Published in final edited form as:

Stat Med. 2008 September 20; 27(21): 4313–4327. doi:10.1002/sim.3293.

Evaluating competing adverse and beneficial outcomes using a mixture model

Bryan Lau^{1,2,*}, Stephen R. Cole², Richard D. Moore^{1,2}, and Stephen J. Gange²

¹Department of Medicine, Johns Hopkins School of Medicine, Baltimore, Maryland, U.S.A.

²Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, U.S.A.

SUMMARY

A competing risk framework occurs when individuals have the potential to experience only one of several mutually exclusive outcomes. Standard survival methods often overestimate the cumulative incidence of events when competing events are censored. Mixture distributions have been previously applied to the competing risk framework to obtain inferences regarding the subdistribution of an event of interest. Often the competing event is treated as a nuisance, but it may be of interest to compare adverse events against the beneficial outcome when dealing with an intervention. In this paper, methods for using a mixture model to estimate an adverse-benefit ratio curve (ratio of the cumulative incidence curves for the two competing events) and the ratio of the subhazards for the two competing events are presented. Both parametric and semi-parametric approaches are described with some remarks for extending the model to include uncertainty in the event type that occurred, left-truncation in order to allow for time-dependent analyses, and uncertainty in the timing of the event resulting in interval censoring. The methods are illustrated with data from a HIV clinical cohort examining whether individuals initiating effective antiretroviral therapy have a greater risk of antiretroviral discontinuation or switching compared to HIV RNA suppression.

Keywords

survival analysis; competing risks; mixture model; cumulative incidence function; subhazard

1. INTRODUCTION

Competing risks occur in survival analyses when the occurrence of one event precludes the occurrence of the remaining set of possible events. The canonical example is cause of death, where a death due to cancer (for example) removes a subject from risk of death due to cardiovascular disease (for a review of competing risks see references [1,2,3]). In competing risk settings, it is well known that the complement of the standard Kaplan Meier survival curve (with competing events treated as censored observations) will overestimate the cause specific cumulative incidence whenever the competing risks are not independent [4,5,6,7]. Generally, the independence of competing risks is not identifiable because only one of the competing risks is observed [8].

In the presence of competing risks, estimation of the cumulative incidence of a particular event of interest often uses Prentice's cumulative incidence estimator [9]. Methods have been developed to compare the hazard for the subdistribution of interest across categorical exposure variables [10,11]. Alternatively, mixture models have been used to compare the subdistribution hazard across categorical or continuous exposure variables [11,12].

Competing risk methods often treat one type of outcome as the event of interest, and the other as a nuisance to be removed. However, therapies generally have the possibility of severe adverse events in addition to the intended beneficial outcome. Among HIV infected men and women in developed countries, combinations of antiretroviral drugs are typically started when the number of measured T cells expressing the CD4 molecule are between 250 and 500 cells/mm³. Based on HIV treatment guidelines [13], individuals who initiate anti HIV medications and are followed will be observed (in the short term) to either (a) suppress the concentration of HIV RNA circulating in the blood to below detectable limits, or (b) alter their anti HIV medication plan. To evaluate the short term benefit-risk profile of these medications and provide an estimate of the expected response to therapy, it might be of interest to consider the time to the first occurrence of these events. The suppression of HIV viral load is a desirable effect of anti HIV medications, while a change in anti HIV medication plan is typically undertaken because of adverse events, such as toxicity or reduced efficacy perhaps due to HIV that is resistant to the specific antiretroviral regimen. This example motivates the use of methods that can estimate the occurrence of the adverse and beneficial pathways acting simultaneously but only having one outcome realized.

In section 2 we extend the mixture model approach of Larson and Dinse [12] to estimate an adverse benefit ratio (ABR) and an adverse benefit subhazard ratio (ABSR). These measures evaluate the balance of the adverse event relative to the beneficial event with 1.0 being the equilibrium point, > 1 and < 1 the balance toward adverse and beneficial events, respectively. In section 3, we apply our approach to contemporary data from the Johns Hopkins HIV Clinic Cohort. We close in section 4 with a discussion of the strengths and limitations of the proposed approach.

2. MIXTURE MODEL

Let the survival data be represented by a bivariate random variable form rather than by the latent failure time notation [2]. Thus (T_i, J_i) for individual $i=1, \dots, N$ with $p \geq 2$ types of failures, where T_i is a non-negative random variable representing time to first event and J_i takes a value j from the set $1, \dots, p$ to indicate the type of failure. As with most survival data, this bivariate random variable will be incomplete should the observation time end prior to any of the failure types being observed. Thus let $J_i = 0$ when no failure of any type is observed over the period of study and T_i is the individual's contributed time at risk. We assume that the censoring mechanism is non informative.

From Larson and Dinse [12], the competing risk for p failure types may be represented by a mixture model which consists of a mixture probability and a conditional failure time distribution. **Thus the cumulative incidence function $F_j^*(t) = \Pr(T_i \leq t, J_i = j)$ for the j th type of failure is decomposed as**

$$\begin{aligned} F_j^*(t) &= \Pr(T_i \leq t | J_i = j) \Pr(J_i = j) \\ &= F_j(t) \Pr(J_i = j) \end{aligned}$$

and therefore the corresponding subdensity function is defined as

$$\begin{aligned} f_j^*(t) &= \frac{dPr(T_i \leq t | J_i = j)}{dt} Pr(J_i = j) \\ &= f_j(t) Pr(J_i = j). \end{aligned}$$

