



Published in final edited form as:

Biometrics. 2010 June ; 66(2): 512–522. doi:10.1111/j.1541-0420.2009.01282.x.

A Generalized Estimator of the Attributable Benefit of an Optimal Treatment Regime

Jason Brinkley^{1,*}, Anastasios Tsiatis^{2,**}, and Kevin J. Anstrom^{3,***}

¹ Department of Biostatistics, East Carolina University, Greenville, NC, U.S.A

² Department of Statistics, North Carolina State University, Raleigh, NC, U.S.A

³ Duke University Medical Center, Duke University, Durham, NC, U.S.A

Summary

For many diseases where there are several treatment options often there is no consensus on the best treatment to give individual patients. In such cases it may be necessary to define a strategy for treatment assignment; that is, an algorithm which dictates the treatment an individual should receive based on their measured characteristics. Such a strategy or algorithm is also referred to as a treatment regime. The optimal treatment regime is the strategy that would provide the most public health benefit by minimizing as many poor outcomes as possible. Using a measure that is a generalization of attributable risk and notions of potential outcomes, we derive an estimator for the proportion of events that could have been prevented had the optimal treatment regime been implemented. Traditional attributable risk studies look at the added risk that can be attributed to exposure of some contaminant, here we will instead study the benefit that can be attributed to using the optimal treatment strategy.

We will show how regression models can be used to estimate the optimal treatment strategy and the attributable benefit of that strategy. We also derive the large sample properties of this estimator. As a motivating example we will apply our methods to an observational study of 3856 patients treated at the Duke University Medical Center with prior coronary artery bypass graft surgery and further heart related problems requiring a catheterization. The patients may be treated with either medical therapy alone or a combination of medical therapy and percutaneous coronary intervention without general consensus on which is the best treatment for individual patients.

Keywords

Attributable Risk; Causal Inference; Influence Function; Optimal Treatment Regime

1. INTRODUCTION

Since its inception by Levin (1953), attributable risk (*AR*) has become a widely used measure of association between risk factors and disease status. Jewell (2004) denotes the population attributable risk as the fraction of all disease cases that can be linked to exposure from some factor. *AR* has been referred to by many other names (e.g. ‘etiologic fraction’ or

*brinkleyj@ecu.edu

**tsiatis@stat.ncsu.edu

***kevin.anstrom@duke.edu

‘attributable fraction’) and has a long history within the medical and epidemiology communities.

Observational data can be used to assess the impact of various treatment strategies on patient health outcomes, where a treatment strategy is defined as an algorithm which dictates the treatment an individual should receive based on their characteristics. Furthermore, we define the optimal treatment strategy to be the strategy that will result in the smallest proportion of poor outcomes (events) among the study population. In this manuscript, we propose a generalization of the attributable risk measure which is defined as the fraction of observed events that could have been prevented had the optimal treatment strategy been implemented. In contrast to traditional *AR* studies where the goal is to measure the added risk attributed to exposure of some contaminant, this estimator is intended to study the positive public health impact of a treatment strategy. As such we will refer to it as attributable benefit (*AB*).

For simplicity we consider the case where there are two treatment options which will be labeled by the treatment indicator $T = 0$ or 1 . We denote the event of interest by the binary indicator Y where $Y = 1$ denotes poor outcome and $Y = 0$ denotes good outcome. Let X denote the vector of covariates measured on each individual that may be used to determine which treatment they receive, we denote a treatment strategy g as a function that maps values of X to 0 or 1. That is, $g(x)$ denotes the treatment T (either 1 or 0) that will be assigned to an individual with covariates $X = x$.

The goals of this paper are to define and estimate the optimal treatment strategy $g_{opt}(x)$; i.e., the treatment strategy which results in the smallest probability of a poor outcome, and to estimate the attributable benefit of using the optimal regime. We will show that our proposed estimator is a generalization of the standard attributable risk measure. In Section 2, we describe the motivating example in more detail. In Section 3, we define the optimal treatment strategy and develop an estimator for attributable benefit that is a function of an individual's covariates. In order to formally define the parameter of interest and the required assumptions, we will use the ideas of potential outcomes and causal inference as described by Neyman (1923) and Rubin (1974). In Section 4, the large sample properties of the estimator are developed and two strategies for developing confidence intervals are derived. The behavior of the proposed estimator is described in the simulation studies of Section 5 and the example in Section 6. We conclude with discussions and limitations of the current method.

2. Motivating Problem

Our motivating problem comes from the Duke Databank for Cardiovascular Diseases (DDCD), a large database housed at the Duke University Medical Center (DUMC). Started in 1969, the DDCD collects data on the in-hospital clinical course of all patients who undergo cardiac catheterization at Duke University Medical Center. A wide variety of baseline patient information is collected and the DDCD attempts to collect follow-up data on patients with at least one diseased vessel six months and one year after hospitalization and yearly thereafter. Data are collected in the form of a mailed questionnaire or surveyed by phone for those who do not respond to the mailing (Al-Khatib *et al.*, 2007). One of the largest and oldest such databases, the DDCD provides essential information to cardiologists about the effectiveness of various treatments on patient health.

Here we focus on the subset of patients with a previous coronary artery bypass graft (CABG) surgery and requiring a later catheterization due to continued symptoms. For technical reasons, the vast majority of patients are not considered to be candidates for a second CABG surgery and therefore their treatment options are limited to either medical therapies (MED) or some combination of medical therapies and percutaneous coronary

intervention (PCI) which is also known as angioplasty. A question of clinical interest is whether PCI provides some benefit (here in the form of lower 1-year mortality rates) above optimal medical therapy for either the overall study population or some subset defined by clinical characteristics.

3. Potential Outcomes and Statistical Causality

Using the ideas and notation developed by Rubin (1974) we define the potential outcomes $Y^*(1)$ and $Y^*(0)$ to denote the indicator of a poor outcome (event) for an arbitrary individual in our population had they received treatment 1 or 0 (respectively). In actuality, at most one of these potential outcomes can be observed for any given individual. In addition, we also collect baseline covariates X on these individuals. Consequently, the potential outcomes for any individual will be defined as $\mathbf{W} = \{Y^*(1), Y^*(0), X\}$ and it is assumed that there is some underlying population distribution for the potential outcomes W .

In contrast, the observable data will be denoted by $\mathbf{O} = (Y, T, X)$ where T denotes the treatment (0 or 1) that was actually given to an individual, Y denotes the outcome indicator that resulted from that treatment, and X is the vector of baseline covariates measured on the individuals. The key assumptions that will allow us to derive important aspects of the distribution of the potential outcomes, which are not directly observable, from the distribution of the observable data were proposed by Rubin (1974).

The first assumption, also known as the Stable Unit Treatment Value Assumption (SUTVA), states that the observed outcome for an individual is the same as the potential outcome for the corresponding treatment that the individual received. An underlying and often overlooked part of this assumption is that under SUTVA, we assume that there is no interference between units. According to Rubin (1986), the SUTVA assumption implies that the value of the potential outcome for a subject does not depend on what treatments other subjects receive. In terms of the current framework, the SUTVA assumption can be written as $Y = Y^*(1)T + Y^*(0)(1 - T)$.

The second assumption is also referred to as the strong ignorability assumption or the no unmeasured confounders assumption. This assumption states that the treatment T assigned to an individual is independent of the potential outcomes $\{Y^*(1), Y^*(0)\}$ conditional on the covariates X . In an observational study the treatment that a physician chooses to give a patient can be reasonably assumed to only be based on characteristics of the patient at the time of treatment and not on the patients potential outcome (which of course is not known at the time of treatment). Consequently, this second assumption will be tenable if the key factors influencing the decision making process for a physician were captured in the data X . If, however, there are additional factors beyond the data X that influence treatment decisions, then this assumption may not hold. We discuss this issue further in Section 7; however, for the remainder of this paper we will assume both SUTVA and strong ignorability hold.

