Education and debate

CONSORT statement: extension to cluster randomised trials
Marion K Campbell, Diana R Elbourne, Douglas G Altman for the CONSORT Group

Reports of cluster randomised trials require additional information to allow readers to interpret them accurately. The effective reporting of randomised controlled trials has received useful attention in recent years. Many journals now require that reports conform to the guidelines in the Consolidated Standards of Reporting Trials (CONSORT) statement, first published in 1996 and revised in 2001. The statement includes a checklist of items that should be included in the trial report. These items are evidence based whenever possible and are regularly reviewed. The statement also recommends including a flow diagram to show the flow of participants from group assignment through to the final analysis.

The CONSORT statement focused on reporting parallel group randomised trials in which individual participants are randomly assigned to study groups. However, in some situations it is preferable to randomly assign groups of individuals (such as families or medical practices) rather than individuals. Reasons include the threat of contamination of some families or medical practices) rather than individuals. As in an earlier discussion paper we considered the implications of the CONSORT statement for the reporting of cluster randomised trials. Here we present updated guidance, based on the 2001 revision of the CONSORT statement.

Methodological issues in cluster randomised trials

Compared with individually randomised trials, cluster randomised trials are more complex to design, require more participants to obtain equivalent statistical power, and require more complex analysis. The methodological issues in cluster randomised trials have been widely discussed. In brief, observations on individuals in the same cluster tend to be correlated (non-independently), and so the effective sample size is less than the total number of individual participants.

The reduction in effective sample size depends on average cluster size and the degree of correlation within clusters, known as the intraclass (or intraclass) correlation coefficient (ρ). The intraclass correlation coefficient is the proportion of the total variance of the outcome that can be explained by the variation between clusters. To retain power, the sample size should be multiplied by \(1+(m−1)\rho\), called the design effect, where \(m\) is the average cluster size. Hayes and Bennett describe a related coefficient of variation, \(k\), between clusters and Connelly considers an economic approach. Software is available to adjust for the intraclass correlation coefficient in an analysis.

The conduct of cluster randomised controlled trials may also differ from that of trials that randomise individuals. For instance, clusters are usually randomised all at once (or in batches) rather than one at a time. After randomisation, individuals in the clusters may be approached for consent, which raises the possibility of post-randomisation selection bias, or they may not, which raises ethical concerns.

An expanded explanation of the methods of cluster randomised trials is available on the CONSORT website (www.consort_statement.org).

Quality of reporting of cluster trials

Surveys of published cluster trials have found that their conduct and reporting have been poor. Of 21 cluster trials identified in two major public health journals, only four (19%) had accounted for clustering in the planning of the trial. Similarly, in a review of physicians’ patient care practices, 70% (38/54) of the identified studies had not appropriately accounted for the clustered nature of their study data in their analysis. Of 16 cluster trials reviewed by Donner et al, only four provided any rationale for adopting a clustered design, only three accounted for clustering in the sample size calculations, and only eight accounted for clustering in the analysis.

Recent studies have shown continuing problems with the design and analysis of cluster trials; 42% (62/149) of trials of implementation research interventions did not account appropriately for the clustering in their design, and 42% (10/24) of trials of clinical decision support systems did not appropriately account for clustering in their analysis (none accounted for clustering in their sample size calculations). A recent review of 51 cluster randomised trials conducted in sub-Saharan Africa...
published in the past 30 years showed only 10 trials
took clustering into account in sample size or power
calculations, and only 19 took clustering into account
in the analysis. Intraclass correlation coefficients and
design effects were reported in only one and three
trials, respectively.23

Extension of CONSORT statement to
cluster trials
To accommodate the reporting of the special features
of the cluster randomised trial, we have extended the
CONSORT statement to include the following
information:
● The rationale for adopting a cluster design
● How the effects of clustering were incorporated into
the sample size calculations
● How the effects of clustering were incorporated into
the analysis
● The flow of both clusters and individuals through
the trial, from assignment to analysis.

Specific changes to the checklist items and the flow
diagram to include these issues are described below.
For items not mentioned here the advice is as for indi-
vidually randomised trials.3

Table 1 shows the modified checklist with the addi-
tions to the standard CONSORT list in italics. In this
section, we discuss the rationale for these extensions
and provide examples of good reporting. In some
elements we have added text in square brackets to
explain the context.

Title and abstract
Item 1: How participants were allocated to interven-
tions (eg random allocation, randomised, or randomly
assigned), specifying that allocation was based on clusters.

