Published in final edited form as:

Clin Trials. 2014 August ; 11(4): 408–417. doi:10.1177/1740774514537727.

Inference about the expected performance of a data-driven dynamic treatment regime

Bibhas Chakraborty^{1,2}, Eric B. Laber³, and Ying-Qi Zhao⁴

¹Centre for Quantitative Medicine, Duke-NUS Graduate Medical School, Singapore

²Department of Biostatistics, Columbia University, New York, NY, USA

³Department of Statistics, North Carolina State University, Raleigh, NC, USA

⁴Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, Madison, WI, USA

Abstract

Background—A dynamic treatment regime (DTR) comprises a sequence of decision rules, one per stage of intervention, that recommends how to individualize treatment to patients based on evolving treatment and covariate history. These regimes are useful for managing chronic disorders, and fit into the larger paradigm of personalized medicine. The *Value* of a DTR is the expected outcome when the DTR is used to assign treatments to a population of interest.

Purpose—The Value of a data-driven DTR, estimated using data from a *sequential multiple assignment randomized trial*, is both a data-dependent parameter and a non-smooth function of the underlying generative distribution. These features introduce additional variability that is not accounted for by standard methods for conducting statistical inference, e.g., the bootstrap or normal approximations, if applied without adjustment. Our purpose is to provide a feasible method for constructing valid confidence intervals for this quantity of practical interest.

Methods—We propose a conceptually simple and computationally feasible method for constructing valid confidence intervals for the Value of an estimated DTR based on subsampling. The method is self-tuning by virtue of an approach called the double bootstrap. We demonstrate the proposed method using a series of simulated experiments.

Results—The proposed method offers considerable improvement in terms of coverage rates of the confidence intervals over the standard bootstrap approach.

Limitations—In this paper, we have restricted our attention to Q-learning for estimating the optimal DTR. However, other methods can be employed for this purpose; to keep the discussion focused, we have not explored these alternatives.

Conclusions—Subsampling-based confidence intervals provide much better performance compared to standard bootstrap for the Value of an estimated DTR.

1 Introduction

Dynamic treatment regimes (DTRs) reflect the increasingly popular theme of personalized medicine in biostatistical research. DTRs facilitate personalized medicine in time-varying settings by allowing treatment selection to depend on time-varying – or *dynamic* – information. Operationally, a DTR consists of a sequence of decision rules, one per stage of clinical intervention, that recommends a treatment based on a patient's available treatment and covariate history. These decision rules offer a vehicle for the personalized management of many chronic conditions, e.g., alcohol and drug abuse [1], tobacco addiction [2], cancer [3, 4, 5, 6], HIV infection [7, 8], and mental illnesses [9, 10, 11], where a patient typically has to be treated at multiple stages. In essence, DTRs constitute operationalized *clinical decision support systems*, a key element of the *chronic care model* [12]; see Chakraborty and Moodie [13] for a book-length treatment of the topic of DTRs. A simple example of a two-stage behavioral DTR for smoking cessation is: “Initially provide a behavioral message with high degree of tailoring (message individually tailored according to the smoker's baseline variables), and provide a booster prevention message after six months as a follow-on intervention (while providing nicotine patch all along, to actively address the pharmacological aspect of smoking cessation).”

Expert opinion is one approach to constructing DTRs [e.g., 14, 15]; however, there has been a recent surge of interest in making DTRs *evidence-based*, i.e., data-driven. The quality of a DTR is usually assessed in terms of its *Value*. The Value of a DTR is the average primary outcome obtained when the DTR is applied to the entire population of interest (see below). A DTR is said to be *optimal* if it yields the highest Value. Most statistical research in the area of DTRs concerns: (a) the comparison of two or more preconceived DTRs in terms of their Value; or (b) the estimation of the optimal DTR, i.e., estimation of the sequence of decision rules that would result in the highest Value, within a certain class.

High-quality data (i.e., data free from causal confounding) for comparing or constructing DTRs can be obtained from *Sequential Multiple Assignment Randomized Trials* (SMARTs) [9, 16, 17, 18]. Methodological research on SMARTs is experiencing a steady growth to accommodate the increasing prevalence of such designs in practice, e.g., in cancer [6, 19, 20, 21, 22, 23], smoking [2], childhood autism [1, 24], childhood attention deficit hyperactivity disorder [25, 26], drug abuse during pregnancy [1, 27], and alcoholism [1]; see also [28] and [29] for comprehensive lists of SMART studies. The increasing popularity of SMARTs is further reflected in the recent NIH program announcements specifically asking for such designs [30]. For a discussion of SMART designs including power, efficiency, and sample sizes, see [17, 18, 25, 31, 32, 33, 34] and references therein. DTRs can also be constructed from longitudinal, observational studies; but the analysis becomes more complex in case of observational studies because the analytic techniques for such studies must proactively address selection bias and time-varying confounding [8, 35, 36]. In this article, for simplicity, we will restrict our attention to data from SMARTs only.

There exist a variety of methods for estimating DTRs from either SMART or observational data, including: *G-estimation* [37]; *Q-learning* [2, 38, 39]; *marginal structural models* [8, 10, 35]; *outcome weighted learning* [40]; and *augmented inverse probability weighting* [41, 42].

Regardless of the estimation method, accurate measures of uncertainty are needed if the estimated DTR will be used to inform clinical practice or guide subsequent research. We consider the problem of constructing a confidence interval for the Value of an estimated optimal DTR. This problem is made complex by the fact that the Value is a data-dependent and non-smooth parameter [37, 43]. Note that the Value of a fixed DTR (i.e., one that is not data-driven) does not suffer from these issues and has been addressed by numerous authors [e.g., 44, 45, 46, 47, 48]. We propose a conceptually simple and computationally feasible method for constructing a confidence interval for the Value of an estimated optimal DTR based on subsampling [e.g., 11, 49].

In Section 2 we review SMART designs. In Section 3 we define the Value of a DTR, and also describe Q -learning for estimating an optimal DTR. We introduce a novel subsampling based confidence interval in Section 4. The finite sample performance of the proposed confidence interval is evaluated using a suite of simulated experiments in Section 5. Section 6 provides a concluding discussion.

2 Setup and Notation

2.1 Sequential Multiple Assignment Randomized Trials

Initially, *sequential randomization* was only used as a conceptual tool to describe conditions needed to identify an optimal DTR from observational data [50, 51, 52, 53]. SMART trials, informed by this work, satisfy the sequential randomization condition by design [16, 17, 18]. SMART designs involve an initial randomization of patients to available treatment options, followed by re-randomizations at each subsequent stage of some or all of the patients to treatment options available at that stage. The re-randomizations and the set of treatment options at each stage may depend on information collected in prior stages such as how well the patient responded to the previous treatment.