Under these assumptions, for example, we can compute the marginal probability $P\{Y^*(0) = 1\}$ from the distribution of the observed data as

$$\begin{aligned} P\{Y^*(0)=1\} &= E_X [P\{Y^*(0)=1|X\}] \\ &= E_X [P\{Y^*(0)=1|X, T=0\}] \\ &= E_X \{P\{Y=1|X, T=0\}\}, \end{aligned}$$

where we use the notation $E_X(\cdot)$ to emphasize that this expectation is taken with respect to the marginal distribution of X . Similarly, we can show that

$$P\{Y^*(1)=1\}=E_X\{P(Y=1|X, T=1)\}.$$

3.1 Optimal Treatment Regime

As mentioned previously, a treatment regime is an algorithm which assigns treatment to a patient according to factors that are measured on the individual. As such, we define a treatment regime to be a function g which maps the values of X to 0 or 1. Hence the treatment regime would assign treatment $g(x)$ (either 0 or 1) to an individual whose covariates $X = x$. We also denote by \mathcal{g} the class of all such treatment regimes. We can now define the potential outcome for an arbitrary individual in our population under the hypothetical situation that treatment regime g was being implemented by $Y^*\{g(X)\} = Y^*(1)I\{g(X) = 1\} + Y^*(0)I\{g(X) = 0\}$ where $I(\cdot)$ denotes the indicator function. The most simple strategies would be g_0 or g_1 where all patients were given treatment 0 or 1 (respectively) regardless of the values of their covariates X .

We now define the optimal strategy as the one which results in the smallest probability of a poor outcome. That is,

$$g_{opt}(X) = \arg \min_{g \in \mathcal{G}} (E[Y^*\{g(X)\}]).$$

Under the SUTVA and no unmeasured confounders assumption, it is straightforward to show that

$$\begin{aligned} E[Y^*\{g(X)\}] &= P\{Y^*\{g(X)\}=1\} \\ &= E_X[P(Y=1|T=1, X)I\{g(X)=1\} \\ &\quad + P(Y=1|T=0, X)I\{g(X)=0\}], \end{aligned} \tag{1}$$

and the optimal treatment regime is given by

$$g_{opt}(X) = I\{P(Y=1|T=1, X) - P(Y=1|T=0, X) < 0\}. \tag{2}$$

That is, the optimal treatment regime would assign treatment 1 to an individual with baseline covariates X if the conditional probability that $Y = 1$ given $T = 1$ and X is smaller than the conditional probability $Y = 1$ given $T = 0$ and X ; otherwise give treatment 0.

Now that we have some notation and insight into the different treatment regimes, we can formally define the parameter of interest. Recall that the objective is to estimate the proportion of events that would have been prevented had the optimal strategy been used. We define this quantity as the attributable benefit of using the optimal treatment regime. We start by defining the attributable benefit for any treatment regime, say $g(X)$. From the notation above we write AB_g as:

$$AB_g = \frac{P(Y=1) - P\{Y^*\{g(X)\}=1\}}{P(Y=1)} \quad (3)$$

$$= 1 - \frac{P\{Y^*\{g(X)\}=1\}}{P(Y=1)}. \quad (4)$$

Thus we define AB_{opt} as:

$$AB_{opt} = 1 - \frac{P\{Y^*(g_{opt})=1\}}{P(Y=1)}. \quad (5)$$

Since $P(Y=1)$ denotes the overall probability of failure in the population using whatever is the current clinical practice, then this can be estimated by the simple sample proportion

$\bar{Y}_n = \sum_{i=1}^n Y_i/n$ without making any additional assumptions. Therefore, we will focus attention on estimating the treatment regime g_{opt} and the proportion of events for that regime $P\{Y^*(g_{opt})=1\}$. Then we want to show how these results can be used to derive the estimator and large sample properties for AB_{opt} .

We note that AB_g is a more general type of adjusted attributable risk that is commonly used in epidemiological research. That is, if we were considering estimating the proportion of disease that can be attributed to exposure ($T=1$), then we would want to estimate

$$\frac{P(Y=1) - P\{Y^*(0)=1\}}{P(Y=1)},$$

which under the assumptions of SUTVA and no unmeasured confounders is equal to

$$\frac{P(Y=1) - E_x\{P(Y=1|T=0, X)\}}{P(Y=1)}, \quad (6)$$

which is what is being estimated using the estimators for adjusted attributable risk given by Benichou (2001) as well as Graubard and Fears (2005).

Returning to the more general case, we now consider how to estimate $P\{Y^*(g_{opt})=1\}$ from the sample of observed data (Y_i, T_i, X_i) , $i = 1, \dots, n$ assumed to be independent and identically distributed (i.i.d.). We begin by positing a statistical model where

$$P(Y=1|T, X) = \mu(T, X, \beta). \quad (7)$$

Because Y is a binary variable, a natural choice is to use logistic regression models which are used to study the impact of treatment, covariates, and their interaction. If we define $d(x, \beta) = q\{\mu(1, x, \beta)\} - q\{\mu(0, x, \beta)\}$ where $q(\cdot)$ is some arbitrary monotone function, then the

optimal treatment regime defined by (2) can also be written as $g_{opt}(x) = I\{d(x, \beta_0) < 0\}$, where β_0 denotes the true value of β . We introduce the monotone function $q(\cdot)$ because, as we will see later, for the logistic regression models that we will be considering it will be more convenient to work with the logit function; i.e., where $q(p) = \log\{p/(1-p)\}$. A natural estimator for $g_{opt}(x)$ is given by $\hat{g}_{opt}(x) = A(x, \hat{\beta}_n)$, where $A(x, \beta) = I\{d(x, \beta) < 0\}$ and $\hat{\beta}_n$ is the maximum likelihood estimator for β obtained by maximizing (in β)

$$\prod_{i=1}^n \{\mu(T_i, X_i, \beta)\}^{Y_i} \{1 - \mu(T_i, X_i, \beta)\}^{1-Y_i}.$$

For many practical situations there may exist a preferred or standard treatment, say treatment 0. As suggested to us by the Editor, for such scenarios, we may only want to give treatment 1 when this treatment is shown to be significantly better than treatment 0. Toward that end we also want to consider the more conservative estimated treatment strategy $\hat{g}_{opt}(x, c) = A(x, \hat{\beta}_n, c)$, where $A(x, \beta, c) = I\{d(x, \beta) < c \times \text{se}\{d(x, \hat{\beta}_n)\}\}$ and $\text{se}\{d(x, \hat{\beta}_n)\}$ denotes the standard error of $d(x, \hat{\beta}_n)$ which can be derived using the usual large sample properties of the maximum likelihood estimator $\hat{\beta}_n$ and the delta method. That is, treatment 1 will be preferred only for values of x where $d(x, \hat{\beta}_n)/\text{se}\{d(x, \hat{\beta}_n)\} < c$, where c (generally a negative number) is chosen according to the strength of desired significance; for example, at the α level of significance we would use $\Phi^{-1}(\alpha)$, where Φ is the standard normal distribution function. Notice that $\hat{g}_{opt}(x, c) = \hat{g}_{opt}(x)$ when $c = 0$ and because $\hat{\beta}_n$ converges in probability to β_0 and $\text{se}\{d(x, \hat{\beta}_n)\}$ converges in probability to zero as n goes to infinity, this implies that $\hat{g}_{opt}(x, c)$ will converge to $g_{opt}(x)$ (the optimal treatment regime) as n goes to infinity for any constant c . However, the smaller the value of c (i.e., the smaller the significance level α), the more conservative the strategy is in the sense that we are less likely to give the non-standard treatment.

