Computational Statistics and Data Analysis xx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

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

Web-based statistical tools for the analysis and design of clinical trials that incorporate historical controls

Nan Chen [a], Bradley P. Carlin [b], Brian P. Hobbs [c,*]

[a] Department of Biostatistics, University of Texas M.D. Anderson Cancer Center, Houston, TX, United States
[b] Division of Biostatistics, University of Minnesota, Minneapolis, MN, United States
[c] Quantitative Health Sciences and The Taussig Cancer Institute, Cleveland Clinic, Cleveland, OH, United States

ARTICLE INFO

Article history:
Received 12 July 2017
Received in revised form 3 May 2018
Accepted 4 May 2018
Available online xxxx

Keywords:
Adaptive design
Bayesian analysis
Clinical trials
Hierarchical modeling
Historical controls
MCMC
Simulation

ABSTRACT

A collection of web-based statistical tools (http://research.mdacc.tmc.edu/SmeeactWeb/) are described that enable investigators to incorporate historical control data into analysis of randomized clinical trials using Bayesian hierarchical modeling as well as implement adaptive designs that balance posterior effective sample sizes among the study arms and thus maximize power. With balanced allocation guided by "dynamic" Bayesian hierarchical modeling, the design offers the potential to assign more patients to experimental therapies and thereby enhance efficiency while limiting bias and controlling average type I error. The tools effectuate analysis and design for static (non-hierarchical Bayesian analysis) and two types of dynamic (hierarchical Bayesian inference using empirical Bayes and spike-and-slab hyperprior) methods for Gaussian data models, as well as a dynamic method for time-to-failure endpoints based on a piecewise constant hazard model. The site also offers interfaces to facilitate calibration of the model hyperparameters. These allow users to test different parameters in the presence of the historical data on the basis of their resultant frequentist properties, including bias and mean squared error. All calculations are performed on a central computational server. The user may upload data, choose trial settings, run computations in real-time, and review the results using only a web browser. The back-end web module, computation module, and MCMC sampling module are developed in the C#, R, and C++ languages, respectively, and a communication module is also available to ensure the continued connection between the client computer and the back-end server during the Bayesian computations. The statistical tools are described and demonstrated with examples.

1. Introduction

The quality of health care evolves through the continual endeavor to enhance the safety and effectiveness of current therapeutic strategies through clinical study. Yet, translating biomedical discoveries into clinical practice is inherently challenging. Beneficial therapeutic strategies are established through a gradual process devised to define the safety and efficacy profiles of new strategies in phases, over the course of a sequence clinical trials. Transitions between phases involve latency periods wherein the next study is designed and reviewed prior to initiation, introducing inefficiency. In oncology, such latency periods span a duration of nearly two years on average (Committee on Cancer Clinical Trials and the NCI

* Correspondence to: Cancer Biostatistics Section Head in The Taussig Cancer Institute; Associate Staff, Department of Quantitative Health Sciences in The Lerner Research Institute, Cleveland Clinic, 9500 Euclid Ave. CA-60, Cleveland, OH 44195, United States.
E-mail address: hobbsb@ccf.org (B.P. Hobbs).

https://doi.org/10.1016/j.csda.2018.05.002
0167-9473/

Cooperative Group Program Board on Health Care Services, 2010). Moreover, each single successful study typically requires several years to achieve the targeted enrollment, and many studies fail due to low recruitment (Williams et al., 2015). This system produces redundancies, whereby similar treatment strategies are replicated, either as experimental or comparator standard-of-care therapies, across development phases and multiple studies. While systemic redundancy is necessary as sequential learning is needed to effectively devise prospective studies, given the nearly prohibitive cost of conducting clinical trials in humans, The Institute of Medicine recently advocated for the need to restructure the entire clinical trials system to avoid such redundancies as well as address other deficiencies that limit the effectiveness and efficiency of trials (Committee on Cancer Clinical Trials and the NCI Cooperative Group Program Board on Health Care Services, 2010). The initiative was recently re-affirmed with the 21st Century Cures Act, recently passed by the U.S. Congress and signed into law by President Obama in late 2016.

