July 12, 2014

Protocol: Consort extension to stepped wedge cluster randomised controlled trial

Version 1

Investigators

Karla Hemming[1] *

Alan Girling[1]

Terry Haines[2]

Richard Lilford[3]

[1] School of Health and Population Sciences, University of Birmingham, Birmingham, B15 2TT, UK

[2] Monash University, Victoria 3800, Australia

[3] University of Warwick, Coventry CV4 7AL, UK

Correspondence email address: k.hemming@bham.ac.uk

Contents

Background

The stepped-wedge cluster randomised trial (SW-CRT) is a novel research study design that is increasingly being used in the evaluation of service delivery type interventions[12]. The SW-CRT is a type of randomised trial in which clusters are randomised to a date at which they initiate the intervention under evaluation. The design involves random and sequential crossover of clusters from control to intervention, until all clusters are exposed[3].

It is a pragmatic study design which can reconcile the need for robust evaluations with political or logistical constraints. Whilst not exclusively for the evaluation of service delivery intervention it is particularly suited to evaluations that do not rely on individual patient recruitment.

The SW-CRT offers a randomised method of evaluation of an intervention delivered at the level of the cluster. In cases where randomisation to either control or intervention arm is precluded, it offers a means of a randomised evaluation in place of a non-randomised evaluation. The design can also be used as an alternative to the conventional parallel cluster design (or one of its variations).

How the stepped wedge design relates to other cluster studies

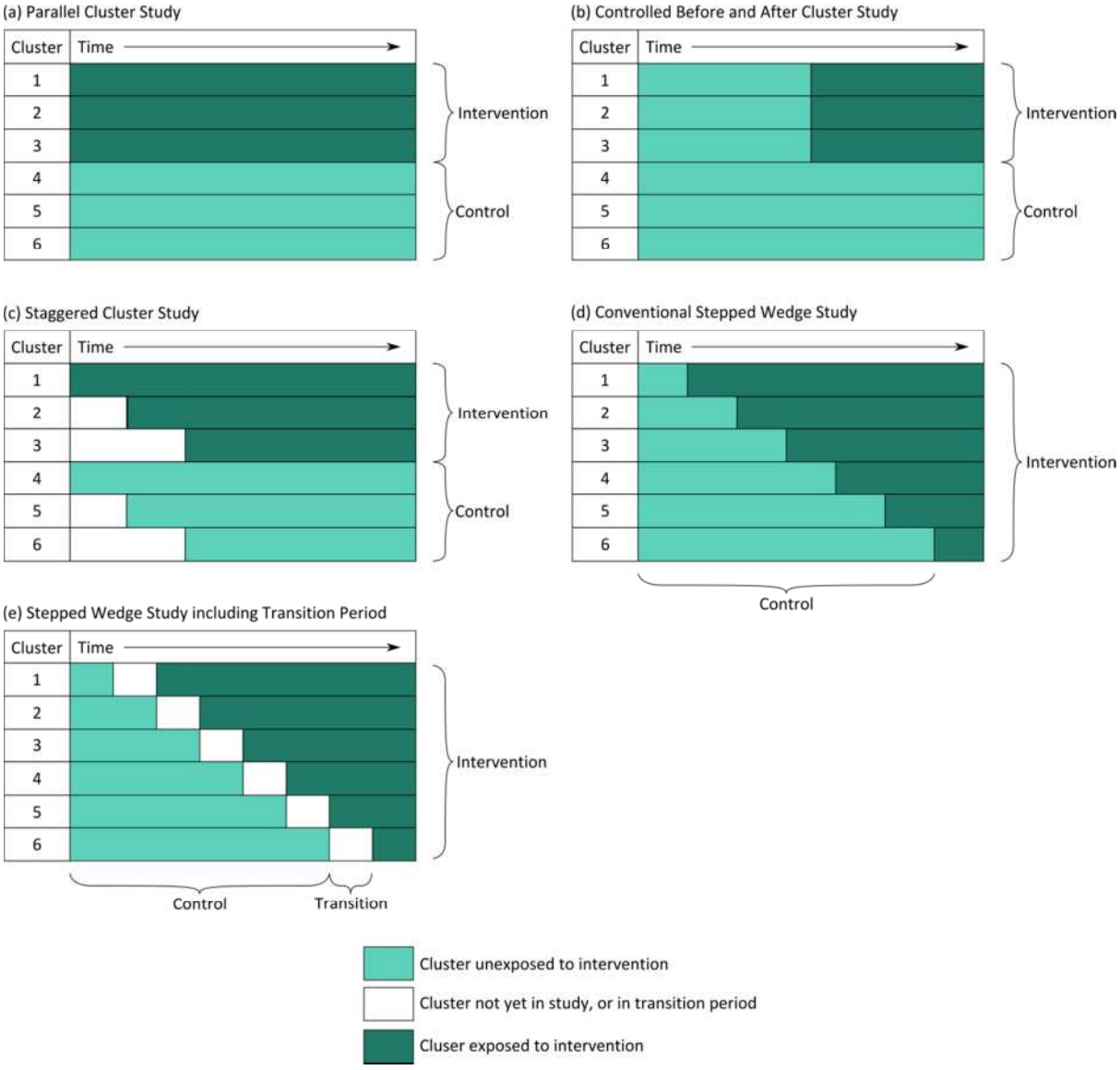
In the evaluation of interventions delivered at the level of a general practice, ward or hospital, in which it is not possible to randomise individuals, randomisation is carried out at the level of the cluster (i.e. ward or general practice). There are broadly three types of cluster trials to choose from, illustrated in Figure 1. In the conventional (parallel) cluster randomised trial, clusters are randomised either to the intervention or control arm at the beginning of the trial and remain in that arm for the duration of the study – we shall refer to this as a 'simple' parallel cluster trial (Figure 1a). This design may be elaborated into a controlled before-and –after cluster (randomised) trial (sometimes called an ANCOVA design) – which was what was effectively used in Example 2, the evaluation of the Mexican health insurance. In such a trial half of the clusters switch from control to intervention arm at one point in time, but with observations taken both before and after the switch takes place (Figure 1b). In the SW-CRT, this is extended so that every cluster provides before and after observations and every cluster switches from control to become exposed to the intervention, but not at the same point in time (Figure 1d). The stepped-wedge study takes its name from the wedge shape stepping apparent in the schematic illustrations.

