

This is a repository copy of *Clustering by Health Professionals in Individually Randomised Controlled Trials*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/154590/>

Version: Published Version

Article:

Waleed, Mohammad, Kazim, Isma, Abdul, Hmaid et al. (4 more authors) (2019) Clustering by Health Professionals in Individually Randomised Controlled Trials. *European Medical Journal*. pp. 53-61.

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial (CC BY-NC) licence. This licence allows you to remix, tweak, and build upon this work non-commercially, and any new works must also acknowledge the authors and be non-commercial. You don't have to license any derivative works on the same terms. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Clustering by Health Professionals in Individually Randomised Controlled Trials

Authors: *Mohammad Waleed,¹ Isma Kazmi,² Mishah Farooq,³ Abdul Hamid,⁴ Fazal Karam,⁵ Victoria Allgar,⁶ Kenneth Y.K. Wong⁷

1. Department of Cardiology, Leeds General Infirmary, Leeds, UK
 2. Department of Renal medicine, St James' University Hospital, Leeds, UK
 3. Bradford Royal Infirmary, Bradford, UK
 4. Department of Anaesthesia, Northern General Hospital, Sheffield, UK
 5. Department of Orthopedics, Saidu Group of Teaching Hospital, Saidu Sharif, Pakistan
 6. York Medical School, University of York, York, UK
 7. Department of Cardiology, Blackpool Teaching Hospitals NHS Foundation Trust (Blackpool Victoria Hospital), Blackpool, UK
- *Correspondence to waleed1280@yahoo.com

Disclosure: The authors have declared no conflicts of interest.

Acknowledgements: This research was started as a part of dissertation project while the authors were completing a master's degree in public health from the University of Sheffield. With the author's contributions, the dissertation was reviewed and amended for publication purposes as a review article. This work has not been presented previously in any scientific meeting. The authors confirm that no financial support or grants were received for this work.

Received: 16.07.18

Accepted: 16.01.19

Keywords: Adjustment for clustering, clustering, health professional, individual randomised controlled trials (iRCT), natural clustering.

Citation: EMJ. 2019;4[4]:53-61.

Abstract

Purpose: The aim of this study was to investigate the prevalence of clustering by health professionals in individually randomised controlled trials (iRCT), and its adjustment in both the sample size calculation estimates and the analysis of the data collected in iRCT (that is, trials that randomise individuals only). As a result, cluster randomised controlled trials will not be the part of this review study. Additionally, the authors aimed to discover the prevalence of the various forms of clustering in iRCT.

Methods: iRCT, in which the intervention was delivered by a health professional, were electronically searched in three medical journals. The dates searched were from 1st January 2000–31st August 2009. The retrieved trials were then screened to exclude those with complex designs and trials with more than two parallel arms. The selected trials were then fully reviewed for the presence of clustering effects and any corresponding adjustment. Data about the sample size calculation in the selected trials were also included. A basic form was generated for the purpose of data extraction from each of the selected trials.

Results: Of the 130 iRCT reviewed, clustering of outcomes was present in 127 (98%) trials. Only 61 trials (47%) had adjusted for the clustering effects in their design and analysis, while 53% of the trials had ignored the clustering effect, and hence no adjustment had been made in the trial design or analysis.

Regarding the various forms of clustering, clustering by centre in multicentre trials was found in 79 trials (60%), followed by natural clustering in 26 trials (20%), and clustering imposed by the design of the study in 23 trials (18%).

Conclusion: Potential clustering of outcomes exists in almost all iRCT; however, this review found that <50% of iRCT took clustering into account and adjusted the sample size calculation and statistical analysis of this data for clustering. Almost half of the reviewed iRCT ignored the clustering effect. As a result, inaccurate and nongeneralisable results could have been generated.

BACKGROUND

Researchers that individually randomised controlled trials (iRCT) assume that the observed outcomes of participants are independent. In practice, there are a number of situations in which there is some doubt about the validity of this assumption. One example is the correlation, or clustering, of the observed outcomes in participants treated by the same health professional.¹ The importance of this issue was emphasised by Lee and Thompson,¹ who assessed clustering in 42 iRCT and concluded that clustering of outcomes exists in almost all iRCT, but is usually ignored in the analysis, which leads to underestimates of uncertainty and overly extreme p values.¹ In another article, Lee and Thompson proposed random effect models to allow for such clustering and investigated their effect on estimation and interpretation of the treatment effect.²

INTRODUCTION

What Does the Term Clustering Mean?

