VARIANCE AND DISSENT

The balanced incomplete block design is not suitable for the evaluation of complex interventions

Jasper Trietsch[a,*], Pieter Leffers[b], Ben van Steenkiste[a], Richard Grol[c], Trudy van der Weijden[a]

[a]Department of General Practice, School for Public health and Primary Care (CAPHRI), Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands
[b]Department of Epidemology, Maastricht University, PO Box 616, 6200 MD, Maastricht, The Netherlands
[c]IQ Healthcare, Radbout University Nijmegen, PO Box 9101 (144), 6500HB Nijmegen, The Netherlands

Abstract

Objectives: In quality of care research, the balanced incomplete block (BIB) design is regularly claimed to have been used when evaluating complex interventions. In this article, we reflect on the appropriateness of using this design for evaluating complex interventions.

Study Design and Setting: Literature study using PubMed and handbooks.

Results: After studying various articles on health services research that claim to have applied the BIB and the original methodological literature on this design, it became clear that the applied method is in fact not a BIB design.

Conclusion: We conclude that the use of this design is not suited for evaluating complex interventions. We stress that, to prevent improper use of terms, more attention should be paid to proper referencing of the original methodological literature. © 2014 Elsevier Inc. All rights reserved.

Keywords: Research methodology; Evaluation studies; Primary care; Health care evaluation mechanisms; Quality of health care; Epidemiologic research design

1. Introduction

Incorrect citing in scientific literature can lead to a chain of erroneous interpretations or the use of incorrect methods or terms. In this short report, we describe the case of the balanced incomplete block (BIB) design. In quality of care research, the BIB design is regularly mentioned in the sections reporting on the applied study design. Study of the literature on the BIB design led us to original work by Cochran and Cox from the 1950s describing a design that appeared to be very different from the designs that were actually used in the quality of care research articles claiming to have applied the BIB. This raised questions about possible consequences of improper reporting of the BIB design for our own research with respect to the interpretation of study results.

In this article, we describe the features of the original BIB design, and we compare it with the design that we used in our earlier research projects. We illustrate where we failed in citing of the literature on the BIB design, and we describe the implications of this error for the validity of conclusions of our research.

1.1. Features of the original BIB design

Comparing the outcomes of multiple interventions under various conditions is a well-known challenge in experimental research. It is often impossible to carry out such studies because of limitations on the number of available research subjects and because of limited resources. In agriculture, a solution for this problem was developed for crop-optimization studies by statisticians involved in combinatorial problems research. These methods were thoroughly described by Cochran and Cox [1]. For instance, when testing several new genetic varieties of corn under different growth conditions, vast areas of land would be needed. The BIB design enables researchers to compare harvest returns of the varieties using plots of land (blocks), which each have different conditions between but uniform conditions within these plots. Not all varieties of corn will be grown on each of these plots: incomplete testing. By balancing the allocation of the varieties over the different

---

What is new?

Key findings

- Contrary to what has been claimed in various publications reporting on quality of care research, the balanced incomplete block (BIB) design has not been used. We show that the use of this design is not suited for this type of research.

What this adds to what was known?

- The BIB design dates from the first half of the last century. Because of its efficiency increasing potential, it can be an attractive design when resources are limited and when certain prerequisites are met. We show that indirect referencing has led to improper use of the name of the design in quality of care research and that the design should also not be applied in the evaluation of complex interventions.

What is the implication and what should change now?

- To increase transparency and enable critical appraisal of evidence, authors and peer reviewers should pay more attention to appropriate use of terms such as BIB design.

Box 1 General theorem for BIB designs

$bk = vr$ and	v = the number of compared
$r(k-1) = \lambda(v-1)$	(genetic) varieties
When,	b = the number of plots of land
$v > k > 0$	(blocks)
$r > 0$	r = the number of plots in which
$\lambda > 0$	each variety is present
	k = the number of varieties per plot
	λ = the number of plots in which
	each pair of varieties is present
	(pair replication)

---

plots, a comparison of outcome (eg, returns in harvest) can be made between varieties that were never really compared under the same conditions. The simplest example is that of difference in harvest between varieties A and C under condition I can be estimated, although they were not directly tested against each other under condition I. Under the assumption of the absence of effect modification by the condition, the harvest difference between A and C can be calculated by comparing varieties A and B under condition I and varieties B and C under condition II. Statistical testing is done using analysis of variance [1,2].

To speak of a BIB design, several conditions have to be met. A theorem is available to test whether a design meets the requirements of a BIB design. The notation of the parameters of a BIB design is {b,r,k,v,λ} (Box 1). Fig. 1 shows a {3,2,2,3,1} design: a design of three blocks with a variation replication rate of 2, two varieties per block, three different varieties, and a pair replication rate of 1, as described previously.

1.2. Features of the BIB design as applied in quality improvement research

According to Cochran and Cox, the BIB design is suitable for situations in which repeated testing of varieties will lead to the same result, as can be expected when conditions can be well controlled such as in agricultural or laboratory sciences. Unfortunately, in most types of clinical research, patients will be permanently influenced by the intervention that is being evaluated and therefore repeated testing cannot be expected to lead to the same result. As a consequence, the BIB design cannot be used for patient-centered research. However, several publications within quality of care research on complex health care interventions report on the use of the BIB [3–6], and it was also advocated as appropriate for complex guideline implementation trials [7].

Testing all components of the complex interventions separately is generally not possible because of limited resources or limitations on the number of available research subjects, let alone that all components can be tested under the various conditions. The applied design is claimed to overcome these limitations while it is also considered to be attractive because it controls for the Hawthorne effect [8]. However, from the published reports, it can be concluded that the BIB design was not applied at all.

Our research group has used the same design in studies on complex quality improvement interventions [9–11]. An example is the work of Verstappen et al. They performed a cluster randomized controlled trial and claimed
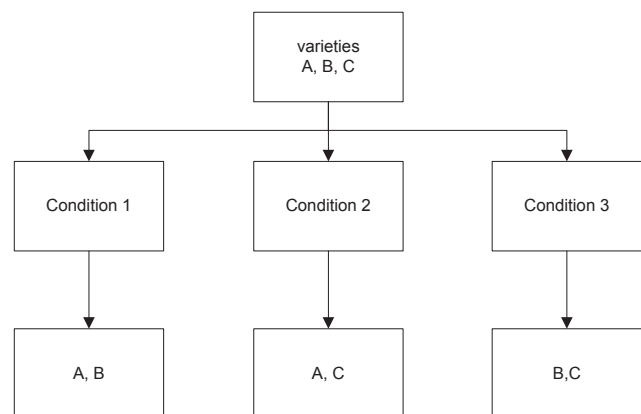

Fig. 1. Example of a balanced incomplete block design with {3,2,2,3,1} design.