

Research article

Open Access

Intracluster correlation coefficients in cluster randomized trials: empirical insights into how should they be reported

Marion K Campbell*¹, Jeremy M Grimshaw² and Diana R Elbourne³

Address: ¹Health Services Research Unit, University of Aberdeen, Aberdeen, UK, ²Ottawa Health Research Institute, Ottawa, Canada and ³Medical Statistics Unit, London School of Hygiene and Tropical Medicine, London, UK

Email: Marion K Campbell* - m.k.campbell@abdn.ac.uk; Jeremy M Grimshaw - jgrimshaw@ohri.ca; Diana R Elbourne - diana.elbourne@lshtm.ac.uk

* Corresponding author

Published: 28 April 2004

Received: 28 January 2004

BMC Medical Research Methodology 2004, 4:9

Accepted: 28 April 2004

This article is available from: <http://www.biomedcentral.com/1471-2288/4/9>

© 2004 Campbell et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: Increasingly, researchers are recognizing that there are many situations where the use of a cluster randomized trial may be more appropriate than an individually randomized trial. Similarly, the need for appropriate standards of reporting of cluster trials is more widely acknowledged.

Methods: In this paper, we describe the results of a survey to inform the appropriate reporting of the intracluster correlation coefficient (ICC) – the statistical measure of the clustering effect associated with a cluster randomized trial.

Results: We identified three dimensions that should be considered when reporting an ICC – a description of the dataset (including characteristics of the outcome and the intervention), information on how the ICC was calculated, and information on the precision of the ICC.

Conclusions: This paper demonstrates the development of a framework for the reporting of ICCs. If adopted into routine practice, it has the potential to facilitate the interpretation of the cluster trial being reported and should help the development of new trials in the area.

Background

In evaluative health care research, the randomized controlled trial is generally considered the gold standard design for assessing the relative effectiveness of alternative interventions, as it ensures that selection bias and other common sources of bias are minimized [1]. In most of these trials, patients are allocated individually to the different treatments. It is being increasingly recognised, however, that there are many situations where randomizing by groups of individuals may be more appropriate [2,3]. For example, when assessing a dietary intervention, it is common to randomize families as an intact unit, to avoid the possibility of different members of the same family being

assigned to different interventions. Trials that randomize by groups are known as cluster randomized trials [3,4].

The adoption of a clustered design is not without cost, however, as the design and conduct of these trials requires special considerations. They are more complex to design, they require more participants than an individually randomized trial, and require more complex analysis. This added complexity arises primarily because observations on individuals within the same cluster may be correlated – that is, the outcomes for individuals within clusters are likely to be more similar than those across clusters. The statistical measure of this 'clustering effect' is known as the

intracluster correlation coefficient, or ICC [5]. The ICC for a particular outcome can be described as the amount of variation in the dataset that can be explained by the variation between clusters.

The need for clear reporting of randomized trials has been widely recognised. This has been highlighted through the publication of the CONSORT statement, which outlines the common standards for the reporting of trials [6], and through the revision to the statement published in 2001 [7]. The majority of the CONSORT work to date, however, has been for individually randomized trials. Recent studies have, however, shown problems with the reporting of cluster trials [8-10]. For example, ICCs were reported in only 6/149 trials in the MacLennan study [8], in only 13/152 trials in the Eldridge study [9] and in only 1/51 trials in the Isaakidis study [10]. Two recent papers have therefore considered extensions to the CONSORT Statement for the reporting of cluster trials [11,12].

To aid the interpretation of the results of cluster trials, it is important that appropriate information is presented in their reports. The ICC for each trial outcome is widely acknowledged to be one of the primary factors which helps researchers interpret the results of the trial. It also helps researchers planning new trials in the area, as estimates of ICCs are used to inform sample size calculations – standard sample size calculations require to be inflated by a factor:

$$1+(n-1)\rho$$

(where n is the average cluster size and ρ is the ICC for the desired outcome) to accommodate for the lack of independence in the data [3]. This factor is commonly called the 'variance inflation factor' or the 'design effect' [3]. In response to this, there have been calls to increase the reporting of ICCs within the reports of cluster trials [13,14]. Whilst the reporting of ICCs would appear to be increasing (for example through several recent publications of collections of ICCs [14-16]), there is, however, little consistency in the information presented to describe the ICC.

