The Royal Society of Medicine

Journal of the Royal Society of Medicine; 2015, Vol. 108(5) 192–198
DOI: 10.1177/0141076815582303

A brief history of the cluster randomised trial design

Jenny Moberg[1] and Michael Kramer[2]
[1]Norwegian Knowledge Centre for the Health Services, N-0130 Oslo, Norway
[2]McGill University, Montreal, Quebec, Canada H3A 0G4
Corresponding author: Jenny Moberg. Email: jenny.moberg@kunnskapssenteret.no

Introduction

The cluster randomised trial is commonly considered a relatively new research study design.[1–3] Here we trace to a few very early reports the idea of comparing interventions applied to groups of individuals, and through the evolution of this idea to the modern-day cluster randomised trial. This has been defined as a comparative study in which the units randomised are pre-existing (natural or self-selected) groups whose members have an identifiable feature in common, and in which outcomes are measured in all, or a representative sample of the individual members of the groups.[1] Summaries of the reports of many of the examples of cluster randomised trials published before the methodological review of this research design by Donner *et al.*[4] can be viewed in the James Lind Library (http://jameslindlibrary.org/topics/allocation-bias/cluster-allocation/ see Appendix for details of our literature search).

What is a cluster?

The groups used in cluster randomised trials vary widely and range in size from families to entire communities. The common feature shared by members of a cluster may be:

- geographical, for example, villages;[5]
- communities;[6]
- administrative areas;[7]
- social, for example, families, households or religious congregations;[8–11]
- educational or occupational, for example, schools or school classes;[12–17]
- hospitals and worksites;[18–20] or
- professional, for example, all students taught by a specific teacher, or patients treated by a specific clinician.[21]

Why use cluster randomised trials?

Cluster randomised trials are well suited and are now commonly used to evaluate public health, health policy and health system interventions. They are ideal for testing interventions when the decision (policy) about whether or not to implement the intervention will be taken on behalf of a group. Cluster randomised trials are also useful when the nature of the intervention carries a high risk of contamination when individuals randomised to different comparison groups are in frequent contact with one another and thus may be influenced ('contaminated'), in either or both directions, by the alternative treatment(s). Contamination is likely to occur in comparisons of public health promotion interventions within the same community, and of different approaches to healthcare provided by the same clinician to patients under his or her care. In addition to these scientific reasons, cluster designs can also have practical advantages over individual randomisation because of lower implementation costs or administrative convenience.

Early examples of group allocation

The earliest mentions of which we are aware of treatment comparisons in which the intervention was assigned to a group, rather than to an individual, are centuries old. In 1648, Van Helmont proposed a trial of his new methods of treating febrile patients without purging and blood-letting, in which the participants would be put into groups then randomised by 'casting lots' to decide which group would receive which of the treatments to be compared.[22] It is unlikely that this trial ever took place, but the idea of cluster randomisation is there.

In 1657, Starkey proposed a trial in defence of van Helmont's treatment methods, in which groups of patients were to be assigned to be treated by Starkey (according to van Helmont's methods), or by van Helmont's critics. Starkey seems to have appreciated that the process of treatment allocation should be designed to prevent confounding by differences between the groups receiving different treatments. He suggested that patients first be divided into groups of 10. Starkey and his opponent should

then alternately divide each 10 into two groups of five, allowing those who did not do the dividing to choose one of the groups of five patients. The 'divider' should then treat the remaining five patients. As in van Helmont's proposed trial, the groups of patients did not exist prior to the trial but were created specifically for the trial.[23]

Celli's 1900 trial may be the first in which pre-existing groups were allocated to treatment – an important step towards the modern cluster randomised trial design.[9,24] Celli studied whether mosquito netting reduced malaria in households of Italian railway workers. The households were selected (although not randomised) to receive or not receive the intervention. Neighbouring households were used as controls. This trial heralds one of the most common uses of cluster randomised trials today: the evaluation of infectious disease control methods and, in particular, of methods to prevent malaria.

