Published in final edited form as:

J Am Stat Assoc. 2007 June ; 102(478): 573–582. doi:10.1198/016214507000000130.

Sensitivity Analyses Comparing Time-to-Event Outcomes Existing Only in a Subset Selected Postrandomization

Bryan E. Shepherd [Assistant Professor],

*Department of Biostatistics, Vanderbilt University, Nashville, TN 37232 (E-mail:
bryan.shepherd@vanderbilt.edu)*

Peter B. Gilbert [Associate Member], and

Fred Hutchinson Cancer Center

Thomas Lumley [Associate Professor]

Department of Biostatistics, University of Washington, Seattle, WA 98195

Abstract

In some randomized studies, researchers are interested in determining the effect of treatment assignment on outcomes that may exist only in a subset chosen after randomization. For example, in preventative human immunodeficiency virus (HIV) vaccine efficacy trials, it is of interest to determine whether randomization to vaccine affects postinfection outcomes that may be right-censored. Such outcomes in these trials include time from infection diagnosis to initiation of antiretroviral therapy and time from infection diagnosis to acquired immune deficiency syndrome. Here we present sensitivity analysis methods for making causal comparisons on these postinfection outcomes. We focus on estimating the survival causal effect, defined as the difference between probabilities of not yet experiencing the event in the vaccine and placebo arms, conditional on being infected regardless of treatment assignment. This group is referred to as the always-infected principal stratum. Our key assumption is *monotonicity*—that subjects randomized to the vaccine arm who become infected would have been infected had they been randomized to placebo. We propose nonparametric, semiparametric, and parametric methods for estimating the survival causal effect. We apply these methods to the first Phase III preventative HIV vaccine trial, VaxGen's trial of AIDSVAX B/B.

Keywords

Acquired immune deficiency syndrome; Causal inference; Kaplan–Meier; Principal stratification

1. INTRODUCTION

In preventative human immunodeficiency virus (HIV) vaccine efficacy trials, it is of interest to assess the effects of vaccines on outcomes that occur only in infected individuals (Nabel 2001; Graham 2002; Gilbert et al. 2005). The current wave of candidate HIV vaccines undergoing efficacy testing have been specifically designed to alter disease progression and postinfection outcomes (www.hvti.org). The effect of the vaccine on time-to-event postinfection outcomes, such as the time from HIV infection diagnosis until acquired immune deficiency syndrome (AIDS), are particularly of interest.

In efficacy trials, such as VaxGen's trial of AIDSVAX B/B, thousands of healthy HIV-uninfected volunteers are randomized to vaccine or placebo. They are followed for a certain amount of time, with the primary endpoint being HIV infection. Those who become infected are then enrolled into a postinfection study and monitored for several more years.

Comparisons of outcomes between those in the vaccine and placebo arms who become infected during the course of the trial can be misleading, because they condition on a postrandomization variable (infection) and thus are susceptible to selection bias (Rosenbaum 1984; Halloran and Struchiner 1995). An intention-to-treat (ITT) analysis could be performed using all randomized individuals, ignoring infection status and defining the outcome of interest as the postinfection outcome (e.g., time from randomization until AIDS). But this is problematic, because most trial participants will not become infected and thus are not followed past the time at which the initial study concludes, τ_0 . Censoring these uninfected individuals at τ_0 induces dependent censoring, and the Kaplan–Meier method provides a biased estimate that underestimates the survival probability for $t > \tau_0$. On the other hand, assuming that the postinfection outcome occurs in none of the individuals not infected at time τ_0 and censoring these individuals at the time when the postinfection study ends, $\tau_0 + \tau$, most likely overestimates the survival probability for $t > \tau_0$, particularly if τ is large. An unbiased ITT analysis would be to censor everyone who does not experience the postinfection event at τ_0 . This is inefficient, because much information is thrown away. From a statistical standpoint, the ideal trial design would not have an initial stopping point, but would follow all individuals for the entire study time, $\tau_0 + \tau$. However, resource constraints make an efficacy trial of this type infeasible.

The method that we propose here is based on principal stratification (Frangakis and Rubin 2002). Our analysis addresses a different question than an ITT analysis: Among those who would have been infected regardless of treatment assignment, does assignment to the vaccine increase/decrease the probability of being AIDS-free t months postinfection? Many have pointed out that comparing an outcome in the subgroup of individuals who would have been infected regardless of treatment assignment is a causal comparison (Kalbfleish and Prentice 1980; Robins 1995; Rubin 2000; Robins and Greenland 2000; Frangakis and Rubin 2002). This subgroup has been referred to as the *always-infected*, or *doomed*, principal stratum (Hudgens, Hoering, and Self 2003; Gilbert, Bosch, and Hudgens 2003; Hudgens and Halloran 2006). An analysis conditioning on membership in the always-infected principal stratum is of particular interest when we want to know whether or not there exists a mechanism through which the vaccine alters a time-to-event outcome in infected individuals.

Because we do not know which participants would have been infected regardless of treatment assignment, assumptions must be made to answer the causal question. Our key assumption is that any participant in the vaccine arm who becomes infected would have been infected if randomized to placebo. Under this assumption, we advocate a sensitivity analysis approach in which the probability of infection if assigned vaccine, given infection in the placebo arm and time from infection diagnosis to some event of interest, is of a known form indexed by an assumed sensitivity parameter.

This work extends the sensitivity analysis methods reported by Hudgens et al. (2003) (HHS), Gilbert et al. (2003) (GBH), and Shepherd, Gilbert, Jemai, and Rotnitzky (2006) (SGJR) to a time-to-event outcome defined postrandomization. In Section 2 we discuss notation, assumptions, and estimands; in Section 3 we discuss estimation; in Section 4 we explore the finite-sample properties of our estimators; and in Section 5 we apply our methods to investigate the effect of vaccination on the time to initiation of antiretroviral therapy in the VaxGen trial of AIDSVAX B/B. In Section 6 we discuss the methods and results, and we provide technical details in the Appendix.

2. NOTATION, CAUSAL ESTIMAND, AND ASSUMPTIONS

Consider a study in which N subjects, independently and randomly selected from a given population of interest, are randomized to either placebo or vaccine. Let $Z_i = 1$ if subject i , $i = 1, \dots, N$, is randomized to vaccine and $Z_i = 0$ if this subject randomized to placebo. Trial

participants are monitored for HIV infection for a predetermined period; the infection status during the study follow-up period for subject i is the indicator S_i , where $S_i = 1$ if infected and $S_i = 0$ if not. Let T_i be the time from infection diagnosis until some event for subject i , and define C_i as the time from infection diagnosis until censoring. If $S_i = 1$, then we observe $Y_i = \min(T_i, C_i)$ and $\Delta_i = I(Y_i = T_i)$. Note that T_i , C_i , Y_i , and Δ_i exist only if $S_i = 1$; otherwise, they are assigned the value $*$.

To define the estimand of interest, we use potential outcomes/counterfactuals (Neyman 1923; Rubin 1978; Robins 1986). Specifically, define $S_i(0)$ to be the infection status indicator if, possibly contrary to fact, subject i is assigned to placebo. Define $S_i(1)$ to be the infection status indicator if subject i is assigned to vaccine. Similarly, define $T_i(0)$ to be the time-to-event outcome if participant i is assigned to placebo and $T_i(1)$ to be the time-to-event outcome if this participant is assigned to vaccine. The potential outcomes $C_i(z)$, $Y_i(z)$, and $\Delta_i(z)$ are defined similarly for $z = 0, 1$. For a subject who does not become infected if assigned treatment z [i.e., $S_i(z) = 0$], we define $T_i(z) = Y_i(z) = C_i(z) = \Delta_i(z) = *$. This notation implicitly assumes that the potential outcomes of each trial participant are not influenced by the treatment assignments of other participants, an assumption known as the stable unit treatment value assumption (SUTVA) (Rubin 1978). SUTVA may be violated in vaccine studies of infectious disease when a local population is treated intensively, producing herd immunity. It is a reasonable assumption when study participants are geographically dispersed, as in our example.

