. 2013 May 24;346(may24 4):f3345–f3345.
2. Nunan D, Heneghan C, Spencer EA. Catalogue of bias: allocation bias. *BMJ Evid-Based Med*. 2018 Feb;23(1):20–1.
3. Schulz KF, Grimes DA. Unequal group sizes in randomized trials: guarding against guessing. *The Lancet*. 2002 16th March;359(9310):966–70.
4. Schulz KF, Chalmers I, Grimes DA, Altman DG. Assessing the Quality of Randomization From Reports of Controlled Trials Published in Obstetrics and Gynecology Journals. *JAMA*. 1994 13th July;272(2):125–8.
5. Efid J. Blocked Randomization with Randomly Selected Block Sizes. *Int J Environ Res Public Health*. 2011 Jan;8(1):15–20.
6. Adetugbo K, Williams H. How Well Are Randomized Controlled Trials Reported in the Dermatology Literature? *Arch Dermatol*. 2000 1st March;136(3):381–5.
7. Maggio L, Dong T, Driessen E, Artino Jr. A. Factors associated with scientific misconduct and questionable research practices in health professions education. *Perspect Med Educ*. 2019 Apr 1;8(2):74–82.
8. Gray R, Gray G, Brown E. A review of prospective registration of trials published in nursing science journals in 2017. *J Adv Nurs*. 2019;75(12):3263–71.
9. RESNIK D, ZENG W. RESEARCH INTEGRITY IN CHINA: PROBLEMS AND PROSPECTS. *Dev World Bioeth*. 2010 Dec;10(3):164–71.
10. Jull A, Aye PS. Endorsement of the CONSORT guidelines, trial registration, and the quality of reporting randomized controlled trials in leading nursing journals: A cross-sectional analysis. *Int J Nurs Stud*. 2015 Jun;52(6):1071–9.
11. Mills EJ, Wu P, Gagnier J, Devereaux PJ. The quality of randomized trial reporting in leading medical journals since the revised CONSORT statement. *Contemp Clin Trials*. 2005 1st August;26(4):480–7.
12. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomized trials. *BMJ*. 2010 Mar 24;340:c869.
13. Park Y-J, Lee S-J, Shin N-M, Shin H, Jeon S, Lee J, et al. Application and Effect of Mobiletype-Bone Health Intervention in Korean Young Adult Women with Low Bone Mass: A Randomized Control Trial. *Asian Nurs Res*. 2017 1st March;11(1):56–64.

-
14. Düzkaça DS, Uysal G, Bozkurt G, Yakut T, Çitak A. Povidone-Iodine, 0.05% Chlorhexidine Gluconate, or Water for Periurethral Cleaning Before Indwelling Urinary Catheterization in a Pediatric Intensive Care: A Randomized Controlled Trial. *J Wound Ostomy Continence Nurs.* 2017;44(1):84–8.
 15. Dönmez AA, Kapucu S. The effectiveness of a clinical and home-based physical activity program and simple lymphatic drainage in the prevention of breast cancer-related lymphedema: A prospective randomized controlled study. *Eur J Oncol Nurs.* 2017 Dec;31:12–21.
 16. Gürkan ÖC, Kömürcü N. The effect of a peer education program on combating violence against women: A randomized controlled study. *Nurse Educ Today.* 2017 Oct;57:47–53.
 17. Zhang H, Huang S, Guo X, Zhao N, Lu Y, Chen M, et al. A randomized controlled trial in comparing maternal and neonatal outcomes between hands-and-knees delivery position and supine position in China. *Midwifery.* 2017 Jul;50:117–24.
 18. Gray R, Bressington D, Jones M, Thompson D. 'Editor's Response'. *Midwifery.* 2021 1st August;99:103029.
 19. Bick PD. Editor's Response. *Midwifery.* 2021 1st August;99:103037.
 20. Weir C, Murray G. Fraud in clinical trials. *Significance.* 2011;8(4):164–8.
 21. Littner Y, Mimouni FB, Dollberg S, Mandel D. Negative Results and Impact Factor: A Lesson From Neonatology. *Arch Pediatr Adolesc Med.* 2005 Nov 1;159(11):1036.