Evaluation of cluster-randomized trials on maternal and child health research in developing countries

Handlos, Line Neerup; Chakraborty, Hrishikesh; Sen, Pranab Kumar

Published in:
Tropical Medicine & International Health

DOI:
10.1111/j.1365-3156.2009.02313.x

Publication date:
2009

Document version
Early version, also known as pre-print

Tropical Medicine and International Health doi:10.1111/j.1365-3156.2009.02313.x

VOLUME 14 NO 8 PP 947–956 AUGUST 2009

Systematic Review

Evaluation of cluster-randomized trials on maternal and child health research in developing countries

Line Neerup Handlos[1,2,3], Hrishikesh Chakraborty[1,2] and Pranab Kumar Sen[1]

1 *Department of Biostatistics, University of North Carolina, Chapel Hill, NC, USA*
2 *Statistics and Epidemiology, RTI International, Research Triangle Park, NC, USA*
3 *University of Copenhagen, Copenhagen, Denmark*

Summary

OBJECTIVE To summarize and evaluate all publications including cluster-randomized trials used for maternal and child health research in developing countries during the last 10 years.

METHODS All cluster-randomized trials published between 1998 and 2008 were reviewed, and those that met our criteria for inclusion were evaluated further. The criteria for inclusion were that the trial should have been conducted in maternal and child health care in a developing country and that the conclusions should have been made on an individual level. Methods of accounting for clustering in design and analysis were evaluated in the eligible trials.

RESULTS Thirty-five eligible trials were identified. The majority of them were conducted in Asia, used community as randomization unit, and had less than 10 000 participants. To minimize confounding, 23 of the 35 trials had stratified, blocked, or paired the clusters before they were randomized, while 17 had adjusted for confounding in the analysis. Ten of the 35 trials did not account for clustering in sample size calculations, and seven did not account for the cluster-randomized design in the analysis. The number of cluster-randomized trials increased over time, and the trials generally improved in quality.

CONCLUSIONS Shortcomings exist in the sample-size calculations and in the analysis of cluster-randomized trials conducted during maternal and child health research in developing countries. Even though there has been improvement over time, further progress in the way that researchers utilize and analyse cluster-randomized trials in this field is needed.

keywords cluster analysis, developing countries, maternal health, child health, evaluation

Introduction

Reducing the worldwide maternal and child mortality ratios from 1990 to 2015 by 75% and 66%, respectively, is a key Millennium Development Goal (United Nations 2008). Given that, in a global perspective, the worst conditions among mothers and their children exist in developing countries, a serious effort should be undertaken in these countries to achieve this goal.

Several interventions have been implemented in maternal and child health care in developing countries throughout the years in order to reduce maternal and child mortality. The majority of them have been assessed by individually randomized controlled trials, but for practical, ethical, or economic reasons, these studies are not always appropriate in developing countries. Using clusters instead of individuals as a randomization unit has, however, proven to be a more efficient and inexpensive alternative, and the method is attractive in settings in which individual randomization is difficult or impossible (Hayes *et al.* 2000). Particularly, in the field of maternal and child health care cluster randomization has proved practical, as interventions which are known to have an impact on clusters of people rather than only individuals are common. Examples of such interventions are immunization strategies and educational and nutritional interventions spread via health service centres or mass media.

Empirical evaluations have shown that methodological shortcomings are common in the sample-size calculations and in the analysis of cluster-randomized trials in fields other than maternal and child health in developing countries. An evaluation of all cluster-randomized trials conducted in sub-Saharan Africa until 2001 by Isaakidis and Ioannidis (2003) showed that only 10 of 51 (20%) trials had accounted for clustering in sample-size calculations, and that only 37% had taken clustering into account

in analysis. Eldridge *et al.* (2004) found that only 20% of 199 trial reports from cluster-randomized trials in primary care had accounted for the clustering in the design phase and 59% of them had accounted for clustering in the analysis. Furthermore, Donner *et al.* (1990) found that only three of 16 (19%) studies concerning non-therapeutic interventions from 1979 to 1989 accounted for cluster-randomization in the design phase and eight of 16 (50%) trials took clustering into account in the analysis. In continuation of the shortcomings found in cluster-randomized trials in other fields, we found that it was justified to expect limitations to be present in maternal and child health research in developing countries also. As no evaluation of trials in this field has been done previously, conducting one was found to be relevant.

The aim of this evaluation was to summarize and evaluate the cluster-randomized trials in maternal and child health research that have been conducted in the developing world. The evaluation reports the results of a methodological assessment of all cluster-randomized trials performed in the past 10 years, and it evaluates the extent to which the pre-requisite design and analysis aspects of cluster randomization have been taken into account and reported properly in the trial publications. To evaluate the trials, two checklists based on the Consolidated Standards of Reporting Trials (CONSORT) statement (Moher *et al.* 2001 & Campbell *et al.* 2004) were used.

