Contemporary Clinical Trials 46 (2016) 48–51

Contents lists available at ScienceDirect

Contemporary Clinical Trials

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

A small n sequential multiple assignment randomized trial design for use in rare disease research☆

Roy N. Tamura [a,*], Jeffrey P. Krischer [a], Christian Pagnoux [b], Robert Micheletti [c], Peter C. Grayson [d], Yeh-Fong Chen [e], Peter A. Merkel [f]

[a] Health Informatics Institute, University of South Florida, Tampa, FL, United States
[b] Department of Rheumatology, University of Toronto, Canada
[c] Departments of Dermatology and Medicine, University of Pennsylvania, Philadelphia, PA, United States
[d] Systemic Autoimmunity Branch, National Institute of Arthritis and Musculoskeletal Diseases, Bethesda, MD, United States
[e] Division of Biometrics 3, Office of Biostatistics, US Food and Drug Administration, Silver Spring, MD, United States
[f] Division of Rheumatology, University of Pennsylvania, Philadelphia, PA, United States

ARTICLE INFO

Article history:
Received 6 August 2015
Received in revised form 9 November 2015
Accepted 12 November 2015
Available online 14 November 2015

Keywords:
Re-randomization
Binary data
Weighted Z statistic

ABSTRACT

Background: Clinical trials in rare diseases are difficult to conduct due to the limited number of patients available with each disorder. We developed a Phase 2 trial which is a small n sequential multiple assignment randomized trial (snSMART) design to test several treatments for a rare disease for which no standard therapy exists.
Purpose: This paper illustrates the design, sample size estimation and operating characteristics of an snSMART.
Methods: We investigate the performance of a class of weighted Z statistics via computer simulations.
Results: We demonstrate the increase in power over traditional single stage designs, and indicate how the power changes as a function of the weight given to each stage.
Conclusion: The snSMART design is promising in a rare disease setting where several alternative treatments are under consideration and small sample sizes are necessary.

1. Background

Conducting clinical research in rare diseases poses numerous challenges due to the small number of patients with any given disorder. Griggs et al. [1] make a number of recommendations for clinical trials in rare diseases, including testing more than one therapy within each patient. Most rare diseases have no approved therapy [2] and several drug therapies are often used even though the efficacy data for any of the drugs is limited. This was the case for a recent situation faced by the Vasculitis Clinical Research Consortium and this paper describes our approach to the design and analysis of a clinical trial for patients with a rare form of vasculitis. The design is a small n sequential multiple assignment randomized trial (snSMART).

Sequential multiple assignment randomized trials (SMARTs) in the literature [3,4,5,6,7] are most often conducted in large patient populations with well established therapeutic options. In these situations, the primary objective of the trials is to determine optimal sequences and combination of sequences of the therapies. In contrast, oftentimes in rare diseases, drugs are used without strong evidence of efficacy. Given this difference and also given the limited population size, our objectives were to first determine the best treatment in the initial stage in terms of response rate and second, to make inference around that chosen treatment in comparison to the other treatments. We use the re-randomization feature of the SMART to borrow information over the different stages of the trial to make conclusions for individual treatments. Re-randomization of non-responders in SMART is also attractive to patients as it allows patients to try a second therapy.

The scarcity of patients with vasculitis makes a placebo controlled trial of one or more drugs difficult. In oncology, Simon et al. [8] argue against independent Phase 2 trials of different drugs because differences in patient selection and other factors make the comparison of results across trials difficult; the same argument holds for most diseases. In our form of vasculitis, a crossover trial is problematic for two reasons: i) the disease activity waxes and wanes as opposed to that seen in a stable chronic condition; and ii) patients and physicians are reluctant to have a subject crossover from a drug that is effective and commercially available. We sought to compare the efficacy of three drugs and the primary goals of the trial were to determine the drug which demonstrates the best efficacy and to compare that drug with the next best drug.

☆ This paper reflects the views of the authors and should not be construed to represent the views or policies of the U.S. Food and Drug Administration.

* Corresponding author at: Health Informatics Institute, University of South Florida, 3650 Spectrum Boulevard, Tampa, FL 33612, United States.
E-mail addresses: roy.tamura@epi.usf.edu (R.N. Tamura), jeffrey.krischer@epi.usf.edu (J.P. Krischer), Cpagnoux@mtsinia.on.ca (C. Pagnoux), Robert.Micheletti@uphs.upenn.edu (R. Micheletti), peter.grayson@nih.gov (P.C. Grayson), YehFong.Chen@fda.hhs.gov (Y.-F. Chen), Peter.Merkel@uphs.upenn.edu (P.A. Merkel).

http://dx.doi.org/10.1016/j.cct.2015.11.010
1551-7144/

A goal of randomized early phase trials of active drugs is to ensure with high probability that the chosen treatment is truly the one with the highest underlying response rate. Simon et al. [8] show how the sample size can be determined such that with high probability, the study would correctly select the drug with the highest response rate when the difference in response rates between that drug and the remaining drugs is assumed to be d. If the response rate for an ineffective treatment is 0.25 and an effective agent of interest has a response rate of 0.50, which we assumed as the base case for our trial, then a sample size of 27 per arm assures with 95% probability that the correct drug will be chosen. This sample size is markedly smaller than the sample size to detect a statistically significant difference between the two correctly determined drugs with response rates of 0.25 and 0.50. In order to achieve 95% power with a two-tailed alpha level of 0.05, the study would require enrollment of 95 subjects per arm.