The stepped wedge study can also be individually randomised. We are not concerned with individually randomised stepped wedge studies here as they require consideration of some quite different issues.

The stepped wedge design

In the SW-CRT there is usually a period of baseline data collection, in which no clusters are exposed to the intervention, then subsequently at periodic time points (called the steps), one or a number of clusters are randomised to cross from control to intervention, whilst the

remaining clusters remain unexposed. The study continues until all clusters have crossed to the intervention arm, and there is usually a period at the end of the study in which all clusters are exposed to the intervention. The SW-CRT, can be viewed an extension of the cluster trial, with clusters followed longitudinally over time but with the addition that clusters are randomised sequentially to cross (at different points in time) from control to intervention.



What is known about the quality of reporting of the SW-CRT

Early cluster randomised controlled trials were poorly reported and under-powered[4]. With hindsight we know that many early cluster trials were conducted and reported as if they were individually randomised controlled trials. This lent itself to trials which over estimated the precision to which the treatment effect was estimated. Over the years the quality of reporting and conducting parallel

cluster trials has improved[5]. This has manifested itself in cluster trials which are designed appropriately allowing for the clustering; and analysed allowing for this clustering. Recent cluster trials are more likely to allow for uncertainty in the estimate of the intra-cluster correlation (ICC) and allowance for varying cluster size.

The SW-CRT is a new form of study design which although a form of the cluster randomised controlled trial, has many distinct features, which mean that some design and analysis issues are different. For example, both the design effect and the method of analysis differ to the conventional parallel trial. The Gambia hepatitis intervention study (Box 1) is the probably the earliest and most widely known stepped wedge study[6]. However, even this study protocol contains very limited information on how the data will be analysed and insufficient detail to allow readers to replicate the power calculation.