The clinical trial reported by Amberson and his colleagues in 1931, challenging the use of gold for treating pulmonary tuberculosis, was an early trial using randomisation by a single coin toss to allocate two matched comparison groups, either to injections of a gold-containing treatment (sanocrysin), or to control injections of distilled water. In addition, patients and the investigators measuring the trial outcomes were blinded to the participants' treatment allocation.[25] This trial has sometimes been considered well designed and conducted, but it falls short of the current standards for a cluster randomised trial; the clusters were created for the trial and only two clusters were randomised.[26]

Many early cluster randomised trials in non-medical fields were school-based evaluations of educational interventions. Indeed, methodological discussion of the cluster randomised design appears to have begun in 1940 with Lindquist's book on methods in education research in schools.[27,28] Much of what Lindquist wrote, however, also applies to clinical and public health interventions.

Recent developments

Only sparse use of cluster randomised trials was evident before the 1980s.[29] However, the last half century has seen a steady increase in the number of cluster randomised trials published in the medical literature: from one a year in the 1960s, to seven in 1990, when Donner, Brown and Brasher published their methodological review of cluster randomised trials;[4] to over 120 in 2008.

Every pre-1960s cluster randomised trial of which we are aware tested some aspect of infectious disease prevention or treatment.[4,30–32] In the 1970s, cluster randomised trials were used extensively for such trials, particularly in low-income countries.[33–35] Cluster randomised trials were also recognised as being suitable for evaluating public health interventions aiming to change health behaviour, such as improving dental care,[36] promoting hand-washing[37] and attending for immunisation.[38]

The risk of 'contamination' between comparison groups is high in studies evaluating screening interventions. In the 1980s two large-scale cluster randomised trials of screening interventions showed how this design can be used to reduce the influence of contamination on the effects of an intervention. A trial by Grant et al.[39] evaluated the effect of routine counting of fetal movement by pregnant women on the likelihood of antepartum stillbirth, and showed how the cluster randomised trial design can be useful for assessing the effects of interventions which would otherwise be compromised by a high risk of contamination.[39] In a Swedish screening mammography study published in 1985 by Tabár et al.[40] contamination among trial communities was reduced by offering screening to communities selected randomly from matched communities, separated by 200 km on average.

The use of cluster can make large scale trials, like trials of screening and other public health interventions, more practicable.[40] These features of cluster randomised trials – reduction of contamination, and practicability of very large-scale public health trials – are also well illustrated in a trial of the effect of vitamin A supplementation on childhood mortality, morbidity, and preschool growth.[41–43]

Other cluster randomised trials conducted around this time showed that the design is useful for evaluating the impact of multifaceted approaches to health improvement, for example, a trial of nutritional supplementation and maternal education in expectant mothers and infants at risk of malnutrition,[44] and a trial of breast cancer screening methods and the nurses who implemented them.[45]

Schools have often been used in public health cluster randomised trials. They are convenient places to implement health education interventions relevant to children and adolescents, such as prevention of tobacco, alcohol and drug use; promotion of sexual health; and primary prevention of chronic disease through promotion of healthy eating and physical activity. Entire schools or classes within schools are ready-made clusters.[46–50] School clusters have also been used to evaluate interventions aimed at helping children to become 'health messengers', as in a trial assessing whether hypertension education of children had an impact on the blood pressure of their parents.[51]

Cluster randomised trials have become recognised as being valuable in evaluating many different types of health system interventions – healthcare delivery,[52–54] governance, financial arrangements and implementation strategies. A cluster randomised trial reported by Vogt et al. in 1983 was used to compare methods to increase reporting of notifiable diseases by doctors.[55] This trial also illustrates an important group of cluster randomised trials in which clusters consist of patients treated by the same clinician. These trials are often called 'professional cluster randomised trials'. A clinician is randomly assigned to the intervention, and the intervention is targeted at individual clinicians – not at his or her patients. Such clinician-targeted interventions often aim to modify the behaviour of healthcare providers in some way, for example, by using clinical guidelines, training or decision support systems.[56–59]

In some such trials, clinicians implement the intervention without involving patients in the decision, for example, in reporting cases of a notifiable disease,[55] or arranging for medical assistants to screen for and manage patients with hypertension.[52] With other clinician-targeted interventions, the intervention is intended to impact on both clinician practices and on patient outcomes, for example, educational programmes to help physicians improve blood pressure control among their patients.[59]

