Computational Statistics and Data Analysis 56 (2012) 1150–1160

Contents lists available at SciVerse ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

An R package for implementing simulations for seamless phase II/III clinical trials using early outcomes for treatment selection☆

Nick Parsons [a,*], Tim Friede [b], Susan Todd [c], Elsa Valdes Marquez [d], Jeremy Chataway [e,f], Richard Nicholas [e], Nigel Stallard [a]

[a] Warwick Medical School, University of Warwick, Coventry, CV4 7AL, UK
[b] Department of Medical Statistics, University Medical Center Göttingen, Göttingen, Germany
[c] Applied Statistics, University of Reading, UK
[d] Clinical Trial Service Unit and Epidemiological Studies Unit, Oxford, UK
[e] Imperial College Healthcare NHS Trust, London, UK
[f] National Hospital for Neurology and Neurosurgery, London, UK

ARTICLE INFO

Article history:
Received 25 June 2010
Received in revised form 20 October 2010
Accepted 29 October 2010
Available online 5 November 2010

Keywords:
Adaptive seamless design
Clinical trial
asd
R package

ABSTRACT

Adaptive seamless phase II/III clinical trial designs allowing treatment selection at an interim analysis have gained much attention because of their potential benefits compared to more conventional drug development programmes with separate trials for individual phases. A scenario of particular interest is that in which the final outcome in the trial is based on long-term follow-up, but the interim analysis can only realistically be based on early (short-term) outcomes. A new software package (`asd`) for the statistical software R implements simulations for designs of this type, in addition to the simpler scenario where treatment selection is based on the definitive (final) outcome. The methodology is briefly described and two examples of proposed trial designs in progressive multiple sclerosis are provided, with R code to illustrate application of the methodology.

1. Introduction

In a conventional clinical programme for the evaluation of a new drug a number of distinct phases are identified. Phase II clinical trials are the first trials conducted to assess the treatment efficacy and safety and to optimize treatment for example by finding a suitable dose, whereas phase III trials provide definitive evidence of the efficacy and safety of the drug required for regulatory approval. Phase III trials are on a larger scale than phase II trials, usually requiring several hundreds or thousands of patients, and are conducted in the target population. The analysis is typically conducted by ignoring information from previous phases, and often uses different outcome measures from phase II trials. In recent years considerable research interest has focussed on accelerating the process of drug development by combining the conventional separate phase II and III trials of a clinical programme into a single trial. Appropriate statistical methodology to achieve the aim of a single seamless phase II/III trial, while still controlling the overall type I error rate at a pre-specified level, has been suggested by a number of authors; see for instance Thall et al. (1988, 1989), Bauer and Kieser (1999), Stallard and Todd (2003), Posch et al. (2005), Bretz et al. (2006), Koenig et al. (2008) and Stallard and Friede (2008).

Generally these adaptive seamless designs (ASDs) commence with a multi-arm stage including a control treatment and several experimental treatments (e.g. several doses of an experimental drug) and in an interim analysis (analogous to phase II

☆ This paper is supplemented by R package `asd` available in CRAN at http://www.R-project.org. The code from the Appendix is available electronically at the journal's webpage.

* Corresponding author. Tel.: +44 2476150541; fax: +44 2476968617.
E-mail address: nick.parsons@warwick.ac.uk (N. Parsons).