Now if we define

$$h(x, \beta, c) = \mu(1, x, \beta)A(x, \beta, c) + \mu(0, x, \beta)\{1 - A(x, \beta, c)\}, \quad (8)$$

then in accordance to equations (1) and (2), $P\{Y^*(g_{opt}) = 1\} = E\{h(X, \beta_0, 0)\}$ for which a natural estimator (allowing for the conservative strategy) would be

$$\widehat{P}_n\{Y^*(g_{opt})=1, c\} = n^{-1} \sum_{i=1}^n h(X_i, \widehat{\beta}_n, c).$$

Finally, the estimator for attributable benefit for the optimal treatment strategy would be

$$\widehat{AB}_{opt}(c) = 1 - \frac{n^{-1} \sum_{i=1}^n h(X_i, \widehat{\beta}_n, c)}{\bar{Y}_n}.$$

The standard estimator for AB_{opt} without using the conservative strategy will also be denoted by \widehat{AB}_{opt} ; that is, $\widehat{AB}_{opt} = \widehat{AB}_{opt}(0)$.

Because the response Y is binary, we generally will posit a logistic regression model,

$$\text{logit}\{P(Y_i=1|X_i, T_i)\} = \beta^T f(X_i, T_i), \quad (9)$$

where $f(X_i, T_i)$ is a $k \times 1$ vector. This will allow sufficient flexibility for the model to include main effect terms involving treatment and the covariates as well as interactions between treatment and covariates.

Remark 1—When considering the optimal treatment strategy we are entertaining the possibility of assigning different treatments depending on the value of the measured covariates X . Statistical models where $\mu(1, x, \beta) < \mu(0, x, \beta)$ (i.e., treatment 1 is preferred) for some values of x and $\mu(1, x, \beta) > \mu(0, x, \beta)$ (i.e., treatment 0 is preferred) for other values of x are said to have qualitative interaction. When building regression models using data from observational studies or clinical trials, we often find significant treatment by covariate interaction terms. Depending on the magnitude of the effect of such interactions these sometimes result in qualitative interactions and sometimes they do not. In the latter case, such interactions are referred to as quantitative interactions.

The issue of qualitative interactions is controversial. It is not unusual in the analysis of a particular study to discover qualitative interactions which are then found to be spurious in subsequent studies. This has caused many warnings by statisticians to be cautious about the interpretation of the results of such models. Formal statistical tests for qualitative interactions have been developed; for example, see Gail and Simon (1985). In fact, Peto (1995) argues that quantitative interactions may be expected in clinical trials but qualitative interactions are not.

Although it may be argued that in well controlled targeted clinical trials qualitative interactions are not to be expected, the setting we are considering in this paper is one in which different treatments are actually being used in practice with no general consensus on which one is the superior treatment. In such settings, it may be the case that different treatments may be better or worse for different subgroups of patients. Moreover, thanks to the recommendation made by the Editor, if there are situations where one treatment is considered to be standard, then we are also employing a conservative treatment strategy where we only recommend the alternative treatment when there is strong statistical evidence that it is significantly better. Such a strategy should alleviate the difficulties that spurious qualitative interactions create to some extent.

There may also be the belief that quantitative interactions need not be considered in a model because the inclusion of such terms will not affect treatment choice. Indeed, leaving out quantitative interactions in a model will generally not affect the optimal treatment regime but, as we will demonstrate later in some of our simulation studies, not including important interaction terms, even if they are quantitative interactions, will lead to biased estimators of the attributable benefit.

4. Large Sample Properties and Standard Errors

Because the estimator for attributable benefit involves the ratio of two estimators, it will be convenient to work with the transformation $\log(1 - AB_{opt})$. We use large sample theory to

show that $\log\{1 - \widehat{AB}_{opt}(c)\} = \log[\widehat{P}_n\{Y^*(g_{opt})=1, c\}] - \log(\bar{Y}_n)$ is asymptotically normal and derive standard errors via the sandwich variance estimator. Most of this work surrounds finding the influence function for $\log\{1 - \widehat{AB}_{opt}(c)\}$, the details are available in the appendix. The reader may see Tsiatis (2006) for a complete overview of the theory of influence functions.

In the case of a regression model (9), we estimate the influence function for $\log\{1 - \widehat{AB}_{opt}(c)\}$ for the i^{th} individual, ψ_i , from observed data with the quantity

$$\widehat{\psi}_i = \frac{\Delta_{1i} + \Delta_{2i}}{\Delta_3} - \Delta_{4i}, \tag{10}$$

where

$$\begin{aligned} \Delta_{1i} &= h(X_i, \widehat{\beta}_n, c) - n^{-1} \sum_{i=1}^n h(X_i, \widehat{\beta}_n, c), \\ \Delta_{2i} &= \frac{\widehat{\partial} E\{h(X, \beta_0, 0)\}}{\widehat{\partial} \beta^T} \left(n^{-1} \sum_{i=1}^n \left[\frac{f_i f_i^T \exp(\widehat{\beta}_n^T f_i)}{\{1 + \exp(\widehat{\beta}_n^T f_i)\}^2} \right] \right)^{-1} \\ &\quad \times f_i \left\{ Y_i - \frac{\exp(\widehat{\beta}_n^T f_i)}{1 + \exp(\widehat{\beta}_n^T f_i)} \right\}, \\ \Delta_3 &= n^{-1} \sum_{i=1}^n h(X_i, \widehat{\beta}_n, c), \\ \Delta_{4i} &= \frac{Y_i - n^{-1} \sum_{i=1}^n Y_i}{n^{-1} \sum_{i=1}^n Y_i}. \end{aligned}$$

Note that the estimated gradient $\widehat{\partial} E\{h(X, \beta_0, 0)\} / \widehat{\partial} \beta^T$ is derived using numerical derivatives as described in the appendix (A.6).

Now that we have derived an estimated influence function for $\log\{1 - \widehat{AB}_{opt}(c)\}$, the estimated variance, using the sandwich variance estimator, is given by

$$\widehat{V} = n^{-1} \sum_{i=1}^n \left(\frac{\Delta_{1i} + \Delta_{2i}}{\Delta_3} - \Delta_{4i} \right) \left(\frac{\Delta_{1i} + \Delta_{2i}}{\Delta_3} - \Delta_{4i} \right)^T.$$

Remark 2

In the appendix, we argue that one of the conditions necessary to prove the asymptotic normality of our estimator is that $E\{h(X, \beta, 0)\}$ be differentiable in β . For most models of $\mu(T, X, \beta)$, if $\mu(1, x, \beta_0) = \mu(0, x, \beta_0)$, then $h(x, \beta, 0)$ will not be differentiable in β at β_0 . Consequently, if $P\{\mu(1, X, \beta_0) = \mu(0, X, \beta_0)\} > 0$, then $E\{\mu(X, \beta, 0)\}$ would not be differentiable in β at $\beta = \beta_0$ and hence the asymptotic theory would not necessarily hold. For example, if $\text{logit}\{\mu(T, X, \beta)\} = \beta_1 + \beta_2 X + \beta_3 T$, then $\mu(1, X, \beta_0) = \mu(0, X, \beta_0)$ when $\beta_3 = 0$. For this model $\text{logit}\{h(X, \beta, 0)\} = \beta_1 + \beta_2 X + \beta_3 I(\beta_3 < 0)$, which is not differentiable at $\beta_3 = 0$. Therefore, we must be careful to avoid such situations. If the strong null hypothesis were

true; i.e., $\mu(1, x, \beta_0) = \mu(0, x, \beta_0)$ for all values of x , as in the example above at $\beta_3 = 0$, then it doesn't matter which treatment is given to any patient and the issue of finding the optimal treatment regime is no longer of interest. Therefore, to avoid this difficulty, we suggest that one first consider testing the null hypothesis for no treatment difference and only if there is strong evidence of treatment difference should one continue the exercise of deriving the optimal treatment regime and its attributable benefit.