It is therefore very likely that without tailored reporting guidelines, the stepped wedge cluster trial is in danger of being poorly reported and undertaken. This would be very unfortunate since the SW-CRT is likely to be the design of choice in very pragmatic situations, when the alternative would have been an observational study design (such as controlled before and after). We therefore propose to tailor the existing reporting guidelines for the parallel cluster randomised trial to produce a guideline for the reporting of the SW-CRT.

There have been two systematic reviews of SW-CRTs1,[2]. These systematic reviews were broadly concerned with identifying the number of breadth of stepped wedge studies, rather than being systematic reviews of quality of reporting. The latest of these reviews identified 10 protocols for SW-CRTs and 15 completed study publications. The breadth of coverage was wide spanning such diverse areas of application such as interventions for public health promotion in developing countries, education and improvements to housing.

Related EQUATOR reporting standards

Guideline name		How this guideline relates to the reporting of the SW-CRT	How the reporting of the SW-CRT differs to the guideline
Consort 2010[7]	Reporting of individually randomised controlled trials	This guideline covers the core standards needed for reporting a randomised controlled trial.	Cluster trials have several distinguishing features – such as defining the cluster, acknowledgement of the cluster in the power and analysis, concealment of allocation from both the cluster and individual, which are important in the reporting. The SW-CRT shares these differences.
Consort 2010 cluster extension[8]	Reporting of parallel cluster randomised	This guideline covers the additional	The SW-CRT has several distinguishing features

	controlled trials	standards needed in the reporting of a parallel cluster trial.	from the parallel cluster trial – such as defining the cluster and the time period; acknowledgement of the cluster and temporal time trends in the power and analysis; concealment of allocation from both the cluster and individual over the duration of the study.
Consort extension for pragmatic trials[9]	Reporting of randomised trials where the aim is to inform whether the intervention works in normal practice.	SW-CRTs are often used to evaluate interventions for which there is already evidence of support that the intervention works under idealised conditions (efficacious) and interest lies in establishing whether the intervention works in normal practice.	The consort extension for pragmatic trials does not explicitly consider cluster trials neither stepped wedge trials.
Consort extension for non-pharmacologic treatments[10]	Reporting of randomised trials in which the intervention is a non-pharmacologic treatments.	SW-CRTs are often used to evaluate non-pharmacological interventions – and so many of the issues in this guideline, such as details of blinding, are relevant.	The consort extension for non-pharmacologic treatments does not explicitly consider cluster randomised trials (though it does consider the aspect of clustering within health care providers) neither stepped wedge trials.
STROBE [11]	Reporting of observational studies	For many interventions evaluated by the SW-CRT, an alternative pragmatic evaluation study design is an observational study.	Observational studies include the controlled before and after study, of which the SW-CRT is closely related, but includes the added component of randomisation and the feature that all clusters eventually become exposed to the intervention.

Aims and Objectives

Our overarching aim is to establish an extension to the reporting guidelines for parallel cluster trials, and so produce recommended reporting guidelines for the stepped wedge cluster randomised controlled trial. To this end, we propose the following objectives:

Objective 1: To conduct a systematic review of the literature to identify methodological papers on the design and analysis of the SW-CRT.

Objective 2: To conduct a Delphi exercise, with methodological experts and trialists, to obtain a consensus on concepts (standards) imperative in the good reporting of the SW-CRT.

Objective 3: To conduct a consensus workshop to finalise the standards, and wording, for inclusion in the extension of the cluster consort guideline for the SW-CRT.

Objective 4: To produce an extension to the consort reporting guideline for cluster trials to the SW-CRT with a document outlining rational.

Methods

Review of literature

We will carry out a systematic review of the literature to identify methodological papers on the SW-CRT. This will include, but will not be limited to, methodological papers and sample size, power, analysis, general design considerations, rationale, recommendations and any existing recommendations for reporting.

Many early SW-CRTs were sometimes described using other terms such as “waiting list designs” or “phased implementations”. We will therefore ensure that our search is sufficiently broad to capture any relevant papers described by a term other than “stepped wedge”. We will not limit our search to the medical literature, but include education, policy and the econometric literature for example.

We will tabulate recommendations for reporting, design and analysis from all identified papers. These recommendations will inform preliminary recommendation standards which will be included in Round 1 of the Delphi exercise (along with other standards identified through the means described below).