Assuming that the study participants make up a random sample from a large population of interest, the outcomes $W_i = (Z_i, S_i(0), S_i(1), T_i(0), T_i(1), C_i(0), C_i(1))$, $i = 1, \dots, N$, are iid copies of a random vector $W = (Z, S(0), S(1), T(0), T(1), C(0), C(1))$, and similarly the observed data $O_i = (Z_i, S_i, Y_i, \Delta_i)$, $i = 1, \dots, N$, are iid copies of $O = (Z, S, Y, \Delta)$. Randomization ensures that

$$(S(0), S(1), T(0), T(1), C(0), C(1)) \perp\!\!\!\perp Z, \quad (1)$$

because $(S(0), S(1), T(0), T(1), C(0), C(1))$ can be considered an unobserved baseline characteristic of each subject. Here, for random variables A , B , and C , $A \perp\!\!\!\perp B|C$ indicates conditional independence of A and B given C .

The four basic principal strata (Frangakis and Rubin 2002) can be defined in terms of the counterfactual pair $(S(0), S(1))$; the never-infected are those with $S(0) = S(1) = 0$, the harmed are those with $S(0) = 0$ and $S(1) = 1$, the protected are those with $S(0) = 1$ and $S(1) = 0$, and the always-infected (ai) are those with $S(0) = S(1) = 1$.

For a subject i who is in the ai principal stratum, a causal effect on his or her time-to-event outcome is some measure of discrepancy between $T_i(0)$ and $T_i(1)$. For example, the difference $T_i(1) - T_i(0)$ was used by HHS and GBH, with the average causal effect in the ai stratum defined as $ACE = E[T(1) - T(0)|S(0) = S(1) = 1]$. With right-censored outcomes, such an estimand may not be identifiable without modeling the distribution of the time-to-event outcome; alternatively, restricted means could be compared (Chen and Tsiatis 2001). However, it is often of primary interest to compare the probabilities of being event-free at time t postinfection diagnosis, defined in the ai stratum as $1 - P(T_i(z) \leq t|S_i(0) = S_i(1) = 1)$ for $z = 0, 1$, where $t = 0$ corresponds to the time of infection diagnosis. We take this approach; our estimand of interest is the “survival” causal effect in the ai stratum:

$$\begin{aligned} SCE(t) &= P(T_i(0) \leq t|S_i(0) = S_i(1) = 1) - P(T_i(1) \leq t|S_i(0) = S_i(1) = 1) \\ &\equiv F_p^{ai}(t) - F_v^{ai}(t). \end{aligned}$$

To estimate $SCE(t)$, we make the following key assumptions:

A1. Monotonicity. $S_i(1) = S_i(0)$ for all i .

A2. Independent censoring. $C_i(z) \perp\!\!\!\perp T_i(z) | S_i(z) = 1$ for $z = 0, 1$.

Assumption A1 implies that no individual is in the harmed principal stratum. This is a strong but plausible assumption in randomized double-blinded, placebo-controlled vaccine trials.

Assumption A2 is common when analyzing time-to-event data, with the only difference being that this independence is at the counterfactual level and is conditional on infection; otherwise, $C_i(z) = T_i(z) = *$.

From these assumptions, $F_v^{ai}(t) = F_v(t) \equiv P(T_i \leq t | Z_i = 1, S_i = 1)$, the distribution of the time-to-event outcome given vaccine and infection, thus is identified. However, to estimate $SCE(t)$, we must also be able to identify $F_p^{ai}(t)$, which remains unidentifiable. The identifiable distribution, $F_p(t) \equiv P(T_i \leq t | Z_i = 0, S_i = 1)$, is a mixture of the distribution of the time-to-event outcome for placebos in the always-infected stratum, $F_p^{ai}(t)$, and in the protected principal stratum, $F_p^{prot}(t)$. The mixing parameter is one minus the probability of being in the always-infected principal stratum given infection in the placebo arm [i.e., $1 - P(S(1) = 1 | S(0) = 1)$], which, under A1, equals $1 - P(S(1) = 1) / P(S(0) = 1) \equiv VE$, an oft-used measure of vaccine efficacy.

The distribution of $F_p^{ai}(t)$ can be represented as a biased sample from the distribution of the time-to-events for subjects with $S(0) = 1$,

$$F_p^{ai}(t) = (1 - VE)^{-1} \int_0^t w(s) dF_p(s),$$

where $w(t) \equiv P(S(1) = 1 | S(0) = 1, T(0) = t)$. Bounds for $F_p^{ai}(t)$ are obtained by setting $w(t) = I_{\{t \leq q^{1-VE}\}}$ and $w(t) = I_{\{t \geq q^{VE}\}}$, with q^{1-VE} and q^{VE} being the $(1 - VE)$ th and VE th quantiles of $T(0)$. This is equivalent to writing

$$F_p^{ai,U}(t) = \min \left\{ \frac{F_p(t)}{1 - VE}, 1 \right\} \quad (2)$$

and

$$F_p^{ai,L}(t) = \max \left\{ \frac{F_p(t) - VE}{1 - VE}, 0 \right\}, \quad (3)$$

as expressions for the bounds of $F_p^{ai}(t)$, given by HHS.

There might be scientific reasons why these sharp bounds are too conservative. A different sensitivity analysis approach would use subject-matter knowledge to restrict the possible range for $F_p^{ai}(t)$ by choosing $w(t)$, resulting in $F_p^{ai}(t)$ somewhere between the bounds $F_p^{ai,U}(t)$ and $F_p^{ai,L}(t)$. This, the approach of GBH, is specified by making the following additional assumption:

A3. $P(S(1) = 1 | S(0) = 1, T(0)) = w(T(0); \beta)$, where $w(t; \beta) = \Phi \{ \alpha + g(t; \beta) \}$, β is fixed and known, $\Phi(\cdot)$ is a known cumulative distribution function, α is an unknown parameter, and for each β , $g(t; \beta)$ is a known function of t .

If one assumes A3, then $F_p^{ai}(t)$ is identified. Of course, the parameter β is not identified by the observed data. It is considered fixed and known, and then varied over a plausible range of

values as a form of sensitivity analysis (Scharfstein, Rotnitzky, and Robins 1999; GBH). One choice for $w(t; \beta)$ is the expit function, that is, $w(t; \beta) = (1 + \exp(-\alpha - \beta t))^{-1}$, used by GBH. For this choice of $w(\cdot)$, the sensitivity parameter β has a log-odds ratio interpretation. Using constant $w(\cdot)$ (e.g., the expit function with $\beta = 0$) is the same as assuming that $T(0) \perp\!\!\!\perp S(1)|S(0) = 1$ or, equivalently, that the distribution of the time-to-event outcome under placebo is the same in the ai and in the protected principal strata. Choosing $\beta = -\infty$ and $\beta = \infty$ corresponds to the sharp bounds given by (2) and (3).

3. ESTIMATION

We first discuss nonparametric inference for the bounds of $SCE(t)$, then discuss parametric and semiparametric inference for $SCE(t)$ under assumed values of β .