The term clustering usually diverts the mind of the reader towards the cluster randomised controlled trials, wherein the groups, or clusters, of patients (rather than individuals) are randomised to a treatment either because of the nature of the treatment or to prevent contamination between treatment groups.³ However, in this review, clustering in iRCT will be analysed.

In iRCT, clustering means that the observation(s) about patients and observed outcomes in iRCT may be correlated due to differences in the behaviours of the health professionals actively delivering the intervention, sociodemographic differences between the patients, or the design of the study.¹ Observed outcomes clustering can also occur by single centres when participating in larger multicentre randomised controlled trials.²

Why is it Important to Consider the Clustering Effect?

Clustering of outcomes in randomised trials reduces the effective sample size, reducing the power of a trial to detect an intervention effect.⁴⁻⁶

Additionally, clustering also affects the generalisability of the results and conclusions.¹ The results obtained and conclusions drawn from a trial cannot be generalised to the whole population if the potential of clustering for outcomes exists in a trial. For example, in therapy trials the sample of therapists in the trial should be representative of those who are going to deliver the intervention in practice,⁷ otherwise, the results obtained cannot be generalised.

In What Forms Does Clustering Exist in Individually Randomised Controlled Trials?

Clustering may be imposed by the design of the trial; this inherent clustering as a result of trial design has been noted in a trial comparing a new one-stop clinic with a dedicated breast clinic for breast cancer screening.⁸ In another form, clustering can be natural rather than imposed either because of the sociodemographic differences between patients or because of the general practitioner's or the practice's influence on delivering the intervention, as observed in a trial comparing fusidic acid cream with placebo for the treatment of impetigo.⁹

Clustering can also appear by centre in a multicentre trial, which was seen in a study comparing the cytological surveillance with immediate referral for colposcopy in management of women with low-grade cervical abnormalities.¹⁰ Observations from the same centre were similar and therefore more correlated and clustered than those from different centres: this phenomenon defines the centre effect.¹¹

How Can Clustering Effects Be Adjusted?

In iRCT the clustering effects can be accounted for by anticipating them at the time of trial design and increasing the sample size accordingly.¹²

Adjustment for the clustering effects in iRCT can also be conducted during the statistical analysis of a trial by using various statistical models.⁴ When analysing data from a multicentre trial, the estimation of the main treatment effect must take into account the differences seen between each centre.¹³⁻¹⁵ This statistical method to limit the effect of clustering is widely accepted, but there is no real consensus on the statistical model to use.¹⁶⁻²⁰ However, the selected method depends on the application of the trial's conclusion.¹¹ If conclusions apply across the participating centres or if the centres cannot be considered as a random sample from a population, the analysis of data will involve a fixed effects regression model. On the contrary, if one wants to extend the results to all the centres that could be concerned by the experimental treatment, the analysis of data will involve a mixed effects model.¹¹

Time Trend of Clustering in Individually Randomised Controlled Trials

At present, there is no obvious trend of clustering prevalence and its accommodation in iRCT. One of the objectives of this review is to identify the time trend of clustering prevalence in iRCT and the according adjustments in such trials. The aim of this study is to conduct a systematic analysis of the iRCT for the potential effects of clustering by health professionals.

Objectives of this study

- To identify the prevalence of clustering in iRCT.
- To identify whether the researchers have allowed for clustering effects or ignored clustering effects in the selected iRCT.
- To explore the different ways used to accommodate for clustering effects in the selected iRCT.
- Analyse the time trend analysis of the presence of clustering in iRCT and any according adjustment.

Trials to be Included in This Study

Only iRCT conducted by healthcare professionals have been included in this study. The term healthcare professional encompasses doctors, nurses, physiotherapists, and acupuncturists. iRCT involving a pharmaceutical and/or health technology have not been included in the final selection of the studies that were reviewed. Likewise, cluster randomised controlled trials were excluded during the final selection of the articles.