The distribution functions $F_j(t)$ and $f_j(t)$ are modeled by a distribution of appropriate choice, and $Pr(J_i = j)$ is modeled using a binary data model (e.g. logistic) when there are only 2 failure types or as a categorical data model (e.g. multinomial logistic model) for $p > 2$. When the scientific question of interest is time to an adverse event and a beneficial event, an adverse-to-benefit ratio may be determined. For simplicity we develop the model for two failure types: $j = 1$ indicate a beneficial outcome and $j = 2$ indicate an adverse outcome. The adverse-benefit ratio (ABR) can then be defined as

$$ABR(t) = \frac{F_{j=2}^*(t)}{F_{j=1}^*(t)} = \frac{Pr(T_i \leq t | J_i = 2) Pr(J_i = 2)}{Pr(T_i \leq t | J_i = 1) Pr(J_i = 1)} \quad (1)$$

The interpretation of the ABR is the ratio of the cumulative incidence of the adverse event relative to that of the beneficial event such that values >1 (<1) indicates a higher risk for the adverse (beneficial) event. This adverse-benefit ratio is a function over time and as $T \rightarrow \infty$

$$\text{then } ABR(t) \rightarrow \frac{Pr(J_i = 2)}{Pr(J_i = 1)}.$$

An alternative measure of interest may involve estimating which of the events is more likely to occur given that individuals have survived to time t without experiencing either the adverse or beneficial outcome. To evaluate this question, we can estimate the ratio of the subhazard functions. The subhazard function $\rho_j(t)$ for event $j = 1, 2$ is defined similarly to the hazard in the non-competing risk framework except the subdensity (or cause-specific probability density) replaces the standard density:

$$\rho_j(t) = \frac{f_j^*(t)}{S(t)}, \quad j=1, 2$$

where $S(t)$ is the overall survival function [1,2]. Therefore, the subhazard function is the instantaneous rate of the j th event. Using these measures, a ratio of the subhazard functions of the competing risks (adverse-benefit subhazard ratio [ABSR]) may be estimated which simplifies to a ratio of the subdensity functions

$$ABSR(t) = \frac{\rho_2(t)}{\rho_1(t)} = \frac{f_2^*(t)/S(t)}{f_1^*(t)/S(t)} = \frac{f_2^*(t)}{f_1^*(t)}. \quad (2)$$

An $ABSR(t) > 1$ indicates the adverse event is at greater risk of occurring at time t , while a value below 1 indicates the beneficial event is more likely to occur.

2.1. Maximum Likelihood Estimation

The i th individual's contribution to the likelihood function for the mixture model for the two outcome competing risk situation when the events are measured from a common origin is:

$$L_i = [\pi_i f_1(t_i)]^{\delta_i} \times [(1 - \pi_i) f_2(t_i)]^{1 - \delta_i} \times [\pi_i S_1(t_i) + (1 - \pi_i) S_2(t_i)]^{(1 - \delta_i - \theta_i)} \quad (3)$$

where t_i is the i th individual's observed time for the two failures indicated by the indicator functions

$$\delta_i = \begin{cases} 0 & \text{if } J_i \neq 1 \\ 1 & \text{if } J_i = 1 \end{cases} \quad \text{and} \quad \theta_i = \begin{cases} 0 & \text{if } J_i \neq 2 \\ 1 & \text{if } J_i = 2 \end{cases}$$

such that individuals are right censored if $\delta_i = \theta_i = 0$ and $\pi_i = \Pr(J_i = 1|X_i)$ is the mixture probability, $f_1(t_i)$ and $f_2(t_i)$ are the probability density function for the parametric distribution being utilized for the beneficial and adverse event, respectively, and similarly $S_1(t_i)$ and $S_2(t_i)$ are the survivor function (i.e. $1 - F_j(t)$) for the beneficial and adverse events, respectively.

To estimate $f_1(t_i)$ and $f_2(t_i)$ and their respective survivor functions, we propose using a flexible parametric distribution. In our application, the conditional failure time distributions for both the adverse and beneficial events were modeled with the three parameter generalized gamma distributions $GG(\beta, \sigma, \lambda)$ with location (β), scale (σ), and shape (λ) parameters [14]. The probability density functions $f_1(t_i)$ and $f_2(t_i)$ are given by:

$$f_j(t_i) = \frac{|\lambda_j|}{\sigma_j t_i \Gamma(\lambda_j^{-2})} \left[\lambda_j^{-2} (e^{-\beta_j t_i})^{\lambda_j / \sigma_j} \right]^{\lambda_j^{-2}} \exp \left[-\lambda_j^{-2} (e^{-\beta_j t_i})^{\lambda_j / \sigma_j} \right]$$

and the survival function is:

$$S_j(t_i) = \begin{cases} 1 - \Gamma \left[\lambda_j^{-2} (e^{-\beta_j t_i})^{\lambda_j / \sigma_j}; \lambda_j^{-2} \right] & \text{if } \lambda_j > 0 \\ \Gamma \left[\lambda_j^{-2} (e^{-\beta_j t_i})^{\lambda_j / \sigma_j}; \lambda_j^{-2} \right] & \text{if } \lambda_j < 0 \end{cases}$$

where $\Gamma[\cdot; \cdot]$ is the cumulative distribution function for the two parameter gamma distribution with mean and variance equal to $\gamma > 0$, that is $\Gamma(t, \gamma) = \int_0^t y^{\gamma-1} e^{-y} dy / \Gamma(\gamma)$ [14]. Thus, the number of parameters to be estimated for our application is >7 (one 3-parameter generalized gamma for each event, plus at least one parameter for π_i , and additional regression parameters for covariates). The generalized gamma was chosen as it is extremely flexible, in which the hazard can take various shapes including declining, increasing, bathtub shape (decreasing then increasing), or arc shaped (increasing then decreasing) hazards. In fact, the generalized gamma distribution can accommodate a whole range of forms including many of the most commonly used distributions, such as exponential (generalized gamma converges to exponential when $\lambda = \sigma = 1$), gamma ($\lambda = \sigma$), log-normal (limiting case when $\lambda = 0$), and Weibull ($\lambda = 1$) distributions as special cases [14].