Methods

Our aim was to summarize and evaluate all publications of cluster-randomized trials in maternal and child health research that implemented cluster-level randomization, made conclusions on an individual level, and were conducted in developing countries in the last 10 years. In March–April 2007 and March–June 2008 available search engines – including PubMed, SCOPUS, and the Cochrane Library – were reviewed for all relevant papers published in English between January 1998 and June 2008. The keywords that were used in the initial search were: Cluster OR group OR community AND randomized OR randomized AND intervention OR trial.

In the initial search all publications were evaluated to detect whether they met the eligibility criteria. The criteria were that: (i) trials should have a cluster-randomized design; (ii) they should have been published between 1998 and 2008; (ii) they should have been conducted in maternal and child health research in what is designated as the developing regions of the world by the United Nations, and (iii) they should draw conclusions at the individual level. We do realize that trials with conclusions on cluster-level when randomized by cluster are numerous and important in the field of maternal and child health care, but we chose to exclude these studies, as they do not require adjustments for clustering, and thereby do not contain the same prospects of making erroneous conclusions as studies with individual-level analysis do (Chakraborty 2008).

Study reports that reflected secondary publications of a main study report were also included in the evaluation, given that those articles reported different variables as outcomes and thereby used methods and analyses different from those used in the primary study. In addition, whenever secondary publications reported additional useful information about the trial design or analysis of the primary publication, this information was recorded and used to give due credit to the trial. The references of each eligible paper were reviewed in order to find additional eligible trials published during that time period. Papers that presented no description of the methods for design or analysis and did not provide any reference to another publication with exposition of these details were excluded from the scope of this study.

In a secondary evaluation, each eligible article was systematically examined and evaluated by two of the authors. From each publication, information concerning study characteristics, sample-size calculations, analysis and conclusions was extracted. More specifically, for each article, the study recorded: (i) whether the trial was identified as cluster randomized in the title; (ii) whether the rationale for using a cluster-randomized design was stated; (iii) whether a description of what level the interventions pertained to was given; (iv) whether stratification or pairing (an extreme form of stratification in which each stratum consists of two clusters which are randomly assigned to different arms) was used and if so, whether any rationale was stated for doing so; (v) whether a description of how sample size was determined was given; (vi) whether the sample-size calculations took clustering into account; (vii) what method (if any) was used to account for cluster randomization in sample size calculations; (viii) whether magnitude of Intracluster correlation coefficient (ICC), design effect or coefficient of variation was stated; (ix) whether the analysis adjusted for confounding; and (x) whether the analysis took clustering into account. Location, primary object, publication year, sample size, number of clusters and cluster size were also recorded. This checklist was inspired by the CONSORT statement (Moher *et al.* 2001) and the extended version of CONSORT that has been specially formulated for cluster-randomized trials (Campbell *et al.* 2004).

In deciding whether clustering had been taken into account in sample-size calculations, theory concerning within-cluster correlation was used. ICC is a measurement that accounts for the degree to which responses from participants within the same cluster are similar. The sample

size of a trial depends on the magnitude of the ICC; the larger the ICC, the more participants and clusters are needed. Consequently, to determine the sample size needed in a cluster-randomized trial, an ICC has to be estimated before the data collection begins. In practice, the ICC is either estimated from previous trials, from data collected preliminary to the final data collection, or from simulation (Chakraborty *et al.* 2009). Based on the ICC a design effect is often calculated, to decide how much a sample size determined to be appropriate for an individually randomized trial should be magnified to agree with a cluster-randomized design (Chakraborty 2008). Another method to determine the sample size in a cluster-randomized trial is the coefficient of variation (Hayes & Bennett 1999).

When analysing cluster-randomized data, conclusions can be made on either cluster or individual levels. When making conclusions on a cluster level, no adjustment for clustering is needed because the unit of randomization is the same as the unit of analysis. On the contrary, when making conclusions on an individual-level analysis, it is necessary to account for within- and between-cluster correlation. There are several methods to make conclusions on an individual level when cluster-level randomization is used; common for the methods is the importance of accounting for clustering. If clustering is not accounted for in analysis, there is an extensive likelihood of false statistical significance (Chakraborty 2008).

Results

Characteristics of studies

The initial search yielded more than 10 000 articles. These were all scanned and evaluated according to the eligible criteria, and 35 papers were found eligible. In the secondary evaluation, the articles were carefully examined, and pertinent information was extracted. An overview of the papers can be seen in Table 1.

Of the 35 papers 11 were published in *Lancet* and five in *BMJ*, three in *BioMed Central*, two in the *American Journal of Tropical Medicine and Hygiene,* and two each in *Tropical Medicine and International Health,* the *American Journal of Clinical Nutrition* and *Pediatrics*. The other journals in which one paper was published in each were the *Journal of Nutrition*, the *New England Journal of Medicine*, the *Journal of the American Medical Association*, *Food and Nutrition Bulletin*, the *journal of the American Society for Nutritional Sciences*, *Midwifery*, *General Obstetrics,* and the *Transactions of the Royal Society of Tropical Medicine and Hygiene*.