Establishment of an international expert steering committee

The steering committee will have three fundamental roles:

Role 1: The steering committee will recommend preliminary standards (Open round one of Delphi exercise) for the reporting guideline. These preliminary standards, will, in conjunction with those identified by the investigators, and through the systematic review, but used to inform the standards included in the Delphi Round 1.

Role 2: The steering committee will be asked to comment on all iterations of the Delphi exercises.

Role 3: The steering committee will be invited attend a consensus workshop in which the standards and wording will be agreed.

We will identify a steering committee which includes the following areas of expertise:

The SW-CRT is a relatively new study design, there are however some methodologists with an interest and expertise in for example the statistical aspects of this design (sample size, power and analysis) and others with methodological expertise in for example the rationale for the SW-CRT. We will seek to include these experts as part of the steering group.

Again because the SW-CRT is a relatively new study design, those non-methodologists, who have designed and reported SW-CRTs will likely be able to provide valuable information on pragmatic aspects such as lessons, learnt, what went well, and what did not. We will therefore seek to include on the steering panel some non-methodologists who have designed and reported SW-CRTs.

Cluster trials: the SW-CRT is a special form of the parallel cluster trial and so therefore shares many of the same issues which are important when reporting a cluster trial. It will therefore be important to draw on those who have expertise in the design and analysis of cluster trials.

Reporting guidelines: the writing of reporting guidelines is in many ways an expertise in itself. It will therefore be important to draw on the opinions of those who have experience in contributing to, or writing other reporting guidelines, especially those for randomised trials, cluster trials or pragmatic evaluations.

We will personally write to authors of the existing related guidelines, and other academics who are already known to us as having the above listed expertise.

Identification of potential participants for the Delphi Exercise

The participants in the Delphi exercise will be invited to provide their opinion on standards to be included in the reporting guideline. They will provide these opinions by online methods only and will not be invited to participate in any face to face meetings. We will therefore include in this group, any people identified as potential participants in the steering panel who declined due to time commitments and additionally seek to include academics who have any of the expertise listed above (under the expertise of the steering panel). This group will be to some extent self-selected, as we will use wider recruitment methods, for example, email circulation lists of relevant groups.

Delphi Exercise

The Delphi exercise will be carried out to obtain a preliminary consensus on standards to be included. This will be carried out in an iterative process: proposing standards, asking for opinions on the standards, modifying the standards and feeding back standards. These preliminary agreed standards will be taken forward to the consensus meeting. The exercise will be carried out electronically using a web-based method. All responses will be kept anonymous.

1. Delphi preliminary round
    - Recommendations for reporting, design and analysis identified through the methodological systematic review will be identified. These recommendations will inform preliminary recommendation standards which will be included in Round One of the Delphi exercise.
2. Delphi Open round exercise
    - The steering committee will be invited to recommend preliminary standards, which will again be included in Round One of Delphi exercise.
    - ? Include the wider group in this open round (i.e. take suggestions from everyone)

3. Synthesis of the open round and preliminary rounds
    - Standards suggested in the preliminary and open rounds will be synthesed into preliminary standards to be included in Round One of Delphi exercise.
4. Delphi Round One
    - The preliminary standards identified in the preliminary and open rounds, will be included in round one of the Delphi exercise.
    - Participants invited to participate in the Delphi exercise will be asked to score each of the potential standards, indicating their level of agreement that the standard is important.
    - Participants will be invited to provide text based comments on the terminology and wording used.
5. Scoring of Round One
    - The scores for each standard included in Round One will be summarised for each standard separately. Any text based comments will also be synthesied.
6. Delphi Round Two
    - In Round Two the summary scores and responses will be feedback to those taking part along with the original standards. Any standards clearly identified through Round One will be removed from Round Two.
    - Participants will again be asked to score each of the potential standards, indicating their level of agreement that the standard is important.
    - Participants will again be invited to provide text based comments on the terminology and wording used.
7. Scoring of Round Two
    - Round Two will be scored and synthesed in the same way as Round One. These standards and scores will be taken to the consensus workshop.