Extension to inclusion of covariates—The methods above may be extended to allow for regression with covariates. Let X_i denote a column vector of M time-stationary covariates (an expansion to time-dependent covariates follows in section 2.1.1). The parameters for the mixture probability, probability density function, and survivor function can be modified to allow for a linear combination of independent variables including a constant term for an intercept. For example, in defining the likelihood function (3) contribution, the π_i **parameter was defined as** $1/1 + e(\beta^T X_i)$ **where** β^T is the vector of parameters to be estimated. Similarly, for implementing a generalized gamma the location (β), scale (σ), and shape (λ) parameters may be a linear combination of the covariates. That is the $GG(\beta, \sigma, \lambda)$ may be extended to $GG(\beta^T X_a, \sigma^T X_b, \lambda^T X_c)$, where X_a , X_b , and X_c are

covariate subsets of the column vector X_i . If the scale and shape parameters are not allowed to vary by covariates, then these generalized gamma distributions are parameterized similarly to a conventional regression model with $GG(\alpha + \beta^T X, \sigma, \lambda)$. This assumes proportional times and is an accelerated failure time model.

Estimation of the parameters for the mixture model can be performed in SAS using the NLMIXED procedure in which the log-likelihood can be defined from equation (3). Furthermore, one can compute general functions of the estimated parameters (possibly depending on time) using the NLMIXED estimate statement. The standard errors for the functions are determined through the delta method with the variance-covariance matrix derived from the inverse of the Hessian matrix estimated using a quasi-Newton Raphson algorithm. The time-varying ABR and ABSR presented in equation (1) and (2), respectively, can be estimated as functions of time using the estimated model parameters.

2.1.1. Extension to uncertainty of event type or timing—It is assumed for the above likelihood function that the event type is known. However, it is possible that the time of the event is known but there is uncertainty about which event occurred. In assessing causes of death, for example, there may be unresolvable uncertainty in the actual cause or incomplete ascertainment of this information. Instead of eliminating or censoring those data where the outcome is uncertain, the above likelihood may be modified to incorporate the uncertainty of event types by summing the $f_j(t)$ for all events that could have occurred, such that the likelihood in (3) would be

$$L = [\pi_i f_1(t_i)]^{\delta_i} \times [(1 - \pi_i) f_2(t_i)]^{\theta_i} \times [\pi_i f_1(t_i) + (1 - \pi_i) f_2(t_i)]^{v_i} \times [\pi_i S_1(t_i) + (1 - \pi_i) S_2(t_i)]^{(1 - \delta_i - \theta_i - v_i)}$$

where v is an indicator for whether $J = 1$ or $J = 2$ and the indicator function δ and θ are slightly modified such that both equal 0 when $v = 1$ [1]. This may be done since competing events are mutually exclusive in the competing risk framework - that is:

$$\begin{aligned} [Pr(T_i=t, J=1) \cup Pr(T_i=t, J=2)] &= Pr(T_i=t, J=1) + Pr(T_i=t, J=2) - \\ &\quad (Pr(T_i=t, J=1) \cap Pr(T_i=t, J=2)) \\ &= Pr(T_i=t, J=1) + Pr(T_i=t, J=2). \end{aligned}$$

Another important consideration is incorporating left-truncation and interval censoring. Left-truncation methods are useful to avoid bias due to late entry into study follow-up or for incorporating time-varying covariates into survival models [16]. Furthermore, interval censoring occurs because of uncertainties in the timing of events [17]. To incorporate the former, the above likelihood function in (3) can be expanded to allow for left-truncation:

$$L = \left[\pi_i \frac{f_1(t_i)}{S(w_i)} \right]^{\delta_i} \times \left[(1 - \pi_i) \frac{f_2(t_i)}{S(w_i)} \right]^{\theta_i} \times \left[\pi_i \frac{S_1(t_i)}{S(w_i)} + (1 - \pi_i) \frac{S_2(t_i)}{S(w_i)} \right]^{(1 - \delta_i - \theta_i)}$$

where $w_i > 0$ and $w_i = 0$ for the i th individuals with and without truncated times, respectively. Therefore such individuals are observed from w_i to t_i . This likelihood can be further extended to include interval censoring:

$$\begin{aligned} L &= \left[\pi_i \frac{f_1(t_i)}{S(w_i)} \right]^{\delta_i(1 - \zeta_i)} \times \left[\pi_i \frac{S_1(t_{2i}) - S_1(t_{1i})}{S(w_i)} \right]^{\delta_i \zeta_i} \times \left[(1 - \pi_i) \frac{f_2(t_i)}{S(w_i)} \right]^{\theta_i(1 - \eta_i)} \\ &\quad \times \left[(1 - \pi_i) \frac{S_2(t_{2i}) - S_2(t_{1i})}{S(w_i)} \right]^{\theta_i \eta_i} \times \left[\pi_i \frac{S_1(t_i)}{S(w_i)} + (1 - \pi_i) \frac{S_2(t_i)}{S(w_i)} \right]^{(1 - \delta_i - \theta_i)} \end{aligned}$$

where $t_{1i} < t_{2i}$ and t_{1i} and t_{2i} indicate the two boundary times of the interval in which the event $J_i = 1$ or $J_i = 2$ occurs, ζ_i is an indicator function for interval censoring for the $J_i = 1$ event, and $\eta_i = 1$ is also an indicator function for identifying individuals who are interval censored for the $J = 2$ event within the t_{1i} to t_{2i} interval.