Through the 1990s, the number of published trials including cluster randomisation increased,[29] and the terms 'group randomized' (Murray 1998), 'community randomized',[1,4] and even 'place randomized' were all used to describe cluster randomised trials. A *BMJ* series on statistics in 1997 and 1998 used the term 'cluster randomised'.[60,61] By the early 2000s, with the published extension of the CONSORT statement on reporting guidelines for cluster randomised trials[62] and several reviews of cluster randomised trials,[35,63,64] the term 'cluster randomised trial' had become the most commonly used term for this design. The publication of important, large-scale, well-conducted cluster randomised trials in this century, such as those evaluating the effects of community groups on birth and other outcomes in poor rural populations,[65–69] can be considered as a 'coming of age' of the cluster randomised trial design.

Current and future challenges

Study design

As experience with cluster randomised trials has increased over time, difficulties and problems have become apparent. The design (especially with respect to blinding), analysis and conduct of cluster randomised trials are often more complicated than for individually randomised trials. Cluster randomised trials are conducted with as few as one intervention and one control cluster, and insufficient numbers of clusters (inadequate sample sizes) are a persistent problem. We have not included cluster randomised trials with fewer than two clusters in each arm in the James Lind Library, or as examples in this article. Stratification and matching have been used to increase the comparability of the clusters, which helps increase precision with small numbers of clusters. While blinding of participants and outcome assessors is ideal in all randomised trials, it is often difficult or even impossible in cluster randomised trials. The units of randomisation and the units of observation may be different, and this affects informed consent, recruitment, sample sizes, randomisation and analysis.

Study analysis

Cluster randomised trials can be analysed in the same way as any individually randomised trial by each cluster providing one data item into the analysis, for example average blood pressure among all patients of a randomised physician. By using all the individual data points in each cluster in the analysis, the statistical power of a trial can be increased. However, the effect of clustering must be considered. As early as 1940, Lindquist recognised the need to account for clustering in the analysis of cluster randomised trials.[27,28] In 1978, Cornfield pointed out the need for special consideration of the statistical features of cluster randomised trials in health research and, in particular, the need to account for between-cluster variation.[70] And as Donner and Klar pointed out in 2000, analysis must also take into consideration variation in cluster size, which is often substantial.[1]

Members of clusters are more likely to have similar outcomes than a randomly selected sample of individuals from the same population, particularly when members self-select into a cluster. The most commonly used measure of the degree of similarity among members of a cluster is the intra-class correlation coefficient. The larger the intra-class correlation coefficient, the larger the number clusters of individuals needed to achieve comparable statistical power to a trial using individual randomisation. Analysing a cluster randomised trial without accounting for clustering yields a falsely low estimate of variance and hence inflates statistical significance. And as shown by Kramer et al.,[71] loss of statistical power is even more dramatic when the outcome measurements are also clustered with treatment ('double jeopardy').

A recent study re-analysing the results from cluster randomised trials of health system interventions using time series methods shows that, if data from cluster randomised trials are analysed without taking account of trends over time, the findings may be misleading.[72] Fretheim and colleagues suggest adding time series approaches to the overall comparison of randomised groups, so as to gauge changes in effect of the intervention over time.

Reporting

The quality of reports of cluster randomised trials has been very variable. Some studies are reported simply as 'randomised trials', leaving readers unaware that the unit of randomisation is anything other than the individual. Specific key words that would identify cluster randomised trials are often not provided in abstracts, so full-text publications have to be retrieved to establish whether or not the study reported was a cluster randomised trial. Despite extension of the CONSORT statement on reporting guidelines for cluster randomised trials,[62] the titles or abstracts of 50% of cluster randomised trials still fail to indicate this.[73] Conversely, other papers report to be cluster randomised trials in their titles but, on careful inspection, are clearly not cluster randomised trials.[29]

Summing up

Cluster randomised trials have a long history in both educational and health research and, over the last several decades, have assumed an increasing role in rigorous evaluations of complex clinical, public health and health system interventions in which individual randomisation is likely to be 'contaminated' by contact among individual participants randomised. Cluster randomised trials can also help overcome the administrative barriers and economic costs inherent in contacting, recruiting and randomising large numbers of individuals. Major challenges include ensuring statistical power by recruiting adequate numbers of clusters, possible use of randomised crossover of clusters,[74–76] minimising intra-cluster correlation of outcomes and measurements, and better reporting.