Figure 1 shows a hypothetical SMART that may be used to estimate the behavioral DTR for smoking cessation discussed in the introduction. In this hypothetical trial, each participant is randomly assigned to one of two possible initial behavioral treatments: message with high degree of tailoring or message with low degree of tailoring (using baseline information). After six months, participants' intermediate outcomes (quit status, number of non-smoking months, number of cigarettes smoked while in the study, etc.) are collected, and participants are re-randomized to one of the two subsequent behavioral treatment options at stage 2: a booster prevention message, or a control message. All participants are provided with nicotine patches to address the pharmacological aspect of smoking cessation. One goal of the study is to construct a DTR that would maximize the mean number of non-smoking months over the 12-month study period. This hypothetical study is a simplified version of a real SMART as described in [2]; see also Strecher et al. [54] for detailed rationale behind considering such behavioral intervention strategies.

2.2 Data Structure

For clarity, we consider SMARTs with two stages of intervention only, as in Figure 1; generalization to more than two stages is relatively straightforward [55]. The longitudinal

data trajectory observed on each patient has the form $(O_1, A_1, O_2, A_2, O_3)$, where $O_j, j = 1, 2$ denotes the vector of covariates measured prior to treatment at the beginning of the j -th stage, O_3 denotes the outcomes measured at the end of stage 2 (end of the study), and $A_j, j = 1, 2$ denotes the treatment assigned at the j -th stage subsequent to observing O_j . The primary outcome is $Y = r(O_1, A_1, O_2, A_2, O_3)$, for a known function r , which is observed at the end of the study. For example, in the smoking cessation study above, O_1 may include addiction severity and co-morbid conditions at baseline, O_2 may include the same variables observed at the end of stage 1 (six months post initial randomization) and adherence to initial treatment, and Y may be the number of non-smoking months over the 12-month study period. The methodology proposed in this article also applies to trial designs in which only a subset of subjects (e.g., nonresponders) are randomized at the second stage; this case is handled by a judicious choice of the outcome; see [55] for details and an example. We briefly illustrate this case via simulation in Section 5.

Define the history at each stage as $H_1 = O_1$ and $H_2 = (O_1^T, A_1, O_2^T)^T$; thus, the history at stage j consists of the information available prior to the assignment of the j -th treatment. The data available to estimate an optimal DTR consists of a random sample of n independent and identically distributed trajectories. For simplicity, assume that there are only two possible treatments at each stage, $A_j \in \{-1, 1\}$ and that they are randomized, conditional on history, with known randomization probabilities. A DTR $d = (d_1, d_2)$ is a pair of functions with d_j mapping the values in the support of H_j to the available treatment options $\{-1, 1\}$. In the next section we discuss methods for estimating a high-quality DTR, i.e., a DTR which could lead to a better benefit for the population compared with current alternatives, using data collected in a SMART.

3 Defining and Estimating an Optimal DTR

3.1 Value of a fixed DTR

The Value of a DTR $d = (d_1, d_2)$ is:

$$V^d = \mathbb{E}_{d_1, d_2}[Y] = \mathbb{E}_{d_1, d_2}[r(O_1, A_1, O_2, A_2, O_3)], \quad (1)$$

where $\mathbb{E}_{d_1, d_2}$ denotes expectation taken with respect to the distribution of the entire data trajectory $(O_1, A_1, O_2, A_2, O_3)$ subject to the restrictions $A_1 = d_1(O_1)$ and $A_2 = d_2(O_1, A_1, O_2)$. A regime d is often said to be *embedded* in the study if d is actually employed to allocate treatments to a subset of subjects in the study, i.e., if the restrictions $A_1 = d_1(O_1)$ and $A_2 = d_2(O_1, A_1, O_2)$ are naturally satisfied for a non-null subset of subjects. There are 4 embedded DTRs in the hypothetical smoking cessation SMART described in Figure 1. One of these embedded DTRs is: “provide message with low degree of tailoring as the initial treatment (this is stage-1 rule, d_1); then at the second stage, give booster prevention message (this is stage-2 rule, d_2).”

When the regime of interest is an embedded DTR in the SMART, estimation of Value is relatively straightforward. The problem becomes more complex if d is not embedded. An

example of a non-embedded regime in the smoking cessation study is: “as the initial treatment, provide the message with high degree of tailoring to subjects without any college education, and the message with low degree of tailoring to others; then at the second stage, give booster prevention message to the non-quitters at six months, and control message to others.” A non-embedded regime can be estimated from the observed data if it is *feasible* [56]; a regime d is feasible if there is a positive probability that a subject in the study would follow d . For a feasible regime d , one can use inverse probability weighting (IPW) to express the Value of the regime in terms of the generative model [57, 58, 59]. The IPW estimator is

$$V^d = \mathcal{E}_{d_1, d_2}[Y] = \mathcal{E} \left[\frac{\mathcal{I}_{d_1(H_1)=A_1} \mathcal{I}_{d_2(H_2)=A_2}}{\pi_1(A_1|H_1)\pi_2(A_2|H_2)} Y \right], \quad (2)$$

where $\mathbb{E}$ denotes expectation with respect to the joint trajectory distribution, $\mathbb{I}_{\text{condition}}$ denotes an indicator function taking the value 1 when the ‘condition’ holds and 0 otherwise, and $\pi_j(a_j|h_j) = P(A_j = a_j|H_j = h_j)$ is the treatment allocation probability at the j th stage, $j = 1, 2$; these are known by design in a SMART, but must be estimated in observational studies. A plug-in estimator of the Value is

$$\hat{V}^d = \frac{1}{n} \sum_{i=1}^n \left[\frac{\mathcal{I}_{d_1(H_{1i})=A_{1i}} \mathcal{I}_{d_2(H_{2i})=A_{2i}}}{\pi_1(A_{1i}|H_{1i})\pi_2(A_{2i}|H_{2i})} Y_i \right], \quad (3)$$

which can be highly variable if the weights in the denominator of the expression are close to zero. Because $\hat{V}^d$ is a plug-in estimator there is potential for upward bias when the same data are used to construct and evaluate a DTR; an alternative approach would be to use cross-validation, however, for simplicity we do not consider this further.