3.1 Nonparametric Estimation: Sharp Bounds

Under SUTVA, (1), A1, and A2, and for $t < \tau$, where τ is the maximum postrandomization follow-up time, the sharp bounds given by (2) and (3) are consistently estimated as

$$\widehat{F}_p^{ai,U}(t) = \min \left\{ \frac{\widehat{F}_p(t)}{1 - \widehat{VE}}, 1 \right\}$$

and

$$\widehat{F}_p^{ai,L}(t) = \max \left\{ \frac{\widehat{F}_p(t) - \widehat{VE}}{1 - \widehat{VE}}, 0 \right\},$$

where $\widehat{VE} = \min\{1 - (n_v N_p)/(n_p N_v), 0\}$ [with $N_z(n_z)$ the number of individuals randomized to (infected in) treatment arm z , where $N_v + N_p = N$] and $\widehat{F}_p(t)$ is the Kaplan–Meier estimator of $F_p(t)$. Because under monotonicity, $F_v^{ai}(t) = F_v(t)$, estimated sharp bounds for $SCE(t)$ are $\widehat{F}_p^{ai,L}(t) - \widehat{F}_v(t)$ and $\widehat{F}_p^{ai,U}(t) - \widehat{F}_v(t)$, where $\widehat{F}_v(t)$ is the Kaplan–Meier estimator of $F_v(t)$.

Under certain conditions, $\widehat{F}_p^{ai,U}(t)$ and $\widehat{F}_p^{ai,L}(t)$ are asymptotically normal. Specifically, $\widehat{VE}$ is asymptotically normal if $0 < VE < 1$ or, equivalently, $0 < P(S(1) = 1) < P(S(0) = 1)$. The Kaplan–Meier estimate, $\widehat{F}_p(t)$, is asymptotically normal under the usual conditions (Kalbfleish and Prentice 1980). Therefore, $\widehat{F}_p^{ai,U}(t)$ and $\widehat{F}_p^{ai,L}(t)$ are asymptotically normal if, in addition to these conditions, $0 < F_p(t) < 1 - VE$ and $VE < F_p(t) < 1$. These latter conditions are a result of (2) and (3). If these conditions are violated then, by definition, $F_p^{ai,U}(t)$ and $F_p^{ai,L}(t)$ are 1 and 0, and hence estimates will not be asymptotically normal. This asymptotic result follows from the asymptotic normality of $(\widehat{F}_p(\cdot), \widehat{VE})$, the Hadamard differentiability of the maps $(\widehat{F}_p(t)/(1 - \widehat{VE}), \widehat{VE})$ and $\widehat{F}_p/(1 - \widehat{VE})/(1 - \widehat{VE})$, and the functional delta method (Andersen, Borgan, Gill, and Keiding 1992). Under these conditions, expressions for the asymptotic variance are

$$\text{var}(\widehat{F}_p^{ai,U}(t)) = \left(\frac{p_0}{p_1}\right)^2 \sigma^2(t) + \left(\frac{F_p(t)}{p_1}\right)^2 \frac{p_0(1 - p_0)}{N_p} + \left(\frac{F_p(t) p_0}{p_1^2}\right)^2 \frac{p_1(1 - p_1)}{N_v}$$

and

$$\text{var}(\widehat{F}_p^{ai,L}(t)) = \left(\frac{p_0}{p_1}\right)^2 \sigma^2(t) + \left(\frac{1 - F_p(t)}{p_1}\right)^2 \frac{p_0(1 - p_0)}{N_p} + \left(\frac{(1 - F_p(t)) p_0}{p_1^2}\right)^2 \frac{p_1(1 - p_1)}{N_v},$$

where $\sigma^2(t)$ is the variance of the Kaplan–Meier estimate, $\hat{F}_p$, $p_0 \equiv P(S = 1|Z = 0)$, and $p_1 \equiv p_0(1 - VE) = P(S = 1|Z = 1)$. From these equations, we can estimate variances in the usual manner by plugging in parameter estimates.

Alternatively, variances may be estimated using a standard bootstrap procedure. Specifically, from $(O_1, \dots, O_N)$, sample with replacement N vectors O_i , creating $(O_1^*, \dots, O_N^*)$. Compute $\widehat{SCE}^*(t)$ based on the bootstrap sample $(O_1^*, \dots, O_N^*)$. Repeat this process K times and estimate the variance of $\widehat{SCE}(t)$ by the sample variance of the $\widehat{SCE}^*(t)$'s. Then $100(1 - \alpha)\%$ -level confidence intervals can be constructed using the resulting variance estimate (Wald intervals), using the $\alpha/2$ and $(1 - \alpha/2)$ quantiles of $\widehat{SCE}^*(t)$ (percentile intervals), or by studentizing with the asymptotic variance estimate.

3.2 Parametric Estimation

In addition to SUTVA, (1), A1, and A2, assume A3 and make the following modeling assumption:

M1. The distribution of $T(1)$ given $S(1) = 1$ is known up to a finite-dimensional parameter η_1 , that is, $f_{T(1)|S(1)=1}(t|S(1)=1) = f_v(t; \eta_1)$, where η_1 is unknown and for each η_1 , $f_v(\cdot; \eta_1)$ is a known density.

Also make one of the following two assumptions on the distribution of $T(0)$:

M2a. The distribution of $T(0)$ given $S(0) = 1$ is known up to a finite-dimensional parameter η_0^a , that is, $f_{T(0)|S(0)=1}(t|S(0)=1) = f_p(t; \eta_0^a)$, where η_0^a is unknown and for each η_0^a , $f_p(\cdot; \eta_0^a)$ is a known density.

M2b. The distribution of $T(0)$ given $S(0) = S(1) = 1$ is known up to a finite-dimensional parameter η_0^b , that is, $f_{T(0)|S(0)=S(1)=1}(t|S(0)=S(1)=1) = f_p^{ai}(t; \eta_0^b)$, where η_0^b is unknown and for each η_0^b , $f_p^{ai}(\cdot; \eta_0^b)$ is a known density.

For ease of reference, we call the model defined by SUTVA, (1), A1, A2, A3, M1, and M2a model M_a . We call the model defined like M_a except replacing M2a with M2b and demanding that $w(t; \beta) > 0$ for all t M_b .

Under M_a , $SCE(t)$ is a function of unknown parameter $(\alpha, \eta_0^a, \eta_1)$; specifically, $SCE(t) = SCE_a(t; \alpha, \eta_0^a, \eta_1)$, where $SCE_a(t; \alpha, \eta_0^a, \eta_1)$

$$\equiv \frac{\int_0^t w(s; \alpha, \beta) f_p(s; \eta_0^a) ds}{\int_0^\infty w(s; \alpha, \beta) f_p(s; \eta_0^a) ds} - \int_0^t f_v(s; \eta_1) ds.$$

Similarly, under model M_b , $SEC(t) = SCE_b(t; \eta_0^b, \eta_1) \equiv \int_0^t f_p^{ai}(s; \eta_0^b) ds - \int_0^t f_v(s; \eta_1) ds$. Thus the maximum likelihood estimators (MLEs) of $SCE(t)$ under M_a and M_b are equal to the functions $SCE_a(t; \cdot, \cdot, \cdot)$ and $SCE_b(t; \cdot, \cdot)$ evaluated at the MLEs of $(\alpha, \eta_0^a, \eta_1)$ and (η_0^b, η_1) under models M_a and M_b .