The aim of this study, therefore, was to determine the most appropriate descriptors for the reporting of an ICC, by surveying researchers working in the field of cluster trials.

Methods

Since 1998, at least three international workshops on cluster randomized trials have been held in the UK, comprising researchers and statisticians involved in the design and conduct of cluster trials. Across the three workshops, 119 researchers and statisticians with an interest in cluster

trials either attended, or expressed a wish to attend. These individuals formed the primary population for the survey.

In addition, the UK Directory of Academic Statisticians 2001 [17] was searched for those expressing a special interest in the field of cluster trials. This identified 11 statisticians with an interest in the area. Ten of these were already identified for inclusion in the survey; hence only one additional statistician was identified through this approach. This gave a final survey population of 120 individuals, representing researchers and statisticians from 11 countries (104 individuals from the UK; 3 from the Netherlands; 2 each from Denmark, Finland and South Africa; and 1 each from Norway, Switzerland, Spain, Canada, USA, Australia and Thailand), which included known leaders in the field.

A small focus group was held to develop the framing of a questionnaire to identify descriptors of an ICC. This group consisted of four statisticians and two triallists who were experienced in the design and conduct of cluster trials. The group felt that the inclusion of examples would facilitate completion of the questionnaire, and generated a small number of factors which might affect the interpretation of the ICC. Examples included:

- the average cluster size (which is linked with the ICC in generating sample size calculations);
- the number of clusters in the dataset;
- the setting (whether community or hospital);
- the country/countries involved in the trial;
- the disease or specialty grouping; and
- the type of outcome – whether process or outcome.

The survey population was then sent a postal questionnaire [see Additional File 1] and participants were encouraged to identify factors which they thought should be used to describe an ICC. All questions in the questionnaire were of 'open' format (that is, allowing free-text input) rather than tick-box format. Multiple suggestions were permissible, and all free-text responses were subsequently coded by the same researcher (MKC). A reminder questionnaire was sent out four weeks after dispatch, if no response was obtained to the initial mailing.

Results

Of the 120 questionnaires dispatched, 78 were completed (and a further questionnaire was returned as the named individual was no longer at the listed address). This gave

