Available online at www.sciencedirect.com
ScienceDirect
Psychiatry Research 160 (2008) 335–345
PSYCHIATRY RESEARCH
www.elsevier.com/locate/psychres

Comparison of data analysis strategies for intent-to-treat analysis in pre-test–post-test designs with substantial dropout rates

Agus Salim *, Andrew Mackinnon, Helen Christensen, Kathleen Griffiths

Centre for Mental Health Research, Australian National University, Acton, Australia

Received 17 January 2007; received in revised form 31 May 2007; accepted 10 August 2007

Abstract

The pre-test–post-test design (PPD) is predominant in trials of psychotherapeutic treatments. Missing data due to withdrawals present an even bigger challenge in assessing treatment effectiveness under the PPD than under designs with more observations since dropout implies an absence of information about response to treatment. When confronted with missing data, often it is reasonable to assume that the mechanism underlying missingness is related to observed but not to unobserved outcomes (*missing at random, MAR*). Previous simulation and theoretical studies have shown that, under MAR, modern techniques such as maximum-likelihood (ML) based methods and multiple imputation (MI) can be used to produce unbiased estimates of treatment effects. In practice, however, ad hoc methods such as last observation carried forward (LOCF) imputation and complete-case (CC) analysis continue to be used. In order to better understand the behaviour of these methods in the PPD, we compare the performance of traditional approaches (LOCF, CC) and theoretically sound techniques (MI, ML), under various MAR mechanisms. We show that the LOCF method is seriously biased and conclude that its use should be abandoned. Complete-case analysis produces unbiased estimates only when the dropout mechanism does not depend on pre-test values even when dropout is related to fixed covariates including treatment group (*covariate-dependent: CD*). However, CC analysis is generally biased under MAR. The magnitude of the bias is largest when the correlation of post- and pre-test is relatively low.

Keywords: Intention to treat (ITT); Trials of psychotherapies; Mixed models; Multiple imputation; Bias

1. Introduction

Missing data are ubiquitous in clinical trials in psychiatry. While observations may be missing because participants are temporarily unavailable or fail to complete a scheduled assessment, withdrawal from a study is by far the most common cause of missingness and is also the cause of greatest concern. Withdrawal threatens the integrity of a trial because it breaks randomization (cf. Peto et al., 1976, 1977): it can never be determined if attrition is related to an unobserved factor associated with outcome. Failure to take this into account appropriately may result in reaching erroneous conclusions about the effectiveness or ineffectiveness of an intervention. This reasoning underpins the intent (ion)-to-treat (ITT) principle (Guyatt and Rennie, 2001). The ITT principle requires that all participants are

* Corresponding author. Centre for Mental Health Research, Australian National University, Building 63, Eggleston Road, Acton ACT 0200, Australia. Tel.: +61 2 6125 8413; fax: +61 2 6125 0733.
E-mail address: agus.salim@anu.edu.au (A. Salim).

0165-1781/$ - see front matter © 2007 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.psychres.2007.08.005

retained in analyses regardless of their path through the trial. Participants are retained in the treatment group they are randomized to ("as randomized"), rather than being grouped post hoc according to the actual treatment they received ("as treated"). In the terminology of Schwartz and Lellouch (1967), ITT analysis is a pragmatic approach, the objective of the analysis being to estimate treatment *effectiveness* rather than *efficacy*, which is the objective of an "as treated" analysis. The *effectiveness* concept is arguably of particular relevance in clinical and public health contexts, since a treatment may not be tolerated even if it is efficacious due to aversive side effects or because of the time and effort involved. These effects are reflected in an ITT analysis but not in an "as treated" analysis, where the focus is on efficacy in people who comply fully with treatment.

When implementing ITT analysis, the missing measurements are required for participants who have withdrawn. These values are frequently imputed by assuming that the participant's status did not change from the last occasion on which he or she was observed to the end of the trial. This is referred to as last observation carried forward (LOCF) imputation. The origins of this approach are unclear. It is frequently portrayed as yielding a conservative estimate of the treatment effect and its statistical significance. When there is a natural declining time trend in the outcome variable as in degenerative conditions such as dementia, however, assuming stability may yield in an overestimation of treatment effect (Little and Yau, 1995). Even if the estimate of treatment effect is conservative, LOCF imputation may affect the variance and covariance of measures in ways that yield optimistic tests of statistical significance. Another frequently used but essentially ad hoc approach is complete-case (CC) analysis. In this approach, only subjects who complete the trial are retained for analysis. Although directly contradicting the principle of ITT, Graham and Donaldson (1993) showed that, in the case of linear regression, under *covariate-dependent* (CD) missingness mechanisms (see Section 2.1 for definitions), CC analysis produces an unbiased estimate of treatment effectiveness. This result was supported by Little (1995), who stated that CC is generally unbiased under the CD missingness mechanism. Molenberghs et al. (2004) show that CC is generally biased when the mechanism underlying the missingness is *missing at random* (MAR), but it is unbiased when the missingness mechanism is *missing completely at random* (MCAR).

In contrast to LOCF and CC, multiple imputation (MI; Rubin, 1987; Little and Rubin, 2002) and maximum-likelihood based methods (ML; e.g. Verbeke and Molenberghs, 2000) are principled and theoretically rigorous approaches to the problem of missing data in randomized trials. Schafer and Graham (2002) compared MI and ML approaches. They found that the performance of both methods was very similar. In particular, both methods are valid when the MAR missingness mechanism (see Section 2.1) holds, but the methods are biased when the mechanism is *non-ignorable* in the sense that the probability of withdrawals depends on the unobserved part of the outcome variable. Growing availability of software is seeing the increasing but far from universal application of MI and ML (see Gueorguieva and Krystal, 2004).

In psychiatric research, a pervasive and crucial limitation of both simulation studies and investigations using real data has been the failure to explicitly consider trials using a pre-test–post-test design (PPD). For example, Houck et al. (2004) analysed a real dataset from a 12-week antidepressant drug trial using each of the four approaches (CC, LOCF, MI and ML) and found that each approach produced different conclusions. In contrast to pharmacological trials such as these, where measurements are taken at regular intervals, the pre-test–post-test design involves only two occasions of measurement: prior to treatment (pre-test) and after the complete intervention has been delivered (post-test). This design is predominant in trials of psychotherapies and in public health interventions. Its popularity may well reflect budgetary constraints on non-commercial trials. However, psychotherapeutic treatments and public health interventions are conceptualized as integrated packages. Hence it is perceived to be meaningful to take measurements only once the complete program has been delivered.

Conceptually, the impact of withdrawal and missing data in the PPD is striking. Where measurements are taken regularly over a trial, available data may characterize individual trajectories reasonably well even for participants who subsequently withdraw. In the PPD, withdrawal implies complete absence of information about response. Applying LOCF in this design (as is frequently done) is actually carrying forward the first and only observation!

The advantages of MI and ML over LOCF have been clearly demonstrated when multiple observations are taken over time and thus when some information about response is available (Mallinckrodt et al., 2001). In the context of PPD, Molenberghs et al. (2004) provide general formulae that demonstrate the fact that LOCF is biased under all missingness mechanisms, while CC is unbiased under MCAR but generally biased under MAR. While the formulae are general ones and analytically simple, the parameters governing the formulae are