Sixty-six per cent of the studies were conducted in Asia (predominantly in Nepal, Bangladesh and India), 20% in Africa and 11% in South America. One of the studies was multisited and conducted in both South America and Asia.

The primary objectives of the trials varied by types of interventions. Ten were nutritional interventions, seven dealt with preventing parasitic diseases such as malaria and helminths, five included training of traditional birth attendants and interventions to improve antenatal health care in general, four dealt with mobilizing or training local communities, four trials had medical trials and immunization as primary objectives, and one trial examined the psychosocial stimulation of children. Two interventions promoted breastfeeding, one hand washing, and one use of primary health care.

The number of clusters in each of the identified trials varied from seven (Jokhio *et al.* 2005) to 88 940 (Bhandari *et al.* 2007). The most commonly used unit of randomization in cluster-randomized trials in the examined publications was community (used in 54% of trials), but wards and health zones were also used as units of randomization (in 11% and 26% of trials, respectively). Households were less commonly (9%) used as clusters. The sample size varied widely from 136 (Hyder *et al.* 2007) to 350 000 (More *et al.* 2008) participants, although the majority of the studies (54%) had less than 10 000 participants. The average sample size was just below 26 000, and only five of the 35 studies had more than 100 000 participants. The mean cluster sample size varied from just above one (Bhandari *et al.* 2007) to around 7300 (More *et al.* 2008). A mean sample size per cluster of more than 200 was less common (26% of trials), while clusters with less than 50 participants were common (43% of trials).

Table 2 outlines how compliant the trials were with selected CONSORT guidelines. Only 51% of the papers identified themselves as cluster-randomized trials in the title, although the majority (61%) mentioned their cluster-randomized study design in the abstract. A few (n = 3) of the publications that were not identified as cluster-randomized trials in the title were instead designated as 'community-randomized trials' or 'community trials'.

Six trials stated a rationale for using the cluster-randomized design, the most commonly used rationale was that by intervening at cluster-level cross-contamination between treatment regimens was avoided (Hyder *et al.* 2007; Jokhio *et al.* 2005). In one trial (Majoko *et al.* 2007) the setting did not allow effective individual randomization whereas in another trial (Powell *et al.* 2004) it was not feasible for the children to receive different treatments within the same clinic. Thus both the intervention and the setting influenced the choice of study design.

In Table 3 some of the main findings of the included trials are listed. To control for confounding in the design phase and increase the power of the trial, all the included

Table 1 List of papers published from 1998 to 2008 using cluster-randomized design and drawing conclusions on the individual level in maternal and child health research in developing countries

Location of study	Reference	No. participants	No. clusters	Primary objective	Controlled for confounding	Accounted for cluster randomization in sample size calculations	Adjusted for cluster randomization in analysis
Argentina, Cuba, Saudi Arabia and Thailand	Villar *et al.* (2001)	24 526	53	Comparison of the standard model of antenatal care with a new model	In design by stratification and in analysis by controlling for various variables	Yes	Yes
Bangladesh	Baqui *et al.* (2008)	47 158	24	Community-based intervention package to reduce neonatal mortality	No confounder control in design. Controlling for various variables in the analysis	Yes	No
Bangladesh	Hyder *et al.* (2007)	136	13	Iron supplements' effect on haemoglobin, serum ferritin and serum transferring receptor levels in children	Neither controlled in design or analysis	Yes	No
Bangladesh	Arifeen *et al.* (2004)	2082	10	Evaluation of the Integrated Management of Childhood Illness strategy	Control in design by pairing. No control in analysis	No	Yes
Bangladesh	Baqui *et al.* (2002)	8070	30	Zinc supplementation to children with diarrhoea	Stratification in design and control for various variables in analysis	Yes	Yes
Bangladesh	Haider *et al.* (2000)	726	40	Peer counselling on breastfeeding	No control in either design or analysis	No	Yes
Ecuador	Cooper *et al.* (2006)	2373	68	Effect of anthelmintic treatments on the prevalence of atopy and allergy	No confounder control in design. Adjustment for various variables in the analysis	Yes	Yes
Ghana	Chandramohan *et al.* (2005)	2485	96	Intermittent preventive treatment for malaria in infants with sulfadoxine-pyrimethamine	No control in design. Adjusted for various variables in the analysis	Yes	Yes
Ghana	Browne *et al.* (2001)	1961	96	Impact of impregnated bed nets on malaria and anaemia in pregnancy	No control for confounding in either design or analysis	Yes	Yes
Ghana	Binka *et al.* (1998)	134 400	96	Impregnated bed nets effect on child mortality	No control in design, adjusted for one variable in analysis	Yes	No
Haiti	Ruel *et al.* (2008)	1500	20	Maternal and child nutrition programme	Pairing used in design and adjustment for various variables in analysis	Yes	Yes
Honduras	Morris *et al.* (2004)	5600	70	Monetary incentives to pregnant women	Blocked and stratified in design. No adjustment in analysis	Yes	Yes