Declarations

Competing interests: None declared

Funding: None declared

Ethical approval: Not applicable

Guarantor: JM

Contributorship: JM wrote the first draft of this article. JM and MK contributed further text and revised the article.

Acknowledgements: We are grateful to Tikki Pang for providing WHO support to assist in the identification of reports of cluster randomised trials, to Marit Johansen for searching for them, and to Allan Donner and Atle Fretheim for helpful comments on an earlier draft.

Provenance: Invited contribution from the James Lind Library.

References

1. Donner A and Klar N. *Design and Analysis of Cluster Randomization Trials in Health Research*. London: Arnold, 2000.
2. Eldridge S and Kerry S. *A Practical Guide to Cluster Randomized Trials in Health Services Research*. Chichester: Wiley, 2012.
3. Murray DM. *Design and Analysis of Group-Randomized Trials*. Vol. 27, New York: Oxford University Press; Monographs in Epidemiology and Biostatistics, 1998.
4. 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.
5. Horwitz O and Magnus K. Epidemiologic evaluation of chemoprophylaxis against tuberculosis. *Am J Epidemiol* 1974; 99: 333–242.
6. Stanton BF and Clemens JD. An educational intervention for altering water-sanitation behaviors to reduce childhood diarrhea in urban Bangladesh. II. A randomized trial to assess the impact of the intervention on hygienic behaviors and rates of diarrhea. *Am J Epidemiol* 1987; 125: 292–301.
7. Job-Spira N, Meyer L, Bouvet E, Janaud A and Spira A. The prevention of sexually transmitted diseases which affect fertility: methodological problems and initial results. *Eur J Obstet Gynecol Reprod Biol* 1988; 27: 157–164.
8. Farr BM, Hendley JO, Kaiser DL and Gwaltney JM. Two randomized controlled trials of virucidal nasal tissues in the prevention of natural upper respiratory infections. *Am J Epidemiol* 1988; 128: 1162–1172.
9. Celli A. The new prophylaxis against malaria in Lazio. *Lancet* 1900; 156: 1603–1606.
10. Puska P, Nissinen A, Pietinen P and Iacono J. Role of dietary fat in blood pressure control. *Scand J Clin Lab Invest Suppl* 1985; 176: 62–69.
11. Gyorkos TW, Frappier-Davignon L, MacLean JD and Viens P. Effect of screening and treatment on imported intestinal parasite infections: results from a randomized, controlled trial. *Am J Epidemiol* 1989; 129: 753–761.
12. Flay BR, Ryan KB, Best JA, Brown KS, Kersell MW, d'Avernas JR, et al. Are social-psychological smoking prevention programs effective? The Waterloo study. *J Behav Med* 1985; 8: 37–59.
13. Walter HJ, Hofman A, Connelly PA, Barrett LT and Kost KL. Primary prevention of chronic disease in