When planning a future trial supplemental information obtained from prior study is needed at the "design stage" to formulate plausible data-generating models that can be used to identify clinically meaningful effect sizes, as well as evaluate trial operating characteristics. Conventionally, data obtained from similar patient cohorts acquired from prior study is utilized formally to facilitate "comparative" evaluations of surrogate endpoints in single-arm phase II trials or conduct retrospective systematic literature reviews. Apprehension pertaining to formally incorporating data from historical studies into the comparative evaluations effectuated by randomized study is well-founded. Intrinsic to randomized design is the desire to infer causal relationships using random allocation strategies that attempt to balance the prognostic determinants (both known and unknown) which obscure the attribution of trends observed in the data to the studied interventions. The classical statistical tests that are conventionally used to compare study arms in randomized trials rely on exchangeable data sampling models. Pooling data from disparate studies using classical statistical tests yields statistical estimators that are sensitive to bias stemming from "trial effects." For example, trial effects stemming from differences in enrollment characteristics, patient surveillance, or clinical-care practice diminish the extent to which one can infer causal pathways, and thereby undermine the purpose of randomized study.

Statistical methods for integrating information from commensurate trials that relax the assumption of inter-cohort data exchangeability and leverage inter-trial redundancy have been developed. Pocock (1976) was first to propose using Bayesian models to incorporate supplemental information into the analysis of a primary data source through static, data-independent shrinkage estimators that require the extent of between-source variability to be prespecified. Numerous models have been discussed since, which involve prespecification of the amount of borrowing under different paradigms related to the power prior (Ibrahim and Chen, 2000; Hobbs et al., 2011; De Santis, 2006; Rietbergen et al., 2011) or inflating the standard error to downweight supplemental cohorts (Goodman and Sladky, 2005; French et al., 2012; Whitehead et al., 2008).

Hierarchical linear models and models which include adaptive down-weighting of data from supplemental cohorts have been extensively explored as well. For these models, the extent of shrinkage towards the supplemental sources is not predetermined but is estimated from the data. More strength is borrowed in the absence of evidence for inter-trial effects, which controls the extent of bias induced from using the supplemental information. One approach is the power prior of Ibrahim and Chen (2000) which can be constructed to discount supplemental sources relative to the primary data. Bayesian (Smith et al., 1995) and frequentist (Doi et al., 2011) methods which utilize hierarchical modeling have been developed to estimate between-source variability with univariate observables or repeated measures. Other authors have considered hyperprior specifications for Bayesian hierarchical models (Daniels, 1999; Natarajan and Kass, 2000; Spiegelhalter, 2001; Gelman, 2006; Browne and Draper, 2006; Kass and Natarajan, 2006). Recently, the use of Bayesian hierarchical modeling to leverage supplemental controls in analyses (Neuenschwander et al., 2010; Pennello and Thompson, 2008; Chen et al., 2011; Neelon and O'Malley, 2010) and trial designs (Hobbs et al., 2013) has been explored. Dynamic approaches to incorporating supplemental information using hierarchical modeling with sparsity inducing spike-and-slab hyperpriors and empirical Bayesian inference have also been described (Hobbs et al., 2011, 2012; Murray et al., 2014).

The impetus for leveraging historical controls, however, is often the desire to use fewer concurrent patients on previously studied control arms. While joint inference may lead to increased precision for estimation of the control (or null) effect, supplementing the control data alone creates imbalances in effective information and thereby impacts frequentist size and power only moderately. Consequently, the most effective utilization of historical control data actually occurs within the context of prospective adaptive trial design with allocation strategies devised to balance the extent of effective information among study arms.

In this article, we describe a collection of web-based statistical tools (http://research.mdacc.tmc.edu/SmeeactWeb/) hosted by M.D. Anderson Cancer Center that enable investigators to incorporate historical control data into conduct of randomized clinical trials using Bayesian methods. The interfaces use Bayesian hierarchical models to produce shrinkage estimators which can be used as the basis for integrating supplementary control data into an analysis of the primary trial source data (Hobbs et al., 2012, 2013). In addition to standard posterior summary statistics, the interfaces output the estimated effective historical sample size as well as the resultant randomization probability which should be used to assign the next trial patient to experimental versus control therapies when targeting balanced effective sample size at the trial's final analysis. This adaptive randomization method is based on the general concept of "multi-source adaptive randomization" first described in Hobbs et al. (2013).

The remainder of this article is organized as follows. Section 2 describes Bayesian models implemented by the interfaces, their derivations of effective historical sample size (EHSS), and the adaptive randomization method. Section 3 describes the computational infrastructure of the web-hosted software, while Section 4 describes implementation of the specific interfaces. Section 5 discusses a simulation study demonstrating the frequentist operating characteristics of multi-source adaptive randomization designs. Finally, Section 6 offers a brief closing discussion. A user manual is appended with the supplementary content.

Please cite this article in press as: Chen N., et al., Web-based statistical tools for the analysis and design of clinical trials that incorporate historical controls. Computational Statistics and Data Analysis (2018), https://doi.org/10.1016/j.csda.2018.05.002.