Study Design

This study is a systematic review of iRCT published in three selected journals from 1st January 2000–31st August 2009. The journals were selected for this study due to their diverse impact factors (at the time of the study) and the ease of access to the fully published articles from the university portal of the University of Sheffield. The retrieved articles were fully reviewed for the presence of various forms of clustering, relevance of the assumptions made while calculating sample size, and adjustment made for the effect of clustering during statistical analyses.

Strategic Plan for the Search

Basic literature search

The electronic databases searched for relevant literature were Medline® via OVIDSP online, the university portal/OVID online, and Google scholar. The database was searched from 1950 to date. The keywords used for electronic searching were “clustering”, “clustering effects”, “statistical models”, “randomised controlled trials”, “randomized controlled trials”, “natural clustering”, “multicentre trial”, “adjustment for clustering”, “health professional”, “individual randomized trials”, “individual randomized trials”, “sample size”, “sample size calculation”, and “random effect model”.

The keyword “clustering” yielded 27,084 papers, while the terms “randomized controlled trials” and “randomized controlled trials” yielded 286,121 and 6,006 papers, respectively. The search was narrowed on the basis the of forms of clustering, models to allow for clustering, adjustment

methods, sample size, and sample size calculation. The search was further narrowed by only selecting iRCT and excluding cluster randomised trials. Finally, iRCT conducted by pharmacists and those with technological subjects, such as *Helicobacter pylori*, cardiac markers, and other tests were also excluded as these studies did not meet the inclusion criteria.

Individual journal search

A separate search strategy was used for accessing and scrutinising each of the three journals. The portal and search strategy used for each of the journals is briefly discussed below.

BMJ

The yearly archives of the online issue of BMJ were accessed via the HighWire Press Free line

using the university library electronic journal database. Within each year, the full reports of the articles regarding the primary care/general practice were accessed, and the retrieved articles were categorically arranged on the basis of their study design. The articles in the education and debate portion of the archives were not accessed.

The Lancet

The electronic search was made using ScienceDirect via the university library electronic journals database. The terms “randomized controlled trials” and “randomised controlled trials” were searched via ScienceDirect.

Articles were then selected or excluded by applying the limitations of the iRCT in the specified time period (1st January 2000–31st August 2009), which reduced the number of the articles to 82.