childhood: changes in risk factors after one year of intervention. *Am J Epidemiol* 1985; 122: 722–781.
14. Connor MK, Smith LG, Fryer A, Erickson S, Fryer S and Drake J. Future Fit: a cardiovascular health education and fitness project in an after-school setting. *J School Health* 1986; 56: 329–333.
15. Biglan A, Severson H, Ary D, Faller C, Gallison C, Thompson R, et al. Do smoking prevention programs really work? Attrition and internal and external validity of an evaluation of a refusals skills training program. *J Behav Med* 1987; 10: 159–171.
16. Bush PJ, Zuckerman AE, Theiss PK, Taggart VS, Horowitz C, Sheridan MJ, et al. Cardiovascular risk factor prevention in black schoolchildren: two-year results of the ''Know Your Body'' program. *Am J Epidemiol* 1989; 129: 466–482.
17. Vartiainen E, Puska P and Tossavainen K. Prevention of non-communicable diseases: risk factors in youth. The North Karelia Youth Project (1984–88). *Health Promotion* 1986; 1: 269–283.
18. Kornitzer M and Rose G. WHO European Collaborative Trial of multifactorial prevention of coronary heart disease. *Prevent Med* 1985; 14: 272–278.
19. World Health Organisation European Collaborative Group. European collaborative trial of multifactorial prevention of coronary heart disease: final report on the 6-year results. *Lancet* 1986; 19;1: 869–872.
20. Mayer JA, Dubbert PM, Scott RR, Dawson BL, Ekstrand ML and Fondren TG. Breast self-examination: the effects of personalized prompts on practice frequency. *Behav Ther* 1987; 18: 135–146.
21. Wilson DM, Taylor DW, Gilbert JR, Best JA, Lindsay EA, Willms DG, et al. A randomized trial of a family physician intervention for smoking cessation. *JAMA* 1988; 260: 1570–1574.
22. Van Helmont JB. *Ortus medicinæ: Id est, initia physic inaudita. Progressus medicinæ novus, in morborum ultionem, ad vitam longam. [The Dawn of Medicine: That Is the Beginning of a New Physic: A New Advance in Medicine, a Victory over Disease, to (Promote) a Long Life]*. Amsterdam: Apud Ludovicum Elzevirium, 1648.
23. Starkey G. *Nature's Explication and Helmont's Vindication, or a Short and Sure Way to a Long and Sound Life*. London: E Cotes for Thomas Alsop at the two Sugar-loaves over against St Antholin's Church at the lower end of Watling Street, 1657.
24. Ferroni E, Jefferson T and Gachelin G. *Angelo Celli and Research on the Prevention of Malaria in Italy a Century Ago*. JLL Bulletin: Commentaries on the history of treatment evaluation (www.jameslindlibrary.org), 2011.
25. Amberson JB, McMahon BT and Pinner M. A clinical trial of sanocrysin in pulmonary tuberculosis. *Am Rev Tuberculosis* 1931; 24: 401–435.
26. Diaz M and Neuhauser D. Lessons from using randomization to assess gold treatment for tuberculosis. JLL Bulletin: Commentaries on the history of treatment evaluation (www.jameslindlibrary.org), 2004.
27. Lindquist EF. *Statistical Analysis in Educational Research*. Boston: Houghton Mifflin, 1940.
28. Klar N and Donner A. The impact of EF Lindquist's 1940 text ''Statistical Analysis in Educational Research'' on cluster randomization. JLL Bulletin: Commentaries on the history of treatment evaluation (www.jameslindlibrary.org), 2004.
29. Bland JM. Cluster randomized trials in the medical literature: two bibliometric surveys. *BMC Med Res Methodol* 2004; 4: 21.
30. Coburn AF. The prevention of respiratory tract bacterial infections. *JAMA* 1944; 126: 88–89.
31. Mellanby H, Andrewes CH, Dudgeon JA and Mackay DG. Vaccination against influenza A. *Lancet* 1948; 251: 978–982.
32. Comstock GW. Isoniazid prophylaxis in an undeveloped area. *Am Rev Respir Dis* 1962; 86: 810–822.
33. Storey J, Rossi-Espagnet A, Mandel SPH, Matsushima T, Lietaert P, Thomas D, et al. Sulfalene with pyrimethamine and chloroquine with pyrimethamine in single-dose treatment of Plasmodium falciparum infections. *Bull World Health Org* 1973; 49: 275–282.
34. Sutter EE and Ballard RC. Community participation in the control of trachoma in Gazankulu. *Soc Sci Med* 1983; 17: 1813–1817.
35. Isaakidis P and Ioannidis JP. Evaluation of cluster randomized trials in Sub-Saharan Africa. *Am J Epidemiol* 2003; 158: 921–926.
36. Reiss ML, Piotrowski WD and Bailey JS. Behavioral community psychology: encouraging low-income parents to seek dental care for their children. *J Appl Behav Anal* 1976; 9: 87–97.
37. Black RE, Dykes AC, Anderson KE, Wells JG, Sinclair SP, Gary GW Jr, et al. Handwashing to prevent diarrhea in day-care centers. *Am J Epidemiol* 1981; 113: 445–451.
38. Yokley JM and Glenwick DS. Increasing the immunization of preschool children; an evaluation of applied community interventions. *J Appl Behav Anal* 1984; 17: 313–325.
39. Grant A, Elbourne D, Valentin L and Alexander S. Routine formal fetal movement counting and risk of antepartum late death in normally formed singletons. *Lancet* 1989; 12: 345–349.
40. Tabár L, Fagerberg CJ, Gad A, Baldetorp L, Holmberg LH, Gröntoft O, et al. Reduction in mortality from breast cancer after mass screening with mammography Randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare. *Lancet* 1985; 13: 829–832.
41. Sommer A, Tarwotjo I, Djunaedi E, West KP Jr, Loeden AA, Tilden R, et al. Impact of vitamin A supplementation on childhood mortality: a randomized controlled community trial. *Lancet* 1986; 24: 1169–1173.