Example
Title: “Self help smoking cessation in pregnancy: cluster
randomised trial.”23
Abstract (design): “Pragmatic cluster randomised
controlled trial with community midwife as the unit of
randomisation,”23

Explanation
The primary reason for identifying the design in the
title or abstract is to ensure appropriate indexing of the
study as a cluster randomised trial in Medline. This
indexing ensures ease of identification of these studies
for inclusion in systematic reviews. In addition, readers
will not be misled by apparently large sample sizes.

Researchers should also consider reporting the
number of clusters in the abstract.

Introduction
Item 2: Scientific background and explanation of
rationale, including the rationale for using a cluster design.

Example
Our intention was to enhance the application of evidence by
the whole labour ward team so, to minimise contamination,
the unit of randomisation and analysis was the obstetric unit.23

Explanation
Under the principles of the Helsinki declaration, it is
unethical to expose people unnecessarily to the risks of
research.23 Because a cluster randomised design
increases the complexity of the research and requires
many more participants than an individually ran-
domised design (to ensure equivalent statistical power),
it is particularly important that the rationale for adopt-
ing a cluster design is outlined in the introduction.23

| Table 1 Checklist of items to include when reporting a cluster randomised trial (adaptations from standard guidelines in italic) |
|---|---|---|
| Paper section and topic | Item | Descriptor |
| Title and abstract | Design | 1* How participants were allocated to interventions (eg random allocation, randomised, or randomly assigned), specifying that allocation was based on clusters |
| Introduction | Background | 2* Scientific background and explanation of rationale, including the rationale for using a cluster design |
| Methods | Participants | 3* Eligibility criteria for participants and clusters and the settings and locations where the data were collected |
| | Interventions | 4* Precise details of the interventions intended for each group, whether they pertain to the individual level, the cluster level, or both, and how and when they were actually administered |
| | Objectives | 5* Specific objectives and hypotheses and whether they pertain to the individual level, the cluster level, or both |
| | Sample size | 7* How total sample size was determined (including method of calculation, number of clusters, cluster size, a coefficient of intraclass correlation (ICC or k), and an indication of its uncertainty) and, when applicable, explanation of any interim analyses and stopping rules |
| Randomisation | Sequence generation | 8* Method used to generate the random allocation sequence, including details of any restriction (eg blocking, stratification, matching) |
| | Allocation concealment | 9* Method used to implement the random allocation sequence, specifying that allocation was based on clusters rather than individuals and clarifying whether the sequence was concealed until interventions were assigned |
| Implementation | 10 Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups |
| Blinding (masking) | 11 Whether participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment. If done, how the success of blinding was evaluated |
| Statistical methods | 12* Statistical methods used to compare groups for primary outcome(s) indicating how clustering was taken into account, methods for additional analyses, such as subgroup analyses and adjusted analyses |
| Results | Participant flow | 13* Flow of clusters and individual participants through each stage (a diagram is strongly recommended). Specifically, for each group report the numbers of clusters and participants randomly assigned, receiving intended treatment, completing the study protocol, and analysed for the primary outcome |
| | Recruitment | 14 Dates defining the periods of recruitment and follow up |
| | Baseline data | 15* Baseline information for each group for the individual and cluster levels as applicable |
| | Numbers analysed | 16* Number of clusters and participants (denominator) in each group included in each analysis and whether the analysis was by intention to treat. State the results in absolute numbers when feasible (eg 10/20 not 50%) |
| Outcomes and estimation | 17* For each primary and secondary outcome, a summary of results for each group for the individual or cluster level as applicable, and the estimated effect size and its precision (eg 95% confidence interval) and a coefficient of intraclass correlation (ICC or k) for each primary outcome |
| Ancillary analyses | 18 Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses, indicating those prespecified and those exploratory |
| Adverse events | 19 All important adverse events or side effects in each intervention group |
| Discussion | Interpretation | 20 Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision and the dangers associated with multiplicity of analyses and outcomes |
| | Generalisability | 21* Generalisability (external validity) to individuals and/or clusters (as relevant) of the trial findings |
| | Overall evidence | 22 General interpretation of the results in the context of current evidence |

*Addition to CONSORT guidelines 2001
Methods

The main difference when reporting a cluster trial, as opposed to an individually randomised trial, is that there are two levels of inference rather than one: the cluster level and the individual level. Thus, to allow readers to interpret the results appropriately, it is important to indicate explicitly the level at which the interventions were targeted, the hypotheses were generated, the outcomes were measured, and randomisation was done.

Participants

Item 3: Eligibility criteria for participants and clusters and the settings and locations where the data were collected.