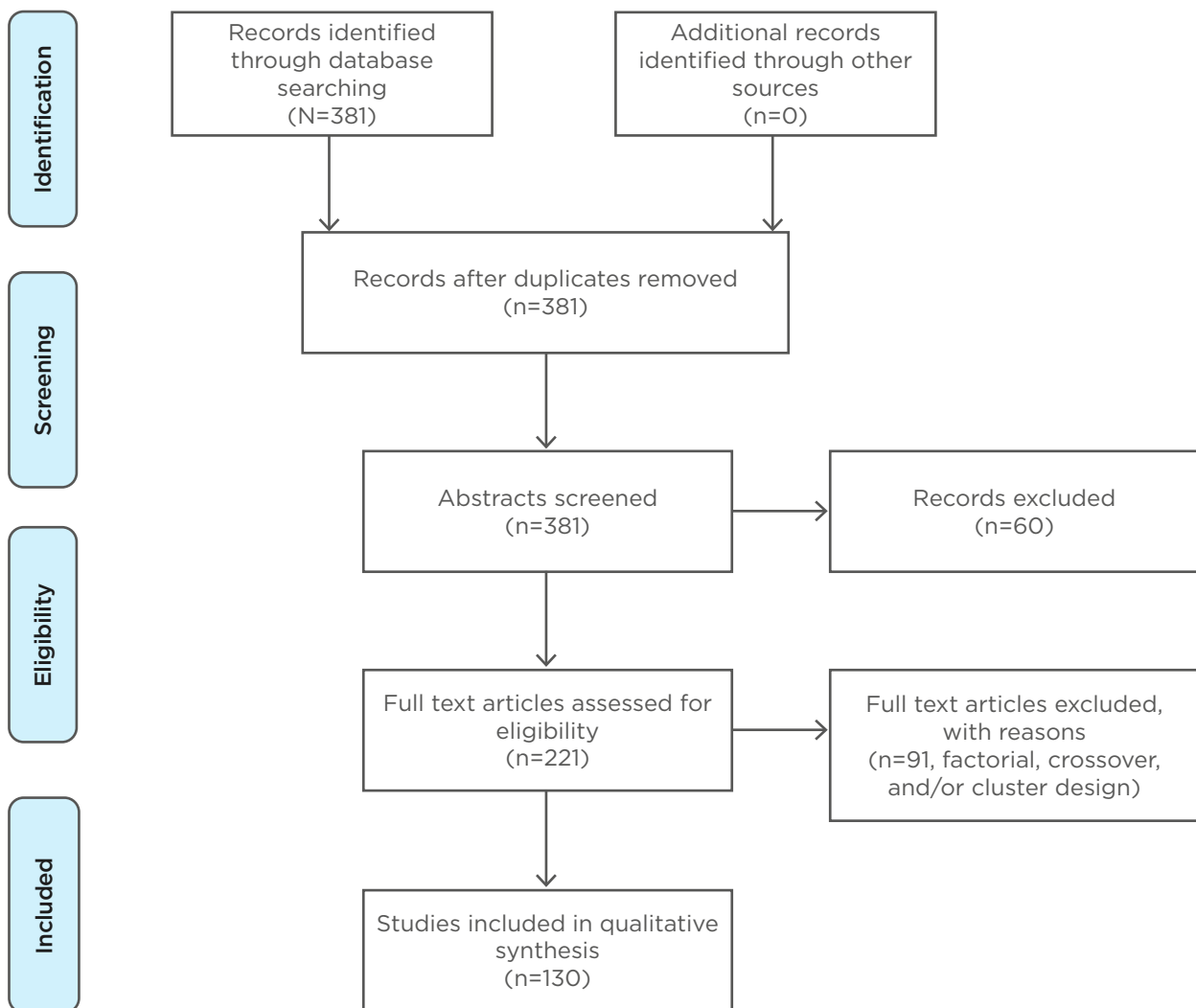


Figure 1: The identification, screening, and selection of the studies detailing individually randomised controlled trials included in this investigation.

Journal of Psychiatry and Neurosciences

PubMed Central, accessed via the electronic journal database of the university library portal, was used to search for JPN articles. The volumes of the JPN were scrutinised in detail from January 2000–July 2009. This search yielded 30 randomised controlled trials.

Screening of the titles and abstracts of the retrieved articles

The full literature search yielded 381 relevant articles. Two independent researchers screened the titles and abstracts of the retrieved articles, and, as a result, 160 articles were excluded.

Full text review of the retrieved articles

A review of the full text of the selected articles was conducted, and, as a result, 91 articles were excluded on the basis of full text review because they did not meet the inclusion criteria. Among these, trials with the factorial, crossover, and/or cluster design were excluded. Additionally, the follow-up studies were excluded from the retrieved articles (Figure 1).

Data Extraction

The iRCT selected for the final analysis were thoroughly reviewed for data extraction. A form was generated and used for the data extraction during the analysis of the selected iRCT. Data was extracted both from full text published iRCT and online extra material and registration websites.

Data extraction from the full text of the articles

The full text articles were studied and data regarding the general characteristics of the iRCT, sample size calculation, and statistical adjustments made for clustering were recorded.

General characteristics of the selected studies

The journal of publication, the year of publication, the type of intervention, number of multicentre trials, the health professionals conducting the trial, and the presence of clustering were recorded.

Sample size calculation

The methodology sections of the selected articles were studied in full detail, with particular focus placed on sample size calculation. All of the parameters used for the calculation of the sample size were collected. Any assumptions made and justification for the assumptions were also recorded. These parameters included Type I error, Type II error or power, one or two tailed tests, type of test, assumptions made in the control group, and predicted treatment effect.

Adjustment during statistical analysis

The results of the trials were studied and analysed for cluster effect adjustments, especially during statistical analysis of the results. Various models used for the statistical adjustment were also recorded.

Data extraction from the online extra material and trial registration websites

The trial registration websites were accessed and the target sample size and all the parameters used for the sample size calculation for the retrieved articles were recorded. Additionally, the extra materials related to the selected articles which were available online were also accessed and searched for the target sample size calculation and the parameters used for the sample size calculation.