Table I (*Continued*)

Location of study	Reference	No. participants	No. clusters	Primary objective	Controlled for confounding	Accounted for cluster randomization in sample size calculations	Adjusted for cluster randomization in analysis
India	More *et al.* (2008)	350 000	48	Support to local women as facilitators in mobilizing communities for better health care	Stratification to control in design. No adjustment in analysis	Yes	Yes
India	Bhandari *et al.* (2007)	94 359	88 940	Zinc, folic acid and iron supplementation to children	Stratified in design phase. Adjusted for various variables in analysis	Yes	Yes
India	Bhandari *et al.* (2004)	1025	8	Promotion of complementary feeding practices and physical growth in infants and young children	Paired in design phase. In analysis is adjusted for various variables	Yes	Yes
India	Bhandari *et al.* (2003)	1115	8	Promotion of breastfeeding and its effect on diarrheal illness and growth	Pair-design. In analysis was adjusted for various variables	Yes	Yes
Indonesia	SUMMIT Study Group (2008)	31 290	262	MMN + IFA supplementation to pregnant women	Stratified in design-phase. No adjustment in analysis	Yes	Yes
Indonesia	Fahdhy and Chongsuvivat-wong (2005)	721	20	To assess the effectiveness of promoting the use of the WHOs partograph by midwives for labour in a maternity home	No adjustment in design. Adjusting for various variables in analysis	No	Yes
Jamaica	Powell *et al.* (2004)	139	18	Psychosocial stimulation for undernourished children	Stratified in design-phase. Adjusted for various variables in analysis	No	No
Kenya	Kuile *et al.* (2003)	1890	60	Impact of insecticide treated bed nets on malaria and morbidity in children	Stratified in design-phase. Adjusted for one variable in the analysis	Yes	Yes
Kenya	Schulman *et al.* (1998)	503	28	Insecticide-treated bed nets to prevent malaria and anaemia in pregnant women.	No adjustment was made in either design or analysis	No	No
Mali	Winch *et al.* (2003)	286	10	Counsel on drug use to treat malaria in children	Pairing in design-phase. No adjustment in analysis	Yes	Yes
Nepal	Tielsch *et al.* (2007)	17 530	413	To assess the efficacy of cleansing of newborn infants with chlorhexidine on neonatal mortality	Stratification in design-phase. Stratified analyses were planned	No	Yes
Nepal	Mullany *et al.* (2006a)	15 123	413	Evaluation of the impact of topical chlorhexidine treatment on cord-separation times	No control in design-phase. Adjustment for various variables in analysis	No	Yes

Table I (*Continued*)

Location of study	Reference	No. participants	No. clusters	Primary objective	Controlled for confounding	Accounted for cluster randomization in sample size calculations	Adjusted for cluster randomization in analysis
Nepal	Mullany *et al.* (2006b)	15 123	413	Topical applications of chlorhexidine to the umbilical cord for prevention of omphalitis and neonatal mortality	Blocked in design-phase and adjusted for various variables in analysis	No	Yes
Nepal	Wade *et al.* (2006)	5400	24	Behaviour change in perinatal care practices among women exposed to a women's group intervention	Paired according to similar topography, ethnicity and population densities before randomization	Yes	Yes
Nepal	Manandhar *et al.* (2004)	28 931	42	Evaluation of a participatory intervention with women's groups on birth outcomes	Pairing in design-phase. No adjustment in analysis	Yes	Yes
Nepal	Christian *et al.* (2003a)	4926	426	Effect of micronutrients given to pregnant women on foetal loss and infant mortality	Randomization in design. No adjustment in analysis	Yes	Yes
Nepal	Christian *et al.* (2003b)	4926	426	Effect of micronutrients given to pregnant women on low birth weight	Blocking in design. Adjustment for one variable in analysis	Yes	Yes
Nepal	Katz *et al.* (2000)	43 559	270	Effect of maternal vitamin A or β-carotene supplementation on foetal loss and survival of infants	Stratified in design and no adjustment in analysis	No	Yes
Nepal	West *et al.* (1999)	44 646	270	Vitamin A supplementation to women of reproductive age	Blocking in design-phase. No adjustment in analysis	No	No
Pakistan	Khan *et al.* (2006)	21 059	60	Mass immunization strategy	Stratification in design-phase. No adjustment in analysis	Yes	Yes
Pakistan	Jokhio *et al.* (2005)	19 557	7	Training of traditional birth attendants	No adjustment for confounding in either design or analyses	Yes	Yes
Pakistan	Luby *et al.* (2004)	4691	36	Promotion of hand washing	No adjustment for confounding in either design or analyses	Yes	No
Zimbabwe	Majoko *et al.* (2007)	13 517	23	Comparison of two antenatal care models.	Stratified in design. No adjustment in analysis	Yes	Yes

MMN, multiple micronutrient; IFA, iron and folic acid.