3. APPLICATION TO EVALUATING TREATMENTS OF HIV

With the introduction of highly active antiretroviral therapy (HAART) around 1996, HIV disease progression to AIDS and death has slowed dramatically [18,19,20,21,22,23]. According to the treatment guidelines published by the Department of Health and Human Services [13], one of the primary goals of antiretroviral therapy is to suppress HIV RNA to undetectable levels. However, HAART regimens have been difficult for patients to adhere to consistently. This has been due to such factors as the number of pills within various regimens [24,25], complex dosing schedules [25], and the possibility of adverse events such as metabolic abnormalities [26,27], liver toxicity [28], renal toxicity [29], and possible increase risk of cardiovascular disease [30]. HAART regimens have been improved since their introduction due to new drugs with lower pill burdens and adverse events and increased clinical experience in treating HIV patients with antiretroviral drugs; however, these issues remain as important clinical concerns.

The Johns Hopkins HIV Clinical Cohort (JHHCC) is a longitudinal, dynamic, clinical cohort of patients receiving care through the Johns Hopkins AIDS Service that provides care for a large proportion of HIV-infected patients in the Baltimore metropolitan area [31]. The JHHCC has followed patients since 1989 and almost 6,000 patients have been enrolled. Recruitment for enrollment into the cohort occurs at the first visit at the HIV clinic and approximately 85% of patients receive all of their health care within the Johns Hopkins Health System. All patients give informed consent and the JHHCC is conducted in accordance with the ethical standards of the Johns Hopkins Institutional Review Board and with the Helsinki Declaration of 1975.

Utilizing the JHHCC, we investigated the cumulative incidence and the ABR for switching or discontinuation of initial HAART regimens compared to HIV RNA suppression according to therapy guideline periods. HAART was defined as three or more antiretroviral drugs taken together. This definition allows for capturing what was initially considered HAART upon introduction of effective treatment through multiple (≥ 3) antiretroviral drugs in 1996 and more current definitions of HAART (e.g. current definitions are 2 nucleoside/nucleotide reverse transcriptase inhibitors and either a non-nucleoside reverse transcriptase inhibitor or a protease inhibitor). Thus, this definition captures temporal changes in HAART regimens.

All individuals within the JHHCC who initiated HAART were eligible for this analysis. Individuals were observed from initiation of HAART to the first occurrence of (i) HIV RNA suppression to less than 400 copies/ml or (ii) discontinuation or switching of antiretroviral drugs within the HAART regimen. Therapy discontinuation and switching in the HIV setting is usually synonymous with treatment failure. Individuals discontinuing or switching antiretroviral drugs within the HAART regimen may do so because of non-tolerance or an adverse event, non-adherence to the drugs, or possibly because of developing HIV resistance mutations to drugs within the regimen. It has been shown that the majority of individuals who discontinue or switch regimens do so because of either adverse events or reduced efficacy of the specific drug regimen [32,33,34,35,36,37]. Furthermore, the failure of the first HAART regimen is a significant adverse event as the success of the first regimen has been associated with the greatest probability of long-term virologic suppression [37,38]. The therapy guideline periods were defined as (a) prior to 1998 for early HAART regimens, (b)

1999 – 2000 which defines the period for the first set of established guidelines by the Department of Health and Human Services, (c) 2001 – 2002 for the introduction of ritonavir-boosted protease inhibitors as apart of the guidelines, and (d) 2003 – 2005 which corresponds to the current guideline period. Indicator functions for the guideline periods and a parameter included for the $\log_{10}$ HIV RNA (centered at $\log_{10}(70000 \text{ cps/ml})$) were used in the mixture component of the log-likelihood function. Additionally, these covariates were included in the estimation of the location parameters for both the generalized gamma distributions for HIV RNA suppression as well as for the HAART switch or discontinuation. Covariates were not included in the scale and shape parameters for either generalized gamma distribution.

Among 1601 individuals initiating HAART, there were 825 (51.5%) individuals who suppressed HIV RNA to below 400 copies/ml prior to a change in their HAART regimen and 703 (43.9%) individuals who either switched or discontinued their HAART regimens prior to HIV RNA suppression. The remaining 73 (4.6%) patients were right censored. The median CD4 count and $\log_{10}$ HIV RNA for the study population upon HAART initiation was 150 (interquartile range: 35-287) cells/ml and 4.86 (interquartile range: 4.32-5.37) $\log_{10}$ copies/ml, respectively. The cause-specific cumulative incidence curves for HAART discontinuation or switching and for HIV RNA suppression by guideline periods are shown in figure 1 along with the non-parametric cumulative incidence function estimator [39]. No covariates besides the calendar period were included in the mixture model to facilitate comparison of parametric fit against that of the non-parametric cumulative incidence function. As is shown by the overlap in the figure, the two generalized gamma mixture model describes the non-parametric cumulative incidence function estimator well. Furthermore while the cumulative incidence of HAART discontinuation or switching exceeded that of HIV RNA suppression in time period (a) (figure 1a), there was a switch for the subsequent eras (figure 1b-d). This indicates a higher risk for HAART discontinuation or switching as compared to HIV RNA suppression before 1998. By 600 days after HAART initiation, 60% (95% CI: 0.56-0.64) of individuals were estimated to have discontinued or switched HAART regimens without ever suppressing their HIV RNA levels. This is in contrast to the 38% (95% CI: 34-42) estimated to have HIV RNA suppression without switching or discontinuing their therapy by 600 days (figure 1a). The opposite is seen for the subsequent eras in which individuals are more likely to have suppressed their HIV RNA without discontinuing or switching HAART regimens once past 33 days after HAART initiation.