Example

The study comprised 41 practices in Wessex ... Inclusion criteria were ≥ 4 medical partners; list size > 7000; a diabetes register with >1% of practice population; and a diabetes service registered with the health authority ... Nurses reported all new cases of diabetes to the trial office. Willing patients aged 30-70 were included in the trial. Patients were excluded if they were private patients, housebound, mentally ill, had severe learning difficulties, or were subsequently found to have been diagnosed previously with, or not to have, diabetes, or were found to have type 1 diabetes. Explanation

Because there are two levels of inference, the eligibility criteria for clusters, as well as participants, need to be reported. In a cluster trial, the primary eligibility criterion is often all the clusters in a defined geographical area.

Intervention

Item 4: Precise details of the interventions intended for each group, whether they pertain to the individual level, the cluster level, or both, and how and when they were actually administered.

Example

We ... paired the 14 [urban sectors of Trujillo, Venezuela] according to the incidence of cutaneous leishmaniasis in the 12 months before the baseline household survey. For each of the seven pairs we randomly allocated one sector ... to the intervention group and the other to the control group ... In the intervention group the windows of all 241 houses ... were covered with loosely hanging polyester curtains impregnated with the pyrethroid insecticide ... In the 222 houses in six of the control sectors the windows were covered with non-impregnated curtains and in one randomly selected control sector with 106 houses no curtains were provided. Explanation

Again, if the intervention was targeted at the cluster level, specific details of how it was administered should be described.

Objectives

Item 5: Specific objectives and hypotheses, and whether they pertain to the individual level, the cluster level, or both.

Example

We aimed to compare the effectiveness of three different interventions for improving the secondary preventive care for patients with coronary heart disease delivered at the level of general practice.
district and less than 50 km apart); and prior STD attendance rates at the health centre.  

Explanation
Cluster randomised trials may use a simple, completely randomised design; a matched cluster design; or a stratified design. In individually randomised trials random assignment generally ensures that any baseline differences in group characteristics are the result of chance rather than some systematic bias. This cannot be assumed, however, for the cluster randomised trial.

Although the assumption holds for cluster specific characteristics (that is, characteristics of the randomly allocated clusters), the researcher has little control over the individuals within each cluster and the number of clusters is usually relatively small. As a result, some form of constraint (matching or stratification) is often imposed on randomisation in a cluster randomised design in an attempt to minimise imbalance across treatment groups. Any constraint imposed on the cluster randomised trial affects the sample size and the analysis and thus should be reported.

Allocation concealment
Item 9: Method used to implement the random allocation sequence, specifying that allocation was based on clusters rather than individuals and clarifying whether the sequence was concealed until interventions were assigned.

Example
Practices agreeing to participate were … assigned by simple random allocation to use the computer decision support system … Randomisation was performed with a table of random numbers by a researcher not involved in the study and who was blind to the identity of the practices.

Explanation
In individually randomised trials, adequate concealment of the treatment allocation is crucial to minimising potential bias. If the person recruiting participants has foreknowledge of the allocation, bias can result. In a cluster randomised trial, allocation of treatment is predetermined for each member of the cluster. Hence the potential for selection bias (selective inclusion of patients into the trial) within clusters is particularly high. It is, therefore, particularly important that authors outline any strategies that were implemented to minimise the possibility of selection bias—for example, whether all patients within a cluster were included, or, if not, whether recruitment of patients was by a person masked to the cluster allocation.

Statistical methods
Item 12: Statistical methods used to compare groups for primary outcome(s) indicating how clustering was taken into account; methods for additional analyses, such as subgroup analyses and adjusted analyses.

Examples
Because we randomised obstetric units … we analysed rates of marker clinical practices by obstetric units. We used cluster specific methods because practices rather than patients were randomised … We used hierarchical logistic regression.

Table 2  Example of baseline information for each group at individual and cluster levels (adapted from Feder et al)

<table>
<thead>
<tr>
<th>Practice factors at baseline</th>
<th>Intervention group</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>25</td>
<td>27</td>
</tr>
<tr>
<td>List size:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low (&lt;1600)</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Medium (1600-2200)</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>High (&gt;2200)</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>Practice nurse:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>Yes</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>Patient factors at baseline</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>172</td>
<td>156</td>
</tr>
<tr>
<td>Mean age (years)</td>
<td>56.5</td>
<td>56.8</td>
</tr>
<tr>
<td>No (%) of smokers</td>
<td>68/161 (42)</td>
<td>49/141 (35)</td>
</tr>
<tr>
<td>No (%) with myocardial infarction</td>
<td>103/172 (60)</td>
<td>91/156 (58)</td>
</tr>
</tbody>
</table>

Results
Participant flow
Item 13: Flow of clusters and individual participants through each stage (a diagram is strongly recommended). Specifically, for each group report the numbers of clusters and participants randomly assigned, receiving intended treatment, completing the study protocol, and analysed for the primary outcome. Describe protocol deviations from study as planned, together with reasons.