For scenarios where there is a standard treatment and the strong null hypothesis is true, then the conservative strategy would with high probability conclude to give everyone the standard treatment. Consequently, the conservative strategy may not be as susceptible to the large sample difficulties under the null hypothesis. This will be investigated further in the simulation studies.

4.1 Confidence intervals

We consider two different methods for constructing $(1 - \alpha)$ confidence intervals for AB_{opt} . The first involves exponentiating the $(1 - \alpha)^{th} \times 100\%$ confidence interval for $\log\{1 - \widehat{AB}_{opt}(c)\}$ to obtain the following interval

$$1 - \exp \left[\log \left\{ \frac{\sum_{i=1}^n h(X_i, \widehat{\beta}_n, c)}{\sum_{i=1}^n Y_i} \right\} \pm z_{\alpha/2} \sqrt{\widehat{V}} \right],$$

where $z_{\alpha/2}$ denotes the $(1 - \alpha/2)^{th}$ quantile of a standard normal distribution. While the second confidence interval is a by-product of the delta theorem, for which we have

$$\widehat{AB}_{opt}(c) \pm z_{\alpha/2} \left\{ \frac{\sum_{i=1}^n h(X_i, \widehat{\beta}_n, c)}{\sum_{i=1}^n Y_i} \right\} \sqrt{\widehat{V}}.$$

It is not immediately clear if one interval is superior to the other, though it has been suggested that the back-transformed intervals are superior. We will explore this through simulation in the next Section.

5. Simulation Studies

We report results of several simulations, each based on 10000 Monte Carlo data sets. For simplicity we consider data of the form $O_i = (Y_i, T_i, X_i)$ where Y_i is binary disease status, T_i is treatment, and X_i will be one potentially confounding covariate generated from a $N(0, 1)$ distribution. For each given X_i we generated a Bernoulli treatment indicator T_i from the logistic regression model:

$$\text{logit}\{P(T_i=1|X_i)\}=\alpha_0+\alpha_1 X_i, \quad (11)$$

and a Bernoulli disease indicator Y_i from the logistic regression model:

$$\text{logit}\{P(Y_i=1|T_i, X_i)\}=\beta_0+\beta_1 T_i+\beta_2 X_i+\beta_3 T_i X_i. \quad (12)$$

Thus different choices of $\gamma = \{\alpha_0, \alpha_1, \beta_0, \beta_1, \beta_2, \beta_3\}$ will lead to different $\omega = [P(T = 1), P(Y = 1), P(Y^*(0) = 1), P(Y^*(1) = 1), P\{Y^*(g_{opt}) = 1\}]$, therefore different combinations of the model parameters allow us to study the different items of interest. From model (12), one can see that the parameter β_3 , the parameter associated with the interaction, will be essential for finding scenarios where there is no one dominant treatment. For this model, the optimal treatment regime is given by $g_{opt}(x) = I(\beta_1 + \beta_3 x < 0)$ and the estimated optimal treatment regime (without using a conservative strategy) is $\hat{g}_{opt}(x) = I(\hat{\beta}_{1n} + \hat{\beta}_{3n}x < 0)$.

Once the data were generated, estimates of $(\beta_0, \beta_1, \beta_2, \beta_3)$ were found for each Monte Carlo dataset using SAS logistic procedure. The generated data and logit model estimates were then input into SAS IML where \widehat{AB}_{opt} , and confidence intervals were calculated for each Monte Carlo dataset.

Table 1 illustrates seven different examples with different combinations of model parameters. The first six are cases where there is treatment difference and hence the asymptotic theory for estimating attributable benefit should hold. In the last example there is no treatment difference and as indicated in our remark 2 the asymptotic theory may break down. Although we did argue that we may not be interested in estimating attributable benefit if there is no evidence of a significant treatment effect, this scenario was included in order to investigate the extent of difficulty that such a scenario presents.

We depict the true γ and ω values along with estimates of \widehat{AB}_{opt} , Monte Carlo standard error, delta-theorem standard error and confidence intervals for both back-transformed and delta-theorem methods averaged over all 10000 Monte Carlo samples. For comparison we also considered the treatment regime where everyone gets the better of the two treatments; i.e., either treatment regime g_0 or g_1 , whichever had the lower proportion of events. We refer to this as the dominant treatment and the corresponding attributable benefit as AB_{dom} . Sample sizes of $n = 1000$ were used in all examples; however, we found that the coverage probabilities were accurate in smaller sample sizes, except in those cases where the probabilities in ω fell below 10%.

Examining table 1 we see that for the first six scenarios the Monte Carlo average of the estimates for \widehat{AB}_{opt} were indeed very close to the true values as was the Monte Carlo standard error and the average of the delta-theorem standard errors. Both sets of confidence intervals provided very good coverage with probabilities around 0.95. In most cases the back-transformed interval had better coverage than its delta-theorem counterpart.

From these examples we see the importance of AB_{opt} as a measure of a regime's benefit when there are even small interactions between treatment and confounders. But in simulation scenario 7 we do see that when there is no treatment difference \widehat{AB}_{opt} overestimates the true attributable benefit of zero (as would be expected) and underestimates the coverage probability.

We also considered the conservative estimate $\widehat{AB}_{opt}(c)$, with $c = -1.96$; i.e., the strategy where we give the nonstandard treatment only to those individuals where the nonstandard treatment is significantly better at the .025 level of significance. Specifically, we considered the conservative strategy $\hat{g}_{opt}(x, c) = I(\hat{\beta}_{1n} + \hat{\beta}_{3n}x < -1.96 \times se)$, where se is the estimated standard error of $\hat{\beta}_{1n} + \hat{\beta}_{3n}x$ which can be easily derived from the estimated standard errors of the parameters for the logistic regression model (12).

The results are given in table 2. In cases where $AB_{dom} = AB_{opt}$ then both the conservative and optimal strategies are also the same. Although asymptotically $\widehat{AB}(c)$ will converge to AB_{opt} as n goes to infinity, we do see that in cases where there is substantial qualitative interaction (i.e., $AB_{opt} > AB_{dom}$) that the conservative strategy underestimates the attributable benefit for the optimal treatment strategy slightly and consequently underestimates the coverage probability. However, for simulation 7, where the strong null hypothesis holds, we find that the estimated attributable benefit using the conservative strategy showed little bias and better coverage probability as compared to the standard estimator.

In addition to the simulations listed in table 1, additional simulations were performed to look at various aspects of model and regime misspecification. Simulations with multiple covariates (see the Web Appendix B) had similar results to the single covariate case in showing the effectiveness of AB_{opt} as a measure of public health interest. Model overfitting was studied in the multiple covariate case and confidence intervals from overfit models maintained good coverage for the true AB_{opt} , as expected the overfit models did lose efficiency versus the correct model.

While model overfitting did not seem to be a problem, model underfitting could lead to biased estimators. Table 3 demonstrates four scenarios where the model was misspecified by leaving out an important interaction term. That is, the data were generated as specified in table 1, but the model that was fit was $\text{logit}\{P(Y_i = 1|T_i, X_i)\} = \beta_0 + \beta_1 T_i + \beta_2 X_i$. When the incorrect model was fit without the interaction term, then such a model can only choose a single treatment to be optimal. Thus the attributable benefit we are measuring is of the strategy that gives everyone the dominant treatment (AB_{dom}).

Consider first simulation 2, the proportion of deaths from the dominant treatment was only slightly larger than the optimal treatment regime resulting in similar AB 's (i.e., $AB_{dom} = .577$, $AB_{opt} = .588$). This corresponds to the case where there is only slight qualitative interaction. However, in this case, the estimated attributable benefit for that dominant treatment had a mean of .458 rather than the true .577. We see similar bias in simulations 3 and 6, along with very poor coverage probabilities. Simulation 4 however illustrates that there may not always be such a bias. Thus a misspecified model not only cannot find the optimal treatment regime but it also can result in a biased estimator for the benefit of the treatment regime it does choose as optimal. Consequently, important interaction terms whether they be quantitative or qualitative interactions need to be considered in the model. Misspecification simulations were not performed for those simulations where $AB_{dom} = AB_{opt}$ or AB_{opt} is small.