In the absence of censoring, the likelihood induced by these assumptions (minus A2) was given by SGJR. This likelihood can be easily modified to account for independent postinfection censoring,

$$\begin{aligned}
L_a(\alpha, \eta_0^a, \eta_1) &\propto \prod_{i=1}^N \left\{ \left[f_v(y_i; \eta_1)^{\delta_i} (1 - F_v(y_i; \eta_1))^{1-\delta_i} p_0 \right. \right. \\
&\quad \times \int w(y; \alpha, \beta) f_p(y; \eta_0^a) dy^{s_i} \\
&\quad \times \left[1 - p_0 \int w(y; \alpha, \beta) f_p(y; \eta_0^a) dy \right]^{1-s_i z_i} \\
&\quad \times \left. \left[f_p(y_i; \eta_0^a)^{\delta_i} (1 - F_p(y_i; \eta_0^a))^{1-\delta_i} p_0 \right]^{s_i} \right. \\
&\quad \times \left. (1 - p_0)^{1-s_i 1-z_i} \right\},
\end{aligned} \tag{4}$$

under model M_a . Under model M_b , the likelihood $L_b(\cdot)$ is defined as $L_a(\cdot)$ but with $f_p^*(y; \alpha, \eta_0^b) \equiv w^{-1}(y; \alpha, \beta) f_p^{ai}(y; \eta_0^b) / (\int w^{-1}(y; \alpha, \beta) f_p^{ai}(y; \eta_0^b) dy)$ replacing $f_p(y; \eta_0^a)$.

Provided that the protected principal stratum $\{S_i(1) = 0, S_i(0) = 1\}$ is nonempty, under sufficiently smooth parameterizations, the MLEs of the model parameters are asymptotically normally distributed. The variance of the normal limiting distribution can be consistently estimated with either the observed or the (estimated) expected information. These in turn can be used in conjunction with the delta method to obtain consistent variance estimators of $SCE(t)$ for each fixed t . Sensitivity analyses are performed by varying β .

It is noteworthy that this likelihood can be easily extended to condition on baseline covariates. Likelihood-based methods when the outcome of interest is not right-censored, with or without baseline covariates, were extensively studied by SGJR; general principles that they stated apply here as well.

3.3 Semiparametric Estimation

Consider estimation under SUTVA, (1), A1, A2, and A3, that is, performing sensitivity analyses by modeling $w(\cdot)$ but leaving the distributions of $T(0)$ and $T(1)$ unspecified. This can be thought of as extending GBH to time-to-event outcomes.

The estimating equations GBH used to estimate $F_p^{ai}(t)$ when $(n_v/N_v < n_p/N_p)$ can be written as

$$\begin{aligned}
0 &= \Psi(p_0, \alpha) \\
&= \begin{cases} \sum_{i=1}^N (1 - Z_i)(S_i - p_0) \\ \sum_{i=1}^N Z_i \left(S_i - p_0 \int_0^\infty w(t; \alpha, \beta) d\hat{F}_p(t) \right). \end{cases}
\end{aligned} \tag{5}$$

When $n_v/N_v \geq n_p/N_p$, $F_p^{ai}(t)$ is estimated with $\hat{F}_p(t)$, the nonparametric estimate of $F_p(t)$. Similar to Section 3.1, here a natural approach would use these same equations, only now estimating $F_p(t)$ with the Kaplan–Meier estimate. In practice, however, this approach may not be feasible, because $\hat{F}_p(t)$ is not well defined for $t > \tau$. The integral in the second estimating equation of (5) cannot be computed for $t > \tau$.

One way to fix this problem would be to assume some distributional form for the tail of $F_p(\cdot)$. This approach would be similar to the parametric methods of Section 3.2, however, and in this section we want to leave $F_p(\cdot)$ unspecified. Another approach would change the form of $w(\cdot)$, making it constant after time τ . For example, consider $w(\cdot)$, defined as

$$w(t; \alpha, \beta) = \begin{cases} (1 + \exp(-\alpha - \beta t))^{-1} & \text{for } t \leq \tau \\ (1 + \exp(-\alpha - \beta \tau))^{-1} & \text{for } t > \tau. \end{cases} \tag{6}$$

Another choice for $w(\cdot)$ could be

$$w(t; \alpha, \beta) = (1 + \exp(-\alpha - \beta I_{t > t_0}))^{-1} \quad (7)$$

for some $\tau_0 \leq \tau$. Both (6) and (7) define $w(\cdot)$ with the expit function, but do so in a manner such that $w(\cdot)$ is constant for $t > \tau$, allowing us to write

$$\int_0^\infty w(s; \alpha, \beta) d\widehat{F}_p(s) = \int_0^\tau w(s; \alpha, \beta) d\widehat{F}_p(s) + w(\tau; \alpha, \beta)(1 - \widehat{F}_p(\tau)).$$

Of course, these choices of $w(\cdot)$ have implications regarding interpretation. Under (6), the interpretation of β is technically as follows. Given infection in the placebo arm, the odds of infection if randomized to the vaccine arm for $T = t_1$ versus $T = t_2$ are $\exp\{\beta[\min(t_1, \tau) - \min(t_2, \tau)]\}$. This more complex interpretation might appear troublesome; however, over the range of t where there are data, β has a standard log-odds ratio interpretation. Under (7), β also has a standard odds ratio interpretation, except now we have dichotomized $w(\cdot)$, assigning a particular probability of being in the ai stratum for $t > t_0$ and another for $t \leq t_0$.

With $w(\cdot)$ modeled by either (6) or (7), an extension of GBH using the Kaplan–Meier estimates for $F_p(t)$ is the semiparametric MLE. In the Appendix we show that under these same assumptions and $0 < VE < 1$, $p_0 > 0$, and for a properly specified well-behaved $w(\cdot)$ (i.e., constant for $t > \tau$, twice differentiable, and bounded), $\widehat{F}_p^{ai}(t)$ is consistent and asymptotically normal for $t \in (0, \tau]$. Therefore, $\widehat{SCE}(t)$ is also consistent and asymptotically normal.

Because $\widehat{SCE}(t)$ is asymptotically normal, pointwise Wald-based confidence intervals based on the bootstrap will be valid for large sample sizes. It is also possible to obtain an analytic form for the asymptotic variance of $\widehat{SCE}(t)$. This variance estimate relies on the ability to approximate $\widehat{F}_p(t)$ with a sum of iid random variables. (Such an approximation can be obtained from Stute 1995.) Using this result, we can augment (5) by including additional estimating equations,

$$\{0, \dots, 0\}^T = \left\{ \sum_{i=1}^N (1 - Z_i) S_i(V_{1i} - F_p(t_1)), \dots, \sum_{i=1}^N (1 - Z_i) S_i(V_{ki} - F_p(t_k)) \right\}^T,$$

where k is the number of distinct failure times in the placebo arm, t_j is the j th ordered failure time, and for a specific j , V_{ji} are iid random variables for $i = 1, \dots, N$. We can then estimate the variance of parameter estimates using a sandwich estimator–type approach, and then estimate the variance of $\widehat{SCE}(t)$ using the delta method. Details are given in the Appendix.

4. SIMULATIONS

To evaluate the small-sample performance of our estimators of $SCE(t)$, we conducted a 2×4 factorial simulation experiment, corresponding to generating data under $VE \equiv P(S(1) = 0 | S(0) = 1) \approx .3$ or $.6$ and $\beta = .1, .2, 1$, or ∞ . Each simulation generated 1,000 vectors W according to the following steps:

Step 1. The first 500 vectors were set at $Z = 0$; the second 500, at $Z = 1$.

Step 2. $S(0)$ was drawn from a Bernoulli(p_0) distribution with $p_0 = .25$. This choice yields an expected number of infections in the placebo arm of 125, which is typical for a Phase III vaccine trial.

Step 3. $T(0)$ was generated for all realizations with $S(0) = 1$ according to the distribution $F_p(t; \eta)$ with $F_p(\cdot)$ a Weibull distribution and $\eta = (\text{shape} = .5, \text{scale} = 25)$. This distribution was chosen to reflect the distribution of the time from infection diagnosis to initiation of

antiretroviral therapy (ART) in the VaxGen trial, in which approximately 50% of infected participants started ART by 24 months postinfection diagnosis.