3.2 Estimation of an Optimal DTR via Q-learning

There exists a variety of methods for estimating an optimal DTR, i.e., the DTR d that leads to the highest Value. Here we focus on a simple method called *Q-learning* [38, 60]. From dynamic programming, it is known that if $Q_2(h_2, a_2) = \mathbb{E}(Y|H_2 = h_2, A_2 = a_2)$ and $Q_1(h_1, a_1) = \mathbb{E}(\max_{a_2} Q_2(H_2, a_2)|H_1 = h_1, A_1 = a_1)$, then the optimal regime satisfies

$d_j^{\text{opt}}(h_j) = \arg \max_{a_j} Q_j(h_j, a_j)$. *Q-learning* mimics the dynamic programming solution by estimating the *Q*-functions, Q_j , $j = 1, 2$, based on the observed data, often assuming linear working models of the form $Q_j(h_j, a_j; \beta_j, \psi_j) = \beta_j^\top h_{j0} + (\psi_j^\top h_{j1}) a_j$, where h_{j0} and h_{j1} are (possibly different) features of h_j . A version of the *Q-learning* algorithm is:

1. Stage 2 regression:

$$(\hat{\beta}_2, \hat{\psi}_2) = \arg \min_{\beta_2, \psi_2} \sum_{i=1}^n (Y_i - Q_2(H_{2i}, A_{2i}; \beta_2, \psi_2))^2.$$

2. Stage 1 dependent variable: $\hat{Y}_{1i} = \max_{a_2} Q_2(H_{2i}, a_2; \hat{\beta}_2, \hat{\psi}_2)$, $i = 1, \dots, n$.

3. Stage 1 regression:

$$(\hat{\beta}_1, \hat{\psi}_1) = \arg \min_{\beta_1, \psi_1} \sum_{i=1}^n \left(\hat{Y}_{1i} - Q_1(H_{1i}, A_{1i}; \beta_1, \psi_1) \right)^2.$$

Note that in step 2 above, the quantity $\hat{Y}_{1i}$ is a predictor of the unobserved random variable $\max_{a_2} Q_2(H_{2i}, a_2)$, $i = 1, \dots, n$. The estimated optimal DTR using Q-learning is given by $(\hat{d}_1, \hat{d}_2)$, where the stage- j optimal rule is specified as $\hat{d}_j(h_j) = \arg \max_{a_j} Q_j(h_j, a_j; \hat{\beta}_j, \hat{\psi}_j)$, $j = 1, 2$. Q-learning with linear regression for two stages has been implemented in the R package qLearn [61].

3.3 Value of an Estimated DTR

For an estimated (optimal) DTR, say $\hat{d}$, a fundamental question is whether or not $\hat{d}$ has a higher Value than standard care. We assume that the Value of standard care is known though this is not essential. One way to compare $\hat{d}$ with standard care is to construct a confidence interval for the Value of $\hat{d}$, and check if the known Value of standard care lies within that interval. This is analogous to constructing a confidence interval for the unknown mean μ of a random variable and then checking if it contains a postulated value μ_0 . Using (2), the Value of $\hat{d}$ is given by

$$V^{\hat{d}} = \mathbb{E}_{\hat{d}_1, \hat{d}_2} Y = \mathbb{E} \left[\frac{\mathcal{I}_{\hat{d}_1(H_1)=A_1} \mathcal{I}_{\hat{d}_2(H_2)=A_2}}{\pi_1(A_1|H_1)\pi_2(A_2|H_2)} Y \right],$$

where $\mathbb{E}$ is taken with respect to the joint trajectory distribution but *not* the data used to construct $\hat{d}$. Thus, $V^{\hat{d}}$ depends both on the unknown generative distribution and on the data used to construct $\hat{d}$ and is therefore a data-dependent parameter [43, 62]. Data-dependent parameters, though somewhat unusual, are appropriate when studying the performance of an estimated predictive model or decision rule since the primary focus is the performance of the estimated predictive model (conditioned on the observed data) rather than the average performance of the estimation procedure averaged across data sets [e.g., 43, 63, 64, 65]. The indicator functions present in the expression of $V^{\hat{d}}$ make $V^{\hat{d}}$ a non-smooth function of the data. Recall that A_j is coded to take values in $\{-1, 1\}$. For the purpose of illustration, we consider linear decision rules of the form $\hat{d}_j(h_j) = \text{sign}(\hat{\psi}_j^\top h_{j1})$, where h_{j1} is a feature constructed from h_j ; Q-learning with linear models yields decision rules of this form. In this case, the Value can be shown to equal

$$V^{\hat{d}} = \mathbb{E} \left[\frac{Y}{\pi_1(A_1|H_1)\pi_2(A_2|H_2)} \mathcal{I}_{\min(A_1\hat{\psi}_1^\top H_{11}, A_2\hat{\psi}_2^\top H_{21}) \geq 0} \right], \quad (4)$$

which can be viewed as a weighted misclassification error with weights $Y/\pi_1(A_1|H_1)\pi_2(A_2|H_2)$ and ‘margin’ $\min(A_1\hat{\psi}_1^\top H_{11}, A_2\hat{\psi}_2^\top H_{21})$ [42, 66]. The misclassification error is a well-known example of a non-regular data-dependent parameter due to the non-smoothness of the indicator function at zero. A consequence is that standard methods for inference, such as the usual n -out-of- n bootstrap [67], cannot be applied without modification (see [43] for details).

In the next section we propose a simple subsampling based approach to construct valid confidence intervals for $V^{\hat{d}}$.

4 Subsampling Confidence Interval for $V^{\hat{d}}$

Let $\hat{V}^{\hat{d}} = n^{-1} \sum_{i=1}^n Y_i \mathcal{I}_{\min(A_{1i} \hat{\psi}_1^T H_{11i}, A_{2i} \hat{\psi}_2^T H_{21i}) \geq 0} / \{\pi_1(A_{1i}|H_{1i})\pi_2(A_{2i}|H_{2i})\}$ denote the plug-in estimator of $V^{\hat{d}}$. To construct a confidence interval we approximate the percentiles of the limiting distribution of $\sqrt{n}(\hat{V}^{\hat{d}} - V^{\hat{d}})$. As mentioned previously, normal approximations or standard bootstrap estimates of this limiting distribution are not consistent [43, 68]. One

approach to consistently estimate the distribution of $\sqrt{n}(\hat{V}^{\hat{d}} - V^{\hat{d}})$ is to use a subsampling procedure called the m -out-of- n bootstrap [69, 70, 71]. The m -out-of- n bootstrap mimics the usual nonparametric bootstrap except that the resample size, typically denoted by m , is allowed to be data-dependent but must satisfy $m \rightarrow_p \infty$ and $m/n \rightarrow_p k \in [0, 1]$ as $n \rightarrow \infty$. Given the resample size m , we can form a $(1 - \eta) \times 100\%$ confidence interval for $V^{\hat{d}}$ as follows. We draw B m -out-of- n bootstrap samples and calculate the bootstrap estimates $V^{\hat{d},b}$, $b = 1, \dots, B$, and find $\hat{l}$ and $\hat{u}$, the $(\eta/2) \times 100$ and $(1 - \eta/2) \times 100$ percentiles of