Table 2 Compliance with selected CONSORT guidelines

Checklist item	Studies that have included the item	Studies that have not included the item
Identification of cluster-randomized design in title	18	17
Rationale for using a cluster-randomized design	6	29
Rationale for stratification, blocking or pairing*	7	16
Description of whether the interventions pertained to cluster-level or individual-level	35	0
Description of how sample size was determined	33	2
Presentation of ICC, magnitude of design effect or coefficient of variation	25	10
Description of how clustering was taken into account in the statistical analyses	22	13

*Only 23 studies used stratification, blocking or pairing, therefore not all studies are represented in this row.
CONSORT, Consolidated Standards of Reporting Trials; ICC, intracluster correlation coefficient.

Table 3 Main findings in the 35 included studies

Number of trials that used stratification or blocking in design	16
Number of trials that used pairing in design	7
Number of trials that accounted for the cluster-level design in sample-size calculations	25
Methods used for accounting for the cluster-level design in the sample-size calculations	
ICC or design effect	18
Coefficient of variation	7
Number of trials that adjusted for counfounding in the analysis	17
Number of trials that accounted for the cluster-level design in analysis	28

ICC, intracluster correlation coefficient.

trials had selected which clusters should receive interventions by randomization. Furthermore, 16 of the 35 trials had either stratified or blocked the clusters before they were randomized to ensure an equal distribution of baseline characteristics in the intervention and the control group. Among the trials that had used this method to prevent the data from being confounded, the clusters had been stratified according to, among other indicators, geographical distribution, access to health care, weight and age of participants, baseline mortality and morbidity rate, population density, ethnicity, and gender. Pairing was also commonly used to avoid confounding; in addition to the 16 trials that used stratification or blocking, seven of the trials utilized pairing before randomization, and the parameters, which determined the pairing, were similar to the factors that were used in stratifications. Only seven of the 23 trials that used stratification, blocking or pairing had stated the rationale for using these methods. Without exception, the rationale described was to ensure baseline balance.

All trials had described whether the intervention pertained to cluster or individual-level and about one-half (49%) of the trials adjusted for confounding in the analysis by controlling for different baseline variables. Six of the 35 trials had not accounted for confounding in either the design or the analysis phase.

Accounting for clustering

Ten of the 35 trials did not use ICC, design effect or coefficient of variation to adjust for clustering in sample-size calculations. Of the 25 trials that took the cluster randomization into account in calculating the sample size, 72% used ICC or design effect in the calculations and 28% used the coefficient of variation method to account for the cluster-randomized design in the sample size calculations.

Of the 18 trials that present ICC values or design effect, eight have estimated the magnitude from data from previous trials, seven have estimated the value from data collected prior to the final data collection, and three do not state the origin of the value. The magnitude of the coefficients of variation was in two of the seven trials determined on the basis of existing data collected prior to the study, whereas two trials had based the magnitude of the coefficient on estimates available at the national level or for the specific area. The remaining three trials have not stated any origin of the coefficients used.

Seven of the 35 included trials did not account for the cluster-randomized design in the analysis; instead, the data were analysed as if they were randomized at the individual level. Of the 28 studies that did account for the cluster design, only 22 had described how they did it.

The countries with the most problems in accounting for cluster randomization in trials were Bangladesh and Nepal. In both, more than one-half of the trials conducted did not take clustering into account. In Nepal, the problem was most present in the sample size-calculations (five of nine trials did not account for clustering in the sample-size calculations, while one of nine did not account for clustering in the analysis), while the trials in Bangladesh had an equal amount of problems in sample size calculations and analysis (two trials did not account for clustering in sample-size calculations, and two neglected to account for it in the analysis).

The distribution of papers according to journal showed no particular trend towards more appropriate conduction or reporting in journals that more commonly had published cluster-randomized trials.

Change over time

As shown in Figures 1 and 2, the per-annum number of cluster-randomized trials in maternal and child health research conducted in developing countries increased over time. Only 8 (23%) of the 35 trials were published in the first 5 years of the period (from 1998 to 2002), while 27 (77%) were published in the last 5 years (between 2003 and 2008). The number of trials making conclusions in a correct manner increased over time. Five of 8 (63%) of the trials published between 1998 and 2002 did not account for cluster randomization in either the design or analysis phases, while only nine of 27 (32%) of the trials published in the period 2003–2008 did not account for cluster randomization in either design or analysis.

Discussion

This report is the first to present a coherent evaluation of all cluster-randomized trials with conclusions on the individual level, that were conducted in maternal and child health in developing countries during the period 1998–2008. The evaluation has found that a large proportion of the included trials use improper methods in sample-size calculations and/or analysis. Fourteen of 35 trials (40%) did not account appropriately for clustering in either sample-size calculations or analysis.