42. West KP, Djunaedi E, Pandji A, Kusdiono, Tarwotjo I and Sommer A. Vitamin A supplementation and growth: a randomized community trial. *Am J Clin Nutr* 1988; 48: 1257–1264.
43. Abdeljaber MH, Monto AS, Tilden RL, Schork MA and Tarwotjo I. The impact of vitamin A supplementation on morbidity: a randomized community intervention trial. *Am J Public Health* 1991; 81: 1654–1656.
44. Waber DP, Vuori-Christiansen L, Ortiz N, Clement JR, Christiansen NE, Mora JO, et al. Nutritional supplementation, maternal education, and cognitive development of infants at risk of malnutrition. *Am J Clin Nutr* 1981; Suppl 4: 807–813.
45. Roberts MM, Alexander FE, Anderson TJ, Forrest APM, Hepburn W, Huggins A, et al. The Edinburgh randomised trial of screening for breast cancer: description of method. *Br J Cancer* 1984; 50: 1–6.
46. Dwyer T, Coonan WE, Leitch DR, Hetzel BS and Baghurst RA. An investigation of the effects of daily physical activity on the health of primary school students in South Australia. *Int J Epidemiol* 1983; 12: 308–313.
47. Lloyd DM, Alexander HM, Callcott R, Dobson AJ, Hardes GR, O'Connell DL, et al. Cigarette smoking and drug use in schoolchildren: III-evaluation of a smoking prevention education programme. *Int J Epidemiol* 1983; 12: 51–58.
48. Simons-Morton BG, Coates TJ, Saylor KE, Sereghy E and Barofsky I. Great Sensations: a program to encourage heart healthy snacking by high school students. *J School Health* 1984; 54: 288–291.
49. Dielman TE, Shope JT, Leech SL and Butchart AT. Differential effectiveness of an elementary school-based alcohol misuse prevention program. *J School Health* 1989; 59: 255–263.
50. Schinke SP, Gilchrist LD, Schilling RF and Senechal VA. Smoking and smokeless tobacco use among adolescents: trends and intervention results. *Public Health Rep* 1986; 101: 373–378.
51. Fors SW, Owen S, Hall WD, McLaughlin J and Levinson R. Evaluation of a diffusion strategy for school-based hypertension education. *Health Educ Quart* 1989; 16: 255–261.
52. Bass MJ, McWhinney IR and Donner A. Do family physicians need medical assistants to detect and manage hypertension? *CMAJ* 1986; 134: 1247–1255.
53. Choi T, Jameson H, Brekke ML, Podratz RO and Mundahl H. Effects on nurse retention. An experiment with scheduling. *Med Care* 1986; 24: 1029–1043.
54. Seto WH, Ching PTY, Fung JPM and Fielding R. The role of communication in the alteration of patient-care practices in hospital – a prospective study. *J Hosp Infect* 1989; 14: 29–37.
55. Vogt RL, Larue D, Klaucke DN and Jillson DA. Comparison of an active and passive surveillance system of primary care providers for hepatitis, rubella, and salmonellosis in Vermont. *Am J Public Health* 1983; 73: 795–797.
56. Chassin MR and Mccue SM. A randomized trial of medical quality assurance. Improving physicians' use of pelvimetry. *JAMA* 1986; 256: 1012–1016.
57. McDonald CJ, Hui SL, Smith DM, Tierney WM, Cohen SJ, Weinberger M, et al. Reminders to physicians from an introspective computer medical record. A two-year randomized trial. *Ann Intern Med* 1984; 100: 130–138.
58. Stross JK, Banwell BF, Wolf FM and Becker MC. Evaluation of an education program on the management of rheumatic diseases for physical therapists. *J Rheumatol* 1986; 13: 374–378.
59. Evans CE, Haynes RB, Birkett NJ, Gilbert JR, Taylor DW, Sackett DL, et al. Does a mailed continuing education program improve physician performance? Results of a randomized trial in hypertensive care. *JAMA* 1986; 255: 501–504.
60. Bland JM and Kerry SM. Statistics notes. Trials randomised in clusters. *BMJ* 1997; 315: 600.
61. Kerry SM and Bland JM. The intracluster correlation coefficient in cluster randomisation. *BMJ* 1998; 316: 1455.
62. Campbell MK, Elbourne DR, Altman DG and CONSORT group. CONSORT statement: extension to cluster randomised trials. *BMJ* 2004; 328: 702–708.
63. 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.
64. Eldridge SM, Ashby D, Feder GS, Rudnicka AR and Ukoumunne OC. Lessons for cluster randomized trials in the twenty-first century: a systematic review of trials in primary care. *Clin Trials* 2004; 1: 80–90.
65. Manandhar DS, Osrin D, Shrestha BP, Mesko N, Morrison J, Tumbahangphe KM, et al. The effect of a participatory intervention with women's groups on birth outcomes in Nepal: cluster randomized controlled trial. *Lancet* 2004; 364: 970–979.
66. Morrison J, Tumbahangphe KM, Budhathoki B, Neupane R, Sen A, Dahal K, et al. Community mobilisation and health management committee strengthening to increase birth attendance by trained health workers in rural Makwanpur, Nepal: study protocol for a cluster randomized controlled trial. *Trials* 2011; 12: 128.
67. Azad K, Barnett S, Banerjee B, Shaha S, Khan K, Rego AR, et al. Effect of scaling up women's groups on birth outcomes in three rural districts in Bangladesh: a cluster-randomized controlled trial. *Lancet* 2010; 375: 1193–1202.
68. Tripathy P, Nair N, Barnett S, Mahapatra R, Borghi J, Rath S, et al. Effect of a participatory intervention with women's groups on birth outcomes and maternal depression in Jharkhand and Orissa, India: a cluster-randomized controlled trial. *Lancet* 2010; 375: 1182–1192.
69. Lewycka S, Mwansambo C, Rosato M, Kazembe P, Phiri T, Mganga A, et al. Effect of women's groups and volunteer peer counselling on rates of mortality, morbidity, and health behaviours in mothers and

