European Journal of Operational Research 248 (2016) 619–633

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Innovative Applications of O.R.

Response-adaptive designs for clinical trials: Simultaneous learning from multiple patients

Vishal Ahuja[1,*], John R. Birge

Booth School of Business, The University of Chicago, 5807 S Woodlawn Ave., Chicago, IL 60637, USA

ARTICLE INFO

Article history:
Received 11 July 2014
Accepted 30 June 2015
Available online 11 July 2015

Keywords:
OR in health services
Adaptive clinical trials
Markov decision process
Bayesian learning
Stents

ABSTRACT

Clinical trials have traditionally followed a fixed design, in which randomization probabilities of patients to various treatments remains fixed throughout the trial and specified in the protocol. The primary goal of this static design is to learn about the efficacy of treatments. Response-adaptive designs, on the other hand, allow clinicians to use the learning about treatment effectiveness to dynamically adjust randomization probabilities of patients to various treatments as the trial progresses. An ideal adaptive design is one where patients are treated as effectively as possible without sacrificing the potential learning or compromising the integrity of the trial. We propose such a design, termed *Jointly Adaptive*, that uses forward-looking algorithms to fully exploit learning from multiple patients simultaneously. Compared to the best existing implementable adaptive design that employs a multiarmed bandit framework in a setting where multiple patients arrive sequentially, we show that our proposed design improves health outcomes of patients in the trial by up to 8.6 percent, in expectation, under a set of considered scenarios. Further, we demonstrate our design's effectiveness using data from a recently conducted stent trial. This paper also adds to the general understanding of such models by showing the value and nature of improvements over heuristic solutions for problems with short delays in observing patient outcomes. We do this by showing the relative performance of these schemes for maximum expected patient health and maximum expected learning objectives, and by demonstrating the value of a restricted-optimal-policy approximation in a practical example.

1. Introduction [2]

The costs of bringing a new drug to market have been estimated to be as high as $5 billion (Forbes, 2013). Clinical trials have been cited as a key factor in raising these costs, with phase III trials now representing about 40 percent of pharmaceutical companies' R&D expenditures (Roy, 2012). The total cost of a clinical trial can reach $300–$600 million (English et al., 2010), potentially an order of magnitude higher when including the value of remaining patent life,[3] and exceed $6000 per enrolled subject (Emanuel, Schnipper, Kamin, Levinson, & Lichter, 2003). Consequently, drug manufacturers face pressure to produce conclusive results faster and reduce the number of subjects required.

Traditionally, clinical trials have followed a non-adaptive or a *fixed* design that randomizes patients to treatments in a constant proportion (probabilistically) throughout the trial. Such a design, in use for several decades, is well-understood by practitioners, and provides a clean way of separating treatments. Common reasons for the prevalence of such designs include a desire to maintain low probabilities of type I error and to protect against bias. However, these designs often result in lengthy trials, poor patient outcomes, and inconclusive results, leading to longer times for drug approval. In recognition of these issues, regulatory bodies, such as the U.S. Food and Drug Administration (FDA), have encouraged the use of adaptive designs (FDA, 2010a, 2010b).

* Corresponding author. Tel.: +1 214 768 3145.
E-mail addresses: vahuja@smu.edu (V. Ahuja), john.birge@chicagobooth.edu (J.R. Birge).

[1] Present address: Cox School of Business, Southern Methodist University, 6212 Bishop Boulevard, Dallas, TX 75275, USA.

[2] This work arose through discussions with our colleagues at The University of Chicago Medical Center, who were interested in a practically implementable adaptive design for a trial such as Clinical Antipsychotic Trials of Intervention Effectiveness Study (CATIE). The NIMH-funded trial, that compared schizophrenia drugs, suffered from several shortcomings, primarily those relating to patient compliance (Lieberman et al., 2005).

[3] Given that the patent for a drug or an intervention is typically filed before clinical trials begin, shortening the trial length can significantly increase potential revenues, not to mention the potential health benefit for the patients outside of the trial. For example, the sales of the drug *Atorvastatin* (trade name: *Lipitor*) decreased by 42 percent, from $2.4 billion to $1.4 billion, after its expiration on November 30, 2011 (Forbes, 2012).

http://dx.doi.org/10.1016/j.ejor.2015.06.077
0377-2217/

There exist several types of adaptive designs (see Chow and Chang, 2008 for a comprehensive list); a commonly used design, and the focus of this work, is the *outcome-* or *response-adaptive design*. Such designs, typically Bayesian in nature, employ learn-and-confirm concepts, accumulating data on patient responses, which is then used to make procedural modifications while the trial is still underway, increasing the likelihood of selecting the *right* treatment for the *right* patient population earlier in a drug development program. Adaptive designs can potentially increase the probability of finding the successful treatment, identify ineffective and unsafe drugs sooner, and require fewer patients in the trial, thereby reducing costs and shortening development timelines. Adaptive designs can also offer a safer alternative to fixed designs, allowing patients, who are initially allocated to a relatively unsafe treatment, to be switched to the safer treatment, as and when it becomes evident during the course of the trial. Henceforth, we will use the term *adaptive* to mean *response-adaptive* design.