RESULTS

Description of the 130 Included Articles

Table 1 describes the characteristics of the 130 selected articles. In 60 trials (46%), the intervention was delivered by the doctors, including general practitioners, physicians, psychiatrists, and surgeons. In the remaining included studies, the intervention was delivered by nurses in 43 trials (33%), physiotherapists in 16 trials (12%), and acupuncturists in the remaining 11 trials (9%). In half of the selected trials, the intervention was pharmacological, with a nonpharmacological intervention used in 45 (35%) trials. In the remaining 20 trials (15%), mixed method interventions were evaluated.

Table 1: Baseline characteristics of the articles and details of sample size calculation.

General characteristics of the 130 included individually randomised controlled trials			
Journal	Number (%)	Health professional	Number (%)
BMJ	75(58)	Doctors	60(46)
The Lancet	40(31)	Nurses	43(33)
Journal of Psychiatry and Neurosciences	15(11)	Physiotherapist	16(12)
		Acupuncturist	11(9)
Year of publication			
2000	13(10)	Intervention	
2001	10(8)	Pharmacological	65(50)
2002	7(5)	Nonpharmacological	45(35)
2003	10(8)	Both	20(15)
2004	8(6)		
2005	11(8)	Clustering	
2006	21(16)	Natural	26(20)
2007	18(14)	Imposed	23(18)
2008	19(15)	By centre in multicentre trials	79(60)
2009	13(10)	Not evident	2(2)
Details of sample size calculations for the 130 selected articles			
Parameter		Number (%)	
Articles not reporting a sample size calculation		16(12)	
Articles reporting sample size calculations		114(88)	
1) Reporting all required parameters		61(47)	
2) Reporting the power of the study		110(85)	
➤ 80%		66(60)	
➤ 85%		7(6)	
➤ 90%		27(26)	
➤ 95%		5(4)	
➤ Other values		5(4)	
3) Reporting the α risk		90(82)	
➤ 0.05		81(90)	
➤ 0.025		4(4)	
➤ Interim analysis		5(6)	
Articles not reporting the target sample size		9(7)	

Types and Prevalence of Clustering

Clustering by centre in multicentre trials was found in 79 trials (60%), followed by natural clustering, which was found in 26 trials (20%). Clustering imposed by the design of the study was noticed in 23 trials (18%), while the form of clustering was not clear in 2 trials (2%).

Allowance for Clustering in the 130 Selected Articles

Table 2 shows the frequency of the trials that corrected for the clustering effect in their study either by making adjustments while calculating the sample sizes or during the statistical analysis. Out of the 130 selected articles, 13 were published during the year 2000. Among these 13 trials, only 2 trials (15%) had adjusted for the clustering effect.

Table 2: Frequency of article allowing and/or ignoring for clustering effect.

Year of study	Number of articles studied	Total Number of iRCT	Number of articles with clustering	Articles with adjustment for clustering, n (%)	iRCT not allowing for clustering, n (%)
2000	39	13	12	2 (15)	11 (85)
2001	32	10	10	2 (20)	8 (80)
2002	20	7	7	2 (29)	5 (71)
2003	28	10	10	3 (33)	7 (67)
2004	24	8	8	3 (38)	5 (62)
2005	32	11	11	5 (45)	6 (55)
2006	63	21	20	12 (57)	9 (43)
2007	56	18	18	12 (67)	6 (33)
2008	48	19	19	12 (63)	7 (37)
2009	39	13	13	8 (62)	5 (38)
Total	381	130	128	61 (47)	69 (53)

iRCT: individually randomised controlled trials.

Likewise, among the 10 trials published during 2001, only 2 trials (20%) had made allowances for the clustering effect. In 2002, 7 articles out of the 130 selected articles were published, and among these 7, only 2 articles (29%) had taken the clustering effect into account. The percentage of trials that took various forms of clustering into account and had made adjustments in their study increased to 45% in 2005 and reached the maximum percentage recorded of 67% during the year 2007. During the year 2008, 12 trials out of 19 (63%) had taken the clustering effect into account. In 2009, of all the trials published up to the search end date, 31st August 2009, 8 trials (62%) showed proper adjustments for various forms of the clustering effect.

Ignoring the Clustering Effects in the 130 Selected Articles

Table 2 summarises the frequency of articles that ignored the clustering effect and made no allowance in their design and analysis. Overall, 69 articles (53%) made no adjustment for the clustering effect in their design and analysis.