$\sqrt{m}(\hat{V}^{\hat{d},b} - \hat{V}^{\hat{d}})$ respectively. The confidence set is then given by

$(\hat{V}^{\hat{d}} - \hat{u}/\sqrt{m}, \hat{V}^{\hat{d}} - \hat{l}/\sqrt{m})$. Choosing the resample size m so that $k = 0$ remedies bootstrap inconsistency for a wide class of nonsmooth estimators [72] and in many cases the coverage of m -out-of- n bootstrap intervals increases as m decreases; however, choosing m too small will reduce efficiency [73]. Thus, m is a potentially important tuning parameter. Here we consider an adaptive choice of m proposed by Chakraborty et al. [11] for use with the Value. Details for choosing m are given in the Appendix.

5 Simulation Study

5.1 Simulation Design

In this section, we present a primary simulation study and a secondary simulation study to provide an empirical evaluation of the proposed confidence intervals. The primary study generates data from a SMART having two stages of treatment and two treatment options per stage irrespective of any intermediate measure of “treatment response” (as in the hypothetical smoking cessation SMART). The secondary study generates data from a more complicated SMART design, wherein only a subgroup of subjects (e.g. non-responders to the initial treatment) are re-randomized at the second stage; see Lei et al. [1] for examples of such studies. The purpose of the secondary study is to illustrate the wider scope of the proposed confidence intervals beyond the specific data structure offered by the motivating smoking cessation SMART.

Generative Model of the Primary Simulation Study—Here we consider a family of generative models generically described as:

$$\begin{aligned}
O_1 &\sim N_3(0, I); O_2 = O_1 + Z, \text{ where } Z \sim N_3(0, I); \\
A_j &\in \{-1, 1\}, \text{ with } P(A_j=1) = \frac{1}{2} \text{ for } j=1, 2; \\
Y &= O_2^T \gamma_1 + A_2 O_2^T \gamma_2 + e, \text{ where } \gamma_1 = \gamma_2 = (c, c, c)^T \text{ and } e \sim N(0, 1).
\end{aligned}$$

The family of models is indexed by a parameter c that represents the effect size and is varied in the set $\{0.5, 1, 1.5, \dots, 5\}$, resulting in 10 example scenarios. The baseline and intermediate set of covariates, denoted by O_1 and O_2 respectively, are three-component vectors; while O_1 is generated from a multivariate normal distribution with zero mean and an identity dispersion matrix I , O_2 is obtained by adding a similar multivariate normal vector Z to O_1 . Following the variables that were actually collected in the smoking cessation study of Strecher et al. [54], the three variables in the above generative model can be potentially conceptualized as standardized versions of continuous scores denoting subjects' motivation to quit smoking, self-efficacy, and pre-treatment level of addiction, measured at baseline and six months post-randomization, respectively.

Generative Model of the Secondary Simulation Study—Here the generative model is described as:

$$\begin{aligned}
O_1 &\sim N_3(0, I); O_2 = (1 + 0.05 A_1) O_1 + 0.5 Z, \text{ where } Z \sim N_3(0, I); \\
A_j &\in \{-1, 1\}, \text{ with } P(A_j=1) = \frac{1}{2} \text{ for } j=1, 2; \\
R &= \mathcal{I}_{\{\sum_{k=1}^3 O_{2k} \leq -1\}}; \\
Y_1 &= O_1^T \gamma_1 + A_1 O_1^T \gamma_2 + e_1; \\
Y_2 &= O_2^T \gamma_1 + A_2 O_2^T \gamma_2 + e_2, \text{ where } \gamma_1 = \gamma_2 = (c, c, c)^T \text{ and } e_1, e_2 \stackrel{i.i.d.}{\sim} N(0, 1); \\
Y &= R \cdot Y_1 + (1 - R) \cdot Y_2.
\end{aligned}$$

Here R is an indicator of treatment response at the end of stage 1; also, Y_1 and Y_2 are the stage-specific outcomes combined to construct the final outcome Y . The treatment at stage 2 is given only to the non-responders of stage 1; responders do not receive any new treatment. The rest of the variables can be conceptualized as in the primary simulation study. The effect size parameter c is set as 0.5.

Q-learning is employed to estimate the optimal DTR. The correctly specified working models for the Q-functions are given by $Q_j(H_j, A_j; \beta_j, \psi_j) = \beta_j^T H_{j0} + \psi_j^T H_{j1} A_j$, $j = 1, 2$, where $H_{20} = (1, O_1^T, A_1 O_1^T A_1)^T$, $H_{21} = (1, O_2^T, A_1)^T$, and $H_{10} = H_{11} = (1 + O_1^T)^T$. Given the observed data, and thus an estimated DTR $\hat{d}$, the true Value of $\hat{d}$, i.e. $V^{\hat{d}}$, is approximated using a separate Monte Carlo evaluation data set of size 10000. We compare the standard percentile bootstrap with the proposed m -out-of- n bootstrap in terms of the mean coverage and mean width of nominal 95% CIs. Comparisons are based on 1000 simulated data sets, each of size $n = 200$, and 1000 bootstrap replications.

5.2 Results

Results of the primary simulation study are shown in the top part of Table 1. The standard bootstrap (n -out-of- n) shows the problem of under-coverage in all the 10 examples tried,

each example denoting a different effect size. The extent of under-coverage is often severe, the lowest observed coverage rate for the nominal 95% bootstrap confidence interval is 83.8%. On the other hand, the proposed subsampling (m -out-of- n , with m chosen via double bootstrap) confidence intervals provide nominal coverage rate in almost all examples. The only exception occurs when the effect size is the smallest in the range we considered ($c = 0.5$); in this example, the coverage of the sub-sampling confidence interval (93.4%) falls marginally below the nominal rate using a binomial test of proportion, while still offering considerable improvement over the standard bootstrap (86.6%). Results of the secondary simulation study are shown in the lower part of Table 1. Here also the subsampling approach offers improved coverage (94.2%) over the standard bootstrap (92.0%). As expected, in both primary and secondary simulation studies, confidence intervals constructed via subsampling are wider than those constructed using the standard bootstrap.