children in rural Malawi (MaiMwana): a factorial, cluster-randomized controlled trial. *Lancet* 2013; 381: 1721–1735.
70. Cornfield J. Randomization by group: a formal analysis. *Am J Epidemiol* 1978; 108: 100–102.
71. Kramer MS, Martin RM, Sterne JA, Shapiro S, Mourad D and Platt RW. The double jeopardy of clustered measurement and cluster randomisation. *BMJ* 2009; 339: 503–505.
72. Fretheim A, Zhang F, Ross-Degnan D, Oxman AD, Cheyne H, Foy R, et al. A reanalysis of cluster randomized trials showed interrupted time-series studies were valuable in health system evaluation. *J Clin Epidemiol* 2015; 68: 324–333.
73. Taljaard M, McGowan J, Grimshaw JM, Brehaut JC, McRae A, Eccles MP and Donner A. Electronic search strategies to identify reports of cluster randomized trials in MEDLINE: low precision will improve with adherence to reporting standards. *BMC Med Res Methodol* 2010; 16: 10–15.
74. Connolly SJ, Philippon F, Longtin Y, Casanova A, Birnie DH, Exner DV, et al. Randomized cluster crossover trials for reliable, efficient, comparative effectiveness testing: design of the Prevention of Arrhythmia Device Infection Trial (PADIT). *Can J Cardiol* 2013; 29: 652–658.
75. Bellomo R, Forbes A, Akram M, Bailey M, Pilcher DV and Cooper DJ. Why we must cluster and cross over. *Crit Care Resusc* 2013; 15: 155–157.
76. Stockwell MS, Catallozzi M, Camargo S, Ramakrishnan R, Holleran S, Findley SE, et al. Registry-linked electronic influenza vaccine provider reminders: a cluster-crossover trial. *Pediatrics* 2015; 135: e75–82.