Explanation
Knowing how many clusters and individuals did not receive the intervention as planned, the proportion dropping out, and the proportion for whom follow up data were not available is essential to interpret the study accurately. For example, individuals excluded after random assignment may not be representative of all participants in the study, and different drop-out rates may be directly related to the treatment received (if, say, one treatment has more severe side effects). The potential for differential adherence and follow up is exacerbated in the cluster randomised design because there are two levels at which drop-outs can occur: whole clusters or individuals in a cluster. It is therefore important to describe the flow of both clusters and individuals when reporting a cluster randomised trial. A flow diagram is usually the best way to present this information (see below).
Table 3  Example of study including data on numbers analysed by cluster and intracluster correlation coefficients (adapted from Feder et al10)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Intervention</th>
<th>Control</th>
<th>Intracluster correlation coefficient</th>
<th>Adjusted odds ratio (95% CI)</th>
<th>Adjusted $\chi^2$ statistic</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of practices</td>
<td>25</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No of patients</td>
<td>172</td>
<td>156</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advice given:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cholesterol</td>
<td>34/81 (67)</td>
<td>32/83 (39)</td>
<td>0.013</td>
<td>4.6 (1.9 to 8.2)</td>
<td>12.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Weight</td>
<td>24/69 (44)</td>
<td>22/73 (31)</td>
<td>0.098</td>
<td>3.0 (1.3 to 6.8)</td>
<td>10.5</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Diet</td>
<td>46/169 (27)</td>
<td>22/154 (14)</td>
<td>0.053</td>
<td>2.4 (1.2 to 4.7)</td>
<td>6.2</td>
<td>&lt;0.05</td>
</tr>
</tbody>
</table>

Example
See table 2.

Explanation
Random assignment by individual ensures that any differences in group characteristics at baseline are the result of chance rather than some systematic bias. This assumption does not hold, however, for cluster randomised trials. It is important to present summary baseline information for both clusters and individuals, most simply as tables of summary data.

Numbers analysed
Item 16: Number of clusters and participants (denominator) in each group included in each analysis and whether the analysis was by intention to treat. State the results in absolute numbers when feasible (eg 10/20 not 50%).

Example
See table 3.

Explanation
The number of participants who contribute to the analysis of a trial is essential to interpreting the results.

Not all participants, however, may contribute to the analysis of each outcome. In a cluster trial, this fact is compounded by the possibility that not all clusters may contribute to a particular analysis. Because the sample size calculation, and hence the power of the study, is calculated on the assumption that all participants and (especially) all clusters will provide information, the number of participants and clusters contributing to a particular analysis should be reported so that any potential drop in statistical power can be assessed. The flow diagram can include this information if there is only one primary outcome; otherwise the numbers of participants and clusters contributing should be summarised for each outcome.

Outcomes and estimation
Item 17: For each primary and secondary outcome, a summary of results for each group for the individual or cluster level as applicable, and the estimated effect size and its precision (eg 95% confidence interval) and a coefficient of intracluster correlation (ICC or k) for each primary outcome.

Example
See table 3.

Explanation
When reporting the results of a cluster randomised trial, point estimates with confidence intervals should be reported for primary outcomes. Given the impact of the extent of the intracluster correlation on the power of the study, the intracluster correlation coefficient or k statistic, for each outcome being analysed should also be provided. This information will allow readers to assess the appropriateness of the original sample size calculations as well as the magnitude of the clustering for each outcome. Showing both adjusted and unadjusted estimates would provide another indication of the extent of the clustering. Several authors have advocated publishing intracluster correlation coefficients to allow them to inform the development of future cluster trials in similar settings.

Discussion
Item 21: Generalisability (external validity) to individuals and/or clusters (as relevant) of the trial findings.

Example
Although our trial was completed successfully from both a methodological and practical point of view, our results may not be generalisable. The 21 participating practices tended to be large, with good nursing support, and may have been particularly committed to improving their quality of care…
Furthermore, the observed intervention effect would probably have been greater if the trial had not taken place in the context of a health authority audit initiative relating to patients with coronary heart disease, backed by a financial incentive.\textsuperscript{80}

\textbf{Explanation}

In the discussion section of any trial report, the external validity of the results should be considered. External validity is more complicated for cluster randomised trials because the results may be generalisable to the clusters, to the individuals in those clusters, or to both, and thus the level at which external validity is addressed should be identified.