6 Discussion

We proposed a subsampling-based confidence interval for the Value of an estimated optimal DTR when estimation is done using Q -learning. The proposed method is adaptive in that it uses a data-driven resample size. The method is conceptually simple, easily implemented without specialized software, and self-tuning via the double bootstrap. In simulated experiments the proposed method delivered improved performance over the standard bootstrap.

While we used the IPW estimator of the Value in our setup, an alternative is to use the Augmented IPW (AIPW) estimator which is more complex but generally more efficient [e.g., 42]. Whether and how the proposed subsampling confidence interval can be applied in conjunction with the AIPW estimator remains an open question, and thus can be an interesting topic for future research.

An important application of the proposed method is the comparison of two competing DTRs. We discussed the case where the Value of a competing regime (say, standard care) was known and compared with an estimated optimal regime. However, the proposed method can also be applied to construct a confidence interval for the difference between the Value of the estimated optimal regime and that of a competing fixed regime. In particular, one can bootstrap the difference between $V^{\hat{d}}$ and the estimator given in (3) to form a confidence interval for the difference in Values.

Acknowledgments

Bibhas Chakraborty acknowledges support from the NIH grant R01 NS072127-01A1, and the startup grant from the Duke-NUS Graduate Medical School, Singapore. Eric Laber acknowledges support from the NIH grant P01 CA142538.

Funding for this conference was made possible (in part) by 2 R13 CA132565-06 from the National Cancer Institute. The views expressed in written conference materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention by trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

Appendix

Data-driven m -out-of- n subsampling

From (4) it can be seen that the stability of V^d across SMART data sets will depend on the proportion of subjects for which $\min(A_1\hat{\psi}_1^\top H_{11}, A_2\hat{\psi}_2^\top H_{21}) \approx 0$. Furthermore, it is important to note that in a SMART (as in any randomized study), treatment effect sizes are often small, in accordance with the principle of *equipoise* [74]. This prevalence of small effect sizes implies that the proportion of subjects for which $\min(A_1\hat{\psi}_1^\top H_{11}, A_2\hat{\psi}_2^\top H_{21}) \approx 0$ is non-negligible. Intuitively, the estimator will ‘jitter’ more across data sets as this proportion increases. Thus, a natural approach is to let the resample size m depend on an estimator of this proportion.

Let $B_{ji} = (H_{j0,i}^\top, H_{j1,i}^\top A_{ji})$ denote the i -th row of the design matrix for the stage- j regression in Q-learning, and θ_j^* be the true value of $\theta_j = (\beta_j^\top, \psi_j^\top)^\top$, for $j = 1, 2$. Then $\hat{\Sigma}_j^*$, the plug-in estimator of the asymptotic covariance of $\hat{\theta}_j$ is given by:

$$\begin{aligned}\hat{\Sigma}_2^* &= n \left(\sum_{i=1}^n B_{2i} B_{2i}^\top \right)^{-1} \sum_{i=1}^n B_{2i} B_{2i}^\top (Y_i - B_{2i}^\top \theta_2^*)^2 \left(\sum_{i=1}^n B_{2i} B_{2i}^\top \right)^{-1}, \\ \hat{\Sigma}_1^* &= n \left(\sum_{i=1}^n B_{1i} B_{1i}^\top \right)^{-1} \sum_{i=1}^n B_{1i} B_{1i}^\top (\hat{Y}_{1i} - B_{1i}^\top \theta_1^*)^2 \left(\sum_{i=1}^n B_{1i} B_{1i}^\top \right)^{-1}.\end{aligned}$$

Finally, let $\hat{\Sigma}_j$ denote the sub-matrix of $\hat{\Sigma}_j^*$ corresponding to elements of ψ_j , thus $\hat{\Sigma}_j$ represent the asymptotic covariance of $\hat{\psi}_j$. Furthermore, let $\chi_{1,1-\nu}^2$ denote the $(1 - \nu) \times 100$ percentile of a Chi-Square distribution with one degree of freedom. Define ψ_j^* , $j = 1, 2$ to be the almost sure limiting values of $\hat{\psi}_j$, $j = 1, 2$ (see [68] for conditions under which these limits exist). Define

$$\hat{p} = \inf_{\nu \in (0,1)} \left[\frac{1}{n} \sum_{i=1}^n \mathbb{I} \left\{ \min \left(\frac{n(H_{11}^\top \hat{\psi}_1)^2}{H_{11}^\top \hat{\Sigma}_1 H_{11}}, \frac{n(H_{21}^\top \hat{\psi}_2)^2}{H_{21}^\top \hat{\Sigma}_2 H_{21}} \right) \leq \chi_{1,1-\nu}^2 \right\} + \nu \right],$$

where $\mathbb{I}\{\cdot\}$ is an indicator function. Then, under mild regularity conditions, $\hat{p}$ is a conservative estimator of $p = P\{\min(A_1\psi_1^{*\top} H_{11}, A_2\psi_2^{*\top} H_{21}) = 0\}$ in the sense that for any fixed value $\epsilon > 0$, $P(\hat{p} - p = \epsilon) \rightarrow 1$; furthermore, when $p = 0$ it follows that $\hat{p} \rightarrow_p p$. We define the class of data-driven resample sizes, indexed by a tuning parameter $\alpha > 0$, as $\hat{m}_\alpha = n^{(1+\alpha(1-\hat{p}))/(1+\alpha)}$. Note that $\hat{m}_0 \equiv n$ and $\hat{m}_\infty = n^{1-\hat{p}}$; see [11] for further discussion of this class. It can be shown that for $0 < \ell < L < \infty$, $\sup_{\ell \leq \alpha \leq L} \hat{m}_\alpha = o_p(n)$ if $p > 0$, and $\inf_{\ell \leq \alpha \leq L} \hat{m}_\alpha/n \rightarrow_p 1$ if $p = 0$ (see [11] for a proof of a similar result). For this method to work, a value of the tuning parameter α must be chosen. A data-driven approach to choosing α can be devised using the double bootstrap; such an algorithm is described below. R code implementing the proposed method is freely available from the authors.

Double bootstrap algorithm for choosing α

To form a $(1 - \eta) \times 100\%$ confidence interval (CI) for $V^{\hat{d}}$, we provide a double bootstrap procedure [see 75, for details] for choosing the tuning parameter α .

Suppose we have estimated $\hat{\psi}_j, j = 1, 2$, subsequently $\hat{d}$ and $V^{\hat{d}}$ from the original data. Consider a grid of candidate values for α ; we used $\{0.025, 0.05, 0.075, \dots, 1\}$. The algorithm is as follows.