The inherent flexibility of a Bayesian adaptive design appears contrary to the established fixed design. Common criticisms of adaptive designs include perceptions of reduced ability to do classical tests of statistical hypotheses, especially control of type I error that FDA requires for regulatory approval (FDA, 2010a, 2010b). Berry and Eick (1995) argues that such objections are either due to a lack of understanding or involve issues that can easily be addressed, for example, by incorporating constraints into the adaptive design (Cheng & Berry, 2007). Berry and Eick (1995) and LeBlond (2010) propose the use of computer simulations to evaluate type I error rate in Bayesian approaches.[4] In fact, the cholesterol-lowering drug Pravigard PAC was the first FDA approval that took a primarily Bayesian focus (Berry, 2006). In addition, the FDA has approved a number of medical devices whose submissions utilized a Bayesian statistical method (LeBlond, 2010). Further, inferential measures such as *predictive probability* make Bayesian approaches better suited for interim analyses as they provide the ability to quantify subsequent trial results given current information (Berry, 1985; 1987; 1993; Lee & Liu, 2008).

Berry and co-authors were among the first to develop a truly Bayesian response-adaptive design (see, for example, Berry, 1978; Berry & Pearson, 1985). In their design, patient randomization to treatments happens sequentially, that is, one at a time and all previous patient response(s) are known and incorporated into the randomization decision(s) for the following patient(s). This design is reasonable for trials where a single patient is randomized at each period, as in the case of individualized therapy trials, or when there is minimal delay in observing outcomes. However, this design is not practically useful when multiple patients need to be randomized simultaneously. One could implement variations of this design, for example by randomizing all the patients at a stage with a probability calculated for a single patient using the existing sequential design. However, such designs are suboptimal in a model that recognizes available information and the timing of opportunities to gather more information, required for updating the policy. We address this gap by developing an adaptive design with multiple simultaneous randomizations to anticipate learning through the trial horizon; we call this the *Jointly Adaptive* design. Following existing literature, we assume that patients are exchangeable, their outcomes are observable before each randomization decision, there is no serial correlation in treatment effects, and the treatment effects remain the same at each stage of the trial.

1.1. Main contributions and organization of the paper

The key contribution of this paper is the development of a Bayesian MDP framework for finite-horizon problems that learns optimally from simultaneous multiple experiments, admits continuous controls, and can be used to evaluate treatments under multiple objectives. In the context of clinical trials, our contributions are the development of a practically implementable response-adaptive design (termed *Jointly Adaptive*) that learns simultaneously from multiple patients and optimally randomizes them to multiple treatments. Further contributions include consideration of a learning objective in addition to the health objective, and evaluation of the relative advantage of the *Jointly Adaptive* design over other implementable response-adaptive designs, the fixed design, and heuristics. We note that our model is generalizable to other MDP settings that involve learning from multiple simultaneous individual experiments, as in the case of customized consumer offers.

The rest of the paper is organized as follows. Section 2 provides a brief overview of the literature. Section 3 presents the underlying model for the *Jointly Adaptive* design. Section 4 describes various adaptive designs and provides some theoretical guarantees. In Section 5, we present numerical results, including application to a recently conducted clinical trial. In Section 6, we summarize and discuss our conclusions as well as the scope and limitations of the adaptive designs.

2. Literature overview

The majority of previous work on trial design appears in the field of statistics. The class of problems involving adaptive designs has its roots in the *multi-armed bandit* problem that balances maximizing reward using knowledge already acquired with undertaking new actions to further increase knowledge, commonly referred to as the *exploitation vs. exploration* tradeoff.

The study of heuristics for the multi-armed bandit problem has a long history. Robbins (1952) is one of the earliest works on this topic that investigated the *play-the-winner* rule in a two-armed bandit problem. Bellman (1956) is one of the first to study the problem of sequential design of experiments using backward induction. Gittins (1979) employs a Dynamic Allocation Index, also called the Gittins Index, to solve bandit problems using forward induction; Katehakis and Veinott (1987) characterizes this index in a way that allows it to be calculated more easily.

Berry (1978) is one of the first studies to fully incorporate a Bayesian learning approach in a two-armed bandit. Extensions to this model include: (a) Berry and Eick (1995), which considers an objective that incorporates the conflicting goals of treating patients as effectively as possible during the trial and, with high probability, correctly identifying the relative efficacy of each treatment, and (b) Cheng and Berry (2007), which proposes a constrained adaptive design to address the "treatment assignment bias" concern raised in the literature (e.g., Chalmers, Celano, Sacks, & Harry Smith, 1983); their constraint ensures that each treatment in the trial has a certain fixed minimum probability of being chosen at each allocation decision. We refer the readers to Berry and Fristedt (1985) for further applications and note that adaptive designs have typically focused on maximizing expected patient health.

A related stream of literature has investigated asymptotically adaptive policies for bandit problems to achieve an optimal rate of regret. Lai and Robbins (1985) is a seminal study whose proposed adaptive policy achieves a $O(\log n)$ lower bound on the regret. Extensions of this study and other examples include Burnetas and Katehakis (1996), Auer, Cesa-Bianchi, and Fischer (2002), and Honda and Takemura (2010). For further details, we direct the readers to these papers and references therein.

Another stream of related literature includes evaluation of adaptive *treatment* strategies, defined by sequences of decision rules on when and how to alter the treatment of a patient in response to outcomes (Murphy, 2005). Such designs share several features with adaptive *trial designs*, for example, the use of past patient responses.

[4] A commercial software that does this simulation is called FACTS™, see www.berryconsultants.com/software/ for details.