Consensus workshop

Members of the steering panel will be invited to participate in a consensus workshop. This work shop will take place in the UK and ideally be funded to allow face to face participation of all members of the steering committee. All standards identified by the Delphi exercise will be verified by the consensus panel. Those standards for which there is contention over either the inclusion of the standard or the wording of the standard will be discussed in detail. To this end, the findings from the Delphi exercise will be reported in a fair and un-prejudiced way – possibly by an independent person should any issue arise which the investigators feel strongly about. Such issues will be scored and revised much in the same way as an electronic Delphi exercise described above. If the steering group and investigators cannot reach a consensus for any of the standards, then this will be reported openly as such in the reporting guideline. This is not only an honest reflection of science, but also in acknowledgement that the history of the SW-CRT is in its infancy and so any reporting guideline produced now cannot be expected to not change in the future.

Production of reporting guideline

The reporting standard will include all of the aspects and wording as agreed at the consensus meeting. We will additionally produce a document explaining the rationale for all of the standards included and illustrate with examples from the literature, using the Consort standard formatting.

This rationale document will be circulated among the steering panel who will be invited to comment and contribute.

Time line

Months	Tasks
July 2014 to September 2014	Write protocol Register protocol with EQUATOR Recruit steering committee Identify participants for the Delphi Exercise
October 2014 to December 2014	Conduct the Open round of the Delphi Exercise
Jan 2016	Apply for funding for consensus workshop
January 2015 to March 2015	Collate responses of the Open Round
April 2015 to June 2015	Conduct Round 1 of the Delphi Exercise
July 2015 to September 2015	Collate responses of Round 1
October 2015 to December 2015	Conduct Round 2 of the Delphi Exercise
January 2016 to March 2016	Collate responses of Round 2
Summer 2016	Consensus Workshop
Autumn 2016	Draft reporting guideline
Winter 2017	Submit for publication

Organisation

Karla Hemming and Richard Lilford will oversee the project. Alan Girling will provide statistical expertise.

Funding

RJL and AJG acknowledges financial support for the submitted work from the National Institute for Health Research (NIHR) Collaborations for Leadership in Applied Health Research and Care for West Midlands (CLAHRC WM). TPH is supported by a Career Development Fellowship from the Australian National Health and Medical Research Council (1069758).

An application for funding will be made in request of support for the consensus workshop.

References

1 Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. *BMC Med Res Methodol.* 2006;**6**:54.

2 Mdege ND, Man MS, Taylor Nee Brown CA, Torgerson DJ. Systematic review of stepped wedge cluster randomized trials shows that design is particularly used to evaluate interventions during routine implementation. *J Clin Epidemiol.* 2011;64(9):936-48.

3 Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials.* 2007;28(2):182-91

4 Bland JM. Cluster randomised trials in the medical literature: two bibliometric surveys. *BMC Med Res Methodol.* 2004;4:21.

5 Ivers NM, Taljaard M, Dixon S, Bennett C, McRae A, Taleban J, Skea Z, Brehaut JC, Boruch RF, Eccles MP, Grimshaw JM, Weijer C, Zwarenstein M, Donner A. Impact of CONSORT extension for cluster randomised trials on quality of reporting and study methodology: review of random sample of 300 trials, 2000-8. BMJ. 2011 Sep 26;343:d5886. doi: 10.1136/bmj.d5886. Review. PubMed PMID: 21948873; PubMed Central PMCID: PMC3180203.

6 The Gambia Hepatitis Intervention Study. The Gambia Hepatitis Study Group. *Cancer Res.* 1987;47(21):5782-7.

7 Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. Ann Int Med 2010;152. Epub 24 March.

8 Campbell MK, Piaggio G, Elbourne DR, Altman DG; for the CONSORT Group. Consort 2010 statement: extension to cluster randomised trials. BMJ. 2012 Sep 4;345:e5661. Full text PDF.PMID: 22951546

9 warenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, Oxman AD, Moher D for the CONSORT and Pragmatic Trials in Healthcare (Practihc) group. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. BMJ 2008; 337;a2390

10 Boutron I, Moher D, Altman DG, Schulz K, Ravaud P, for the CONSORT group. Methods and Processes of the CONSORT Group: Example of an Extension for Trials Assessing Nonpharmacologic Treatments. Ann Intern Med. 2008:W60-W67.

11 von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies.