Comparing competing risk outcomes within principal strata, with application to studies of mother-to-child transmission of HIV

Dustin M. Long and **Michael G. Hudgens**[*]

Abstract

In randomized trials to prevent breast milk transmission of human immunodeficiency virus (HIV) from mother to infant, investigators are often interested in assessing the effect of a treatment or intervention on the cumulative risk of HIV infection by time (age) t in infants who are alive and uninfected at a certain time point $\tau_0 < t$. Such comparisons are challenging for two reasons. First, infants are typically randomized at birth (time $0 < \tau_0$) such that comparisons between trial arms among the subset of infants alive and uninfected at τ_0 are subject to selection bias. Second, in most mother-to-child transmission (MTCT) trials competing risks are often present, such as death or cessation of breastfeeding prior to HIV infection. In this paper we present methods for assessing the causal effect of a treatment on competing risk outcomes within principal strata. In MTCT trials, the causal effect of interest is that of treatment on the risk of HIV infection by time $t > \tau_0$ within the principal stratum of infants who would be alive and uninfected by τ_0 regardless of randomization assignment. Large sample non-parametric bounds and a semi-parametric sensitivity analysis model are developed for drawing inference about this causal effect. A simulation study is presented demonstrating that the proposed methods perform well in finite samples. The proposed methods are applied to a large, recent MTCT trial.

Keywords

Causal inference; Infectious diseases; Principal stratification; Sensitivity analysis

1. Introduction

Every year approximately 200,000 infants become infected with HIV through breastfeeding; in the absence of treatment, half of these infants will die within two years of birth [1, 2]. In clinical trials to prevent MTCT of HIV through breast milk, investigators are often interested in comparing interventions conditional on the infant being alive and uninfected up to a certain time point during the trial [3–6]. Specifically, when randomization occurs at birth (time 0), a time point $\tau_0 > 0$ is often chosen prior to the beginning of the trial and only randomized infants alive and uninfected at τ_0 are considered for analysis. For example, in the Breastfeeding, Antiretroviral, and Nutrition (BAN) study [3, 4] infants were randomized at birth but the primary analysis included only infants HIV uninfected and alive at $\tau_0 = 2$ weeks. Infants infected prior to 2 weeks were excluded because these transmissions likely occurred in utero or during labor and delivery, whereas the primary objective of the trial was to assess the effects of interventions to prevent infection due to breast milk. Similar exclusions were made in the primary analysis of the SWEN and PEPI trials [6, 7].

[*]Correspondence to: Department of Biostatistics, University of North Carolina at Chapel Hill, 3101 McGavran-Greenberg Hall, CB #7420 Chapel Hill, NC 27599-7420. mhudgens@bios.unc.edu.

There are two aspects of the analysis described above that are the focus of this paper. First, an analysis comparing risk of HIV infection between trial arms among infants who are alive and uninfected at time τ_0 after randomization is subject to selection bias. One method to protect against selection bias in this scenario entails principal stratification [8]. Principal stratification uses the potential outcomes of a variable collected post-randomization to define strata of individuals. In the MTCT trial setting, the principal stratum of interest is infants who would be alive and uninfected by time τ_0 under either treatment assignment. Because principal stratum membership is not affected by treatment assignment, comparisons between trial arms within a particular principal stratum are not subject to selection bias. For a recent discussion of the strengths and weaknesses of principal stratification, see Pearl [9] and subsequent responses such as VanderWeele [10].

The second aspect in the analysis of the effect of treatment on the risk of HIV infection in MTCT trials is the presence of competing risks [11]. In particular, death or weaning prior to HIV transmission are competing risks for HIV infection since these events (death, weaning) can preclude HIV infection from occurring. Likewise, HIV infection precludes the possibility of an HIV-free death or weaning prior to HIV infection. One analytical approach that avoids the complication of competing risks is to use a composite endpoint, such as time until HIV infection or death. Using a composite endpoint simplifies analysis and has the advantage of providing a single measure of the overall effect of treatment. However, such an analysis does not provide inference about whether the treatment is having an effect on the risk of HIV infection, death, or both endpoints. Another common approach in the analysis of MTCT trials is to treat infants experiencing HIV-free death as right censored, e.g., when computing the Kaplan-Meier estimator of the cumulative probability of HIV infection (for instance, see Figure 2a of Kumwenda *et al.* [6]). It is well known that computing the Kaplan-Meier estimator by right censoring competing events does not in general yield a consistent estimator of the cumulative risk of the event of interest [12, 13]; in the MTCT setting such Kaplan-Meier estimators will tend to overestimate the risk of HIV infection when there is a non-zero probability of death prior to HIV infection. A third approach, adopted in this paper, is to estimate the cumulative incidence functions of each competing event, namely HIV, death, and weaning. The resulting estimates have a straightforward interpretation as the cumulative risk of each event in settings such as the trial where the other events may occur. Contrasts is the estimated risks between trial arms can then be used to assess treatment effects on each of the competing events.

Previous work on estimating treatment effects within principal strata has considered binary outcomes (e.g., Hudgens and Halloran [14]), continuous outcomes (e.g., Gilbert *et al.* [15]) and survival outcomes (e.g., Hayden *et al.* [16] and Shepherd *et al.* [17]). In this paper we develop methods for estimating treatment effects within principal strata for a survival outcome in the presence of competing risks. In the absence of competing risks the developed methods essentially reduce to those of Shepherd *et al.* [17]. The outline of the remainder of the paper is as follows. In Section 2 notation and assumptions are discussed. In Section 3 inferential methods for the causal effect of interest are presented. The finite sample performance of the methods are assessed in a simulation study in Section 4. These simulations also illustrate how misleading inferences can arise if selection bias are ignored. In Section 5 the methods are applied to investigate the effect of infant antiretroviral therapy (ART) on the cumulative risk of HIV infection in the BAN trial. A brief discussion is given in Section 6.

2. Notation and Assumptions

Suppose *n* individuals are randomly assigned one of two treatments, 0 or 1, at baseline (birth or time 0). For $i = 1, \ldots, n$, let $Z_i = 0$ if subject *i* is assigned treatment 0 and $Z_i = 1$ otherwise.

Let $n_0 = \Sigma\,(1 - Z_i)$ and $n_1 = \Sigma\,Z_i$, where here and throughout $\sum = \sum_{i=1}^{n}$. Without loss of generality, assume $Z_i = 0$ corresponds to placebo or control, and $Z_i = 1$ corresponds to active treatment. In the BAN study analysis, $Z_i = 1$ will refer to the infant ART arm and $Z_i = 0$ will refer to the control arm. Suppose the primary objective is to assess the effect of treatment on the time T_i (from baseline) until some particular event occurs. Assume there are k possible causes or types of events and let J_i denote the event type for individual i with $J_i \in \{1, \ldots, k\}$. In the BAN study there are $k = 3$ competing risks: HIV infection ($J_i = 1$), death prior to HIV infection or weaning ($J_i = 2$), or cessation of breastfeeding prior to HIV infection ($J_i = 3$).

Suppose in the analysis of the effect of treatment Z_i on (T_i, J_i) we would like to condition on some binary post-randomization variable S_i (taking on values 0 or 1) measured at some pre-specified post-randomization time $\tau_0 > 0$. For instance, in the analysis of BAN it is desired to assess the effect of treatment in infants alive and uninfected at time τ_0; in this case we let $S_i = 1$ if an infant becomes infected or dies by τ_0 and $S_i = 0$ otherwise. Note for the BAN example that $S_i = I(T_i \leq \tau_0, J_i \neq 2)$ where $I(\cdot)$ is the usual indicator function, however in the methods developed below S_i need not be defined in terms of T_i or J_i.

Define C_i to be a potential right censoring time and assume $\tau_0 \leq C_i$, i.e., no individuals drop out of the study prior to τ_0 such that S_i is always observed. Let τ_1 denote the maximum length of follow-up for the study such that any individual who has not had an event or dropped out of the study by time τ_1 is administratively censored at that time, i.e., $C_i \leq \tau_1$. Let $Y_i = \min\{T_i, C_i\}$ and $\Delta_i = I(Y_i = T_i)$. Due to censoring, instead of (T_i, C_i, J_i) we only observe $(Y_i, J_i\Delta_i)$; i.e., T_i and J_i are observed if and only if individual i is not right censored.