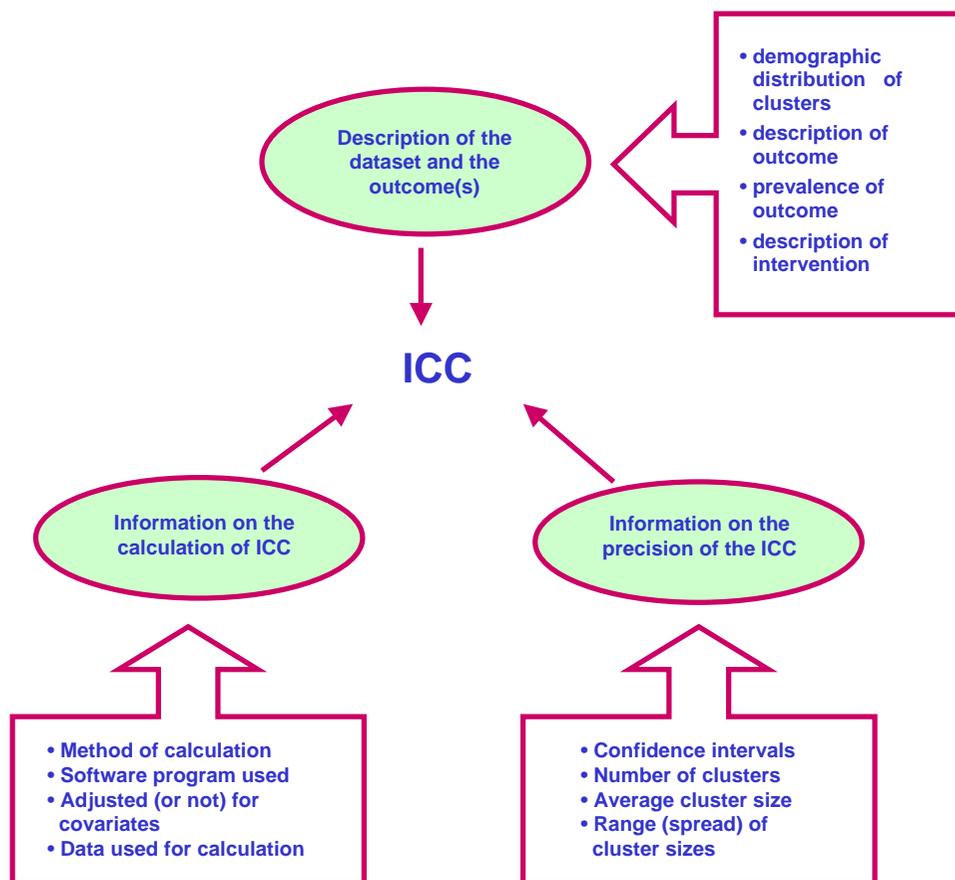


Figure 1
 Framework for describing ICCs

an adjusted response rate to the questionnaire of 66% (78/119).

One hundred suggestions for appropriate descriptors of an ICC were received (representing 34 separate components). From these descriptors, it was possible *a posteriori* to conceptualize the majority of these into three distinct areas (Figure 1):

- a) a description of the dataset and the outcome(s) (67 responses);
- b) information on the method of calculation of the ICC (18 responses); and
- c) information on the precision of the ICC (12 responses).

Description of the dataset

Respondents identified three main elements that they thought should be used to describe the dataset:

- the demographic distribution within and between clusters (e.g. age, sex and ethnic distribution). It was thought that different distributions across clusters could affect the ICC.
- a description of the outcome – was it binary or continuous, what was the underlying prevalence of the outcome and was the outcome measured subjectively or objectively? It was thought that outcomes that were measured subjectively (eg physician assessment of well-being) were likely to display greater clustering than those measured objectively (eg laboratory-processed blood results).
- a description of the intervention. Individuals felt that the interventions should be well described, in order to identify whether specific components of an intervention might be expected to have undue influence on the ICC. For example, if health professional were the unit of allocation and an intervention was designed to make health professionals' practice more consistent, one might expect

that an ICC based on data in the post-intervention phase would be smaller than an ICC based on the pre-intervention data (if the intervention was successful, it would induce less variation between clusters, as health professionals would be more consistent after the intervention).

Calculation of the ICC

The second important dimension identified was the provision of information about how the ICC was calculated. Within this, people felt it would be important to know:

- what method had been used to calculate the ICC (for example, whether by ANOVA or by some other method);
- the software program used to calculate the ICC, as different packages can give different results (not all packages use the same definition of average cluster size [13]);
- what data were used to calculate the ICC – for example whether the ICC was calculated from control data only, from both control and intervention data, using pre-intervention data only or post intervention; and
- whether any adjustment had been made for covariates when the ICC had been calculated. Adjusting for covariates leads, in general, to smaller ICCs, as some of the between cluster variation may be explained by cluster level factors [18].

Precision of the ICC

The final dimension suggested was some information on the precision of the ICC estimate. Confidence intervals were suggested as the primary mechanism for providing information on the precision of the ICC, but the number of clusters together with the average cluster size and the range of cluster sizes (or some other measure of spread) were also put forward as appropriate descriptors.