Step 4. Given $T(0)$, for each realization with $Z = 1$ and $S(0) = 1$, $S(1)$ was drawn from a Bernoulli($w(T(0); \alpha, \beta)$) distribution. For $\beta = .1, .2$, and 1 , $w(t; \alpha, \beta)$ was defined as in (6) with $\tau = 24$ months. To ensure that $VE \approx .3$, α was set at $-.2, -.9$, or -3.6 when β was set at $.1, .2$, or 1 ; and to ensure that $VE \approx .6$, α was set at $-1.8, -3.4$, or -20 when $\beta = .1, .2$, or 1 . For $\beta = \infty$, $w(t) = I_{\{t \geq q^{VE}\}}$, as discussed in Section 2, where $q^3 = 3.18$ and $q^6 = 21.0$.

Step 5. For realizations with $Z = 1$ and $S(1) = 1$, $T(1)$ was set equal to $T(0)$.

Step 6. For the realizations with $S = 1$, C_1 was generated from a Weibull distribution with shape and scale parameters 3 and 35. Then C was set as $\min(\tau, C_1)$.

Step 7. For all realizations with $S = 1$, Y was chosen as $\min(C, T)$ and $\delta = I_{\{Y=T\}}$.

These steps result in simulating data under SUTVA, (1), A1, A2, A3, and model (6), with $SCE(t) = 0$ for all t .

For each simulated dataset, we computed $\widehat{F}_p^{ai}(t)$ and $\widehat{SCE}(t)$ for $t = 24$ months, assuming the true model for $w(\cdot)$ and the true value for β . We constructed Wald-based 95% confidence intervals by estimating the standard error of estimates using both the bootstrap (with 200 bootstrap replications) and asymptotic variance estimates.

Table 1 reports the performance of $\widehat{SCE}(t)$ based on 1,000 simulation iterations. Because $\widehat{SCE}(t)$ is the difference between $\widehat{F}_p^{ai}(t)$ and $\widehat{F}_v(t)$, examining the performance of the estimates of $\widehat{F}_p^{ai}(t)$ is also useful. Table 2 does this using the same simulations and analyses reported in Table 1. In addition to presenting the coverage of the untransformed 95% Wald-confidence intervals for $\widehat{F}_p^{ai}(t)$, Table 2 presents confidence intervals by transforming the symmetric confidence limits for $\log[-\log\{F_p^{ai}(t)\}]$.

In most cases, bias is minimal and coverage is good using either the bootstrap or the asymptotic variance estimate. The only exceptions are when $VE \approx .6$ and β is large. The poor coverage probabilities here are due to a boundary issue. First, consider the simulations with $\beta = \infty$. As discussed in Section 3.1, as $N \rightarrow \infty$, under the usual assumptions and if $VE < F_p(t)$, then $\widehat{F}_p^{ai,L}(t)$ [which is equivalent to $\widehat{F}_p^{ai}(t)$ with $\beta = \infty$] will be asymptotically normal, and Wald-based confidence intervals will cover at their nominal level. However, in these simulations, $F_p(t) = .625$ for $t = 24$ months, which is very close to $VE = .600$. Therefore, due to stochastic variation and our relatively small sample size, $\widehat{VE}$ is often greater than $\widehat{F}_p(t)$, resulting in $\widehat{F}_p^{ai}(t) = 0$ in nearly half of the simulations. Consequently, the distribution of $\widehat{F}_p^{ai}(t)$ is far from normal; thus these confidence intervals that assume normality have poor coverage. (Note that in this particular setting, Wald-based confidence intervals extended outside the $[0, 1]$ range.) Interestingly, the 2.5- and 97.5-bootstrapped percentile confidence intervals for $\widehat{F}_p^{ai}(t)$ and $\widehat{SCE}(t)$ have coverage probabilities of .938 and .964, for the simulations with $VE \approx .6, \beta = \infty$.

The same logic explains why coverage and bias were also poor for the simulations with $VE \approx .6, \beta = 1$. Under these settings, with time measured in months, a value of $\beta = 1$ is quite large; for example, the odds of being infected in the vaccine arm given infection in the placebo arm from $t = 12$ and $t = 24$ (a difference of 1 year) increase multiplicatively, $\exp(12) \approx 163,000$! Analyses based on the assumption that $\beta = 1$ are very similar to analyses assuming that $\beta = \infty$. Again, the distribution of $\widehat{F}_p^{ai}(t)$ is far from normal. Figure 1 shows a histogram of $\widehat{F}_p^{ai}(t=24)$ for the simulations with $VE \approx .6, \beta = 1$, along which a similar histogram with $VE \approx .3, \beta = 1$

(where the method worked well) for purposes of comparison. Figure 1 also shows the true values of $F_p(t)$ and $F_p^{ai}(t)$ under these simulation settings. Note how close $F_p^{ai}(t)$ is to 0 at $t = 24$ months.

5. EXAMPLE

Here we illustrate our methods using data from the VaxGen vaccine trial. This randomized, double-blind, placebo-controlled Phase III trial of AIDSVAX B/B conducted between 1998 and 2003 recruited 5,403 HIV-negative, at-risk individuals from 61 sites spanning large cities of North America and the Netherlands. The ratio of vaccine to placebo assignment was 2:1. Overall, the vaccine was not found to protect against HIV infection ($\widehat{VE} = .048$), although interaction tests suggested that the vaccine might partially prevent infection for nonwhites. Among nonwhites, estimated vaccine efficacy was .469. Detailed study results have been given by Flynn et al. (2005). Here we compare the time from infection diagnosis to the initiation of antiretroviral therapy (ART) between the vaccine and placebo arms among participants (overall and within the non-white subgroup) who would have become infected regardless of randomization assignment. Specifically, we perform sensitivity analyses to test the hypothesis $H_0 : SCE(t) = 0$ for $t = 2$ years. A vaccine effect to delay ART is beneficial to an individual because it delays exposure to drug toxicities, drug resistance, and the depletion of future therapy options.

A total of 368 subjects were infected during the trial; of these, 347 enrolled into the postinfection study phase (225 in the vaccine arm). There was presumably little interaction between trial participants, so SUTVA seemed reasonable. As discussed by SGJR, assumption A1 was also thought to be plausible because (a) this was a double-blinded placebo-controlled trial, so behavior should not be different relative to assignment to the other treatment, and (b) the vaccine was designed in such a way that it could not become the virus, causing infection. Monotonicity is not consistently testable, but no evidence from the trial or from pretrial experiments suggested that it was violated. In addition, the censoring mechanism did not appear to be informative based on similar dropout rates for participants with different levels of behavioral risk; thus there is no evidence that assumption A2 was violated.

Figure 2 illustrates the results of analyses looking at the time from infection diagnosis to the initiation of ART. Figure 2(a) shows the Kaplan–Meier estimates for both the vaccine and placebo arms for the probability of not yet starting ART. The plot also includes the estimates of the upper and lower bounds of $F_p^{ai}(t)$, described in Section 3.1. Figure 2(b) shows a semiparametric sensitivity analyses reported looking at $SCE(t)$ for $t = 2$ years. The plot contains both the estimate for $SCE(t)$ and 95% Wald confidence intervals (constructed using the asymptotic variance approximation and the bootstrap, with 500 bootstrap replications). In these analyses (and all other sensitivity analyses reported in this section), we modeled the probability of infection in the vaccine arm given infection in the placebo arm, $w(t; \alpha, \beta)$, with (6). Before performing these analyses, we elicited a plausible range for the sensitivity parameter β from a subject matter expert, Dr. Marc Gurwith of VaxGen. His “best guess” for a range for $\exp(\beta)$ was .70–1.1, corresponding to β of $-.36$ –.1. (Note that β is defined here in terms of years.) Figure 2(b) shows the estimates of $SCE(t)$ over a much larger range, for β from -3 to 3 . The open circles (and + and $\times$) in the plots represent the sharp bounds of $SCE(t)$ (and 95% Wald and percentile confidence intervals for these bounds based on 500 bootstrap replications), corresponding to an analysis with $\beta = \pm\infty$. (Confidence intervals based on the analytic variance were similar.) Regardless of the range, Figure 2 shows little evidence that the vaccine has a causal effect on initiation of ART.