6. Example

Recall our motivating problem from the Duke Databank for Cardiovascular Diseases (DDCD) and that we are focused on the subset of patients with a previous bypass surgery and requiring a later catheterization due to continued symptoms. For technical reasons, the vast majority of patients are limited to some form of medical therapy (MED) or a combination of medical therapy and percutaneous coronary intervention (PCI). The question of clinical interest is whether PCI provides some benefit in lowering 1-year mortality rates above optimal medical therapy for either the overall study population or some subset defined by clinical characteristics.

In studying this question we will explore the effects that covariates such as patient demographics (i.e., age and gender), health status (i.e., body mass index, pulse, patient and family history of coronary artery disease), and cardiovascular measurements (i.e., ejection

fraction, number of diseased vessels) play as potential confounders. To see the benefit of using AB_{opt} we will compare estimators from models with zero, one, and several confounding variables. The goal of which is to explore the impact of an optimal treatment strategy while simultaneously examining the effects of confounding. The benefit of the optimal treatment regime will be estimated by the proportion of one-year deaths that could have been prevented if that optimal regime had been followed. The study population includes 3856 individuals with a follow-up catheterization between the years 1986–2001 and a history of CABG surgery who did not qualify for a second bypass surgery.

In the entire sample there were 267 deaths within 1 year of the follow-up catheterization, 195 first year deaths for those with medical therapy alone (7.83%) and 72 first-year deaths from patients with a combination of medical therapy and angioplasty (5.27%). We will define treatment in terms of a patient obtaining the poorer naive treatment. That is $T = \{0$ if the patient received angioplasty and medical therapy, 1 if the patient received only medical therapy}. In trying to find the optimal treatment regime we used logistic regression models to predict the probability of death at or before 1 year. In this example, we will assume that the optimal treatment regime will come from the “best” model of one year death rates based on treatment and covariates.

For this discussion we will consider many different models, starting with the simplest model: first year deaths modeled as a function of treatment with no confounding variables. Table 4 has several different logistic regression outputs, starting with treatment as the sole predictor in model 1. With the complete data we have an estimate of the probability of first year death $\hat{P}(Y = 1) = 0.0692$. Under this model the treatment regime would be g_0 , because without interactions the model favors a combination of angioplasty and medical therapy.

From the data we calculate an estimate of $\widehat{AB}_{opt} = 0.2393$ with confidence intervals $\{0.0787, 0.3720\}$ (back-transformed method) and $\{0.0936, 0.3851\}$ (delta method).

To begin to see the effect of confounding, model 2 introduces a statistically significant covariate ($p < 0.05$). This model again favors g_0 as the regime of choice and we find $\widehat{AB}_{opt} = 0.2042$ with confidence intervals $\{0.0377, 0.3418\}$ (back-transformed method) and $\{0.0530, 0.3553\}$ (delta method). We see a 3.5% drop in risk reduction with the addition of a single confounder, and note the agreement between the AB_{opt} confidence intervals and logistic model p-value for treatment effect. For model 3, backward selection regression techniques were used to include main effects into the model. Individual tests for interactions between treatment and main effects were performed to help arrive at the “best” model for this data. Ejection fraction values were missing for 561 patients, 420 of which had angioplasty and medical therapy as their treatment. It was determined that a missing ejection fraction value was informative to treatment decision and as such was itself a potential confounding variable. In order to keep all observations in the analysis an indicator variable was defined for presence of ejection fraction value, then an interaction term was used in the model to determine the effect of an ejection fraction measurement. Note that if a patient did not have an ejection fraction measurement then the indicators value is zero. This is an important aspect to the results because this model now has different confounders working in both directions; that is some confounders result in a higher probability of dying within the first year of treatment, while others lower that probability.

Model 3 has an estimate of $\widehat{AB}_{opt} = 0.2157$ with confidence intervals $\{0.0356, 0.3621\}$ (back-transformed method) and $\{0.0536, 0.3777\}$ (delta method). Therefore the “best” model for this data has AB_{opt} estimates very close to the naive model, but there would have been no way to know apriori that the confounders would have worked in opposing directions. Model 3 again favors the optimal treatment regime where every patient is treated with a

combination of angioplasty and medical therapy. To examine the effect of extraneous variables in the model we consider model 4 which has a non-significant interaction term added (interacting treatment with congestive heart failure status). Model 4 had an estimate of $\widehat{AB}_{opt}=0.1950$ with confidence intervals $\{.0066, 0.3476\}$ (back-transformed method) and $\{0.0257, 0.3642\}$ (delta method). So we see that adding an insignificant variable to the model lowered \widehat{AB}_{opt} by roughly 2% while increasing the delta theorem standard error. Therefore, we conclude from this analysis that all patients should be treated with a combination of angioplasty and medical therapy. We have restricted our discussion to calculating estimates of AB_{opt} using the non-conservative estimator \widehat{AB}_{opt} , but since there was no significant treatment by covariate interactions then $AB_{dom} = AB_{opt}$, in which case, a conservative treatment strategy would also lead to the same answer.

7. Discussion

In this paper we have introduced the notion of attributable benefit associated with a treatment strategy. This measure can be of public health interest and represents the fraction of events that could have been prevented by using the optimal treatment regime. While similar in spirit to adjusted attributable risk, the measure is designed to study the positive public health impact of various treatment strategies. This estimator is most useful in cases where there are qualitative interactions between treatment and patient covariates. In studying the optimal strategy two alternate strategies were also discussed; first a strategy involving giving everyone the dominant treatment, the second involving a conservative strategy where one assigns a standard treatment unless a competing treatment is shown to be significantly better in lowering the risk of an adverse outcome.

In this paper we have focused on deterministic treatment regimes g which mapped the values of x to either 0 or 1. We could have also considered random treatment regimes where $g(x) = p$ for $0 \leq p \leq 1$ to denote the strategy where patients are randomized to treatment 1 with probability p and to treatment 0 with probability $(1-p)$ whenever $X = x$. In fact, under the assumptions of this paper, the observed data can be viewed as resulting from a random treatment regime $g_{curr}(x) = P(T = 1|X = x)$. Another interesting strategy proposed to us by the Editor is to use $g_{opt}(x)$ whenever $|d(x, \hat{\beta}_n)/se\{d(x, \hat{\beta}_n)\}| > c$ (i.e., one of the treatments is significantly better than the other) and to use current practice $g_{curr}(x)$ otherwise. Such a strategy is the random treatment regime

$$g^*(x) = A(x, \widehat{\beta}_n, c) + \{1 - B(x, \widehat{\beta}_n, c)\} + C(x, \widehat{\beta}_n, c)P(T=1|X=x),$$

where $A(x, \hat{\beta}_n, c) = I[d(x, \hat{\beta}_n)/se\{d(x, \hat{\beta}_n)\} < -c]$, $B(x, \hat{\beta}_n, c) = I[d(x, \hat{\beta}_n)/se\{d(x, \hat{\beta}_n)\} > c]$, and $C(x, \hat{\beta}_n, c) = 1 - A(x, \hat{\beta}_n, c) - B(x, \hat{\beta}_n, c)$. An estimator for $E\{Y^*(g^*(X))\}$ could be obtained by

$$n^{-1} \sum_{i=1}^n \left[A(X_i, \widehat{\beta}_n, c)\mu(1, X_i, \widehat{\beta}_n) + \{1 - B(X_i, \widehat{\beta}_n, c)\}\mu(0, X_i, \widehat{\beta}_n) + C(X_i, \widehat{\beta}_n, c)Y_i \right]$$

without having to posit a model for $p(T = 1|X)$. We did not formally evaluate this treatment strategy but we expect it will have properties similar to the conservative strategy discussed in the paper.

Two types of confidence intervals were obtained; the back-transformed interval and the delta method interval. Both had good coverage in the simulation studies we displayed, although the back-transformed generally has more accurate coverage probabilities especially in simulation studies we conducted with smaller sample sizes (not shown here).

With regards to the role of quantitative versus qualitative interactions of treatment and patient covariates and the concern of over-interpreting these interactions, we found that leaving out important interactions, whether qualitative or quantitative, can lead to biased estimators for the attributable benefit of the optimal treatment strategy AB_{opt} and that models that were overfit generally led to better estimators for AB_{opt} . Of course, overfitting may also lead to treatment strategies that are more complicated than are necessary. Therefore, we believe a reasonable strategy is to build models using model selection techniques that tend to overfit but then use the conservative estimated treatment strategy where the standard treatment is used except in cases where there is overwhelming statistical evidence to the contrary.

There is also the concern about sample size, especially because the sample size necessary to detect important interactions with sufficient statistical power may be prohibitively large. Even though we do not have much control over the sample size of observational databases that are used for this type of research, we believe estimating the optimal treatment strategy, especially using a conservative strategy, is worthwhile. The utility and the accuracy of such a treatment strategy can then be evaluated by the estimator for the attributable benefit and the corresponding confidence interval.

We believe that estimating the attributable benefit of a treatment regime from observational data can have positive public health impact, but only in those situations where it is appropriate to do so. As we mentioned earlier, the key assumption that allows us to derive estimators for causal parameters is that of no unmeasured confounders. This, unfortunately, is also the most difficult to verify in practice and hence some have argued against using medical databases to make decisions on treatment; see, for example some excellent discussions on this point by Green and Byar (1984) and Yusuf et al. (1986). But we believe that these databases can offer valuable insight if the assumption of no unmeasured confounders was tenable. In order for that this assumption be plausible we must have some confidence that all important information that may affect the treatment decision process be captured in the database. For that reason, it is important during the creation of such a database that discussions with investigators be carried out to identify all the factors that they believe would influence their treatment decisions and all possible effort be made to capture such information. The Duke Databank for Cardiovascular Diseases (DDCD) used in our analysis has been a rich source of data for cardiovascular research in evaluating treatments for many years and has been refined to include what are believed to be the most important factors used in treatment decision making. Even so, there may be utility in conducting sensitivity analyses that look at the effect that violations of the no unmeasured confounders assumption would have on treatment effect. Such methods have been proposed, for example, by Rosenbaum and Rubin (1983) and Brumback *et al.* (2004). A sensitivity analysis in the spirit of Brumback *et al.* (2004) is proposed in the Web Appendix A.

One potential problem with this estimator, and all model based adjusted attributable risk estimators, is in model misspecification. This entire methodology is predicated on having the correct statistical model. In future work we will consider using doubly robust inverse

propensity score estimators for the *AB* that will allow greater flexibility in finding unbiased estimators. For those interested in the current estimator, a SAS macro has been developed to calculate estimates along with back-transform and delta-theorem confidence intervals. Coding and details of the macro can be found in Web Appendix C.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by NIH grants R37 AI031789 and R01 CA051962. The authors are indebted to the editor for the suggestion to consider the conservative estimator for the optimal treatment regime which is now an integral part of this paper.

References

- Al-Khatib SM, Shaw LK, O'Connor C, Kong M, Califf RM. Incidence and Predictors of Sudden Cardiac Death in Patients with Diastolic Heart Failure. *Journal of Cardiovascular Electrophysiology*. 2007; 18:1231–1235. [PubMed: 17883404]
- Brumback BA, Hernan MA, Sebastien JP, Robins JM. Sensitivity Analyses for Unmeasured Confounding Assuming a Marginal Structural Model for Repeated Measures. *Statistics in Medicine*. 2004; 23:749–767. [PubMed: 14981673]
- Benichou J. A review of adjusted estimators of attributable risk. *Statistical Methods in Medical Research*. 2001; 10:195–216. [PubMed: 11446148]
- Gail M, Simon R. Testing for Qualitative Interactions Between Treatment Effects and Patient Subsets. *Biometrics*. 1985; 41:361–372. [PubMed: 4027319]
- Graubard BI, Fears TR. Standard Errors for Attributable Risk for Simple and Complex Sample Designs. *Biometrics*. 2005; 61:847–855. [PubMed: 16135037]
- Green SB, Byar DP. Using observation data from registries to compare treatments: the fallacy of omnimetrics. *Statistics in Medicine*. 1984; 3:361–370. [PubMed: 6396792]
- Jewell, NP. *Statistics for Epidemiology*. London, U.K: Chapman & Hall; 2004.
- Levin ML. The occurrence of lung cancer in man. *Acta Unio Internationalis Contra Cancrum*. 1953; 9:531–541. [PubMed: 13124110]
- Neyman J. Sur les applications de la thar des probabilités aux expériences Agaricales: Essay de principe. English translation of excerpts by Dabrowska, D. and Speed, T. (1990). *Statistical Science*. 1923; 5:465–480.
- Peto, R. Clinical Trials. In: Price, P.; Siko, K., editors. *Treatment of Cancer*. London, U.K.: Chapman & Hall; 1995. p. 1039-1043.
- Rosenbaum PR, Rubin DB. Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society, Ser B*. 1983; 45:212–218.
- Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*. 1974; 66:688–701.
- Rubin DB. Statistics and Causal Inference: Comment: Which If's Have Causal Answers. *Journal of the American Statistical Association*. 1986; 81:961–962.
- Tsiatis, AA. *Semiparametric Theory and Missing Data*. New York: Springer; 2006.
- Yusuf Y, Wittes J, Bailey K, Furberg C. Digitalis—a new controversy regarding an old drug. The pitfalls of inappropriate methods. *Circulation*. 1986; 73:14–18. [PubMed: 3510086]
- van der Vaart, AW. *Asymptotic Statistics*. New York: Cambridge; 1998.

Appendix

In this Section we will demonstrate that $\widehat{AB}_{opt}(c)$ is asymptotically normal and derive an estimator for its asymptotic variance by finding its influence function. Specifically, if we denote the observed data by the i.i.d. random vectors $\mathbf{O}_1, \dots, \mathbf{O}_n$ then an estimator $\hat{\alpha}_n(\mathbf{O}_1, \dots, \mathbf{O}_n)$ of a vector value parameter α is asymptotically linear if

$$n^{1/2}(\hat{\alpha}_n - \alpha_0) = n^{-1/2} \sum_{i=1}^n \psi_{\hat{\alpha}_n}(\mathbf{O}_i) + o_p(1), \quad (\text{A.1})$$

where $\psi_{\hat{\alpha}_n}(\mathbf{O}_i)$ is a mean zero random vector with finite variance matrix and $o_p(1)$ are terms that converge in probability to zero as n goes to infinity. The random vector $\psi_{\hat{\alpha}_n}(\mathbf{O}_i)$ is referred to as the influence function of the estimator $\hat{\alpha}_n$. The representation given in equation (A.1) immediately implies that the estimator is asymptotically normal; that is, $n^{1/2}(\hat{\alpha}_n - \alpha_0) \rightarrow N(0, E(\psi\psi^T))$. Furthermore the asymptotic variance can be estimated consistently using the sandwich estimator

$$\widehat{E}(\psi\psi^T) = n^{-1} \sum_{i=1}^n \psi_{\hat{\alpha}_n}(\mathbf{O}_i) \psi_{\hat{\alpha}_n}(\mathbf{O}_i)^T.$$

Here we will determine the influence function for estimators of several parameters and show how the estimated influence functions will be used to derive the sandwich estimator for the asymptotic variance of our attributable benefit estimator.