The maximum likelihood estimates for the two generalized gamma mixture model are shown in Table 1. Thus an individual initiating HAART in the ≥ 1998 -2001 era with 70000

cps/ml had a $\pi_i = \frac{e^{-0.52+0.95}}{1+e^{-0.52+0.95}} = 0.61$ probability of experiencing HIV RNA suppression. The parameters shown in table 1 for each of the outcomes describe the generalized gamma distribution in conjunction with the covariate data. Utilizing these parameters of the mixture probability and the two generalized gamma distributions, the adverse-benefit ratio or the adverse-benefit subhazard ratio may be calculated. The crude adverse-benefit ratio is shown in figure 2 to contrast it with the non-parametric estimate. Additionally ABR curves are shown for individuals with the 25th, 50th, and 75th percentiles of HIV RNA. The non-parametric estimation of the ABR is rather unstable at early times but stabilizes as events

accumulate approaching the ABR limit ($\lim_{t \rightarrow \infty} \text{ABR}(t) = \frac{\Pr(J_i=2|\mathbf{X}_i)}{\Pr(J_i=1|\mathbf{X}_i)}$). A trend towards a lower ABR in each of the periods is supported by the estimates and 95% confidence interval from the mixture model.

The ABR suggests that for the earliest HAART era (<1998, figure 2a), individuals were most likely to have HAART switching or discontinuation prior to HIV RNA suppression throughout follow-up. Had the curve been consistently below 1.0, the ABR would then indicate that HIV RNA suppression prior to HAART discontinuation or switch was more likely to occur than the adverse event. However, this was not seen and rather the ABR was initially very high (albeit with a wide confidence interval) as the estimated cumulative incidence for HIV RNA suppression is small. As time lengthened, the ABR estimates declined and stabilized but remained above 1.0 for individuals initiating HAART prior to 1998. In contrast, for the ≥ 1998 -2001 period (figure 2b), initially, the ABR is >1.0 indicating a greater probability of experiencing the adverse event compared to the beneficial event up to 33 days after HAART initiation. At 33 days, individuals had an equivalent probability of experiencing HIV RNA suppression or the adverse event of HAART switching or discontinuation prior to this time point. This can be seen in the cumulative incidences in figure 1b as the two event curves cross at 33 days from HAART initiation. This is further supported by the most two recent HAART eras in which the equivalent time point is slightly earlier (24 and 32 days, respectively). Once HIV RNA is included in the model the ABR shifts to the right with increasing HIV RNA levels (figure 2, dotted curves). Furthermore, for individuals with the median HIV RNA, the ABR limit was 1.69 [95% 1.40-1.97], 0.65 [95% 0.53-0.77], 0.52 [95% 0.39-0.65], and 0.42 [95% 0.29-0.56] for the earliest to the most recent HAART periods, respectively. This suggests a downward trend indicating that individuals are ultimately more than twice as likely to have a HIV RNA suppression prior to either a switch in their HAART regimen or discontinuation of therapy in the most recent HAART era. Together, the cumulative incidence provides the probability of one event occurring prior to time t , and the ABR gives the relative magnitude of one event compared to the other with pointwise confidence intervals to assess significant relative differences between the events. In our application, the results demonstrate how the probability of HAART discontinuation or switching relative to HIV RNA suppression has changed over time. In the current HAART era, one would expect a greater probability of HAART switching or discontinuation soon after therapy initiation but ultimately resulting in more individuals successfully suppressing HIV RNA prior to HAART switching or discontinuation (after 32 days).

The subhazard functions and ABSR curves are given in figure 3 (left and right hand columns, respectively). The subhazards (figure 3: left column) are arc-shaped for HIV RNA suppression indicating a rapid rise in the rate of HIV RNA suppression that peaks prior to 100 days in all eras and then gradually declines over time. Conversely, the subhazard for HAART discontinuation declines over time, which is initially rapid but slows to a steady decline soon after HAART initiation. The ABSR curves (figure 3: right column) for each of the HAART periods follow the same general J-shaped curve, but shift down the y-axis over time periods. The <1998 era was only just below 1.0 between 51 and 111 days. Before 51 days and after 111 days, the ABSR estimate suggests that the relative rate of HAART switching or discontinuation was greater than that of HIV RNA suppression. This curve suggests that individuals that had not suppressed prior to 111 days in the <1998 era that they were more likely to have the adverse event. At 1 year after HAART initiation, the $ABSR=2.24$ (95% CI: 1.48-2.99) indicating that the risk of the HAART switching or discontinuation was 2.24 fold higher than that of HIV RNA suppression. In comparison, the ≥ 1998 -2001 and ≥ 2001 -2003 eras had a larger proportion of the curve below 1.0 (between 16-254 and 13-268 days, respectively) and the ≥ 2003 era remained below 1.0 after 15 days. Taken together, these data show the improvement of HAART regimen success over sequential guideline periods. This may be in part attributed to the introduction of new antiretroviral drugs and increased clinical experience in prescribing therapy.

4. DISCUSSION

Our aim was to describe methods that could evaluate the balance of two competing events (an adverse and a beneficial event) due to an intervention therapy by comparing either the cumulative incidences or the subhazards for the competing events by using a mixture model. As the ABR and the ABSR are ratios of the adverse and beneficial cumulative incidences and subhazards, respectively, these ratios are interpretable as the relative risk of the adverse event compared to the beneficial event. Thus, an ABR of 2.0 constant over time would indicate that at any point in time one would expect twice the number of individuals to incur the adverse event. Conversely, an ABR of 0.5 would indicate twice the number of individuals would be expected to incur the beneficial event. These methods were applied to HIV data in which it was shown that individuals starting antiretroviral therapy in the initial period of HAART were more likely to experience therapy discontinuation or switching rather than HIV RNA suppression. This relationship changed over the subsequent time periods such that individuals initiating therapy were more likely to have HIV RNA suppression rather than HAART discontinuation or switching.

The mixture model for competing events allows for the comparison of the competing outcomes in addition to exploring the effect of covariates on each event. An alternative is to test the null hypothesis, $H_0 : h_1(t) = h_2(t)$ where $h_j(t)$ are the cause specific hazards (or subhazard) for two competing events indicated by $j = 1, 2$, since the cumulative incidence functions for the two events is defined by

$$F_j^*(t) = \int_0^t S(u) h_j(u) du, \quad j=1, 2$$

where $S(u)$ is the probability of being free of event 1 and event 2 at time u [40,2]. Thus $H_0 : h_1(t) = h_2(t)$ is equivalent to the null hypothesis $H_0 : F_1^*(t) = F_2^*(t)$, for $t \geq 0$. To test this hypothesis a weighted log-rank type statistic of the form $L_n(t) = \int_0^t w(u) d(\widehat{\Lambda}_2 - \widehat{\Lambda}_1)(u)$ may be used, in which $\Lambda_j(t) = \int_0^t h_j(u) du$ is the cumulative cause-specific hazard rate for event $j = 1, 2$ and $w(u)$ is a weight function to impart more emphasis on the difference between the cause-specific hazard rates at time u [40,41,42]. However, the log-rank statistic approach gives evidence that the curves are significantly different at some time point between 0 and t , whereas our approach allows for a value of the ABR to be calculated at each time point which may then be plotted to evaluate how the CIFs converge or diverge over time. Furthermore, our methods allow for the ratio of the subhazards to be investigated allowing for the subtle but important change in focus of scientific question, in which the interest is in knowing whether individuals who have survived to time t are likely to have one event over another.

In addition to investigating the ratio of the subhazards, one could also examine the ratio of the hazards for the subdistributions. The hazard of the subdistribution as defined by Gray

$$[43] \text{ is } \tau_j(t) = \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(t < T \leq t + \Delta t, J = j | T > t \text{ or } (T \leq t \text{ and } J \neq j))}{\Delta t} \right\} = \frac{f_j^*(t)}{1 - F_j^*(t)}, j = 1, 2,$$

which is the marginal hazard rate. However, it has previously been recognized that the removal of the competing event (marginal distribution) derived from the joint distribution may not be equivalent to observing the event in isolation [1,44,45].

The mixture model presented in the application is fully parametric. A concern with all parametric methods is whether the chosen distribution is appropriate to describe the data. The mixture of generalized gamma distributions chosen here allows for the model to be

extremely flexible as some of the most common distributions such as exponential, log-normal, and Weibull distributions are special cases of the generalized gamma family. While the generalized gamma distribution provides great flexibility in fitting the data, an alternative approach that could give greater flexibility is by adapting the methods presented to a piecewise constant hazard. Whether or not a better fit could be obtained is dependent on the number of intervals defined which may have diminishing returns in terms of lack of convergence and interpretability challenges due to increasing number of parameters with each interval added. Furthermore, like many survival analysis methods, the mixture model assumes that eventually one of the competing events will occur (i.e. $\sum_{j=1}^P \Pr(J=j) = 1$). This may be a reasonable assumption for some situations (e.g. competing causes of mortality where all causes are considered such as AIDS-related causes versus all other causes of mortality), but this assumption may be inappropriate in other cases in which $\sum_{j=1}^P \Pr(J=j) < 1$ indicating that some individuals are immune to the risks under investigation. In this situation, the model of Maller and Zhou [46] may be of interest. Finally, the methods presented utilized the delta method for estimation of the variance in the ABR and ABSR. However, because the ABR and ABSR are ratio measures, bootstrap methods [47,48] or Fieller's theorem [49,50] may be preferred to the delta method approximation.

In conclusion, the balance of the benefits and potential adverse consequences of therapy is an important consideration in assessing their clinical utility. We have described methods that may be appropriate for examining the balance between adverse and beneficial outcomes for therapies in observational cohorts. We believe these methods may also be useful for intervention studies in which in addition to evaluating efficacy, an adverse-benefit ratio threshold could be defined *a priori* as a study objective for the intervention arm and compared to the adverse-benefit ratio for the control arm. How to adapt these methods into licensure and post-licensure evaluation of therapies remains an important area for future research.

Acknowledgments

We receive funding from the National Institutes of Health (R01-DA011602 for the Johns Hopkins HIV Clinical Cohort; U01-AI-42590 for the Women's Interagency HIV Study; U01-AI069918 for the North American AIDS Cohort Collaboration on Research and Design, which is a part of the International Epidemiologic Databases to Evaluate AIDS (IeDEA); and K01-AI071754 (to B.L.)). The funding sources have had no involvement with this manuscript. We declare that we have no conflict of interest.

Contract/grant sponsor: NIH; contract/grant number: R01-DA011602, U01-AI069918, U01-AI-42590 and K01-AI071754

REFERENCES

1. Crowder, M. Classical competing risks. Chapman & Hall/CRC; New York: 2001.
2. Pintilie, M. Competing Risks A Practical Perspective. John Wiley and Sons, Inc.; New York: 2006.
3. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: Competing risks and multi-state models. *Statistics in Medicine*. 2007; 26:2389–2430. DOI: 10.1002/sim.2712. [PubMed: 17031868]
4. Korn EL, Dorey FJ. Applications of crude incidence curves. *Statistics in Medicine*. 1992; 11:813–829. DOI: 10.1002/sim.4780110611. [PubMed: 1594819]
5. Pepe MS, Kaplan-Meier Mori M. marginal or conditional probability curves in summarizing competing risks failure time data? *Statistics in Medicine*. 1993; 12:737–751. DOI: 10.1002/sim.4780120803. [PubMed: 8516591]
6. Gaynor JJ, Feuer EJ, Tan CC, Wu DH, Little CR, Strauss DJ, Clarkson BD, Brennan MF. On the use of cause-specific failure and conditional failure probabilities: examples from clinical oncology data. *Journal of the American Statistical Association*. 1993; 88:400–409.

7. Lin DY. Non-parametric inference for cumulative incidence functions in competing risks studies. *Statistics in Medicine*. 1997; 16:901–910. DOI: 10.1002/(SICI)1097-0258(19970430)16:8<901::AIDSIM543>3.0.CO;2-M. [PubMed: 9160487]
8. Tsiatis A. A Nonidentifiability Aspect of the Problem of Competing Risks. *Proc. Nat. Acad. Sci. USA*. 1975; 72:20–22. [PubMed: 1054494]
9. Prentice RL, Kalbfleisch JD, Peterson AV Jr. et al. The analysis of failure times in the presence of competing risks. *Biometrics*. 1978; 34:541–554. [PubMed: 373811]
10. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association*. 1999; 94:496–509.
11. Jeong JH, Fine JP. Direct parametric inference for the cumulative incidence function. *Appl. Statist.* 2006; 55:187–200. DOI: 10.1111/j.1467-9876.2006.00532.x.
12. Larson MG, Dinse GE. A mixture model for the regression analysis of competing risks data. *Applied Statistics*. 1985; 34:201–211.
13. Department of Health and Human Services. Guidelines for the Use of Antiretroviral Agents in HIV-1-Infected Adults and Adolescents. Department of Health and Human Services; Washington, DC: 2006. Panel on Clinical Practices for Treatment of HIV Infection. (<http://aidsinfo.nih.gov>). Accessed January 26, 2007
14. Cox C, Chu H, Schneider MF, Muñoz A. Parametric survival analysis and taxonomy of hazard functions for the generalized gamma distribution. *Statistics in Medicine*. 2007; 26:4352–4374. DOI: 10.1002/sim.2836. [PubMed: 17342754]
15. Hougaard, P. *Analysis of Multivariate Survival Data*. Springer; New York: 2001.
16. Klein, JP.; Moeschberger, ML. *Survival Analysis: Techniques for Censored and Truncated Data*. second edition. Springer; New York: 2003.
17. Lindsey J, Ryan LM. Methods for Interval-Censored Data. *Statistics in Medicine*. 1998; 17:219–238. DOI: 10.1002/(SICI)1097-0258(19980130)17:2<219::AID-SIM735>3.0.CO;2-O. [PubMed: 9483730]
18. Egger M, Hirschel B, Francioli P, Sudre P, Wirz M, Flepp M, Rickenbach M, Malinverni R, Vernazza P, Battegay M. Impact of new antiretroviral combination therapies in HIV infected patients in Switzerland: prospective multicentre study. *BMJ*. 1997; 315:1194–1199. [PubMed: 9393221]
19. Palella FJ Jr, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, Aschman DJ, Holmberg SD. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. *N Engl J Med*. 1998; 338:853–860. [PubMed: 9516219]
20. Detels R, Muñoz A, McFarlane G, Kingsley LA, Margolick JB, Giorgi J, Schragger LK, Phair JP. Effectiveness of potent antiretroviral therapy on time to AIDS and death in men with known HIV infection duration. *JAMA*. 1998; 280:1497–1503. [PubMed: 9809730]
21. McNaghten AD, Hanson DL, Jones JL, Dworkin MS, Ward JW. Effects of antiretroviral therapy and opportunistic illness primary chemoprophylaxis on survival after AIDS diagnosis. *AIDS*. 1999; 13:1687–1695. [PubMed: 10509570]
22. Mocroft A, Madge S, Johnson AM, Lazzarin A, Clumeck N, Goebel FD, Viard JP, Gatell J, Blaxhult A, Lundgren JD. A comparison of exposure groups in the EuroSIDA study: starting highly active antiretroviral therapy (HAART), response to HAART, and survival. *J Acquir Immune Defic Syndr*. 1999; 22:369–378. [PubMed: 10634199]
23. Moore RD, Chaisson RE. Natural history of HIV infection in the era of combination antiretroviral therapy. *AIDS*. 1999; 13:1933–1942. [PubMed: 10513653]
24. Altice FL, Friedland GH. The Era of Adherence to HIV Therapy. *Annals of Internal Medicine*. 1998; 129:503–505. [PubMed: 9735090]
25. Ickovics JR, Meisler AW. Adherence in AIDS Clinical Trials: A Framework for Clinical Research and Clinical Care. *Journal of Clinical Epidemiology*. 1997; 50:385–397. [PubMed: 9179096]
26. Brown TT, Cole SR, Li X, Kingsley LA, Palella FJ, Riddler SA, Visscher BR, Margolick JB, Dobs AS. Antiretroviral therapy and the prevalence and incidence of diabetes mellitus in the Multicenter AIDS Cohort Study. *Archives of Internal Medicine*. 2005; 165:1179–1184. [PubMed: 15911733]
27. Brown TT, Li X, Cole SR, Kingsley LA, Palella FJ, Riddler SA, Chmiel JS, Visscher BR, Margolick JB, Dobs AS. Cumulative exposure to nucleoside analogue reverse transcriptase