Reporting sample size calculations

Table 1 summarises the data about sample size calculation; 16 articles (12%) did not report the

sample size calculation, while 9 trials (7%) did not mention the target sample size, neither in the full text report nor on the trial registration database. Even though a sample size calculation was reported by the majority of articles (88%), some of the required parameters for sample size calculation were frequently absent in reports.

In total, 61 articles (47%) included the required parameters for sample size calculations, including the assumptions made for the treatment effect and the control group. Moreover, 110 trials (85%) reported the power of the study. Finally, the α risk was mentioned by 90 trials (82%), which mostly mentioned the two-tailed test.

DISCUSSION

Principal findings

In this study of 130 iRCT published in three medical journals during the 10-year period from 1st January 2000–31st August 2009, the potential for clustering was found to be very common (98%) and only 47% (61/130) of trials studied in this time period have made allowances and adjustments for the clustering effect in their study design and analysis. The time trend analysis showed that the trend of taking clustering into account in iRCT has

increased over recent years, from 15% of studies implementing anticlustering measures in 2000 to 67% in 2007, and 63% in 2008. This trend still remains high (62%) for the selected trials published up to the 31st August 2009 in the three selected medical journals.

A plateau in the time trend analysis graph during the years 2008 and 2009 warrants further investigation into the importance of clustering adjustment to further increase the percentage of iRCT that adjust for clustering.

Sample size calculations were reported in 88% (114/130) of articles. Reporting of the sample size calculation has greatly increased in the past decades, from 4% of reports describing a calculation in 1980 to 83% of reports in 2002.^{21,22} However, some of the required parameters for replication of the sample size calculation are frequently absent in reports.

STRENGTHS OF THE STUDY

Familiarity with the data

Unlike secondary data analysis, in which the data is collected by others and a period of familiarisation is necessary, in this study, the data has been extracted and collected by the researcher, and the data set is self-generated so there is marked familiarity with the structure and contours of the data.

Presence of key variables

Secondary analysis entails the analysis of data collected by others for their own purposes, so one or more key variables may not be present.²³ In this study the data has been collected by the researcher, so the data about the key variables (potential clustering effect, sample size, accommodation for clustering) was collected with special attention.

Long study span and multiple medical journals reviewed

In this study, all the papers published in the three medical journals within the 10 years has been reviewed. Two of them are general medical journals with high impact factors: BMJ and The Lancet, while JPN is a specialist medical journal with a low impact factor.

Time trend analysis

It was difficult to have a long enough study span to analyse the change in trend with time on the potential clustering and its accommodation in the trials.

LIMITATIONS OF THE STUDY

It is difficult to assess whether the assumptions made in the iRCT during the sample size calculation to adjust for the clustering effect had been manipulated or not, as only the published data about the sample size has been used for the study. To obtain feasible sample sizes, the assumptions could be manipulated during the study.²⁴ Additionally, the sample size calculations can be manipulated after the completion of the study, as recently shown by Chan et al.²⁵ by comparing protocols to final articles.

The trials with complex designs, such as factorial designs, crossover trials, and trials with >2 parallel arms, were excluded during the screening process to obtain a homogeneous sample of articles. Therefore, the clustering effect and its adjustment in trials with more complex trials have not been assessed. This may limit the generalisability of the results.

Another limitation is that the treatment-by-centre interaction is not considered. However, as the main objective of a trial is often to assess the overall treatment effect, it is recommended to investigate the treatment effect using a model that only contains the centre effect.^{14,26}

CONCLUSION

The issue of clustering by health professionals in iRCT has gained attention in the last few years, and there is need for further research in this field to elicit some more facts about this matter and to provide further guidelines about the anticipation of, and accommodation for, potential clustering. A simulation study will be helpful to demonstrate the clustering affect on the results of an iRCT. Although each and every form of the clustering may not need to be accounted for at the analysis stage,²⁷ this paper highlights the existence of this issue to the readers and reviewers and the need for analysis adjustment in certain cases.