In order to find the estimated influence function for $\widehat{AB}_{opt}(c)$, it suffices to find the estimated influence function of $\log[\hat{P}_n\{Y^*(g_{opt}) = 1, c\}] - \log(\bar{Y}_n)$. By a simple application of the delta method, when estimating $\log(\alpha)$, then $\psi_{\log(\hat{\alpha})}(\mathbf{O}_i) = \psi_{\hat{\alpha}}(\mathbf{O}_i)/\alpha$, and when estimating $\alpha_1 - \alpha_2$, $\psi_{\hat{\alpha}_1 - \hat{\alpha}_2}(\mathbf{O}_i) = \psi_{\hat{\alpha}_1}(\mathbf{O}_i) - \psi_{\hat{\alpha}_2}(\mathbf{O}_i)$. For the simple sample average estimator \bar{Y}_n of $E(Y)$, the influence function is $Y_i - E(Y)$, the estimated influence function is $Y_i - \bar{Y}_n$, and the estimated influence function of $\log(\bar{Y}_n)$ is $(Y_i - \bar{Y}_n)/\bar{Y}_n$. Therefore deriving the influence function for $\widehat{AB}_{opt}(c)$ amounts to finding the influence function for $\hat{P}_n\{Y^*(g_{opt}) = 1, c\}$.

If we denote by β_0 the true value of β from equation (7) then we note that

$$\widehat{P}_n\{Y^*(g_{opt})=1, 0\} = n^{-1} \sum_{i=1}^n h(X_i, \beta_0, 0), \quad (\text{A.2})$$

where $h(X_i, \beta, c)$ is defined by (8), is an empirical average of i.i.d random variables which converges to and has mean

$$E\{h(X_i, \beta_0, 0)\} = P\{Y^*(g_{opt})=1\}. \quad (\text{A.3})$$

By definition, $h(X_i, \beta, c)$ involves indicator functions and hence may not be differentiable for all β . Thus, we can not be guaranteed of the usual Taylor series expansion to derive the influence function for $\hat{P}_n\{Y^*(g_{opt}) = 1, c\}$. Instead we note that

$$n^{1/2} \left[\widehat{P}_n\{Y^*(g_{opt})=1, c\} - P\{Y^*(g_{opt})=1\} \right] = n^{-1/2} \sum_{i=1}^n [h(X_i, \beta_0, 0) - E\{h(X_i, \beta_0, 0)\}] \tag{A.4}$$

$$+ n^{-1/2} \sum_{i=1}^n \{h(X_i, \widehat{\beta}_n, c) - h(X_i, \beta_0, 0)\}. \tag{A.5}$$

Note that (A.4) is a sum of mean zero i.i.d. random variables Using empirical process theory as described by van der Vaart (1998), if we can show that the functions $h(x, \beta, 0)$ in a neighborhood of β_0 form a Donsker class, then we can use Lemma 19.24 of van der Vaart to show that (A.5) can be written as

$$\frac{\partial E\{h(X, \beta_0, 0)\}}{\partial \beta^T} n^{1/2} (\widehat{\beta}_n - \beta_0) + o_p(1),$$

in which case, the asymptotic results will go through as long as $E\{h(X, \beta, 0)\}$ is differentiable in β even if $h(X, \beta, 0)$ is not. For the models we are considering the functions will generally satisfy a Lipschitz condition; namely, that $|h(x, \beta_1, 0) - h(x, \beta_2, 0)| \leq m(x) \|\beta_1 - \beta_2\|$ for β_1 and β_2 in a neighborhood of β_0 , where $m(x)$ is a square integrable function, in which case, Example 19.7 of van der Vaart (1998) can be used to show that the functions form a Donsker class. There are some cases however when even $E\{h(X, \beta, 0)\}$ is not differentiable, in which case, the asymptotic results may not hold and we may want to avoid such situations (see remark 2).

Standard results for the maximum likelihood estimator can be used to show that

$$n^{1/2} (\widehat{\beta}_n - \beta_0) = n^{-1/2} \sum_{i=1}^n \psi_{\widehat{\beta}_n}(\mathbf{O}_i, \beta_0) + o_p(1),$$

where

$$\psi_{\widehat{\beta}_n}(\mathbf{O}_i) = \left[E \left\{ \frac{(\partial \mu_i(\beta) / \partial \beta)(\partial \mu_i(\beta) / \partial \beta^T)}{\mu_i(\beta)(1 - \mu_i(\beta))} \right\} \right]^{-1} \left[\frac{(\partial \mu_i(\beta) / \partial \beta)}{\mu_i(\beta)(1 - \mu_i(\beta))} \right] \{Y_i - \mu_i(\beta)\},$$

and $\mu_i(\beta) = \mu(T_i, X_i, \beta)$. For the special case of the logistic regression model (9), the influence function is given by

$$\psi_{\widehat{\beta}_n}(Y_i, T_i, X_i, \beta) = \left(E \left[\frac{f_i f_i^T \exp(\beta^T f_i)}{\{1 + \exp(\beta^T f_i)\}^2} \right] \right)^{-1} \times f_i \left\{ Y_i - \frac{\exp(\beta^T f_i)}{1 + \exp(\beta^T f_i)} \right\},$$

where $f_i = f(X_i, T_i)$ and the estimated influence function is given by

$$\widehat{\psi}_{\widehat{\beta}_n, i} = \left(n^{-1} \sum_{i=1}^n \left[\frac{f_i f_i^T \exp(\widehat{\beta}_n^T f_i)}{\{1 + \exp(\widehat{\beta}_n^T f_i)\}^2} \right] \right)^{-1} \times f_i \left\{ Y_i - \frac{\exp(\widehat{\beta}_n^T f_i)}{1 + \exp(\widehat{\beta}_n^T f_i)} \right\}.$$

Combining these results, we have

$$\begin{aligned} & n^{1/2} \left[\widehat{P}_n\{Y^*(g_{opt})=1, c\} - P\{Y^*(g_{opt})=1\} \right] \\ &= n^{-1/2} \sum_{i=1}^n (h(X_i, \beta_0, 0) - E\{h(X_i, \beta_0, 0)\}) \\ &+ \left[\partial E\{h(X, \beta_0, 0)\} / \partial \beta^T \right] \psi_{\widehat{\beta}_n}(O_i, \beta_0) + o_p(1). \end{aligned}$$

Let us denote the i^{th} term in this summation; i.e., the influence function of $\widehat{P}_n\{Y^*(g_{opt}) = 1, c\}$, as $q_i = q(O_i, \beta_0)$. Therefore, the estimated influence function is given by

$$\widehat{q}_i = \left\{ h(X_i, \widehat{\beta}_n, c) - n^{-1} \sum_{j=1}^n h(X_j, \widehat{\beta}_n, c) \right\} + \left[\widehat{\partial E}\{h(X, \beta_0, 0)\} / \partial \beta^T \right] \widehat{\psi}_{\widehat{\beta}_n, i},$$

where the gradient $\partial E\{h(X, \beta_0, 0)\} / \partial \beta^T$ is estimated using numerical derivatives; namely, where the j^{th} element of the vector $\widehat{\partial E}\{h(X, \beta_0, 0)\} / \partial \beta^T$ is obtained by a numerical derivative,

$$\frac{n^{-1} \sum_{i=1}^n \{h(X_i, \widehat{\beta}_n + \varepsilon_j 1_j, c) - h(X_i, \widehat{\beta}_n - \varepsilon_j 1_j, c)\}}{2\varepsilon_j}, \tag{A.6}$$

$j = 1, \dots, k$, where k is the dimension of the parameter β , and 1_j is a k -dimensional vector where the j^{th} element is 1 and all the other elements are 0. We chose $\varepsilon_j = \widehat{\sigma}(\widehat{\beta}_j) / 100$, where $\widehat{\sigma}(\widehat{\beta}_j)$ is the estimated standard error of the j^{th} element of $\widehat{\beta}$; however, we found that the result was insensitive to the choice of ε_j as long as ε_j was small.