It may be of more interest to look at the effect of vaccination on the time-to- event outcomes for the nonwhite subgroup, for which the vaccine appeared to partially protect against HIV

infection. These analyses are shown in Figures 2(c) and 2(d). Because the estimate of VE is much larger in the nonwhite cohort, the bounds for $F_p^{ai}(t)$ are farther apart, as shown in Figure 2(c). This is also reflected in Figure 2(d), with the estimates for $SCE(t)$ covering a wider range. Of course, the smaller sample size ($N = 914$ nonwhites, of whom 59 became infected) also inflates the length of the confidence intervals. It should be noted that 95% Wald-based confidence intervals of $SCE(t)$ at $\beta = \infty$ [where $\widehat{F}_p^{ai,L}(t) = 1$] using the analytic variance expression were much wider than the bootstrap confidence intervals.

Note that for these analyses in the nonwhite cohort, if $\beta > 0$, then $H_0 : SCE(t) = 0$ for $t = 2$ years is rejected at the .05 level. This means that given infection in the placebo arm, if the odds of infection if randomized to vaccine are greater for someone who has a longer time to the initiation of ART, then there is evidence that the vaccine is causing nonwhite participants to have a higher probability of starting ART by 2 years postinfection diagnosis. This would imply that among nonwhites, the vaccine has a detrimental effect, causing more rapid postinfection progression. Interestingly, this value of β is just inside Dr. Gurwith's plausible range for nonwhites ($-.92-.18$); therefore, both the null and the alternative are favored in this range. However, it is worth noting that nowhere in this range is the point estimate positive; thus the analysis provides no support for the effectiveness of the vaccine in nonwhites.

6. DISCUSSION

As candidate vaccines continue to be developed and enter clinical trials, there is particular interest in investigating the effect of vaccination on postinfection outcomes. In this article we proposed sensitivity analysis methods for evaluating the causal effect of vaccination on outcomes defined as the time from infection diagnosis to some postinfection event. We developed nonparametric, parametric, and semiparametric estimation methods and applied them to investigate the causal effect of VaxGen's AIDSVAX B/B vaccine on the time from infection diagnosis to initiation of ART.

From our VaxGen analysis in the nonwhite cohort, we can draw different conclusions over the chosen range. The fact that our analyses produce more than one answer may not be attractive. But we believe that it is an honest way to present the data and allows scientists to look at results and draw their own conclusions. Based on our experience with Dr. Gurwith, eliciting a range for sensitivity parameters is feasible (even though, admittedly, different subject matter experts may choose different ranges for β). Nevertheless, from our communication with other HIV vaccine experts, those with worse postinfection outcomes in the placebo arm have generally been considered those most likely to become infected if randomized to vaccine (in our example, corresponding to $\beta \leq 0$), because they are thought to either have a weaker immune system or to have been exposed to a more virulent virus (Shepherd, Gilbert, and Mehrotra, 2007).

The approach discussed here ignores the time of infection diagnosis and any variation in the amount of time that noninfected participants were followed. An alternative approach would be to use principal strata of time until infection diagnosis, which could be a better approach if there were reasons to believe that the causal effect of the vaccine on the postinfection outcome could differ based on the timing of infection. This could happen, for example, if those infected early in the trial were different immunologically than those infected later. Suppose that U is the discrete failure time from randomization to infection diagnosis, taking values $\{u_1, \dots, u_K, u_{K+}\}$, where $K+$ indicates not being diagnosed with infection by time u_K . Principal strata could be defined as the sets $PS_{jk} = \{U(0) = u_j, U(1) = u_k\}$, for $j, k = 1, \dots, K, K+$; one set of causal estimands of potential interest would be $SCE(t|PS_{jj}) \equiv P(T_i(0) \leq t|PS_{jj}) - P(T_i(1) \leq t|PS_{jj})$, $j = 1, \dots, K$. [Note that in the absence of preinfection censoring, the principal stratum focused on in this article, $\{S(0) = S(1) = 1\}$, is the union of the time-dependent principal strata,

$U_{j,k=1,\dots,K} PS_{jk}$.] Due to the sparseness of data within some principal strata, we might need to group times. For example, if there were $K = 6$ discrete failure times, then we could focus on the principal strata defined by $\{U(0) \leq u_2, U(1) \leq u_2\}$, $\{u_2 < U(0) = u_4, u_2 < U(1) \leq u_4\}$, and $\{u_4 < U(0) \leq u_6, u_4 < U(1) \leq u_6\}$. Alternatively, we could focus on the principal strata defined by $PS(u) \equiv \{U(0) \leq u, U(1) \leq u\}$. For fixed u , this latter principal stratum is equivalent to the always-infected principal stratum that would arise by redefining $S = 1$ as the indicator of infection by time u postrandomization. Estimation of $SCE(t|PS(u))$ when there is censoring before time u warrants further study.

It should be noted that the asymptotic normality of our estimators relies on the assumption that $VE > 0$. For VE near the boundary $VE = 0$, Wald-based confidence intervals may have poor coverage (see Jemai and Rotnitzky 2005 for further discussion). In the VaxGen trial, using the entire cohort the estimate for VE was .048. We performed an additional simulation investigating the performance of Wald-based 95% confidence intervals of $SCE(t = 24 \text{ months})$ under the same settings as described in Section 5, but generating data under the assumption that $VE = .05$ and $\beta = .1$. Coverage using the bootstrap and analytic variance estimates was good (.941 and .938). Of course, we do not know the true VE , in the VaxGen trial so it may be misleading to read too much into these simulations. This is not an issue for the analysis of the nonwhite cohort, where $\widehat{VE} = .469$; however, the asymptotic normality of $\widehat{F}_p^{ai,L}(t)$ for $t = 2$ years could be questioned in the nonwhite cohort, because $\widehat{VE} > \widehat{F}_p(t)$.

Although these methods were designed specifically for HIV vaccine trials, they also may apply in other situations in which causal comparisons conditioning on postrandomization variables are of interest.

APPENDIX: TECHNICAL DETAILS

A. 1 Asymptotic Normality Semiparametric Estimator

Suppose that data are collected over a finite interval $[0, \tau]$ with τ fixed as $N \rightarrow \infty$. Assume SUTVA, (1), A1, A2, A3, $0 < VE < 1$, $p_0 > 0$, and let $t \in [0, \tau]$. Furthermore, assume that $w(t; \alpha, \beta) \in (0, 1)$ for all $t > 0$, α , and β ; is constant for $t > \tau$; and is twice continuously differentiable with respect to α , with a bounded second derivative. Then $\widehat{F}_p^{ai}(t)$ is consistent and asymptotically normal.