\textbf{Flow diagrams}

We previously presented three options for modifying the CONSORT flow diagram for presenting clustered data: presenting the flow of data based only on clusters, only on individual participants, or both.\textsuperscript{4} Further experience suggests that the type of diagram should depend on the type of analysis because different approaches to analysis require information at different levels of the clustered design.

For example, if the analysis is aggregated at the level of the cluster, the flow diagram should relate to the cluster level data (fig 1). To allow meaningful interpretation, a measure of the cluster size (and an indication of how variable cluster sizes are) also needs to be included in the diagram. If, however, the analysis is multilevel or hierarchical, the flow diagram should present data flow for both clusters and individual participants (fig 2).

Although we recommend this diagram for communicating the flow of clusters and participants throughout the study, the exact form and content of the flow diagram should be varied to present the specific features of a trial.

\textbf{Comments}

Reports of cluster trials should include key information on the design and analysis to allow readers to accurately interpret the results. This information is also particularly important for meta-analysts attempting to extract data from such reports.\textsuperscript{40} We therefore recommend that journal editors include our guidelines in their instructions to authors.

Inadequate methodological reporting of trials has been shown to be associated with bias in the estimate of treatment effects.\textsuperscript{82} Use of the CONSORT statement for the reporting of two group parallel trials is associated with improved reporting quality.\textsuperscript{101} We believe that the routine use of this proposed extension to the CONSORT statement will result in similar improvements for cluster trials.

The CONSORT group is also developing modified recommendations to help improve the quality of reporting of clinical trials of other designs, including equivalence trials, crossover trials and multi-arm trials. The most up to date versions of all CONSORT recommendations can be found at www.consortstatement.org. We thank the members of the CONSORT Group, especially David Moher, Ken Schulz, Tom Lang, David Sackett, Peter Gotzsche, Matthias Egger, and John Ioannidis for helpful comments on earlier drafts. The Health Services Research Unit is funded by the Chief Scientist’s Office of the Scottish Executive Health Department. DGA is supported by Cancer Research UK. The views expressed are not necessarily those of the funding bodies.

Competing interests: None declared.
Education and debate

Answer

Ireland general practitioner, Connemara, Republic of John McCormack,
Blowing up a balloon is great for cooperative kids.

Solihull

What options are available to relieve ear blockage due to flight

addition to relieving that dreaded pain, it also keeps you

11 Connelly LB. Balancing the number and size of sites: an economic
research

13 Hutton J. Are distinctive ethical principles required for cluster

15 Simpson JM, Klar N, Donner A. Accounting for cluster randomization: a

14 Edwards S Braunholtz DA, Lilford RJ, Stevens AJ. Ethical issues in the

16 Divine GW, Brown JT, Frazier LM. The unit of analysis error in studies

17 Donner A, Brown KS, Brasher P. A methodological review of


23 World Medical Association. Recommendations guiding physicians in

24 Atienza AA, King AC. Community-based health intervention trials: an


(Question

What options are available to relieve ear blockage due to flight
descents, apart from chewing gum, taking Sudafed tablets if you
have a cold, or yawning?

Parthasarathy Ramesh, consultant physician, Solihull Hospital, Solihull

Answer

Blowing up a balloon is great for cooperative kids.

John McCormack, general practitioner, Connemara, Republic of Ireland

Answer

Nothing scientific here—drink water. I find that, if I drink water as
I ascend or descend on my flights, it helps dramatically. In
addition to relieving that dreaded pain, it also keeps you
hydrated. Eureka.

Aaron J Baxter, medical student, Saba University, Dutch Antilles

Answer

The “pinch and blow” method is used by descending scuba divers and is just as effective for descending airline passengers. Occlude both nostrils by squeezing the end of your nose, then blow hard
through your nose (no air or, er, mucus will escape from your
nose if you’re squeezing hard enough). You’ll feel one ear pop,
through your nose (no air or, er, mucus will escape from your
nose if you’re squeezing hard enough). You’ll feel one ear pop,

Theo Fenton, consultant paediatrician, Mayday Hospital, Croydon

http://bmj.com/cgi/q&a-display/short/bmj_el;42930

These exchanges were posted on the Q&A section of bmj.com. If you want to respond to the question, or ask a new question of your own, follow the link above or go to http://bmj.com/q&a