Now that we have derived the influence function for $\widehat{P}_n\{Y^*(g_{opt}) = 1, c\}$, we note that the i^{th} influence function for $\log[\widehat{P}_n\{Y^*(g_{opt}) = 1, c\}] - \log(\bar{Y}_n)$, call it ψ_i , is equal to

$$\psi_i = \frac{h(X_i; \beta_0, 0) - E\{h(X_i; \beta_0, 0)\}}{P\{Y^*(g_{opt})=1\}} + \frac{[\partial E\{h(X; \beta_0, 0)\} / \partial \beta^T] \psi_{\beta_0}(O_i; \beta_0)}{P\{Y^*(g_{opt})=1\}} - \frac{Y_i - E(Y)}{E(Y)},$$

and the estimated influence function $\hat{\psi}_i$ is given by (10).

Table 1

Simulation Results for $\hat{A}B_{opt}$: BT represent back-transformed confidence intervals while DT represents delta-theorem based confidence intervals

Simulation	1	2	3	4	5	6	7
n	1000	1000	1000	1000	1000	1000	1000
Reps	10000	10000	10000	10000	10000	10000	10000
α_0	-1.25	-1.25	-1.25	-1.25	-1	0.25	-1.25
α_1	2	2	2	0	-1	1	2
β_0	-2.5	-2.5	-2.5	-2.5	1	-2	-2.5
β_1	1.25	1.25	1.25	1.25	1	0.75	0
β_2	1.1	1.1	1.1	1.1	1	1.2	1.1
β_3	0	2	-1	-2	-1	1.5	0
$P(T=1)$	0.318	0.318	0.318	0.223	0.303	0.551	0.318
$P(Y=1)$	0.184	0.263	0.122	0.143	0.786	0.304	0.111
$P\{Y^*(0)=1\}$	0.111	0.111	0.111	0.111	0.697	0.164	0.111
$P\{Y^*(1)=1\}$	0.267	0.363	0.223	0.2545	0.881	0.169	0.111
$P\{Y^*(g_{opt})=1\}$	0.111	0.108	0.1	0.068	0.691	0.348	0.111
AB_{dim}	0.396	0.577	0.094	0.223	0.113	0.446	0
AB_{opt}	0.396	0.588	0.184	0.523	0.121	0.461	0
$\widehat{AB}_{opt}(0)$	0.3953	0.5897	0.1982	0.5277	0.1251	0.4627	0.1073
$\widehat{SE}\{AB_{opt}(0)\}mc$	0.0799	0.0584	0.0605	0.0538	0.0168	0.0632	0.0773
$\widehat{SE}\{AB_{opt}(0)\}dt$	0.0786	0.0578	0.0593	0.0536	0.0169	0.0626	0.0944
BT Lower 95%	0.2195	0.4587	0.0720	0.4098	0.0914	0.3245	-0.1024
BT Upper 95%	0.5310	0.6885	0.3055	0.6218	0.1574	0.5723	0.2715
BT Coverage	0.9415	0.9428	0.9477	0.9492	0.9579	0.9455	0.9320
DT Lower 95%	0.2413	0.4764	0.0820	0.4226	0.0920	0.3400	-0.0776
DT Upper 95%	0.5493	0.7030	0.3144	0.6328	0.1581	0.5855	0.2922
DT Coverage	0.9384	0.9352	0.9420	0.9418	0.9568	0.9420	0.8897

Table 2

Simulation Results for $\hat{A}B_{cons}$; BT represent back-transformed confidence intervals while DT represents delta-theorem based confidence intervals

Simulation	2	3	4	5	6	7
AB_{opt}	0.588	0.184	0.523	0.121	0.461	0
$\hat{A}B_{opt}(-1.96)$	0.5826	0.1482	0.5044	0.1192	0.4551	0.0117
$\widehat{SE}\{AB_{opt}(-1.96)\}mc$	0.0581	0.0752	0.0564	0.0159	0.0638	0.0099
$\widehat{SE}\{AB_{opt}(-1.96)\}dt$	0.0570	0.0995	0.0460	0.0146	0.0631	0.1070
BT Lower 95%	0.4541	-0.0725	0.4054	0.0901	0.3159	-0.2225
BT Upper 95%	0.6804	0.3216	0.5866	0.1472	0.4551	0.2006
BT Coverage	0.9352	0.9644	0.8545	0.9275	0.9368	0.9505
DT Lower 95%	0.4709	-0.0469	0.4143	0.0906	0.3314	-0.1981
DT Upper 95%	0.6943	0.3433	0.5944	0.1477	0.5788	0.2216
DT Coverage	0.9356	0.9731	0.8709	0.9287	0.9421	0.9437

Table 3

Misspecification Simulation Results; BT represent back-transformed confidence intervals while DT represents delta-theorem based confidence intervals

Simulation	2	3	4	6
AB_{dom}	0.577	0.094	0.223	0.446
$\hat{A}B_{dom}$	0.458	0.298	0.223	0.339
$SE(AB_{dom})_{mc}$	0.0484	0.0941	0.0495	0.0575
$SE(AB_{dom})_{dt}$	0.0479	0.0933	0.0489	0.0570
BT Lower 95%	0.3557	0.0887	0.1210	0.2171
BT Upper 95%	0.5445	0.4588	0.3130	0.4417
BT Coverage	0.2505	0.5088	0.9474	0.4685
DT Lower 95%	0.3645	0.1151	0.1272	0.2272
DT Upper 95%	0.5522	0.4809	0.3188	0.4507
DT Coverage	0.3130	0.4121	0.9471	0.5258

Table 4
Logistic Regression Models, analysis performed with SAS 9.1 logistic procedure.

Response: Death Within First Year						
Model 1	DF	Estimate	SE	Wald χ^2	P-value	
Intercept	1	-2.8896	0.1211	569.52	<.0001	
Treatment	1	0.4245	0.1422	8.91	0.0028	
Model 2						
	DF	Estimate	SE	Wald χ^2	P-value	
Intercept	1	-5.0064	0.3285	232.20	<.0001	
Treatment	1	0.3565	0.1434	6.18	0.0129	
Pulse	1	0.0286	0.0040	52.22	<.0001	
Model 3						
	DF	Estimate	SE	Wald χ^2	P-value	
Intercept	1	-4.8861	0.4385	124.16	<.0001	
Treatment	1	0.3999	0.1601	6.24	0.0125	
Congestive Heart Failure	1	0.6040	0.1420	18.10	<.0001	
Number of Disease Vessels	1	0.3487	0.1013	11.86	0.0006	
Mitral Insufficiency	1	0.2744	0.0619	19.64	<.0001	
Pulse	1	0.0182	0.0042	18.81	<.0001	
I(EJFX Present)	1	0.2215	0.2744	0.65	0.4194	
Ejection Fraction*(EJFX Present)	1	-0.0231	0.0049	22.08	<.0001	
Model 4						
	DF	Estimate	SE	Wald χ^2	P-value	
Intercept	1	-4.9472	0.4446	123.82	<.0001	
Treatment	1	0.5454	0.2163	6.36	0.0117	
Congestive Heart Failure	1	0.8205	0.2545	10.40	0.0013	
Number of Disease Vessels	1	0.3468	0.1014	11.71	0.0006	
Mitral Insufficiency	1	0.2765	0.0620	19.90	<.0001	
Pulse	1	0.0180	0.0042	18.32	<.0001	
I(EJFX Present)	1	0.2225	0.2738	0.66	0.4165	
Ejection Fraction*(EJFX Present)	1	-0.0235	0.0049	22.68	<.0001	

Response: Death Within First Year					
Model 1	DF	Estimate	SE	Wald χ^2	P-value
Treatment*CHF	1	-0.3048	0.2987	1.04	0.3075