Proof

Let $(\hat{p}_0, \hat{\alpha})$ be the solutions to the estimating equations given by (5). From the first equation, $\hat{p}_0 = n_p/N_p$, and the second equation can be written as

$$U_N(\alpha) \equiv \int_0^\infty w(t; \alpha, \beta) d\widehat{F}_p(t) - (1 - \widehat{VE}) = 0.$$

Using the empirical distribution of O_i , the process $\widehat{F}_p(\cdot)$ in $D[0, \tau]$ and $\widehat{VE}$ are jointly asymptotically normal, where $D[0, \tau]$ is the space of cadlag functions (i.e., right-continuous, with limits from the left) on $[0, \tau]$. The map $(\widehat{F}_p(\cdot), \widehat{VE}) \mapsto U_N(\alpha)$ is Hadamard differentiable (as follows from problem 7 of van der Vaart 1998, chap. 20); therefore, by the functional delta method, $U_N(\alpha)$ is asymptotically normal. From Taylor's expansion, write

$$\sqrt{N}(\hat{\alpha} - \alpha) = \frac{-\sqrt{N}U_N(\alpha)}{U'_N(\alpha) + 1/2U''_N(\alpha^*)(\hat{\alpha} - \alpha)^2},$$

where α^* is some value between α and $\hat{\alpha}$, and $U_N''(\alpha^*) = \int_0^\infty w''(t; \alpha^*, \beta) d\hat{F}_p(t)$. Note that $U_N''(\alpha^*)$ is bounded because $|w''(t; \alpha, \beta)|$ is bounded. Therefore, because $U_N'(\alpha) \rightarrow^P U'(\alpha) \equiv \int_0^\infty w'(t; \alpha, \beta) dF_p(t)$ by the continuous mapping theorem, $\hat{\alpha} \rightarrow^P \alpha$ by lemma 5.10 of van der Vaart (1998), $U_N''(\alpha^*)$ is bounded, $U_N(\alpha)$ is asymptotically normal, and $\hat{\alpha}$ is asymptotically normal. Finally, consider the map

$$(\hat{\alpha}, \hat{F}_p(\cdot)) \mapsto \frac{\int_0^t w(s; \hat{\alpha}, \beta) d\hat{F}_p(s)}{\int_0^\infty w(t; \hat{\alpha}, \beta) d\hat{F}_p(t)} \equiv \hat{F}_p^{ai}(t).$$

That the process $\hat{F}_p^{ai}(t)$ in $D[0, \tau]$ is asymptotically normal follows from Hadamard differentiability of this map, the chain rule, and the functional delta method.

A.2 Asymptotic Variance of Semiparametric Estimator

Following Stute (1995), $\int \phi(t) d\hat{F}_p(t)$ can be written as a sum of iid terms plus a remainder term, R_{n_p} , where $|R_{n_p}| = o(n_p^{-1/2})$ and $\phi(t)$ is a well-behaved function of t . Define k as the number of distinct failure times and let $t_1, \dots, t_k$ represent the distinct ordered failure times. We can consider that there are $k+2$ parameters to estimate, $(p_0, \alpha, F_p(t_1), F_p(t_2), \dots, F_p(t_k)) \equiv \theta$, adding k estimating equations to (5):

$$\Psi_i(\theta) = \begin{cases} (1 - Z_i)(S_i - p_0) \\ Z_i(S_i - p_0 \int_0^\infty w(s; \alpha, \beta) dF_p(s)) \\ (1 - Z_i)S_i(V_{1i} - F_p(t_1)) \\ \vdots \\ (1 - Z_i)S_i(V_{ki} - F_p(t_k)), \end{cases}$$

with V_{ji} for $j = 1, \dots, k$ and $i = 1, \dots, N$, defined as

$$V_{ji} = \phi_j(Y_i) \gamma_0(Y_i) \delta_i + \gamma_{j1}(Y_i)(1 - \delta_i) - \gamma_{j2}(Y_i),$$

where

$$\begin{aligned} \phi_j(Y_i) &= 1_{\{Y_i \leq t_j\}}, \\ \gamma_0(Y_i) &= \exp\left(\int_{-\infty}^{Y_i} \frac{H^0(dz)}{1 - H(z)}\right), \\ \gamma_{j1}(Y_i) &= \frac{1}{1 - H(Y_i)} \int I_{\{Y_i < \omega\}} \phi_j(\omega) \gamma_0(\omega) H^1(d\omega), \\ \gamma_{j2}(Y_i) &= \int \int \frac{I_{\{Y_i < Y_i, v < \omega\}} \phi_j(\omega) \gamma_0(\omega)}{(1 - H(v))^2} H^0(dv) H^1(d\omega), \\ H^0(y) &= \mathbb{P}_N(Y \leq y, \delta = 0) = \frac{\sum (1 - z_i) s_i I_{\{Y_i \leq y\}}}{\sum (1 - z_i) s_i}, \\ H^1(y) &= \mathbb{P}_N(Y \leq y, \delta = 1) = \frac{\sum (1 - z_i) s_i \delta_i I_{\{Y_i \leq y\}}}{\sum (1 - z_i) s_i}, \end{aligned}$$

and

$$H(y) = \mathbb{P}_N(Y \leq y) = \frac{\sum (1 - z_i) s_i I_{\{Y_i \leq y\}}}{\sum (1 - z_i) s_i}.$$

To be clear, and to simplify further notation,

$$\int_0^\infty w(s; \alpha, \beta) dF_p(s) = \sum_{j=1}^{k+1} w_j(\alpha) F_p(t_j) - F_p(t_{j-1}),$$

where $w_j(\alpha) = w(t_j; \alpha, \beta)$, $w_{k+1}(\alpha) = w(\tau; \alpha, \beta)$, $F_p(t_0) = 0$, and $F_p(t_{k+1}) = 1$. From Section A.1, $\sqrt{N}(\widehat{\theta} - \theta) \rightarrow^d N(0, \Psi)$, where

$$\Psi = E \left[\frac{\partial}{\partial \theta} \psi(\theta) \right]^{-1} E[\psi(\theta) \psi(\theta)^T] E \left[\left(\frac{\partial}{\partial \theta} \psi(\theta) \right)^T \right]^{-1}.$$

Define

$$g(\theta) = \frac{\sum_{j=1}^l w_j(\alpha) (F_p(t_j) - F_p(t_{j-1}))}{\sum_{j=1}^{k+1} w_j(\alpha) (F_p(t_j) - F_p(t_{j-1}))},$$

where $t_l = \sup(t_j)$ such that $t_j < t$. Consequently, $g(\widehat{\theta}) = \widehat{F}_p^{ai}(t)$. By the delta method,

$$\sqrt{N}(g(\widehat{\theta}) - g(\theta)) \rightarrow^d N(0, g'(\theta) \Psi g'(\theta)^T).$$

We can estimate $g'(\theta)$ with $g'(\widehat{\theta})$ and ψ with

$$\widehat{\psi} = \left[\frac{1}{N} \sum_{i=0}^N \frac{\partial}{\partial \theta} \psi_i(\widehat{\theta}) \right]^{-1} \times \frac{1}{N} \sum_{i=1}^N [\psi_i(\widehat{\theta}) \psi_i(\widehat{\theta})^T] \left[\frac{1}{N} \sum_{i=1}^N \left(\frac{\partial}{\partial \theta} \psi_i(\widehat{\theta}) \right)^T \right]^{-1}.$$

Acknowledgements

The authors thank VaxGen for allowing use of their data, and particularly Marc Gurwith for providing ranges for the sensitivity parameter. This article supported in part by National Institutes of Health grant 2 RO1 AI054165-04.