Let $T_i(z)$ be the potential survival time when assigned treatment z for $z = 0,1$ such that $T_i = (1 - Z_i)T_i(0) + Z_i T_i(1)$. Define $C_i(z)$, $S_i(z)$, and $J_i(z)$ similarly. Assume the treatment assignment of individual i does not affect the potential outcomes of other individuals (i.e., there is no interference) and there are not multiple forms of treatment, i.e., the stable unit treatment value assumption (SUTVA) holds [18]. Let $W_i = (S_i(0), S_i(1), T_i(0), T_i(1), J_i(0), J_i(1), C_i(0), C_i(1))$ denote the vector of potential outcomes and $O_i = (Z_i, S_i, Y_i, J_i\Delta_i)$ denote the vector of observable random variables. Assume individuals in the study are a random sample from a larger population such that $W_1, \ldots, W_n$ and $O_1, \ldots, O_n$ are iid copies of W and O respectively.

Principal strata can be defined by sets of individuals with the same potential outcome pair $(S_i(0) = s_0, S_i(1) = s_1)$. Define the never infected (NI) principal stratum to be individuals with $S_i(0) = S_i(1) = 0$, i.e., individuals who would be alive and uninfected at τ_0 regardless of treatment assignment. Similarly define the harmed stratum as those individuals with $S_i(0) = 0$, $S_i(1) = 1$; the protected stratum as those individuals with $S_i(0) = 1$, $S_i(1) = 0$; and the doomed stratum as those individuals with $S_i(0) = S_i(1) = 1$. Motivated by MTCT studies of HIV, we focus on drawing inference about causal effects in the NI principal stratum. For example, in the BAN study we are interested in the principal stratum of infants who would be alive and not infected with HIV by $\tau_0 = 2$ weeks under either randomization assignment.

In the presence of competing risks, a quantity of interest is the cumulative incidence function (CIF) or subdistribution function of (T, J). Let $F(t, j) = P(T \leq t, J = j)$ denote the CIF, i.e., the probability of having event j at or before time t. Define the causal estimand of interest to be $CE(t, j) = F_1^{NI}(t, j) - F_0^{NI}(t, j)$ for $t \in [\tau_0, \tau_1]$ where $F_z^{NI}(t, j) = \Pr[T_i(z) \leq t, J_i(z) = j \mid S_i(0) = S_i(1) = 0]$ for $z = 0, 1$. In words, $CE(t, j)$ is the difference in the probability of having an event of type j by time t for treatment 0 compared to treatment 1 within the NI principal stratum. For example, in the BAN study (where $j = 1$ corresponds to HIV infection), $CE(28, 1)$ is the difference in the probability of HIV

infection by 28 weeks between the two study arms among infants who would be alive and HIV negative by τ_0 weeks regardless of treatment assignment. In the analysis of BAN, $CE(28, 1)$ was of particular interest because per protocol a primary endpoint of the trial was HIV infection by 28 weeks [3].

To draw inference about $CE(t, j)$ we make the following assumptions:

Assumption 2.1   Independent treatment assignment: $Z_i \perp W_i$

Assumption 2.2   Monotonicity: $S_i(1) \quad S_i(0)$ for all i

Assumption 2.3   Independent censoring: $C_i(z) \perp \{T_i(z), J_i(z), S_i(z)\}$ for z = 0, 1

Assumption 2.1 is plausible in randomized clinical trials. Assumption 2.2 is a strong assumption that must be considered carefully and is discussed further in Section 5 in the context of the BAN study. Methods not requiring the monotonicity assumption are discussed in Section 6. Assumption 2.3 is a common assumption when analyzing competing risks data. In the infant ART and control arms of BAN, 15% of participants were administratively censored at τ_1 = 28 weeks and 12% were censored at earlier time points due to drop-out from the study prior to week 28.

Under Assumptions 2.1 and 2.2, $Z_i = 0$ and $S_i = 0$ imply $S_i(0) = S_i(1) = 0$; i.e., individuals who are alive and uninfected by τ_0 when assigned control must be members of the NI principal stratum. Letting $F_0(t, j) = \Pr[T_i(0) \quad t, J_i(0) = j|S_i(0) = 0]$, it follows under Assumptions 2.1 – 2.2 that $F_0^{NI}(t, j) = F_0(t, j)$, which is identifiable from the observable data under Assumption 2.3. However $F_1^{NI}(t, j)$ remains unidentifiable under Assumptions 2.1 – 2.3 because individuals who are alive and uninfected by τ_0 when assigned treatment ($Z_i = 1$) are a mixture of individuals from the NI and protected principal strata. In particular, following Gilbert *et al.* [15], one can show

$$F_1(t, j) = \gamma F_1^{NI}(t, j) + (1-\gamma) F_1^{prot}(t, j), \quad (1)$$

where $\gamma = \Pr[S_i(0) = 0|S_i(1) = 0]$ is the probability an individual is uninfected under control given they would be uninfected under treatment, $F_1(t, j) = \Pr[T_i(1) \quad t, J_i(1) = j|S_i(1) = 0]$ and $F_1^{prot}(t, j) = \Pr[T_i(1) \le t, J_i(1) = j|S_i(0) = 1, S_i(1) = 0]$.

To proceed, one can introduce an additional assumption about the selective effect of conditioning on S_i which renders $F_1^{NI}(t, j)$ identifiable. For example, following Hudgens and Halloran [14], large-sample upper and lower bounds can be obtained by considering extreme selection bias models. The upper bound selection model is given by assuming either $F_1^{prot}(t, j) = 0$ or $F_1^{NI}(t, j) = 1$, while the lower bound selection model is given by assuming either $F_1^{prot}(t, j) = 1$ or $F_1^{NI}(t, j) = 0$. By (1), these models are equivalent to assuming either

$$F_1^{NI}(t, j) = \min\{\gamma^{-1} F_1(t, j), 1\}, \quad (2)$$

or

$$F_1^{NI}(t, j) = \max\left\{\frac{F_1(t, j) - (1-\gamma)}{\gamma}, 0\right\}. \quad (3)$$

Estimating $CE(t, j)$ under (2) or (3) is useful in bounding the estimate of the causal effect above and beyond any possible selective effects induced by conditioning on $S_i = 0$.

The true degree of selection bias may be considerably less than that assumed by (2) or (3). Therefore, we consider a class of selection models that includes the extreme models above as special cases. Through sensitivity analysis over the entire class (as in Robins *et al.* [19] and Gilbert *et al.* [15]), the relationship between the assumed degree of selection bias and inference about $CE(t, j)$ can be explored. These selection models are semiparametric in the sense that no additional restrictions are placed on the distribution of the observable random variables $O_1, \ldots, O_n$ but an unidentifiable parameter (β_j in the model below) is used to quantify the selection bias. One possible selection model is:

Assumption 2.4

$$\exp(\beta_j) = \frac{F_1^{NI}(t, j)/\{1 - F_1^{NI}(t, j)\}}{F_1^{prot}(t, j)/\{1 - F_1^{prot}(t, j)\}}. \quad (4)$$

The parameter β_j equals the log odds ratio of having an event of type j by time t under treatment assignment $z = 1$ in the NI principal stratum versus the protected principal stratum. Note Assumption 2.4 allows for the log odds to differ across event types as indicated by the subscript on β. Also note (4) is unverifiable since β_j is not identifiable from the observable data. For fixed β_j, under Assumptions 2.1 – 2.4 $F_1^{NI}(t, j; \beta_j)$ is identifiable from the observable data and $CE(t, j)$ can be estimated as described in Section 3 below. The extreme models (2) and (3) can be viewed as special cases of Assumption 2.4 as $\beta_j \to \infty$ and $\beta_j \to -\infty$. We refer to $\beta_j = 0$ as the no selection bias model because in this case the odds of having an event of type j by time t are the same in the NI and protected principal strata. Sensitivity analysis of inference about $CE(t; j)$ can be conducted by letting β_j range from $-\infty$ to ∞. Gains in power or precision may be achieved by restricting the range of β_j based on prior information about β_j elicited from subject matter experts [20, 21].

3. Inference

In this section we first consider nonparametric estimation of $CE(t, j)$ under the extreme selection models (2) and (3). Then inference for $CE(t, j)$ under the semiparametric selection model (4) given some value of β_j is discussed in Section 3.2. The construction of uncertainty intervals about $CE(t, j)$ is considered in Section 3.3.

3.1. Nonparametric Estimation: Bounds

Under Assumptions 2.1 – 2.3 consistent estimators of $F_1^{NI}(t, j)$ assuming (2) or (3) are given, respectively, by

$$\widehat{F}_1^{NI,up}(t, j) = \min\left\{\widehat{\gamma}^{-1}\widehat{F}_1(t, j), 1\right\} \text{ and } \widehat{F}_1^{NI,low}(t, j) = \max\left\{\frac{\widehat{F}_1(t, j) - (1 - \widehat{\gamma})}{\widehat{\gamma}}, 0\right\}, \quad (5)$$

where

$$\widehat{\gamma} = \min\left\{\frac{\sum(1-S_i)(1-Z_i)/n_0}{\sum(1-S_i)Z_i/n_1}, 1\right\},$$

and $\hat{F}_1(t, j)$ is the Aalen-Johansen estimator [22] of $F_1(t, j)$ calculated using $(Y_i, J_i\Delta_i)$ for individuals with $Z_i = 1$ and $S_i = 0$. It can be shown that $\hat{\gamma}$ and $\hat{F}_1(t, j)$ are nonparametric maximum likelihood estimators (NPMLEs) of γ and $F_1(t, j)$. Thus the estimators in (5) can be viewed as NPMLEs of $F_1^{NI}(t, j)$. Because Assumptions 2.1 and 2.2 imply $F_0^{NI}(t, j) = F_0(t, j)$, consistent estimators of $CE(t, j)$ assuming either (2) or (3) are $\widehat{CE}^{up}(t, j) = \widehat{F}_1^{NI,up}(t, j) - \widehat{F}_0(t, j)$ or $\widehat{CE}^{low}(t, j) = \widehat{F}_1^{NI,low}(t, j) - \widehat{F}_0(t, j)$, where $\hat{F}_0(t, j)$ is the Aalen-Johansen estimator of $F_0(t, j)$ calculated using $(Y_i, J_i \Delta_i)$ for individuals with $Z_i = S_i = 0$. In the nomenclature of Vansteelandt *et al.* [23], the interval $[\widehat{CE}^{low}(t, j), \widehat{CE}^{up}(t, j)]$ is an estimated ignorance region of $CE(t, j)$.

If $0 < \gamma < 1$, then $\hat{\gamma}$ is asymptotically normal. The Aalen-Johansen estimators $\hat{F}_z(t, j)$, for $z = 0, 1$ are asymptotically normal assuming $0 < F_z(t; j) < 1$ and certain regularity conditions [24]. Therefore, $\widehat{F}_1^{NI,up}(t, j)$ is asymptotically normal if, in addition to these conditions,

$$F_1(t, j) < \gamma. \quad (6)$$

If (6) does not hold, then $\widehat{F}_1^{NI,up}(t, j) \xrightarrow{p} 1$ and hence is not asymptotically normal. Under conditions where $\widehat{F}_1^{NI,up}(t, j)$ is asymptotically normal, a consistent estimator of the variance of $\widehat{F}_1^{NI,up}(t, j)$ is

$$\widehat{\mathrm{var}}\{\widehat{F}_1^{NI,up}(t, j)\} = \frac{\widehat{\mathrm{var}}\{\widehat{F}_1(t, j)\}}{\widehat{\gamma}^2} + \left\{\frac{\widehat{F}_1(t, j)}{\widehat{\gamma}}\right\}^2\left(\frac{1}{N_0} - \frac{1}{n_0} + \frac{1}{N_1} - \frac{1}{n_1}\right), \quad (7)$$

where $\widehat{\mathrm{var}}\{\widehat{F}_1(t, j)\}$ is a consistent estimator of the variance of $\hat{F}_1(t, j)$ (e.g., see Aalen *et al.* [24], Section 3.4.5) and $N_z = \sum I(S_i = 0, Z_i = z)$. Similarly $\widehat{F}_1^{NI,low}(t, j)$ is asymptotically normal if, in addition to the conditions above,

$$1 - \gamma < F_1(t, j). \quad (8)$$

If (8) does not hold, $\widehat{F}_1^{NI,low}(t, j) \xrightarrow{p} 0$ and hence is not asymptotically normal. If $\widehat{F}_1^{NI,low}(t, j)$ is asymptotically normal, the variance can be consistently estimated by

$$\widehat{\mathrm{var}}\{\widehat{F}_1^{NI,low}(t, j)\} = \frac{\widehat{\mathrm{var}}\{\widehat{F}_1(t, j)\}}{\widehat{\gamma}^2} + \left\{\frac{1 - \widehat{F}_1(t, j)}{\widehat{\gamma}}\right\}^2\left(\frac{1}{N_0} - \frac{1}{n_0} + \frac{1}{N_1} - \frac{1}{n_1}\right). \quad (9)$$

Derivations of (7) and (9) are given in the appendix. When (6) and (8) hold, pointwise Wald-type confidence intervals for $CE(t, j)$ can be constructed in the usual manner. Alternatively, the bootstrap percentile method can be used for computing confidence intervals of $CE(t, j)$. If (6) and (8) do not hold, then $\widehat{F}_1^{NI,up}(t, j) \xrightarrow{p} 1$ and $\widehat{F}_1^{NI,low}(t, j) \xrightarrow{p} 0$,

i.e., the bounds are non-informative. Note that conditions (6) and (8) can be assessed based on observed data by comparing $\hat{\gamma}$ and $\hat{F}_1(t, j)$.

3.2. Semiparametric Estimation

Under Assumptions 2.1 – 2.4, for fixed β_j a semiparametric estimator of $F_1^{NI}(t, j)$ can be constructed by plugging $\hat{F}_1(t, j)$ and $\hat{\gamma}$ into equation (1) and then simultaneously solving (1) and (4) for $F_1^{NI}(t, j)$. This can be accomplished by expressing $F_1^{prot}(t, j)$ as a function of β_j and $F_1^{NI}(t, j)$ using (4), replacing $F_1^{prot}(t, j)$ by this expression in (1), and finding the solution to (1) using a one-dimensional line search. Define the solution as $\widehat{F}_1^{NI}(t, j;\beta_j)$ and let the corresponding estimator of the causal effect be $\widehat{CE}(t, j;\beta_j)=\widehat{F}_1^{NI}(t, j;\beta_j)-\widehat{F}_0(t, j)$. Without a closed form for $\widehat{F}_1^{NI}(t, j;\beta_j)$, confidence intervals of $F_1^{NI}(t, j)$ and $CE(t, j)$ for an assumed value of β_j can be constructed using the bootstrap percentile method; alternatively, Wald-type confidence intervals can be constructed based on bootstrap estimates of $\mathrm{var}\{\widehat{F}_1^{NI}(t, j;\beta_j)\}$ and $\mathrm{var}\{\widehat{CE}(t, j;\beta_j)\}$.

Note $\lim_{\beta_j\to\infty}\widehat{CE}(t, j;\beta_j)=\widehat{CE}^{up}(t, j)$ and $\lim_{\beta_j\to-\infty}\widehat{CE}(t, j;\beta_j)=\widehat{CE}^{low}(t, j)$, i.e., the estimators that arise from the extreme selection models (2) and (3) are special cases of the estimators from the semiparametric bias model (4). Under the no selection model $\beta_j = 0$, $\widehat{CE}(t, j;\beta_j)=\widehat{F}_1(t, j)-\widehat{F}_0(t, j)$, i.e., the causal effect is estimated by the difference in Aalen-Johansen estimators from the two treatment groups as in a standard competing risks analysis. In other words, assuming the no selection model gives rise to a naive or "net" estimator [8] which simply compares subsets of the two randomization groups conditional on being observed HIV free and alive at τ_0.

3.3. Uncertainty Regions

The pointwise confidence intervals described in Sections 3.1 and 3.2 will contain $CE(t, j)$ with the stated coverage probability provided the correct value of β_j is assumed. However, the true value of β_j is not identifiable from the observed data. Therefore, following Vansteelandt *et al.* [23], it is useful to also construct a $(1 - a)100\%$ uncertainty interval which contains $CE(t, j)$ with probability $1 - a$ without conditioning on any assumption about the value of β_j. Under the assumptions given in Section 3.1 where $\widehat{CE}^{up}(t, j)$ and $\widehat{CE}^{low}(t, j)$ are consistent and asymptotically normal, a large sample $(1 - a)100\%$ pointwise uncertainty interval for $CE(t, j)$ is given by

$$[\widehat{CE}^{low}(t, j)-c_{\alpha/2}^*\widehat{\mathrm{var}}\{\widehat{CE}^{low}(t, j)\}^{1/2}, \quad \widehat{CE}^{up}(t, j)+c_{\alpha/2}^*\widehat{\mathrm{var}}\{\widehat{CE}^{up}(t, j)\}^{1/2}]$$

where $c_{\alpha/2}^*$ can be computed using equation (4.3) of Vansteelandt *et al.* [23], $\widehat{\mathrm{var}}\{\widehat{CE}^{low}(t, j)\}=\widehat{\mathrm{var}}\{\widehat{F}_1^{NI,low}(t, j)\}+\widehat{\mathrm{var}}\{\widehat{F}_0(t, j)\}$ and $\widehat{\mathrm{var}}\{\widehat{CE}^{up}(t, j)\}=\widehat{\mathrm{var}}\{\widehat{F}_1^{NI,up}(t, j)\}+\widehat{\mathrm{var}}\{\widehat{F}_0(t, j)\}$

4. Simulation Study

Simulations were conducted to evaluate the performance of the methods described in Section 3 for drawing inference about $CE(t, j)$. Data were simulated based on the BAN study

under five models: $\beta_j = -\infty, -1, 0, 1, \infty$ for fixed j. These five choices of β_j correspond to the two extreme selection models ($\beta_j = -\infty, \infty$), two intermediate selection models ($\beta_j = -1$, 1), and the no selection bias model ($\beta_j = 0$). The Gompertz distribution was used to simulate competing risks data [25]. Under the Gompertz distribution the CIF can be expressed as $F(t, j) = 1 - \exp[\lambda_j\{1 - \exp(a_j t)\}/a_j]$ where $\{a_1, \ldots, a_k, \lambda_1, \ldots, \lambda_k\}$ are chosen such that

$$\sum_{j=1}^{k}\Pr[J=j]=\sum_{j=1}^{k}F(\infty, j)=1.$$ For the simulation study $k = 3$ and the parameters $\{a_1, a_2, a_3, \lambda_1, \lambda_2, \lambda_3\}$ were selected such that $F_1(28, 1) = 0.02$, $F_1(28, 2) = 0.02$, $F_1(28, 3) = 0.70$,

and $\sum_{j=1}^{3}F_1(\infty, j)=1.$ These probabilities correspond roughly to the estimated risk of HIV infection ($j = 1$), death ($j = 2$) prior to HIV infection or weaning, and cessation of breastfeeding prior to HIV infection ($j = 3$) at 28 weeks in the BAN study among infants randomized to the infant ART arm who were HIV negative and alive at 2 weeks.

Simulations were conducted under two scenarios (for each of the five models). For the first scenario we let $\gamma = 0.9884$, corresponding to the estimated value of γ from the BAN study. In this scenario we considered estimating $CE(28, 1)$, i.e., the effect of treatment on risk HIV infection at 28 weeks. Note (6) and (8) hold in this scenario for $t = 28$ and $j = 1$ such that the estimators of the bounds are asymptotically normal. Because $\gamma = 0.9884$ is near the boundary value of 1, for the second scenario we let $\gamma = 0.75$. In order for (6) and (8) to hold in the second scenario, we considered estimating $CE(28, 3)$, i.e., the effect of treatment on weaning at 28 weeks. For the first scenario simulations were conducted under the alternative hypothesis $CE(28, 1) = -0.05$, i.e., the risk of HIV infection is lowered by 5% due to treatment. For the second scenario simulations were conducted where $CE(28, 3) = 0.05$, i.e., women are more likely to breastfeed at 28 weeks when the infant receives ART. For each model and each scenario, data sets of $n = 1520$ iid copies of W were simulated according to the following steps. The description below is for the first scenario where $j = 1$ is the event of interest; simulations were conducted analogously for the second scenario where $j = 3$ is the event of interest.

Step 1 $S_i(1)$ was drawn from a Bernoulli(0.0458), where 0.0458 was the estimated risk of infection or death at two weeks in the infant ART arm of BAN.

Step 2   If $S_i(1) = 1$, then by monotonicity $S_i(0) = 1$. In this case we let $T_i(0) = J_i(0) = T_i(1) = J_i(1) = *$ because the survival time and failure type for individuals with $S_i = 1$ are not used by any of the estimators of $CE(t, j)$.

Step 3   If $S_i(1) = 0$, then $(T_i(1), J_i(1))$ were generated according to the Gompertz models described above. In particular, first $J_i(1)$ was generated from a multinomial distribution with cell probabilities $1 - \exp(\lambda_j/a_j)$ for $j = 1, 2, 3$. Then $T_i(1)$ was set equal to $\tau_0 + U_i$ where U_i was randomly generated from the conditional distribution $\Pr[T_i(1) \leq t | J_i(1) = j] = F(t, j)/\Pr[J_i(1) = j]$ using the inverse probability transformation. Generating $T_i(1)$ in this fashion guarantees that $T_i(1) > \tau_0 = 2$ whenever $S_i(1) = 0$.

Step 4   If $S_i(1) = 0$, $S_i(0)$ was generated as follows. For $\beta_1 = -\infty$,

$S_i(0)=I(T_i(1)<q_1^{(1-\gamma)}, J_i(1)=1)$ where $q_j^{(1-\gamma)}$ is defined in general such that $\Pr[T_i(1) \leq q_j^{(1-\gamma)}, J_i(1)=j|S_i(1)=0]=1-\gamma$. Note for the first scenario (8) holds for $t = 28$ and $j = 1$, guaranteeing the existence of $q_1^{(1-\gamma)}$. For, $\beta_1 = -1, 0, 1$, the value of $F_1^{prot}(28, 1;\beta_1)$

was found by solving (1) and (4) simultaneously, and then $S_i(0) \sim$ Bernoulli(p_{β_1}) where $p_{\beta_1}=(1-\gamma)I(T_i(1)<28, J_i(1)=1)F_1^{prot}(28, 1;\beta_1)/F_1(28, 1)+(1-\gamma)\{1-I(T_i(1)<28, J_i(1)=1)\}\{1-F_1^{prot}(28, 1;\beta_1)\}/\{1-F$

. For $\beta_1 = \infty$, $S_i(0) \sim$ Bernoulli(p_∞) where $p_\infty = (1 - \gamma)\{1 - I(T_i(1) < 28, J_i(1) = 1)\}/\{1 - F_1(28, 1)\}$. Note for the first scenario (6) holds for $t = 28$ and $j = 1$, implying $1 - \gamma < 1 - F_1(28, 1)$ thus ensuring $p_\infty < 1$.

Step 5   If $S_i(0) = 0$, then we let $J_i(0) = J_i(1)$. If $S_i(0) = 0$ and $J_i(0) = 1$, then $T_i(0) = T_i(1)/\varepsilon$, where ε was chosen such that $CE(28, 1) = -0.05$. If $S_i(0) = 0$ and $J_i(0) = 1$, then $T_i(0) = T_i(1)$. If $S_i(0) = 1$, then we set $T_i(0) = J_i(0) = *$.

Step 6 $C_i(0)$ and $C_i(1)$ were generated from exponential distributions with means 29 weeks and 18 weeks respectively.

Step 7 Z_i was randomly assigned such that $n_1 = 852$ and $n_0 = 668$.

Step 8   Given Z_i, we set $Y_i = \min\{T_i(Z_i), C_i(Z_i)\}$, $\Delta_i = I(Y_i = T_i(Z_i))$, $J_i = J_i(Z_i)$, and $S_i = S_i(Z_i)$.

These steps resulted in simulated data sets satisfying Assumptions 2.1 – 2.4 with $CE(28, 1) = -0.05$ for the first scenario. For each data set simulated, $\widehat{CE}(28, 1; \beta_j)$ was computed for $\beta_j = -\infty, -1, 0, 1, -$. Bootstrap percentile and Wald 95% confidence intervals as well as the uncertainty intervals described in Section 3.3 were also computed for each simulated data set, assumed value of β_1, and estimator of $CE(28, 1)$.

Table 1 reports the mean relative bias of $\widehat{CE}(28, j; \beta_j)$ based on 10,000 simulated data sets for both scenarios ($\gamma = 0.9884$, $j = 1$, and $\gamma = 0.75$, $j = 3$) and each model ($\beta_j = -\infty, -1, 0, 1, \infty$). The proposed estimator $\widehat{CE}(28, j; \beta_j)$ is approximately unbiased when β_j is correctly specified; for incorrectly specified β_j the relative bias can be quite large. For example, if β_1 is (incorrectly) assumed to be zero, corresponding to the naive analysis that simply compares infants HIV free and alive at two weeks from each study arm, when in fact $\beta_1 = -\infty$, then the relative bias of $\widehat{CE}(28, 1; \beta_1)$ is 23%. This demonstrates how a naive analysis that ignores the potential for selection bias can yield incorrect inference. This is demonstrated further in the scenario where $\gamma = 0.75$, in which case misspecifying β_3 leads to even greater relative bias.

Table 2 shows the empirical coverage probabilities of 95% pointwise bootstrap confidence intervals based on 500 bootstrap replications per simulated data set. When the correct β_j is specified, the confidence intervals associated with $\widehat{CE}(28, j; \beta_j)$ have approximately 95% coverage. Similar results were found using Wald confidence intervals (results not shown). Because β_j is not identifiable from the observable data, coverage of the uncertainty regions is perhaps of more practical interest. For the 50,000 simulated data sets from the first scenario (i.e., combining across the 10,000 data sets for each of the five values of β_j), the empirical coverage of the 95% pointwise uncertainty regions was 97%. Similarly for the second scenario, the empirical coverage of the uncertainty intervals was 97%.

5. Application to BAN Study

The BAN study was a randomized clinical trial to assess interventions for the prevention of breast milk transmission of HIV in 2369 HIV infected mothers and their infants in Lilongwe, Malawi [3,4]. There were three arms in the BAN study: daily ART for the infant, daily ART for the mother, or control. While the primary analysis of the study considered comparisons of both ART arms to control, we will focus on comparing the infant ART and control arms only. In March 2008 the data and safety monitoring board stopped the control arm due to efficacy but recommended continued enrollment of mother/infant pairs into the two active treatment arms. This led to an imbalance in the final number of infants randomized to the three arms, with 852 infants in the infant ART arm and 668 in the

control arm. In the infant ART arm there were 37 HIV infections and 2 deaths before $\tau_0 = 2$ weeks, while the control arm had 36 HIV infections and 2 infant deaths prior to 2 weeks. Thus $\hat{\gamma} = (630/668)/(813/852) = 0.9884$, as in the first scenario of the simulations in Section 4. Among infants HIV free and alive at 2 weeks, in the infant ART (control) arm 12 (32) became HIV infected, 588 (384) weaned prior to HIV infection, and 5 (6) died prior to HIV infection or weaning by 28 weeks. Figure 1 shows the Aalen-Johansen estimates of the cumulative risk of HIV, death prior to HIV infection or weaning, and cessation of breastfeeding prior to HIV infection for infants who were alive and uninfected at 2 weeks as in a standard analysis, i.e., assuming the no selection model $\beta_j = 0$ holds for all j. Figure 1(a) suggests a difference in the risk of HIV infection between the infant ART arm and the control arm, however direct comparison between the arms is subject to selection bias.

Figure 2 shows the semiparametric sensitivity analysis described in Section 3.2. The plot depicts $\widehat{CE}(28, 1; \beta_1)$ and pointwise 95% Wald confidence intervals for each value of β_1 (using bootstrap variance estimates). Note for the infant ART arm $\hat{F}_1(28, 1) = 0.0141$, suggesting (6) and (8) hold for $t = 28$ and $j = 1$. The estimated ignorance region for $CE(28, 1)$ equals $[-0.056, -0.044]$ and the estimated 95% uncertainty interval equals $[-0.078, -0.025]$. This estimated uncertainty interval was computed using bootstrap variance estimates; using the analytical variance estimates (7) and (9) yielded a slightly wider uncertainty interval of $[-0.084, -0.025]$. In either case, because the uncertainty interval excludes 0, we conclude there is evidence of a causal effect of infant ART on the cumulative incidence of HIV at 28 weeks in the NI stratum. Moreover, without any assumptions about the selection bias mechanism, we are 95% confident daily infant ART lowers the risk of HIV infection at 28 weeks between 3% and 8%.

The veracity of these results relies on several key assumptions. While interference between infants was not likely, SUTVA could have been violated by changes in the infant ART regimen. Per protocol, if an infant on ART had an adverse event due to the study drug (nevirapine), the ART was changed (to lamivudine) and the infant remained in the study. Thus not all infants were on the same treatment for the duration of the study. Therefore, the effect of ART being estimated can be viewed as an average causal effect over all administered ARTs [26]. While this interpretation answers the hypothesis proposed for the BAN study, it does not indicate which particular ART causes the greatest reduction in risk of HIV infection. Assumption 2.1 seems reasonable because treatment was randomized. While mothers were not blinded, they were counseled to breastfeed their infants regardless of randomization assignment and self-reported frequency of exclusive breastfeeding was comparable between study arms [4]. The BAN study principal investigator, Dr. Charles van der Horst, indicated that monotonicity (Assumption 2.2) is reasonable (personal communication). Dr. van der Horst conjectured that an infant could have an adverse reaction to ART leading to increased susceptibility to HIV infection but he felt this was "highly unlikely." Monotonicity is also supported by the estimated risk of HIV infection or death at two weeks being lower in the infant ART arm than in the control arm.

Finally, note that two of the three endpoints in BAN were interval censored. In particular, the HIV infection times of the infants were interval censored, known only to be between the last negative and first positive HIV tests. Similarly, the actual timing of weaning is known only to be visits where the mother reported still breastfeeding and weaning. On the other hand, the time of death was known exactly for all infants. Other analyses of the BAN data have found that formally accounting for interval censoring almost always gives nearly the same result as using the midpoint or right endpoint of the interval. This is not surprising given the visits in the BAN study were fairly close together, typically two to four weeks apart. In settings where the intervals are wider, midpoint or right endpoint imputation may yield misleading results. Instead, a non-parametric estimator of $F_1(t, j)$ that allows for

interval censored event times [27] can be employed in place of the Aalen-Johansen estimator. Inference that formally accounts for interval censoring is challenging however, owing to slow rates of convergence and non-standard limiting distributions of non-parametric estimators (for continuous time models) [28, 29].

6. Discussion

The objective of many MTCT trials is to determine differences in the cumulative risk of breastfeeding transmission of HIV between study arms conditional on infants being HIV free and alive by some time point $\tau_0 > 0$. Here we have presented methods for evaluating the effect of treatment on the cumulative risk of HIV within a principal stratum when death and weaning are competing risks. Large sample non-parametric bounds and a semi-parametric sensitivity analysis model were developed, and the methods were applied to the BAN study, a large, recent MTCT trial. A simulation study was presented demonstrating that the proposed methods perform well in finite samples similar to the BAN study. The simulations also illustrated how analyses that ignore the potential for selection bias by simply conditioning on being HIV free and alive at τ_0 can give misleading results in settings similar to the BAN study.

The analysis of the BAN study indicates infant ART reduces the risk of HIV infection by 28 weeks in infants who would be HIV free and alive at two weeks regardless of treatment assignment. The proposed methods could be applied in other settings as well. For example, BAN investigators (personal communication) were interested in comparing the risk of HIV infection or death by 48 weeks conditional on infants being HIV free and alive at 28 weeks; here $\tau_0 = 28$ weeks is further from time 0 and the potential for selection bias is even greater than the analysis presented in Section 5. Another example is given by the Zambia Exclusive Breastfeeding (ZEB) study, a randomized MTCT study conducted to evaluate whether abrupt weaning at four months compared with continued breastfeeding increases survival of children of HIV-infected mothers [30]. Randomization occurred at one month postpartum in the ZEB study, however Kuhn *et al.* [30] presented a comparison of the randomized groups conditional on infants being HIV free and breastfeeding at four months.

A key assumption of the methods described in this paper is monotonicity, which implies that the treatment is no worse than control for any individual in terms of the intermediate variable *S.* This assumption seems reasonable in the analysis of the BAN study presented in Section 5, but in other settings it may be unrealistic. For example, monotonicity might be considered dubious in an analysis comparing the two active arms of the BAN trial, i.e., maternal ART versus infant ART. In such settings methods that relax or do not require this assumption would be needed. Following Zhang and Rubin [31], nonparametric bounds analogous to those in Section 2 can be derived without assuming monotonicity. Specifically, note that $F_0(t, j) = \varphi F_0^{NI}(t, j) + (1-\varphi) F_0^{harm}(t, j)$, where $\varphi = \Pr[S_i(1) = 0 | S_i(0) = 0]$ and $F_0^{harm}(t, j) = \Pr[T_i(0) \leq t, J_i(0) = j | S_i(0) = 0, S_i(1) = 1]$. If γ and φ were identifiable, then bounds for $F_0^{NI}(t, j)$ can be constructed analogous to (2) and (3) and combined with bounds for $F_1^{NI}(t, j)$ to obtain the following bounds on $CE(t, j)$:

$$CE^{low}(t, j) = \max \left\{ \frac{F_1(t, j) - (1-\gamma)}{\gamma}, 0 \right\} - \min \left\{ \frac{F_0(t, j)}{\varphi}, 1 \right\} \quad (10)$$

and

$$CE^{up}(t,j) = \min \left\{ \frac{F_1(t,j)}{\gamma}, 1 \right\} - \max \left\{ \frac{F_0(t,j)-(1-\varphi)}{\varphi}, 0 \right\}. \quad (11)$$

However, without the monotonicity assumption γ and φ are not identifiable. Let $\pi = \Pr[S_i(0) = 0, S_i(1) = 1]$ and note that

$$\gamma = \Pr[S_i(0)=0|S_i(1)=0] = \frac{\Pr[S_i(0)=0, S_i(1)=0]}{\Pr[S_i(1)=0]} = \frac{\Pr[S_i(0)=0]-\pi}{\Pr[S_i(1)=0]}$$

and

$$\varphi = \Pr[S_i(1)=0|S_i(0)=0] = \frac{\Pr[S_i(0)=0, S_i(1)=0]}{\Pr[S_i(0)=0]} = \frac{\Pr[S_i(0)=0]-\pi}{\Pr[S_i(0)=0]}$$

are identifiable from the observed data for a fixed value of π. Thus, the lower bound of $CE(t, j)$ is found by minimizing (10) over π where $\max\{0, \Pr[S_i(0) = 0] - \Pr[S_i(1) = 0]\} \leq \pi \leq \min\{\Pr[S_i(0) = 0], \Pr[S_i(1) = 1]\}$. Likewise, the upper bound of $CE(t, j)$ is found by maximizing (11) over the same range of π. Sensitivity analysis could be performed by adapting the methods of Shepherd *et al.* [32]. For instance, similar to Assumption 2.4, a selection model for $F_0^{NI}(t, j)$ could be assumed, such as:

Assumption 2.5:

$$\exp(\eta_j) = \frac{F_0^{NI}(t,j)/\{1-F_0^{NI}(t,j)\}}{F_0^{harm}(t,j)/\{1-F_0^{harm}(t,j)\}}. \quad (12)$$

Sensitivity analysis under Assumptions 2.1, 2.3, 2.4, and 2.5 would be performed by varying π over $\max\{0, \Pr[S_i(0) = 0] - \Pr[S_i(1) = 0]\} \leq \pi \leq \min\{\Pr[S_i(0) = 0], \Pr[S_i(1) = 1]\}$ and η_j, β_j each over $(-\infty, \infty)$. The resulting inference will be more precise if the ranges of π, η_j, and β_j can be further restricted based on prior information elicited from subject matter experts.

For the MTCT research motivating this work, interest focused on the principal stratum of infants HIV free and alive at τ_0 under either treatment assignment. The methods developed could also be applied to infants HIV infected and alive at τ_0 under either treatment where T might denote the time until death from various causes. Beyond MTCT trials, the methods developed could be applied in other settings where inference about treatment effects within principal strata is of interest (e.g., truncation-by-death or non-compliance) and the endpoint is a time-to-event outcome subject to competing risks. Further research might entail allowing the cumulative incidence functions to depend on baseline covariates (e.g., as in Jeong and Fine [33]).

Acknowledgments

References

1. World Health Organization. HIV transmission through breastfeeding: a review of available evidence. Geneva: 2007.

2. UNAIDS. AIDS epidemic update. Joint United Nations Programme on HIV/AIDS; Geneva: Dec. 2007

3. van der Horst C, Chasela C, Ahmed Y, Hoffman I, Hosseinipour M, Knight R, Fiscus S, Hudgens M, Kazembe P, Bentley M, et al. Modifications of a large HIV prevention clinical trial to fit changing realities: A case study of the Breastfeeding, Antiretroviral, and Nutrition (BAN) protocol in Lilongwe, Malawi. Contemporary Clinical Trials. 2009; 30:24–33.10.1016/j.cct.2008.09.001 [PubMed: 18805510]

4. Chasela C, Hudgens M, Jamieson D, Kayira D, Hosseinipour M, Kourtis A, Martinson F, Tegha G, Knight R, Ahmed Y, et al. Maternal or infant antiretroviral drugs to reduce HIV-1 transmission. New England Journal of Medicine. 2010; 362(24):2271–2281.10.1056/NEJMoa0911486 [PubMed: 20554982]

5. Kilewo C, Karlsson K, Ngarina M, Massawe A, Lyamuya E, Swai A, Lipyoga R, Mhalu F, Biberfeld G. Mitra Plus Study Team. Prevention of mother-to-child transmission of HIV-1 through breastfeeding by treating mothers with triple antiretroviral therapy in Dar es Salaam, Tanzania: The Mitra Plus Study. Journal of Acquired Immune Deficiency Syndromes. 2009; 52(3):406–416.10.1097/QAI.0b013e3181b323ff [PubMed: 19730269]

6. Kumwenda N, Hoover D, Mofenson L, Thigpen M, Kafulafula G, Li Q, Mipando L, Nkanaunena K, Mebrahtu T, Bulterys M, et al. Extended antiretroviral prophylaxis to reduce breast-milk HIV-1 transmission. New England Journal of Medicine. 2008; 359(2):119–129.10.1056/NEJMoa0801941 [PubMed: 18525035]

7. Bedri A, Gudetta B, Isehak A, Kumbi S, Lulseged S, Mengistu Y, Bhore AV, Bhosale R, Varadhrajan V, Gupte N, et al. Extended-dose nevirapine to 6 weeks of age for infants to prevent HIV transmission via breastfeeding in Ethiopia, India, and Uganda: an analysis of three randomised controlled trials. Lancet. Jul; 2008 372(9635):300–313.10.1016/S0140-6736(08)61114–9 [PubMed: 18657709]

8. Frangakis CE, Rubin DB. Principal stratification in causal inference. Biometrics. 2002; 58:21–29.10.1111/j.0006-341X.2002.00021.x [PubMed: 11890317]

9. Pearl J. Principal stratification – a goal or a tool? The International Journal of Biostatistics. 2011; 7(1):20.10.2202/1557–4679.1322 [PubMed: 21556288]

10. VanderWeele T. Principal stratification–uses and limitations. The International Journal of Biostatistics. 2011; 7(1):1–14.10.2202/1557–4679.1329

11. Alioum A, Dabis F, Dequae-Merchadou L, Haverkamp G, Hudgens M, Hughes J, Karon J, Leroy V, Newell M, Richardson B, et al. Estimating the efficacy of interventions to prevent mother-to-child transmission of HIV in breast-feeding populations: development of a consensus methodology. Statistics in Medicine. 2001; 20:3539–3556.10.1002/sim.1076 [PubMed: 11746336]

12. Tsiatis, AA. Competing risks. In: Armitage, P.; Colton, T., editors. Encyclopedia of Biostatistics. Wiley: New York; 1998. p. 824-834.