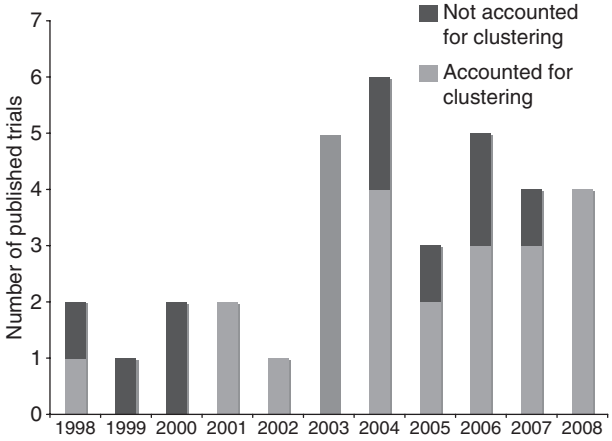

Figure 1 Distribution of published cluster-randomized trials in maternal and child health in developing countries by year. Have the trials accounted for the cluster design in the sample size calculations?

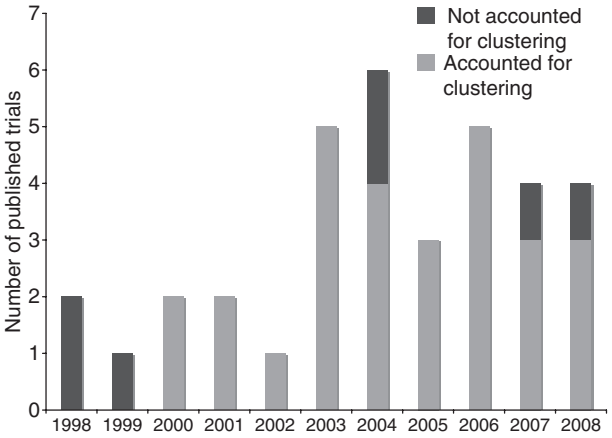

Figure 2 Distribution of published cluster-randomized trials in maternal and child health in developing countries by year. Have the trials accounted for the cluster design in the analyses?

In several trials, authors do not make the right type of analysis for the level on which they draw conclusions. For example, Browne *et al.* (2001) does analyses on the cluster level, even though the conclusions about incidence of *Plasmodium falciparum* infections, haemoglobin levels and delivery outcomes are made on the individual level. This is also seen in the trial conducted by Schulman *et al.* (1998), in which the analysis is presented as being at the community level but the conclusions are made at the individual level. This lack of distinction between cluster- and individual-level analyses can potentially lead to false inferences of significant associations between exposure and outcome.

Another recurring problem in the articles is that there is no justification for choice of magnitude of adjustment for clustering in sample-size calculations. For example, Hyder *et al.* (2007) mentions that a design effect of two was used to account for the clustering effect, but no reasons were presented to support this value. Browne *et al.* (2001) adjusts the sample size by 15% to allow for clustering without giving any explanation for the choice of ICC values or design effect, and Luby *et al.* (2004) double the sample size to account for the effect of clustering without presenting any rationale. Furthermore, as many as 17% of the trials that present an ICC do not state the origin of the value. The shortcomings in the documentation of the sample-size calculations make it difficult (and, in some cases, even impossible) to evaluate whether an appropriate sample-size has been used. However, our guess is that if the authors did not justify their choice of magnitude of the adjustment for clustering in the sample size calculation, an underpowered trial was designed as a value too small was most probably used.

Six of the 35 trials did not adjust for confounding in either design or analysis. Whether no confounding factors were present at these study sites is not within the scope of this evaluation, but only one of the six trials mentions that there was a search for confounders and none was identified. The remaining five do not account for any considerations concerning confounding factors. However, this might not be an issue as the design of all the trials included randomization.

The findings of this evaluation show slightly more frequent use of correct methods to account for clustering than those from previously conducted empirical evaluations of cluster-randomized trials in other fields. This evaluation has shown a tendency towards an improvement over time in the percentage of trials that use appropriate designs and analyses when drawing conclusions on an individual level. This improvement can partly explain the better methodological findings among trials included in this evaluation compared with the findings from earlier evaluations.

Despite the demonstrated improvement, this evaluation proves that a need still exists for further progress in the way that researchers use and analyse cluster-randomized trials in maternal and child health research in developing countries. Especially better reporting and sharing of ICC values are needed, as the literature currently contains only few examples of ICC coefficients in the field of maternal and child health in developing countries. Thus, progress in several areas is essential for the research in this field to create valid results and thereby change the problems with which the developing countries are confronted in the field of maternal and child health.

References

Arifeen SE, Blum LS, Hoque DME *et al.* (2004) Integrated Management of Childhood Illness (IMCI) in Bangladesh: early findings from a cluster-randomised study. *Lancet* **364**, 1595–1602.

Baqui AH, Black RE, Arifeen SE *et al.* (2002) Effect of zinc supplementation started during diarrhoea on morbidity and mortality in Bangladeshi children: community randomized trial. *BMJ* **325**, 1059–1066.

