ELSEVIER

Journal of Clinical Epidemiology 64 (2011) 936–948

Journal of Clinical Epidemiology

REVIEW ARTICLES

Systematic review of stepped wedge cluster randomized trials shows that design is particularly used to evaluate interventions during routine implementation

Noreen D. Mdege[a,*], Mei-See Man[a], Celia A. Taylor (nee Brown)[b], David J. Torgerson[a]

[a]*York Trials Unit, Department of Health Sciences, University of York, Heslington, York YO10 5DD, UK*
[b]*College of Medical and Dental Sciences, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK*

Accepted 3 December 2010

Abstract

Objective: To describe the application of the stepped wedge cluster randomized controlled trial (CRCT) design.

Study Design and Setting: Systematic review. We searched Medline, Embase, PsycINFO, HMIC, CINAHL, Cochrane Library, Web of Knowledge, and Current Controlled Trials Register for articles published up to January 2010. Stepped wedge CRCTs from all fields of research were included. Two authors independently reviewed and extracted data from the studies.

Results: Twenty-five studies were included in the review. Motivations for using the design included ethical, logistical, financial, social, and political acceptability and methodological reasons. Most studies were evaluating an intervention during routine implementation. For most of the included studies, there was also a belief or empirical evidence suggesting that the intervention would do more good than harm. There was variation in data analysis methods and insufficient quality of reporting.

Conclusions: The stepped wedge CRCT design has been mainly used for evaluating interventions during routine implementation, particularly for interventions that have been shown to be effective in more controlled research settings, or where there is lack of evidence of effectiveness but there is a strong belief that they will do more good than harm. There is need for consistent data analysis and reporting.

Keywords: Systematic review; Cluster randomized controlled trial; Stepped wedge design; Implementation research; Research methods; Phased implementation

1. Introduction

Randomized controlled trials (RCTs) are considered the most robust research design for establishing a cause–effect relationship between an intervention and an outcome [1]. Cluster randomized controlled trials (CRCTs) are being increasingly used in the health care context to evaluate interventions in a pragmatic fashion, when it is either not possible or inappropriate to use individual randomization [2]. One variant of the cluster trial design is the stepped wedge trial design. In this approach, clusters, for example, geographical areas, general practitioner surgeries, and hospitals, are identified to take part in a cluster randomized trial and, before the trial begins, are randomly allocated a time when they are given the intervention. In essence, this is a one-way crossover cluster trial where all the groups will receive the intervention, but the time when they receive this is randomly ordered [3]. The clusters crossover at regular intervals, typically from control to intervention.

In Fig. 1, for example, randomization occurs before the start of the trial. None of the clusters receive the intervention at time point 1. Cluster 1 is randomly allocated to receive the intervention at time point 2, while all the remaining clusters contribute to the control at this time point. Cluster 2 is randomly allocated to start receiving the intervention at time point 3 and so on. At time point 6, all clusters will be receiving the intervention. It is also possible to randomize multiple clusters to receive the intervention at each time point.

There are a number of advantages to using this approach compared with the normal parallel group or crossover CRCT. A central tenet of parallel or crossover RCTs is that there must be equipoise, that is, a genuine uncertainty of whether one intervention is better than another [4]. Where

The authors declare that they have no conflicts of interest and financial interests.

* Corresponding author. Tel.: +44-1904-321-836; fax: +44-1904-321-387.

E-mail address: nm619@york.ac.uk (N.D. Mdege).

0895-4356/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi: 10.1016/j.jclinepi.2010.12.003

What is new?

- This review found that the stepped wedge cluster randomized controlled trial (CRCT) design has been particularly applied in routine practice for evaluating interventions that have been shown to be effective in more controlled research settings or where there is lack of evidence of effectiveness but there is a strong belief that they will do more good than harm.
- There is variation in data analysis methods and insufficient quality of reporting in studies using the stepped wedge CRCT design.
- The stepped wedge CRCT design is a potentially useful research design for evaluating interventions during implementation in routine practice, not just in health research but also in non–health research fields.
- Researchers should explore ways of enhancing internal validity through blinding of outcome assessors where possible and use of adequate sequence generation and allocation concealment when planning stepped wedge CRCT.
- Data analysis and reporting of stepped wedge CRCTs should be standardized.

there is no equipoise, it may be unethical to randomize patients to an intervention believed to be inferior or to withdraw an intervention believed to be superior compared with other interventions in the trial [3,4]. The stepped wedge design addresses this concern because the intervention is rolled out to all individuals or clusters in phases. Data collection is conducted at each point where a new cluster receives the intervention. The intervention effect is determined by comparing data points in the intervention section of the wedge with those in the control section. The stepped wedge design is also useful where phased implementation is preferable because of logistical, practical, and financial constraints [3]. Some studies use the stepped wedge design for scientific reasons such as to allow detection of underlying trends or control for time effects. The stepped wedge CRCT design may require fewer clusters than a parallel CRCT design [5]. A stepped wedge CRCT design maximizes statistical power compared with a parallel-group CRCT design [6], because the intervention effect is estimated not only by between-cluster comparisons as in the parallel group design but also by within-cluster comparisons.

Program implementation in routine practice is known to be problematic, and programs are not always effectively evaluated [6]. RCTs that randomize individuals are

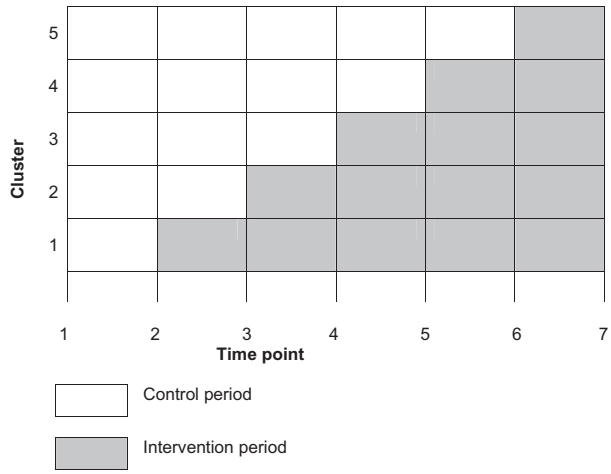

Fig. 1. Diagrammatic illustration of the stepped wedge design.

difficult to implement in routine practice and may not reflect intervention effectiveness at a population level. Weaker nonrandomized designs such as pre- and postintervention evaluations tend to overestimate the intervention effect [1]. The stepped wedge CRCT design offers an alternative to other designs such as the parallel-group CRCT that can be applied in routine practice for intervention evaluation. The phased implementation in the stepped wedge design also allows for improvement of the intervention or its delivery where necessary before the next implementation phase [6].

In an earlier systematic review of the stepped wedge trial design, Brown and Lilford [3] advocated use of the design in intervention evaluation. Their review included 12 studies evaluating a wide range of interventions. They included both randomized and nonrandomized studies, both individual and cluster allocations, and limited the review to the health sector. In this present review, we have expanded the search to include non–health care trials as these often provide lessons that can influence the design of health care trials. In contrast to the previous review [3], the present review will focus on randomized studies and cluster allocations. Thus, the present review aims to describe the application of the stepped wedge CRCT design in evaluating the effectiveness of interventions. The review explores (1) the areas of research in which the stepped wedge CRCT design has been used, (2) the motivation for using the design, (3) the general characteristics of the stepped wedge design, and (4) the methods of data analysis. The review also explores the quality of reporting in CRCT using the stepped wedge design.

2. Methods

A review protocol is included in the supplemental materials (see Supplement 1).