1. Draw B_1 n -out-of- n first-stage bootstrap samples from the data and calculate the bootstrap estimates $V^{\hat{d},(b_1)}, b_1 = 1, \dots, B_1$. Fix α at the smallest value in the grid.
2. Compute the corresponding values of $\hat{m}_{\alpha}^{(b_1)} = n^{(1+\alpha(1-\hat{p}^{(b_1)}))/(1+\alpha)}, b_1 = 1, \dots, B_1$.
3. Conditional on each first-stage bootstrap sample, draw B_2 $\hat{m}^{(b_1)}$ -out-of- n second-stage (nested) bootstrap samples and calculate the double bootstrap versions of the estimate $V^{\hat{d},(b_1,b_2)}, b_1 = 1, \dots, B_1, b_2 = 1, \dots, B_2$.
4. For $b_1 = 1, \dots, B_1$, compute the $(\eta/2) \times 100$ and $(1 - \eta/2) \times 100$ percentiles of $\left\{ \sqrt{\hat{m}^{(b_1)}} \left(\hat{V}^{\hat{d},(b_1,b_2)} - \hat{V}^{\hat{d},(b_1)} \right), b_2 = 1, \dots, B_2 \right\}$; say $\hat{l}_{DB}^{(b_1)}$ and $\hat{u}_{DB}^{(b_1)}$ respectively. Construct the *double centered percentile bootstrap* [67] CI from the b_1 -th first-stage bootstrap data as $(\hat{V}^{\hat{d},(b_1)} - \hat{u}_{DB}^{(b_1)} / \sqrt{\hat{m}^{(b_1)}}, \hat{V}^{\hat{d},(b_1)} - \hat{l}_{DB}^{(b_1)} / \sqrt{\hat{m}^{(b_1)}}), b_1 = 1, \dots, B_1$.
5. Estimate the coverage rate of the double bootstrap CI from all the first-stage bootstrap data sets as

$$\frac{1}{B_1} \sum_{b_1=1}^{B_1} \mathcal{I} \left\{ \hat{V}^{\hat{d},(b_1)} - \hat{u}_{DB}^{(b_1)} / \sqrt{\hat{m}^{(b_1)}} \leq \hat{V}^{\hat{d}} \leq \hat{V}^{\hat{d},(b_1)} - \hat{l}_{DB}^{(b_1)} / \sqrt{\hat{m}^{(b_1)}} \right\}.$$
6. If the above coverage rate is at or above the nominal level, then pick the current value of α as the final value. Otherwise, increment α to the next highest value in the grid.
7. Repeat steps 2 – 6, until the coverage rate of the double bootstrap CI, attains the nominal coverage rate, or the grid is exhausted.

References

1. Lei H, Nahum-Shani I, Lynch K, Oslin D, Murphy SA. A SMART design for building individualized treatment sequences. *The Annual Review of Psychology*. 2012; 8:21–48.
2. Chakraborty B, Murphy SA, Strecher V. Inference for non-regular parameters in optimal dynamic treatment regimes. *Statistical Methods in Medical Research*. 2010; 19:317–343. [PubMed: 19608604]

3. Thall PF, Logothetis C, Pagliaro LC, Wen S, Brown MA, Williams D, et al. Adaptive Therapy for Androgen-Independent Prostate Cancer: A Randomized Selection Trial of Four Regimens. *Journal of the National Cancer Institute*. 2007; 99:1613–1622. [PubMed: 17971530]
4. Miyahara S, Wahed AS. Weighted Kaplan-Meier estimators for two-stage treatment regimes. *Statistics in Medicine*. 2010; 29:2581–2591. [PubMed: 20799259]
5. Zhao Y, Zeng D, Socinski MA, Kosorok MR. Reinforcement learning strategies for clinical trials in nonsmall cell lung cancer. *Biometrics*. 2011; 67:1422–1433. [PubMed: 21385164]
6. Wang L, Rotnitzky A, Lin X, Millikan RE, Thall PF. Evaluation of viable dynamic treatment regimes in a sequentially randomized trial of advanced prostate cancer. *Journal of the American Statistical Association*. 2012; 107:493–508. [PubMed: 22956855]
7. Khalili S, Armaou A. An extracellular stochastic model of early HIV infection and the formulation of optimal treatment policy. *Chemical Engineering Science*. 2008; 63:4361–4372.
8. Robins JM, Orellana L, Rotnitzky A. Estimation and extrapolation of optimal treatment and testing strategies. *Statistics in Medicine*. 2008; 27:4678–4721. [PubMed: 18646286]
9. Dawson R, Lavori PW. Placebo-free designs for evaluating new mental health treatments: The use of adaptive treatment strategies. *Statistics in Medicine*. 2004; 23:3249–3262. [PubMed: 15490427]
10. Shortreed SM, Moodie EEM. Estimating the optimal dynamic antipsychotic treatment regime: Evidence from the sequential-multiple assignment randomized CATIE Schizophrenia Study. *Journal of the Royal Statistical Society, Series C*. 2012; 61:577–599.
11. Chakraborty B, Laber EB, Zhao Y. Inference for optimal dynamic treatment regimes using an adaptive m -out-of- n bootstrap scheme. *Biometrics*. 2013; 69:714–723. [PubMed: 23845276]
12. Wagner EH, Austin BT, Davis C, Hindmarsh M, Schaefer J, Bonomi A. Improving chronic illness care: Translating evidence into action. *Health Affairs*. 20:64–78.
13. Chakraborty, B.; Moodie, EEM. *Statistical Methods for Dynamic Treatment Regimes: Reinforcement Learning, Causal Inference, and Personalized Medicine*. New York: Springer; 2013.
14. Marlowe DB, Festinger DS, Arabia PL, Dugosh KL, Benasutti KM, Croft JR, et al. Adaptive Interventions in Drug Court: A Pilot Experiment. *Criminal Justice Review*. 2008; 33:343–360. [PubMed: 19724664]
15. Marlowe DB, Festinger DS, Arabia PL, Dugosh KL, Benasutti KM, Croft JR. Adaptive interventions may optimize outcomes in drug courts: A pilot study. *Current Psychiatry Reports*. 2009; 11:370–376. [PubMed: 19785978]
16. Lavori PW, Dawson R. A design for testing clinical strategies: Biased adaptive within-subject randomization. *Journal of the Royal Statistical Society, Series A*. 2000; 163:29–38.
17. Lavori PW, Dawson R. Dynamic treatment regimes: Practical design considerations. *Clinical Trials*. 2004; 1:9–20. [PubMed: 16281458]
18. Murphy SA. An experimental design for the development of adaptive treatment strategies. *Statistics in Medicine*. 2005; 24:1455–1481. [PubMed: 15586395]
19. Tummarello D, Mari D, Graziano F, Isidori P, Cetto G, Pasini F, et al. A randomized, controlled phase III study of cyclophosphamide, doxorubicin, and vincristine with etoposide (CAV-E) or teniposide (CAV-T), followed by recombinant interferon- α maintenance therapy or observation, in small cell lung carcinoma patients with complete responses. *Cancer*. 1997; 80:2222–2229. [PubMed: 9404698]
20. Matthey KK, Villablanca JG, Seeger RC, Stram DO, Harris RE, Ramsay NK, et al. Treatment of high-risk neuroblastoma with intensive chemotherapy, radiotherapy, autologous bone marrow transplantation, and 13-cis-retinoic acid. *New England Journal of Medicine*. 1999; 341:1165–1173. [PubMed: 10519894]
21. Habermann TM, Weller EA, Morrison VA, Gascoyne RD, Cassileth PA, Cohn JB, et al. Rituximab-CHOP versus CHOP alone or with maintenance rituximab in older patients with diffuse large B-cell lymphoma. *Journal of Clinical Oncology*. 2006; 24:3121–3127. [PubMed: 16754935]
22. Auyeung SF, Long Q, Royster EB, Murthy S, McNutt MD, Lawson D, et al. Sequential multiple-assignment randomized trial design of neurobehavioral treatment for patients with metastatic malignant melanoma undergoing high-dose interferon-alpha therapy. *Clinical Trials*. 2009; 6:480–490. [PubMed: 19786415]