0167-9473/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2010.10.027

of a conventional programme) individual experimental treatment arms can be dropped for futility. One or more of the promising treatments are then carried forward along with the control and typically the null hypothesis of no difference between selected treatment(s) and the control treatment is tested in a final analysis (analogous to phase III) that combines information from both design stages. Such adaptive designs have been shown to be more efficient than the conventional phases II and III for confirmation of the efficacy of the selected treatments (Bretz et al., 2006). The majority of the works in this area have focussed on ASDs where the definitive (final) endpoint is available in the interim analysis, for treatment selection, and also in the final analysis. Some authors have investigated alternate scenarios; for instance Todd and Stallard (2005) and Stallard (2010) have proposed designs which assume that data are available for both early and final outcomes for some patients and only early outcomes for other patients. However, it is often the case that only data from an early outcome are available for treatment selection, rather than the more well studied case based purely on final outcome data. Friede et al. (submitted for publication) considered the scenario where only phase II type information, what we call an early outcome, is available in the interim analysis for treatment selection and confirmatory testing is exclusively based on definitive phase III type (final) outcomes. This scenario is particularly important for designing trials to assess treatment efficacy in chronic disabling conditions, where final outcomes are based on long-term follow-up, but the interim analysis can only realistically be based on early (short-term) outcomes. In progressive multiple sclerosis, for instance, the final outcome measure would typically be patient's disability after three years, but it may be desirable to conduct an interim analysis before three years when no final endpoint data are available. Such an interim analysis could be based on an early outcome obtained from a functional MRI scan at one year for example. The early outcomes here are viewed as a marker for treatment efficacy; they are *biologically plausible* outcomes (Chataway et al., 2010) in that they give some indication as to whether the mechanism of action of test treatments are working as anticipated in the interim analysis. That is, conceptually, the early outcomes are expected to indicate the ordering of treatment effects for the final outcome measures but not necessarily the size of treatment effects.

Friede et al. (submitted for publication) develop methodology and simulation models for this scenario based on an assumed normal distribution for the test statistics, for both early and final outcome measures, comparing test treatments to a single control treatment and undertake simulations to illustrate some of the properties of the models. However, they provide few details as to how the simulations were undertaken or how the models might be used in practice, particularly for non-normal data. In this paper we give full details of the algorithms used to implement the simulations in this setting and give a number of examples to illustrate the use of these simulation models to aid clinical trial design. The simulation models have been developed into an R (R Development Core Team, 2009) package called `asd` (available from the Comprehensive R Archive Network; http://www.cran.r-project.org/), so we also document this package and illustrate its use. The methods and the software applications described in this paper focus only on situations where early outcomes are used for treatment selection in the interim analysis. However, the R functions described here can also accommodate the simpler case where the definitive (final) endpoint is available in the interim analysis. We consider this as a special case of the more general early outcome problem, that can be implemented by appropriate modification of the correlations between the early and final outcome measures (see details in Section 3.2).

A brief review of software developments for adaptive designs is given by Wassmer and Vandemeulebroecke (2006). A number of features discussed here, such as weighted inverse normal combination functions and simulation tools for assessing the performance of a number of design options, are of course available in some widely used commercial packages for planning and simulation of adaptive group-sequential clinical trials (e.g. ADDPLAN; http://www.addplan.com/). The latest version of ADDPLAN (version 5) provides simulation tools for performing treatment arm selection designs, but not for the specific case of selection based on early outcomes, although for instance it does include other features such as simulations from binomial distributions that are not required for the R package `asd`. As far as we are aware, none of the small number of other software or R packages currently available (e.g. MSToolkit; https://r-forge.r-project.org/projects/mstoolkit/) is able to deal simply with the case where early outcome measures are used for treatment selection.

A brief review of the methodology originally described by Friede et al. (submitted for publication) is given in Section 2, Sections 2.1 and 2.3 describe the setting where early and final outcomes are available and Section 2.2 methods for combining data based on final outcomes. The algorithms and software used to implement the simulations are described in detail in Section 3. Section 4 introduces two typical examples of clinical trial designs to illustrate how simulations can be conducted for normal and non-normal data. Finally we summarise the use of package `asd` in Section 5.

2. Methods

2.1. ASD with treatment selection based on early outcomes

We consider a trial conducted in two distinct stages, involving at least two experimental test treatments and a single control treatment (e.g. standard or placebo). In stage 1, patients are recruited and randomised to receive one of the k_1 experimental treatments ($T_1, \ldots, T_{k_1}$) or the control treatment (T_0). The treatments are then compared using an early outcome measure in an interim analysis at the end of stage 1. The trial is either stopped for futility or one or more of the experimental treatments are selected to continue into stage 2 of the trial for further testing against the control treatment, with other treatments dropped from the trial. The trial continues with randomisation of newly recruited patients to either the control treatment group or to the remaining experimental treatment group(s). Patients from the dropped treatment