13. Andersen P, Abildstrom S, Rosthøj S. Competing risks as a multi-state model. Statistical Methods in Medical Research. 2002; 11(2):203–215.10.1191/0962280202sm281ra [PubMed: 12040697]

14. Hudgens MG, Halloran ME. Causal vaccine effects on binary post-infection outcomes. Journal of the American Statistical Association. 2006; 101:51–64.10.1198/016214505000000970 [PubMed: 19096723]

15. Gilbert PB, Bosch R, Hudgens MG. Sensitivity analysis for the assessment of causal vaccine effects on viral load in HIV vaccine trials. Biometrics. 2003; 59:531–541.10.1111/1541-0420.00063 [PubMed: 14601754]

16. Hayden D, Pauler DK, Schoenfeld D. An estimator for treatment comparisons among survivors in randomized trials. Biometrics. 2005; 61:305–310.10.1111/j.0006-341X.2005.030227.x [PubMed: 15737107]

17. Shepherd BE, Gilbert PB, Lumley T. Sensitivity analyses comparing time-to-event outcomes existing only in a subset selected postrandomization. Journal of the American Statistical Association. 2007; 102(478):573–582.10.1198/016214507000000130 [PubMed: 19122791]

18. Rubin DB. Discussion of "Randomization analysis of experimental data in the Fisher randomization test" by D. Basu. Journal of the American Statistical Association. 1980; 75:591–593.10.2307/2287653

19. Robins, JM.; Rotnitzky, A.; Scharfstein, DO. Sensitivity analysis for selection bias and unmeasured confounding in missing data and causal inference models. In: Halloran, M.; Berry, D., editors. Statistics in Epidemiology, Environment and Clinical Trials. Springer-Verlag; 2000.

20. Scharfstein DO, Halloran ME, Chu H, Daniels MJ. On estimation of vaccine efficacy using validation samples with selection bias. Biostatistics. 2006; 7:615–629.10.1093/biostatistics/kxj031. [PubMed: 16556610]

21. Shepherd BE, Gilbert PB, Mehrotra DV. Eliciting a counterfactual sensitivity parameter. The American Statistician. 2007; 61(1):56–63.10.1198/000313007X163213

22. Aalen O, Johansen S. An empirical transition matrix for non-homogeneous markov chains based on censored observations. Scandinavian Journal of Statistics. 1978; 5(3):141–150.

23. Vansteelandt S, Goetghebeur E, Kenward M, Molenberghs G. Ignorance and uncertainty regions as inferential tools in a sensitivity analysis. Statistica Sinica. 2006; 16(3):953–979.

24. Aalen, O.; Borgan, O.; Gjessing, H. Suvival and Event History Analysis: A Process Point of View. Springer; New York: 2008.

25. Jeong J, Fine J. Direct parametric inference for the cumulative incidence function. Journal of the Royal Statistical Society, Series C-Applied Statistics. 2006; 55(2):187–200.10.1111/j.1467-9876.2006.00532.x

26. VanderWeele, TJ.; Hernan, MA. Causal inference under multiple versions of treatment; COBRA Preprint Series. 2011. p. 77http://biostats.bepress.com/cobra/ps/art77/

27. Hudgens MG, Satten GA, Longini IM. Nonparametric maximum likelihood estimation for competing risks survival data subject to interval censoring and truncation. Biometrics. 2001; 57:74–80.10.1111/j.0006-341X.2001.00074.x [PubMed: 11252621]

28. Groeneboom P, Maathuis MH, Wellner JA. Current status data with competing risks: Consistency and rates of convergence of the MLE. Annals of Statistics. 2008; 36:1031–1063.10.1214/009053607000000974

29. Groeneboom P, Maathuis MH, Wellner JA. Current status data with competing risks: Limiting distribution of the MLE. Annals of Statistics. 2008; 36:1064–1089.10.1214/009053607000000983 [PubMed: 19888358]

30. Kuhn L, Aldrovandi G, Sinkala M, Kankasa C, Semrau K, Mwiya M, Kasonde P, Scott N, Vwalika C, Walter J, et al. Effects of early, abrupt weaning on HIV-free survival of children in Zambia. N Engl J Med. 2008; 359:130–141.10.1056/NEJMoa073788 [PubMed: 18525036]

31. Zhang JL, Rubin DB. Estimation of causal effects via principal stratification when some outcomes are truncated by "death". Journal of Educational and Behavioral Statistics. 2003; 28:353–368.10.3102/10769986028004353

32. Shepherd B, Gilbert P, Dupont C. Sensitivity analyses comparing time-to-event outcomes only existing in a subset selected postrandomization and relaxing monotonicity. Biometrics. 2011; 67(3):1100–1110.10.1111/j.1541-0420.2010.01508.x [PubMed: 21114663]

33. Jeong J, Fine J. Parametric regression on cumulative incidence function. Biostatistics. 2007; 8(2):184–196.10.1093/biostatistics/kxj040 [PubMed: 16636138]

34. Agresti, A. Categorical Data Analysis. 2. John Wiley and Sons; New York: 2002.

7. Appendix

Asymptotic Variances of $\widehat{F}_1^{NI,up}(t, j)$ and $\widehat{F}_1^{NI,low}(t, j)$

To derive the asymptotic variances of $\widehat{F}_1^{NI,up}(t, j)$ and $\widehat{F}_1^{NI,low}(t, j)$, we first derive the large sample variance of $\hat{\gamma}$. Under monotonicity, it is straightforward to show

$\widehat{\gamma}-(N_0/n_0)/(N_1/n_1) \xrightarrow{p} 0$, implying $\hat{\gamma}$ and $(N_0/n_0)/(N_1/n_1)$ have the same limiting distribution; therefore for the derivation below we can assume $\hat{\gamma} = (N_0/n_0)/(N_1/n_1)$. For z, $s = 0, 1$, define $p_{zs} = \Sigma\, I[Z_i = z, S_i = s]/n$ and $\pi_{zs} = \Pr[Z_i = z, S_i = s]$, and let $\mathbf{p} = (p_{00}, p_{01}, p_{10}, p_{11})'$ and $\boldsymbol{\pi} = (\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11})'$. Define the function g as $g(\boldsymbol{\pi}) = \pi_{00}(\pi_{10} + \pi_{11})/\{\pi_{10}(\pi_{00} + \pi_{01})g$ and note that $g(\mathbf{p}) = \hat{\gamma}$ and $g(\boldsymbol{\pi}) = \gamma$. Then by the multivariate central limit theorem and the delta method (e.g., see Agresti 2002 [34], page 580),

$\sqrt{n}(\widehat{\gamma}-\gamma) = \sqrt{n}\{g(\mathbf{p})-g(\pi)\} \xrightarrow{\mathbb{D}} N(0, \sigma_\gamma^2)$ where $\sigma_\gamma^2 = \sum_{z,s=0}^{1}\pi_{zs}(\nabla g_{zs})^2 - (\sum_{z,s=0}^{1}\pi_{zs}\nabla g_{zs})^2$ and $\nabla g_{zs} = g(\boldsymbol{\pi})/\ \pi_{zs}$. It follows from straightforward algebra that

$\sigma_\gamma^2 = \gamma^2[\pi_{01}/\{\pi_{00}(\pi_{00}+\pi_{01})\}+\pi_{11}/\{\pi_{10}(\pi_{10}+\pi_{11})\}]$ for which a consistent estimator is $\widehat{\sigma}_\gamma^2 = \widehat{\gamma}^2 n(1/N_0-1/n_0+1/N_1-1/n_1)$.

For fixed t and j, let $\boldsymbol{\theta}_{tj} = (F_1(t, j); \gamma)'$ and $\hat{\boldsymbol{\theta}}_{tj} = (\hat{F}_1(t, j); \hat{\gamma})'$. Under the conditions stated in Section 3.1 of the main text, in particular assuming equation (6), it is straightforward to show $\widehat{F}_1^{NI,up}(t, j)-\widehat{F}_1(t, j)/\widehat{\gamma} \xrightarrow{p} 0$, implying $\widehat{F}_1^{NI,up}(t, j)$ and $\hat{F}_1(t, j)/\hat{\gamma}$ have the same limiting distribution. Therefore we can assume $\widehat{F}_1^{NI,up}(t, j) = \widehat{F}_1(t, j)/\widehat{\gamma}$ and, analogously, by equation (8) of the main text we can assume $\widehat{F}_1^{NI,low}(t, j) = \{\widehat{F}_1(t, j)-(1-\widehat{\gamma})\}/\widehat{\gamma}$. Define the vector of functions $\boldsymbol{h}(x, y) = (x/y, \{x - (1 - y)\}/y)'$ such that $\boldsymbol{h}(\widehat{\theta}_{tj}) = (\widehat{F}_1^{NI,up}(t, j), \widehat{F}_1^{NI,low}(t, j))'$. Because $\hat{F}_1(t, j)$ and $\hat{\gamma}$ are consistent and asymptotically normal, by the delta method $\sqrt{n}\{\boldsymbol{h}(\widehat{\theta}_{tj})-\boldsymbol{h}(\theta_{tj})\} \xrightarrow{\mathbb{D}} N(0, \nabla\boldsymbol{h}(\theta_{tj})\sum_{tj}\nabla\boldsymbol{h}(\theta_{tj})')$ where