References

1. Lee K, Thompson SG. Clustering by health professional in individually randomized trials. *BMJ*. 2005;330:142-4.
2. Lee K, Thompson SG. The use of random effect models to allow for clustering in individually randomized trials. *Clinical Trial*. 2005;2(2):12.
3. Donner A, Klar N, Design and analysis of cluster randomization trials in health research (2000), London: Arnold.
4. Roberts C. The implications of variation in outcome between health professionals for the design and analysis of randomized control trials. *Stat Med*. 1999;18(19):2605-15.
5. Martindale C. The therapist as fixed effect fallacy in psychotherapy research. *J Consult Clin Psychol*. 1978;46:1526-30.
6. Greenfield S et al. Profiling care provided by different groups of physicians: Effects of patient case-mix (bias) and physician-level clustering on quality assessment results. *Ann Int Med*. 2002;136:1(2)11-21.
7. Crits-Christoph P, Mintz J. Implications of therapists effect for the design and analysis of comparative studies of psychotherapies. *J Consult Clin Psychol*. 1991;59(1):20-6.
8. Dey P et al. Costs and benefits of a one stop clinic compared with a dedicated breast clinic: randomized controlled trial. *BMJ*. 2002;324:507-11.
9. Koning S et al. Fusidic acid cream in the treatment of impetigo in general practice: Double blind randomized placebo controlled trial. *BMJ*. 2002;324:203-6.
10. TAMBOLA Group. Cytological surveillance compared with immediate referral for colposcopy in management of women with low grade cervical abnormalities: Multicentre randomised controlled trial. *BMJ*. 2009;339:b2546.
11. Vierron E, Giraudeau B. Sample size calculation for multicenter randomized trial: Taking the center effect into account. *Contemp Clin Trials*. 2007;28(4):451-8.
12. Hoover DR. Clinical trials of behavioral interventions with heterogeneous teaching subgroup effects. *Stat Med*. 2002;21:1351-64.
13. Anello C et al. Multicentre trials: A US regulatory perspective. *Stat Methods Med Res*. 2005;14(3):303-18.
14. European Medicines Agency. ICH Topic E 9. Note for guidance on statistical principles for clinical trials (CPMP/ICH/363/96) (1998). Available at: https://www.ema.europa.eu/documents/scientific-guideline/ich-e-9-statistical-principles-clinical-trials-step-5_en.pdf. Last accessed: 23 January 2019.
15. Localio AR et al. Adjustments for center in multicenter studies: An overview. *Ann Intern Med*. 2001;135(2):112-23.
16. Fedorov V et al. Estimation of the treatment difference in multicenter trials. *J Biopharm Stat*. 2004;14(4):1037-63.
17. Jones B et al. A comparison of various estimators of a treatment difference for a multi-centre clinical trial. *Stat Med*. 1998;17:1767-77 [discussion 1799-800].
18. Moerbeek M et al. A comparison between traditional methods and multilevel regression for the analysis of multicenter intervention studies. *J Clin Epidemiol*. 2003;56(4):341-50.
19. Permutt T. Probability models and computational models for ANOVA in multicenter clinical trials. *J Biopharm Stat*. 2003;13(3):495-505.
20. Senn S. Some controversies in planning and analysing multi-centre trials. *Stat Med*. 1998;17(15-16):1753-65 [discussion 1799-800].
21. Mills EJ et al. The quality of randomized trial reporting in leading medical journals since the revised CONSORT statement. *Contemp Clin Trials*. 2005;26(4):480-7.
22. Meinert CL et al. Content of reports on clinical trials: A critical review. *Control Clin Trials*. 1984;5(4 Suppl 1):328-47.
23. Bryman A, Burgess A, *Social Research Methods (Fourth Canadian Edition)*, 2008, Oxford: Oxford University Press.
24. Charles P et al. Reporting of sample size calculations in randomized controlled trials: Review. *BMJ*. 2009;338:b1732.
25. Chan AW et al. Discrepancies in sample size calculations and data analyses reported in randomised trials: Comparison of publications with protocols. *BMJ*. 2008;337:a2299.
26. Brown H, Prescott R, *Applied mixed models in medicine*, 3rd edition (1999), Chichester: John Wiley and Sons.
27. Kahan BC, Morris TP. Improper analysis of trials randomised using stratified blocks or minimisation. *Stat Med*. 2011;31(4):328-40.

FOR REPRINT QUERIES PLEASE CONTACT: +44 (0) 1245 334450