Baqui AH, Arifeen SE, Darmstadt GL *et al.* (2008) Effect of community-based newborn-care intervention package implemented through two service-delivery strategies in Sylhet district, Bangladesh: a cluster-randomised controlled trial. *Lancet* **371**, 1936–1944.

Bhandari N, Bahl R, Mazumdar S *et al.* (2003) Effect of community-based promotion of exclusive breastfeeding on diarrhoeal illness and growth: a cluster randomised controlled trial. *Lancet* **361**, 1418–1423.

Bhandari N, Mazumder S, Bahl R *et al.* (2004) An educational intervention to promote appropriate complementary feeding practices and physical growth in infants and young children in rural Haryana, India. *American Society for Nutritional Sciences*, **134**, 2342–2348.

Bhandari N, Taneja S, Mazumder S *et al.* (2007) Adding zinc to supplemental iron and folic acid does not affect mortality and severe morbidity in young children. *Journal of Nutrition* **137**, 112–117.

Binka NF, Indome F & Smith T (1998) Impact of spatial distribution of permethrin-impregnated bed nets on child mortality in rural northern Ghana. *American Journal of Tropical Medicine and Hygiene* **59**, 80–85.

Browne ENL, Maude GH & Binka FN (2001) The impact of insecticide-treated bednets on malaria and anaemia in pregnancy in Kassena-Nankana district, Ghana: a randomized controlled trial. *Tropical Medicine and International Health* **6**, 667–676.

Campbell MK, Elbourne DR, Altman DG & CONSORT group (2004) CONSORT statement: extension to cluster randomised trials. *BMJ* **328**, 702–708.

Chakraborty H (2008) The design and analysis aspects of cluster randomized trials. In: *Statistical Advances in the Bio-medical Sciences: Clinical Trials, Epidemiology, Survival Analysis, and Bioinformatics* (eds A Biswas, S Datta, Fine JP & MR Segal) John Wiley & Sons, Inc., New York, pp. 67–80.

Chakraborty H, Moore J, Carlo AW *et al.* (2009) A simulation based technique to estimate intracluster correlation for a binary variable. *Contemporary Clinical Trials* **30**, 71–80.

Chandramohan D, Owusu-Agyei S, Carneiro I *et al.* (2005) Cluster randomised trial of intermittent preventive treatment for malaria in infants in area of high, seasonal transmission in Ghana. *BMJ* **331**, 727–733.

Christian P, West KP, Khatry SP *et al.* (2003a) Effects of maternal micronutrient supplementation on fetal loss and infant mortality: a cluster-randomized trial in Nepal. *American Journal of Clinical Nutrition* **78**, 1194–1202.

Christian P, Khatry SK, Katz J *et al.* (2003b) Effects of alternative maternal micronutrient supplements on low birth weight in rural Nepal: double blind randomised community trial. *BMJ* **326**, 571–578.

Cooper PJ, Chico ME, Vaca MG *et al.* (2006) Effect of albendazole treatments on the prevalence of atopy in children living in communities endemic for geohelminth parasites: a cluster-randomised trial. *Lancet* **367**, 1598–1603.

Donner A, Brown KS, Brasher P *et al.* (1990) A methodological review of non-therapeutic intervention trials employing cluster randomization 1979–1989. *International Journal of Epidemiology* **19**, 795–800.

Eldridge SM, Ashby D, Feder GS *et al.* (2004) Lessons for cluster randomized trials in the twenty-first century: a systematic review of trials in primary care. *Clinical Trials* **1**, 80–90.

Fahdhy M & Chongsuvivatwong V (2005) Evaluation of World Health Organization partograph implementation by midwives for maternity home birth in Medan, Indonesia. *Midwifery* **21**, 301–310.

Haider R, Ashworth A & Kabir I (2000) Effect of community-based peer counsellors on exclusive breastfeeding practices in Dhaka, Bangladesh: a randomised controlled trial. *Lancet* **356**, 1643–1648.

Hayes RJ & Bennett S (1999) Simple sample size calculation for cluster-randomized trials. *International Journal of Epidemiology* **28**, 319–326.

Hayes RJ, Alexander NDE, Bennet S *et al.* (2000) Design and analysis issues in cluster-randomized trials of interventions against infectious diseases. *Statistical Methods in Medical Research* **9**, 95–116.

Hyder SMZ, Haseen F, Rahman M *et al.* (2007) Effect of daily versus once-weekly home fortification with micronutrient sprinkles on hemoglobin and iron status among young children in rural Bangladesh. *Food and Nutrition Bulletin* **28**, 156–165.

Isaakidis P & Ioannidis JPA (2003) Evaluation of cluster randomized controlled trials in Sub-Saharan Africa. *American Journal of Epidemiology* **158**, 921–926.

Jokhio AH, Winter HR, Cheng KK *et al.* (2005) An intervention involving traditional birth attendants and perinatal and maternal mortality in Pakistan. *The New England Journal of Medicine* **352**, 2091–2099.