References

- Andersen, PK.; Borgan, O.; Gill, RD.; Keiding, N. Statistical Models Based on Counting Processes. Berlin: Springer; 1992.
- Chen PY, Tsiatis AA. Causal Inference on the Difference of the Restricted Mean Lifetime Between Two Groups. Biometrics 2001;57:1030–1038. [PubMed: 11764241]
- Flynn NM, Forthal DN, Harro CD, Judson FN, Mayer KH, Para MF. the rgp120 HIV Vaccine Study Group. Placebo-Controlled Phase 3 Trial of a Recombinant Glycoprotein 120 Vaccine to Prevent HIV-1 Infection. Journal of Infectious Diseases 2005;191:647–649. [PubMed: 15688276]
- Frangakis CE, Rubin DB. Principal Stratification in Causal Inference. Biometrics 2002;58:21–29. [PubMed: 11890317]
- Gilbert PB, Ackers ML, Berman PW, Francis DP, Popovic V, Hu DJ, Heyward WL, Sinangil F, Shepherd BE, Gurwith M. HIV-1 Virologic and Immunologic Progression and Antiretroviral Therapy Initiation Among HIV-1–Infected Participants in an Efficacy Trial of Recombinant Glycoprotein 120 Vaccine. Journal of Infectious Diseases 2005;192:974–983. [PubMed: 16107949]
- Gilbert PB, Bosch RJ, Hudgens MG. Sensitivity Analysis for the Assessment of Causal Vaccine Effects on Viral Load in HIV Vaccine Trials. Biometrics 2003;59:531–541. [PubMed: 14601754]
- Graham BS. Clinical Trials of HIV Vaccines. Annual Review of Medicine 2002;53:207–221.
- Halloran ME, Struchiner CJ. Causal Inference for Infectious Diseases. Epidemiology 1995;6:142–151. [PubMed: 7742400]
- Hudgens MG, Halloran ME. Causal Vaccine Effects on Binary Postinfection Outcomes. Journal of the American Statistical Association 2006;101:51–64. [PubMed: 19096723]

- Hudgens MG, Hoering A, Self SG. On the Analysis of Viral Load Endpoints in HIV Vaccine Trials. *Statistics in Medicine* 2003;22:2281–2298. [PubMed: 12854093]
- Jemai Y, Rotnitzky A. Asymptotic Properties of an Estimator of Treatment Effects on an Outcome Only Existing if a Post-Randomization Event Has Occured. Yannis Jemai doctoral dissertation, Harvard School of Public Health, Department of Biostatistics. 2005
- Kalbfleish, JD.; Prentice, RL. *The Statistical Analysis of Failure Time Data*. New York: Wiley; 1980.
- Nabel GJ. Challenges and Opportunities for Development of an AIDS Vaccine. *Nature* 2001;410:1002–1007. [PubMed: 11309631]
- Neyman J. *Statistical Science* 1923;5:465–480.
- Robins JM. A New Approach to Causal Inference in Mortality Studies With Sustained Exposure Periods, With Application to Control of the Healthy Worker Survivor Effect. *Mathematical Modeling* 1986;7:1393–1512.
- Robins JM. An Analytic Method for Randomized Trials With Informative Censoring: Part I. Lifetime Data Analysis 1995;1:241–254. [PubMed: 9385104]
- Robins, JM.; Greenland, S. Comment on Causal Inference Without Counterfactuals. In: Dawid, AP., editor. *Journal of the American Statistical Association*. 95. 2000. p. 431–435.
- Rosenbaum PR. The Consequences of Adjustment for a Concomitant Variable That Has Been Affected by the Treatment. *Journal of the Royal Statistical Society* 1984;147:656–666.A
- Rubin DB. Bayesian Inference for Causal Effects: The Role of Randomization. *The Annals of Statistics* 1978;6:34–58.
- Rubin, DB. Comment on Causal Inference Without Counterfactuals. In: Dawid, AP., editor. *Journal of the American Statistical Association*. 95. 2000. p. 435–437.
- Scharfstein DO, Rotnitzky A, Robins JM. Adjusting for Non-ignorable Dropout Using Semiparametric Nonresponse Models. *Journal of the American Statistical Association* 1999;94:1096–1146.
- Shepherd BE, Gilbert PB, Jemai Y, Rotnitzky A. Sensitivity Analyses Comparing Outcomes Only Existing in a Subset Selected Post-Randomization, Conditional on Covariates, With Application to HIV Vaccine Trials. *Biometrics* 2006;62:332–342. [PubMed: 16918897]
- Shepherd BE, Gilbert PB, Mehrotra DV. Eliciting a Counterfactual Sensitivity Parameter. *The American Statistician* 2007;61:56–63.
- Stute W. The Central Limit Theorem Under Random Censorship. *The Annals of Statistics* 1995;23:422–439.
- van der Vaart, AW. *Asymptotic Statistics*. Cambridge, U.K.: Cambridge University Press; 1998.

Figure 1.

Histograms of $\widehat{F}_p^{ai}(t)$ for $t = 24$ Months and Plots of $F_p(t)$ (---) and $F_p^{ai}(t)$ (—) at Different Levels of VE for $\beta = 1$.

Figure 2.

Sensitivity Analyses of the Causal Effect of Vaccination on the Probability of not yet Initiating Antiretroviral, Therapy in the VaxGen Trial [(a) and (b)] and in the Trial's Nonwhite Cohort [(c) and (d)]. In (a) and (c), — $1 - \hat{F}_p(t)$; $\cdots$, $1 - \hat{F}_p^{ai,U}(t)$; $\cdots$, $1 - \hat{F}_p^{ai,L}(t)$, and - - - $1 - \hat{F}_v(t)$. In (b) and (d), — is $\widehat{SCE}(t=2 \text{ years})$, $\cdots$ is the 95% bootstrap Wald confidence interval (CI), - - - is the 95% CI based on the asymptotic variance, $\circ$ is $\widehat{SCE}(t)$ at $\beta = \pm\infty$, $+$ is 95% bootstrap Wald CI at $\beta = \pm\infty$, $\times$ is 95% bootstrap percentile CI at $\beta = \pm\infty$, and [] is Dr. Gurwith's plausible range for β .

Table 1
Bias of Estimates and Coverage Probabilities of Wald-Based 95% Confidence Intervals for SCE(t) With t = 24 Months

VE	β	Bias		Coverage probability		
		Mean	Median	Bootstrap	Analytic	
~.3	.1	-.002	0	.946	.948	
	.2	-.007	-.007	.943	.949	
	1	-.006	-.003	.943	.946	
~.6	∞	-.007	.005	.948	.953	
	.1	.003	.006	.939	.940	
	.2	.013	.015	.945	.945	
	1	.042	.020	.933	.912	
	∞	.035	0	.935		

NOTE: Analytic coverage probabilities when $VE \approx .6$, $\beta = \infty$ are not given, because in many simulations there were no failures in the vaccine arm.

Table 2
Bias of Estimates and Coverage Probability of Wald-Based 95% Confidence Intervals for $F_p^{ai}(t)$ With $t = 24$ Months

VE	β	Coverage probability					
		Bias		Standard		Log-log transformation	
		Mean	Median	Bootstrap	Analytic	Bootstrap	Analytic
~.3	.1	-.001	-.002	.939	.941	.946	.942
	.2	-.006	-.002	.932	.939	.938	.949
	1	.001	-.005	.940	.953	.959	.973
	∞	-.007	.002	.934	.947	.957	.967
~.6	.1	.001	.003	.939	.942	.947	.947
	.2	.015	.011	.929	.927	.938	.937
	1	.045	.010	.903	.819	.902	.859
	∞	.036	-.016	.889	.948		

NOTE: Analytic coverage probabilities when $VE \approx .6$, $\beta = \infty$ are not given because in many simulations, $\log[-\log]$ -transformed confidence intervals could not be computed because the estimated value of $F_p^{ai}(t)$ was 0.