Discussion

Cluster randomized trials are increasingly being used in the health care field, and there is a need to ensure that the information presented in their reports is useful and aids interpretation of the trial. Within cluster trials, ICCs for trial outcomes are widely recognised as key items of information that should be reported [3,13]. To date, however, there has been little information to guide researchers on how to describe the ICC, resulting in different researchers providing different information in trial reports. In many trials, if ICCs have been reported at all, authors simply provide point estimates of the observed ICCs and no other information. In contrast a few authors have provided comprehensive information, allowing readers to interpret the data easily. For example, Gulliford and colleagues reported extensive information to aid the interpretation of ICC estimates presented including the setting, the type of

cluster, the average cluster size, the number of clusters, the prevalence of the outcome, the method of calculation, the separate variance components and the design effect [15]. Similarly, Smeeth and Ng [16] provided extensive information on the dataset their ICCs were calculated from and also included information on prevalence of the outcomes, the standard error of each ICC, the average cluster size for each outcome, and the separate within- and between-components of variation.

The results of the survey outlined in this paper have shown that there is consensus amongst researchers in the field as to which factors are important when describing an ICC. We have presented a potential framework for the reporting of ICCs, linking three main dimensions – descriptions of the dataset, the method of ICC calculation and the precision of the ICC estimate. The framework is also likely to be straightforward to adopt in practice as the data requirements outlined within the framework are generally readily available. The inclusion of extensive data about the dataset may require additional journal reporting space but with the advent of web-based supplementation of journal articles this should not pose additional burden on authors or journal editors.

The study reported here is not without limitations, however. Whilst every effort was made to identify a comprehensive sampling frame, we accept that our survey population was relatively opportunistic, although there was widespread representation from different countries, and known leaders in the field were surveyed.

Open-text format was used in the questionnaire to allow respondents to think expansively and not be constrained by a potentially non-representative framework. By using an open-text format, however, we cannot be sure that each respondent detailed an exhaustive list of suggestions. We also cannot attribute any ranking to the relative importance of each suggestion both within respondent and across respondents. If time and resources could have allowed, it would have been beneficial if all the suggestions received could have been listed on a second questionnaire and dispatched to all participants for consensus and ranking. In addition, a single researcher coded the suggestions received. This could have inadvertently resulted in some bias being introduced into the interpretation of free-text responses. By having a single coder, however, we can be reassured that suggestions were coded in a consistent fashion.

The focus of the survey was on the reporting of the traditional ICC rather than the related statistic, k , which is sometimes used in sample size calculations and analyses of cluster randomized trials. This k -statistic has been suggested for outcomes that are expressed in terms of patient-

years [19]. The traditional ICC was the focus of this survey as it is the more commonly used statistic in the reporting of cluster randomized trials. A further survey of statisticians for whom the k -statistic is particularly relevant could, however, yield different results.

The introduction of the CONSORT statement has been instrumental in improving the standards of reporting in clinical trials [20,21]. It has been widely adopted by medical journals worldwide including both general and disease-specific journals. It is important, therefore, that the factors identified from this survey are considered in any statement relating to cluster randomized trials.

Conclusions

In conclusion, this paper has demonstrated the development of a comprehensive framework for the reporting of ICCs. If adopted into routine practice, it has the potential to facilitate the interpretation of the cluster trial being reported and should help the development of new trials in the area. Further research into the descriptors most appropriate to the reporting of the related k -statistic would be beneficial.

Competing interests

None declared.

Authors' contributions

MKC and JMG conceived the idea for the study. MKC carried out the survey, coded responses and analysed the data. MKC wrote the first draft of the paper. MKC, JMG and DRE were all involved in the development and refinement of subsequent drafts. MKC is the guarantor for the study.

Additional material

Additional File 1

Questionnaire sent to participants

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2288-4-9-S1.doc>]

Acknowledgements

The Health Services Research Unit is funded by the Chief Scientist Office of the Scottish Executive Health Department. The views expressed are not necessarily those of the funding body.