23. Mateos MV, Oriol A, Martínez-López J, Gutiérrez N, Teruel AI, de Paz R, et al. Bortezomib, melphalan, and prednisone versus bortezomib, thalidomide, and prednisone as induction therapy followed by maintenance treatment with bortezomib and thalidomide versus bortezomib and prednisone in elderly patients with untreated multiple myeloma: A randomised trial. *The Lancet Oncology*. 2010; 11:934–941. [PubMed: 20739218]
24. Kasari, C. ClinicalTrials.gov database, updated April 26, 2012. National Institutes of Health; Bethesda, MD: 2009. Developmental and augmented intervention for facilitating expressive language (CC-NIA). <http://clinicaltrials.gov/ct2/show/NCT01013545> [accessed on July 24, 2013]
25. Nahum-Shani I, Qian M, Almiral D, Pelham W, Gnagy B, Fabiano G, et al. Experimental design and primary data analysis methods for comparing adaptive interventions. *Psychological Methods*. 2012a; 17:457–477. [PubMed: 23025433]
26. Nahum-Shani I, Qian M, Almiral D, Pelham W, Gnagy B, Fabiano G, et al. Q-Learning: A data analysis method for constructing adaptive interventions. *Psychological Methods*. 2012b; 17:478–494. [PubMed: 23025434]
27. Jones, H. Clinical-Trials.gov database, updated October 19, 2012. National Institutes of Health; Bethesda, MD: 2010. Reinforcement-based treatment for pregnant drug abusers (HOME II). <http://clinicaltrials.gov/ct2/show/NCT01177982> [accessed on July 24, 2013]
28. Methodology Center. [accessed on May 4, 2014] Projects involving SMART design. <http://methodology.psu.edu/ra/adap-treat-strat/projects>
29. Laber, EB. [accessed on May 4, 2014] List of SMART studies. <http://www4.stat.ncsu.edu/~eblaber/smart>
30. Methodology Center. [accessed on May 4, 2014] NIH Funding for SMART. <http://methodology.psu.edu/ra/adap-inter/NIHfunding>
31. Dawson R, Lavori PW. Sample size calculations for evaluating treatment policies in multistage designs. *Clinical Trials*. 2010; 7:643–652. [PubMed: 20630903]
32. Oetting, AI.; Levy, JA.; Weiss, RD.; Murphy, SA. Statistical methodology for a SMART design in the development of adaptive treatment strategies. In: Shrout, PE.; Keyes, KM.; Ornstein, K., editors. *Causality and Psychopathology: Finding the Determinants of Disorders and Their Cures*. New York: Oxford University Press; 2011. p. 179-205.
33. Almiral D, Compton SN, Gunlicks-Stoessel M, Duan N, Murphy SA. Designing a pilot sequential multiple assignment randomized trial for developing an adaptive treatment strategy. *Statistics in Medicine*. 2012; 31:1887–1902. [PubMed: 22438190]
34. Almiral D, Lizotte D, Murphy SA. SMART design issues and the consideration of opposing outcomes: Discussion of “Evaluation of viable dynamic treatment regimes in a sequentially randomized trial of advanced prostate cancer” by Wang et al. *Journal of the American Statistical Association*. 2012 In press.
35. Orellana L, Rotnitzky A, Robins JM. Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, Part I: Main Content. *The International Journal of Biostatistics*. 2010; 6
36. Moodie EEM, Chakraborty B, Kramer MS. Q-learning for estimating optimal dynamic treatment rules from observational data. *Canadian Journal of Statistics*. 2012; 40:629–645. [PubMed: 23355757]
37. Robins, JM. Optimal structural nested models for optimal sequential decisions. In: Lin, DY.; Heagerty, P., editors. *Proceedings of the Second Seattle Symposium on Biostatistics*. New York: Springer; 2004. p. 189-326.
38. Murphy SA. A generalization error for Q-learning. *Journal of Machine Learning Research*. 2005b; 6:1073–1097. [PubMed: 16763665]
39. Zhao Y, Kosorok MR, Zeng D. Reinforcement learning design for cancer clinical trials. *Statistics in Medicine*. 2009; 28:3294–3315. [PubMed: 19750510]
40. Zhao YQ, Zeng D, Rush AJ, Kosorok MR. Estimating individual treatment rules using outcome weighted learning. *Journal of the American Statistical Association*. 2012; 107:1106–1118. [PubMed: 23630406]
41. Zhang B, Tsiatis AA, Laber EB, Davidian M. A robust method for estimating optimal treatment regimes. *Biometrics*. 2012; 68:1010–1018. [PubMed: 22550953]