Katz J, West KP, Khatry SK *et al.* (2000) Maternal low-dose vitamin A or β-carotene supplementation has no effect on fetal loss and early infant mortality: a randomized cluster trial in Nepal. *American Journal of Clinical Nutrition* **71**, 1570–1576.

Khan MI, Ochiai RL, Hamza HB *et al.* (2006) Lessons and implications from a mass immunization campaign in squatter settlements of Karachi, Pakistan: an experience from a cluster-randomized double-blinded vaccine trial. *BioMed Central Trials* **7**, 1–10.

Kuile FOT, Terlouw DJ, Phillips-Howard PA *et al.* (2003) Impact of permethrin-treated bed nets on malaria and all-cause morbidity in young children in an area of intense perennial malaria transmission in Western Kenya: cross-sectional survey. *American Journal of Tropical Medicine and Hygiene* **68**, 100–107.

Luby SP, Agboatwalla M, Painter J *et al.* (2004) Effect of intensive handwashing promotion on childhood diarrhea in high-risk communities in Pakistan. *JAMA* **291**, 2547–2554.

Majoko F, Munjanja SP, Nystrøm L, Mason E & Lindmarka G (2007) Randomised controlled trial of two antenatal care models in rural Zimbabwe. *General Obstetrics* **114**, 802–811.

Manandhar DS, Osrin D, Shrestha BP *et al.* (2004) Effect of a participatory intervention with women's groups on birth outcomes in Nepal: cluster-randomised controlled trial. *Lancet* **364**, 970–979.

Moher D, Schulz KF & Altman D (2001) The CONSORT Statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* **285**, 1987–1991.

More NS, Bapat U, Das S *et al.* (2008) Cluster-randomised controlled trial of community mobilisation in Mumbai slums to improve care during pregnancy, delivery, postpartum and for the newborn. *BioMed Central Trials* **9**, 1–11.

Morris SS, Flores R, Olinto P & Medina JM (2004) Monetary incentives in primary health care and effects on use and coverage of preventive health care interventions in rural Honduras: cluster randomised trial. *Lancet* **364**, 2030–2037.

Mullany LC, Darmstadt GL, Khatry SK *et al.* (2006a) Impact of umbilical cord cleansing with 4.0% chlorhexidine on time to cord separation among newborns in southern Nepal: a cluster-randomized, community-based trial. *Pediatrics* **118**, 1864–1871.

Mullany LC, Darmstadt GL & Khatry SK (2006b) Topical applications of chlorhexidine to the umbilical cord for prevention of omphalitis and neonatal mortality in southern Nepal: a community-based, cluster-randomised trial. *Lancet* **367**, 910–918.

Powell C, Baker-Henningham H, Walker S *et al.* (2004) Feasibility of integrating early stimulation into primary care for undernourished children: cluster randomised controlled trial. *BMJ* **329**, 89–94.

Ruel MT, Menon P, Habicht JP *et al.* (2008) Age-based preventive targeting of food assistance and behaviour change and communication for reduction of childhood undernutrition in Haiti: a cluster randomised trial. *Lancet* **371**, 588–595.

Schulman CE, Dorman EK, Talisuna AO *et al.* (1998) A community randomized controlled trial of insecticide-treated bednets for the prevention of malaria and anaemia among primigravid women on the Kenyan coast. *Tropical Medicine and International Health* **3**, 197–204.

SUMMIT Study Group (2008) Effect of maternal multiple micronutrient supplementation on fetal loss and infant death in Indonesia: a double-blind cluster-randomised trial. *Lancet* **371**, 215–227.

Tielsch JM, Darmstadt GL, Mullany LC *et al.* (2007) Impact of newborn skin-cleansing with chlorhexidine on neonatal mortality in Southern Nepal: a community-based, cluster-randomized trial. *Pediatrics* **119**, 330–342.

United Nations (2008) A gateway to the UN System's work on the MDGs. www.un.org/millenniumgoals/.

Villar J, Ba'aqeel H, Piaggio G *et al.* (2001) WHO antenatal care randomised trial for the evaluation of a new model of routine antenatal care. *Lancet* **357**, 1551–1564.

Wade A, Osrin D, Shrestha BP *et al.* (2006) Behaviour change in perinatal care practices among rural women exposed to a women's group intervention in Nepal. *BMC Pregnancy and Childbirth* **6**, 1–10.

West KP, Katz J, Khatry SK *et al.* (1999) Double blind, cluster randomised trial of low dose supplementation with vitamin A or β-carotene on mortality related to pregnancy in Nepal. *BMJ* **318**, 570–575.

Winch PJ, Bagayoko A, Diawara A *et al.* (2003) Increases in correct administration of chloroquine in the home and referral of sick children to health facilities through a community-based intervention in Bougouni District, Mali. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **97**, 481–490.

Corresponding Author **Line Neerup Handlos,** Department of Biostatistics, University of North Carolina, Chapel Hill, NC, USA.
E-mail: lneerup@gmail.com