2. Trial description

In psychiatric diseases, the sequential parallel design [9]is used to augment the placebo controlled trial with a second stage in which placebo non-responders are re-randomized to drug or placebo. The purpose of the sequential parallel design is to pool the placebo versus drug information from placebo non-responders in the second stage with the placebo versus drug information from the first stage. Adopting a similar approach, we propose a trial design comparing efficacy of three medications in which non-responders of each drug are re-randomized to one of the other two drugs. The resulting design is shown in Fig. 1. In Stage 1, patients are randomized equally to one of the three drugs. Patients who respond at the end of Stage 1 are continued on the drug. Patients who do not respond to drug in Stage 1 are re-randomized equally to receive one of the other two drugs. Response is again assessed at the end of Stage 2. The transition from Stage 1 to Stage 2 happens immediately for all patients and the analysis is conducted when all patients complete the trial. The definition of response is established at the design stage of the trial and documented within the protocol. At the end of the trial, we identify the drug with the best response rate in Stage 1 and then compare this drug with the best of the two remaining drugs. Re-randomizing patients in the trial increases power in the inference between the observed best treatment versus the best of the other treatments. For example, if the two best treatment groups in Stage 1 are the Drugs A and B, then we compare these two groups using both the data in Stage 1 and also the data from Stage 2 when those two groups are randomized to subjects who failed to respond to Drug C. The idea of first selecting groups and then randomizing additional subjects to those groups is the characteristic feature of two stage selection and testing designs [10]; the difference in our trial is that the two stages of the trial are being implemented simultaneously as opposed to in parallel for the traditional two stage selection and testing designs. A second difference is that the second stage of interest for the comparison of Drugs A and B consists of subjects who have first failed to respond on Drug C.

This design can easily be extended to more than 3 treatment groups. In this case, the number of Stage 2 re-randomizations increases and at least theoretically, the number of stages can also increase. As the number of groups increases, the necessary total sample size also increases which may force some limitations on the number of treatment groups for a study in a rare disease. The definition of response could also be extended to jointly consider both efficacy and toxicity.

3. Simulation study of test statistics

If we consider stage/previous drug as a stratification factor, the analysis of binary data in treatment comparisons usually involves pooling of the inference across strata. In our analysis, the two treatment groups from Stage 1 are chosen as the two best treatment groups. Kim [11] has shown that for normal data, under the null hypothesis when the best and next best treatment groups are from the same normal distribution, the t-statistic of these two groups has Type 1 error less than or equal to the nominal level. For binary data, a similar argument shows that the score test from Stage 1 has appropriate Type 1 error.

Mehotra and Railkar [12] investigated several weighted average methods for stratified binary data where the score test statistics from each stratum are weighted. The most commonly used weights are based on the sample size harmonic mean (SSIZE) or based on estimated variance (INVAR) from the strata. Appendix A gives the formulas for the INVAR and SSIZE weights. When the two treatment groups in Stage 1 are chosen based on the data, it is unclear whether these weights are still appropriate. Hence, we also investigate several fixed weights in addition to SSIZE and INVAR weights. The weights considered here are similar to those used in the analysis of a sequential parallel design [5] in that the weight is for the standardized difference for a stage and is made to optimize the power of the test statistic. They differ in the weights utilized in the SMART literature [13] which are used to estimate

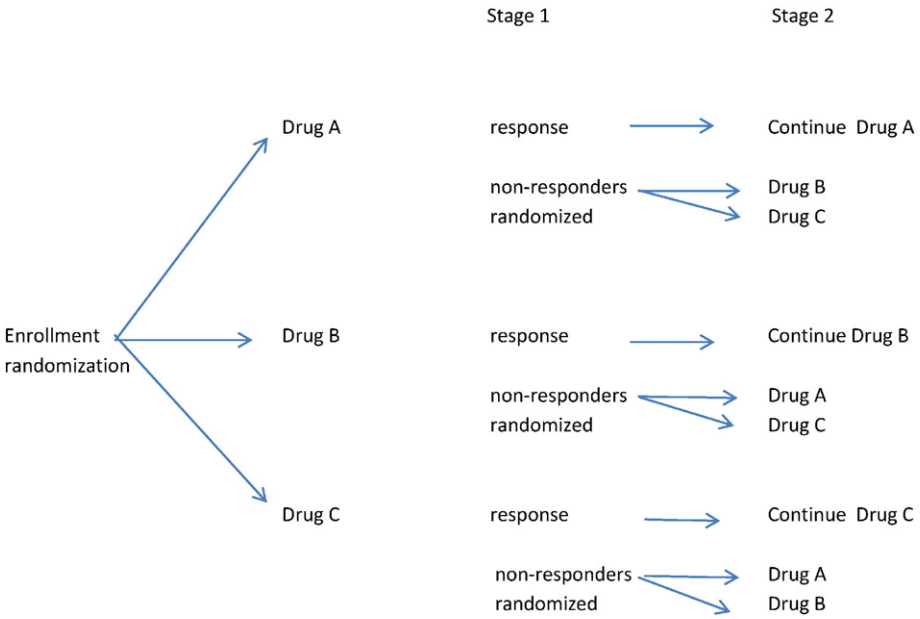

Fig. 1. Diagram of proposed small n sequential multiple assigned randomized trial.