$$\nabla\boldsymbol{h}(x, y) = \left[\begin{array}{cc} 1/y & -x/y^2 \\ 1/y & (1-x)/y^2 \end{array}\right], \qquad \sum_{tj} = \left[\begin{array}{cc} \sigma_{tj}^2 & 0 \\ 0 & \sigma_\gamma^2 \end{array}\right],$$

and σ_{tj}^2 is the asymptotic variance of $\sqrt{n}\{\widehat{F}_1(t, j)-F_1(t, j)\}$ such that in large samples $\mathrm{var}\{\widehat{F}_1(t, j)\} = \sigma_{tj}^2/n$. It follows that $\widehat{F}_1^{NI,up}(t, j)$ and $\widehat{F}_1^{NI,low}(t, j)$ are asymptotically normal with variances

$$\mathrm{var}\{\widehat{F}_1^{NI,up}(t, j)\} = \frac{\mathrm{var}\{\widehat{F}_1(t, j)\}}{\gamma^2} + \frac{F_1(t, j)^2\sigma_\gamma^2}{n\gamma^4}, \quad (13)$$

and

$$\mathrm{var}\{\widehat{F}_1^{NI,low}(t, j)\} = \frac{\mathrm{var}\{\widehat{F}_1(t, j)\}}{\gamma^2} + \frac{\{1-F_1(t, j)\}^2\sigma_\gamma^2}{n\gamma^4}. \quad (14)$$

Replacing $\mathrm{var}\{\hat{F}_1(t, j)\}$, γ, $F_1(t, j)$, and σ_γ^2 in (A.1) and (A.2) with $\widehat{\mathrm{var}}\{\widehat{F}_1(t, j)\}$, $\hat{\gamma}$, $\hat{F}_1(t, j)$, and $\widehat{\sigma}_\gamma^2$ yields equations (7) and (9) from the main text.
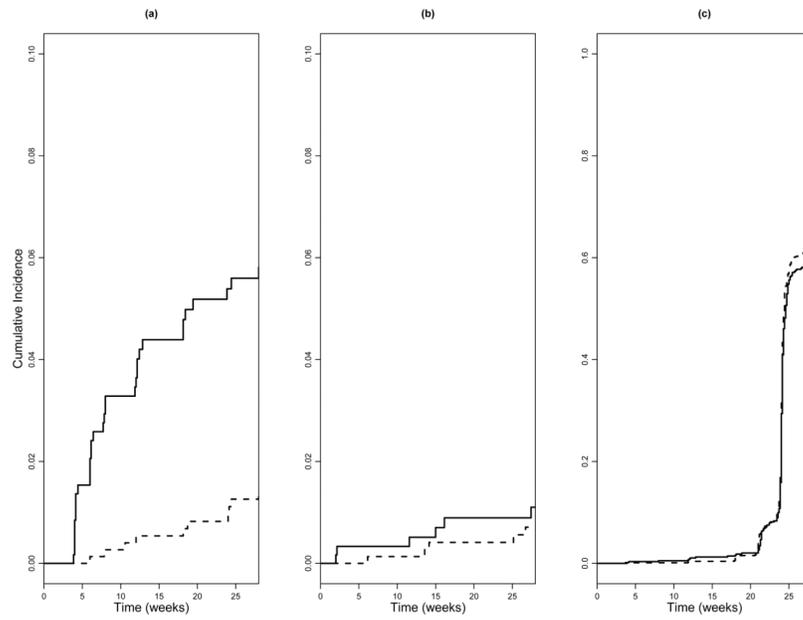
Figure 1.
Estimated cumulative incidence functions, $\hat{F}_z(23, j)$, for the three events from the BAN study: (a) HIV infection, (b) HIV-free death prior to weaning, and (c) cessation of breastfeeding prior to HIV infection. For eachpanel, $Z_i = 0$ (control) is represented by the solid line (—) and $Z_i = 1$ (infant ART) is represented by the dashed line (– – –).
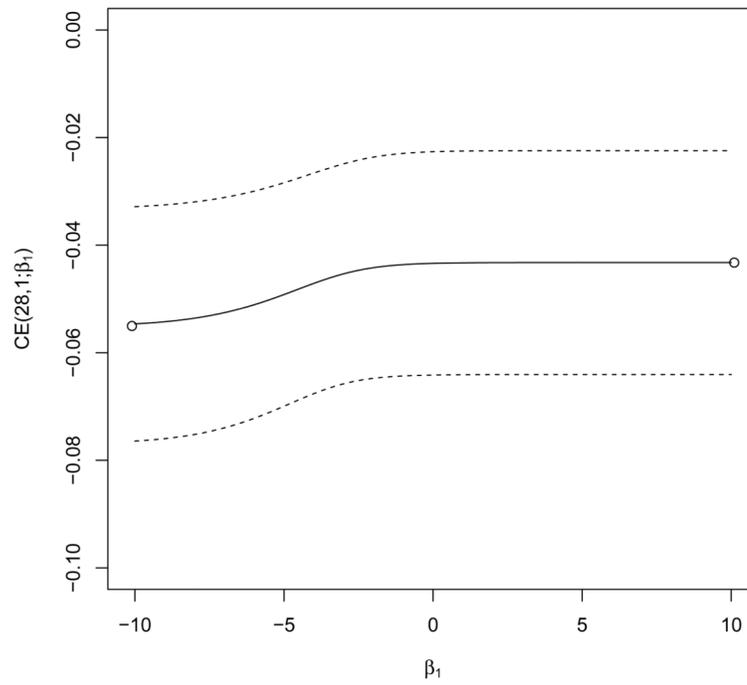
Figure 2.
Sensitivity analysis of the effect of infant ART on the cumulative incidence of HIV at 28 weeks for the BAN study. The solid line — denotes $\widehat{CE}(28, 1; \beta_1)$ and the dotted lines $\cdots$ denote pointwise 95% confidence intervals. The estimated non-parametric bounds corresponding to $\beta_1 = -\infty$ and $\beta_1 = \infty$ are given by $\bigcirc$.

Table 1

Empirical relative bias of estimates of *CE*(28, *j*) from simulation study described in Section 4 for both scenarios. Bold entries correspond to estimates where the assumed β_j was correct. Relative bias of $\widehat{CE}(28, j;\beta_j)$ defined as $\{\widehat{CE}(28, j;\beta_j)-CE(28, j)\}/CE(28, j)$.

True parameters			Assumed β_j				
γ	*CE*	β_j	$-\infty$	-1	0	1	∞
0.9884	−0.05	$-\infty$	**0.02**	0.22	0.23	0.24	0.24
		−1	−0.19	**0.01**	0.01	0.02	0.02
		0	−0.21	−0.01	**0.00**	0.00	0.01
		1	−0.20	−0.01	0.00	**0.00**	0.00
		∞	−0.21	−0.01	−0.00	0.00	**0.00**
0.75	0.05	$-\infty$	**−0.01**	−1.10	−2.01	−3.13	−6.69
		−1	1.02	**−0.01**	−0.88	−1.96	−5.32
		0	2.00	0.91	**−0.01**	−1.12	−4.67
		1	2.66	1.77	0.98	**0.00**	−2.93
		∞	6.50	5.44	4.55	3.45	**0.00**

Table 2

Empirical coverage of pointwise 95% bootstrap percentile confidence intervals of $CE(28, j)$ from simulation study described in Section 4 for both scenarios. Bold entries correspond to estimates where the assumed β_j was correct.

True parameters			Assumed β_j				
γ	CE	β_j	$-\infty$	-1	0	1	∞
0.9884	−0.05	$-\infty$	**0.95**	0.81	0.80	0.80	0.79
		−1	0.90	**0.95**	0.94	0.94	0.94
		0	0.88	0.94	**0.94**	0.94	0.94
		1	0.89	0.95	0.95	**0.95**	0.95
		∞	0.89	0.95	0.95	0.95	**0.95**
0.75	0.05	$-\infty$	**0.95**	0.55	0.06	0.00	0.00
		−1	0.66	**0.94**	0.65	0.06	0.00
		0	0.14	0.66	**0.94**	0.47	0.00
		1	0.01	0.12	0.55	**0.94**	0.01
		∞	0.00	0.00	0.00	0.00	**0.94**