References

- Pocock SJ: *Clinical trials: a practical approach* Chichester: Wiley; 1983.
- Campbell MK, Grimshaw JM: **Cluster randomised trials: time for improvement.** *BMJ* 1998, **317**:1171-2.
- Donner A, Klar N: *Design and analysis of cluster randomization trials in health research* New York: Arnold; 2000.
- Murray DM: *The design and analysis of group randomized trials* Oxford: Oxford University Press; 1998.
- Donner A, Koval JJ: **Design considerations in the estimation of the intraclass correlation.** *Ann Hum Genet* 1982, **46**:271-277.
- Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, Rennie D, Schulz KF, Simel DL, Stroup DF: **Improving the quality of reporting of randomized controlled trials. The CONSORT Statement.** *JAMA* 1996, **276**:637-9.
- Moher D, Schultz K, Altman DG, for the CONSORT Group: **The CONSORT statement: revised recommendation for improving the quality of reports of parallel-group randomised trials.** *Lancet* 2001, **357**:1191-94.
- MacLennan GS, Ramsay CR, Mollison J, Campbell MK, Grimshaw JM, Thomas RE: **Room for improvement in the reporting of cluster randomised trials in behaviour change research [abstract].** *Control Clin Trials* 2003, **24**:695-705.
- Eldridge SM, Ashby D, Feder G, Rudnicka AR, Ukoumunne OC: **Lessons for cluster randomized trials in the twenty-first century: a systematic review of trials in primary care.** *Clinical Trials* 2004, **1**:80-90.
- Isaakidis P, Ioannidis JPA: **Evaluation of cluster randomized controlled trials in sub-Saharan Africa.** *Am J Epidemiol* 2003, **158**:921-6.
- Elbourne D, Campbell MK: **Extending the CONSORT statement to cluster randomised trials: for discussion.** *Stat Med* 2001, **20**:489-496.
- Campbell MK, Elbourne DR, Altman DG, for the CONSORT group: **The CONSORT statement: extension to cluster randomised trials.** *BMJ* 2004, **328**:702-8.
- Ukoumunne OC, Gulliford MC, Chinn S, Sterne JA, Burney PG: **Methods for evaluating area-wide and organization-based interventions in health and health care: a systematic review.** *Health Technol Assess* 1999, **3**(5):iii-92.
- Campbell MK, Grimshaw JM, Steen N, for the Changing Professional Practice in Europe Group: **Sample size calculations for cluster randomised trials.** *J Health Serv Res Policy* 2000, **5**:12-16.
- Gulliford M, Ukoumunne O, Chinn S: **Components of variance and intraclass correlations for the design of community based surveys and intervention studies: data from the health survey for England 1994.** *Am J Epidemiol* 1999, **149**:876-883.
- Smeeth L, Ng E: **Intraclass correlation coefficients for cluster randomized trials in primary care: data from the MRC trial of the assessment and management of older people in the community.** *Control Clin Trials* 2002, **23**:409-421.
- Directory of Academic Statisticians** [<http://www.swan.ac.uk/statistics/das/>]
- Feng Z, Diehr P, Yasui Y, Evans B, Beresford S, Koepsell TD: **Explaining community-level variance in group randomized trials.** *Stat Med* 1999, **18**:539-556.
- Hayes RJ, Bennett S: **Simple sample size calculation for cluster-randomized trials.** *Int J Epidemiol* 1999, **28**:319-26.
- Moher D, Jones A, Lepage L, for the CONSORT Group: **Use of CONSORT Statement and quality of reports of randomized trials: a comparative before-and-after evaluation.** *JAMA* 2001, **285**:1992-5.
- Egger M, Juni P, Bartlett C, for the CONSORT Group: **Value of flow diagrams in reports of randomized controlled trials.** *JAMA* 2001, **285**:1996-9.

Pre-publication history

The pre-publication history for this paper can be accessed here:

<http://www.biomedcentral.com/1471-2288/4/9/prepub>