42. Zhang B, Tsiatis AA, Laber EB, Davidian M. Robust estimation of optimal dynamic treatment regimes for sequential treatment decisions. *Biometrika*. 2013; 100:681–694.
43. Laber EB, Murphy SA. Adaptive confidence intervals for the test error in classification. *Journal of the American Statistical Association*. 2011; 106:904–913. [PubMed: 22053123]
44. Lunceford JK, Davidian M, Tsiatis AA. Estimation of survival distributions of treatment policies in two-stage randomization designs in clinical trials. *Biometrics*. 2002; 58:48–57. [PubMed: 11890326]
45. Wahed AS, Tsiatis AA. Optimal estimator for the survival distribution and related quantities for treatment policies in two-stage randomized designs in clinical trials. *Biometrics*. 2004; 60:124–133. [PubMed: 15032782]
46. Wahed AS, Tsiatis AA. Semiparametric efficient estimation of survival distributions in two-stage randomisation designs in clinical trials with censored data. *Biometrika*. 2006; 93:163–177.
47. Thall PF, Millikan RE, Sung HG. Evaluating multiple treatment courses in clinical trials. *Statistics in Medicine*. 2000; 30:1011–1128.
48. Thall PF, Sung HG, Estey EH. Selecting therapeutic strategies based on efficacy and death in multi- course clinical trials. *Journal of the American Statistical Association*. 2002; 97:29–39.
49. Politis, DN.; Romano, JP.; Wolf, M. Subsampling. New York: Springer Verlag; 1999.
50. Robins JM. A new approach to causal inference in mortality studies with sustained exposure periods – Application to control of the healthy worker survivor effect. *Mathematical Modelling*. 1986; 7:1393–1512.
51. Robins, JM. The analysis of randomized and nonrandomized AIDS treatment trials using a new approach to causal inference in longitudinal studies. In: Sechrest, L.; Freeman, H.; Mulley, A., editors. *Health Service Research Methodology: A Focus on AIDS*. New York: NCHSR, U.S. Public Health Service; 1989. p. 113-159.
52. Robins JM. Information recovery and bias adjustment in proportional hazards regression analysis of randomized trials using surrogate markers. *Proceedings of the Biopharmaceutical Section, American Statistical Association*. 1993:24–33.
53. Robins, JM. Causal inference from complex longitudinal data. In: Berkane, M., editor. *Latent Variable Modeling and Applications to Causality: Lecture Notes in Statistics*. New York: Springer-Verlag; 1997. p. 69-117.
54. Strecher V, McClure J, Alexander G, Chakraborty B, Nair V, Konkel J, et al. Web-based smoking cessation components and tailoring depth: Results of a randomized trial. *American Journal of Preventive Medicine*. 2008; 34:373–381. [PubMed: 18407003]
55. Schulte PJ, Tsiatis AA, Laber EB, Davidian M. Q- and A-learning methods for estimating optimal dynamic treatment regimes. *Statistical Science*. 2014 In press.
56. Robins JM. Correcting for non-compliance in randomized trials using structural nested mean models. *Communications in Statistics*. 1994; 23:2379–2412.
57. Robins, JM. Marginal structural models versus structural nested models as tools for causal inference. In: Halloran, ME.; Berry, D., editors. *Statistical Models in Epidemiology, the Environment, and Clinical Trials*, vol 116 of IMA. New York: Springer; 1999. p. 95p. 134
58. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000; 11:550–560. [PubMed: 10955408]
59. Murphy SA, der Laan MJV, Robins JM. CPRG. Marginal mean models for dynamic regimes. *Journal of the American Statistical Association*. 2001; 96:1410–1423. [PubMed: 20019887]
60. Sutton, RS.; Barto, AG. *Reinforcement Learning: An introduction*. Cambridge: MIT Press; 1998.
61. Xin, J.; Chakraborty, B.; Laber, EB. [accessed on May 4, 2014] qLearn. 2012. <http://cran.r-project.org/web/packages/qLearn/index.html>
62. Dawid A. Selection paradoxes of Bayesian inference. *Lecture Notes – Monograph Series*. 1994:211–220.
63. Efron B. Estimating the error rate of a prediction rule: improvement on cross-validation. *Journal of the American Statistical Association*. 1983; 78:316–331.
64. Hand DJ. Recent advances in error rate estimation. *Pattern Recognition Letters*. 1986; 4:335–346.

65. Hastie, T.; Tibshirani, R.; Friedman, J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd. New York: Springer; 2009.
66. Zhao YQ, Zeng D, Laber EB, Kosorok MR. New statistical learning methods for estimating optimal dynamic treatment regimes. 2013 Submitted.
67. Efron, B.; Tibshirani, R. An Introduction to the Bootstrap. Chapman & Hall/CRC; 1993.
68. Laber EB, Qian M, Lizotte D, Murphy SA. Statistical inference in dynamic treatment regimes. 2011 Submitted.
69. Bretagnolle, J. Annales de l'IHP Probabilit'es et statistiques. Vol. 19. Elsevier; 1983. Lois limites du bootstrap de certaines fonctionnelles; p. 281-296.
70. Swanepoel JWH. A note on proving that the (modified) bootstrap works. Communications in Statistics – Theory and Methods. 1986; 15:3193–3203.
71. Dümbgen L. On nondifferentiable functions and the bootstrap. Probability Theory and Related Fields. 1993; 95:125–140.
72. Shao J. Bootstrap sample size in nonregular cases. Proceedings of the American Mathematical Society. 1994; 122:1251–1262.
73. Bickel P, Gotze F, van Zwet W. Resampling fewer than n observations: Gains, losses and remedies for losses. Statistica Sinica. 1997; 7:1–31.
74. Freedman B. Equipoise and the ethics of clinical research. New England Journal of Medicine. 1987; 317:141–145. [PubMed: 3600702]
75. Davison, AC.; Hinkley, DV. Bootstrap Methods and Their Application. Cambridge, UK: Cambridge University Press; 1997.

Figure 1. Hypothetical SMART design schematic for the smoking cessation example. An “R” within a circle denotes randomization.

Table 1

Coverage rate and mean width of nominal 95% CIs for $V_{\hat{d}}$ based on 1000 Monte Carlo simulations ($n = 200$). Coverage rates significantly smaller than the nominal rate are written in bold font.

c	Coverage		Mean Width	
	Bootstrap	Subsampling	Bootstrap	Subsampling
Primary Simulation Study				
0.5	0.866	0.934	0.976	1.197
1	0.894	0.966	1.616	1.968
1.5	0.838	0.948	2.324	2.829
2	0.882	0.952	3.036	3.771
2.5	0.894	0.956	3.791	4.642
3	0.876	0.954	4.508	5.481
3.5	0.838	0.944	5.323	6.616
4	0.858	0.944	5.898	7.188
4.5	0.854	0.942	6.686	8.138
5	0.870	0.940	7.421	9.063
Secondary Simulation Study				
0.5	0.920	0.942	0.704	0.730