- inhibitors is associated with insulin resistance markers in the Multicenter AIDS Cohort Study. *AIDS*. 2005; 19:1375–1383. [PubMed: 16103768]
28. Sulkowski MS, Mehta SH, Chaisson RE, Thomas DL, Moore RD. Hepatotoxicity associated with protease inhibitor-based antiretroviral regimens with or without concurrent ritonavir. *AIDS*. 2004; 18:2277–2284. [PubMed: 15577540]
 29. Mocroft A, Kirk O, Gatell J, Reiss P, Gargalianos P, Zilmer K, Beniowski M, Viard JP, Staszewski S, Lundgren JD. Chronic renal failure among HIV-1 patients. *AIDS*. 2007; 21:1119–1127. [PubMed: 17502722]
 30. DAD Study Group. Friis-Møller N, Reiss P, Sabin CA, Weber R, Monforte A, El-Sadr W, Thiebaut R, De Wit S, Kirk O, Fontas E, Law MG, Phillips A, Lundgren JD. Class of antiretroviral drugs and the risk of myocardial infarction. *New England Journal of Medicine*. 2007; 356:1723–1735. [PubMed: 17460226]
 31. Moore RD. Understanding the Clinical and Economic Outcomes of HIV Therapy: The Johns Hopkins HIV Clinical Practice Cohort. *J Acquir Immune Defic Syndr Hum Retrovirol*. 1998; 17(Suppl 1):S38–41. [PubMed: 9586651]
 32. d'Arminio Monforte A, Lepri AC, Rezza G, Pezzotti P, Antinori A, Phillips AN, Angarano G, Colangeli V, De Luca A, Ippolito G, Caggese L, Soscia F, Filice G, Gritti F, Narciso P, Tirelli U, Moroni M. Insights into the reasons for discontinuation of the first highly active antiretroviral therapy (HAART) regimen in a cohort of antiretroviral naïve patients. *AIDS*. 2000; 14:499–507. [PubMed: 10780712]
 33. Bonfanti P, Valsecchi L, Parazzini F, Carradori S, Pusterla L, Fortuna P, Timillero L, Alessi F, Ghiselli G, Gabbuti A, Di Cintio E, Martinelli C, Faggion I, Landonio S, Quirino T. Incidence of Adverse Reactions in HIV Patients Treated With Protease Inhibitors: A Cohort Study. *J Acquir Immune Defic Syndr*. 2000; 23:236–245. [PubMed: 10839659]
 34. Häansel A, Bucher HC, Nüesch R, Battegay M. Reasons for Discontinuation of First Highly Active Antiretroviral Therapy in a Cohort of Proteinase Inhibitor-Naïve HIV-Infected Patients. *J Acquir Immune Defic Syndr*. 2001; 26:191–193. [PubMed: 11242189]
 35. Dorrucci M, Pezzotti P, Grisorio B, Minardi C, Muro MS, Vullo V, Monforte A. Italian Cohort of Antiretroviral-Naïve Patients Study. Time to discontinuation of the first highly active antiretroviral therapy regimen: a comparison between protease inhibitor- and non-nucleoside reverse transcriptase inhibitor-containing regimes. *AIDS*. 2001; 15:1733–1736. [PubMed: 11546952]
 36. Cozzi-Lepri A, Phillips AN, d'Arminio Monforte A, Piersantelli N, Prani A, Petrosillo N, Leoncini F, Scerbo A, Tundo P, Abrescia N, Moroni M, Italian Cohort Naïve Antiretrovirals (I.Co.N.A.) Study Group. Virologic and Immunologic Response to Regimens Containing Nevirapine or Efavirenz in Combination with 2 Nucleoside Analogues in the Italian Cohort Naïve Antiretrovirals (I.Co.N.A.) Study. *J Infect Dis*. 2002; 185:1062–1069. [PubMed: 11930316]
 37. O'Brien ME, Clark RA, Besch CL, Myers L, Kissinger P. Patterns and Correlates of Discontinuation of the Initial HAART Regimen in an Urban Outpatient Cohort. *J Acquir Immune Defic Syndr*. 2003; 34:407–414. [PubMed: 14615659]
 38. Bini T, Testa L, Chiesa E, Adorni F, Abeli C, Castelnuovo B, Melzi S, Sollima S, Bongiovanni M, d'Arminio Monforte A. Outcome of a second-line protease inhibitor-containing regimen in patients failing or intolerant of a first highly active antiretroviral therapy. *J Acquir Immune Defic Syndr*. 2000; 24:115–122. [PubMed: 10935686]
 39. Kalbfleisch, JD.; Prentice, RL. *The Statistical Analysis of Failure Time Data*. John Wiley and Sons, Inc.; New York: 1980.
 40. Kochar SC, Lam KF, Yip PSF. Generalized Supremum Tests for the Equality of Cause Specific Hazard Rates. *Lifetime Data Analysis*. 2002; 8:277–288. DOI: 10.1023/A:1015806021930. [PubMed: 12182123]
 41. Kochar SC, Deshpandé JV, Koul H. A review of some distribution-free tests for the equality of cause specific hazard rates. *Analysis of Censored Data - the IMS-LNMS*. 1995; 27:117–162.
 42. Carriere KC, Kochar SC. Comparing sub-survival functions in a competing risks model. *Lifetime Data Analysis*. 2000; 6:85–97. DOI: 10.1023/A:1009697802491. [PubMed: 10763563]
 43. Gray RJ. A class of k-sample tests for comparing the cumulative incidence of a competing risk. *Annals of Statistics*. 1988; 16:1141–1154.

44. Makeham WM. On an application of the theory of the composition of decremental forces. *Journal Institute of Actuaries*. 1874; 18:317–322.
45. Cornfield J. The estimation of the probability of developing a disease in the presence of competing risks. *American Journal of Public Health*. 1957; 47:601–607. [PubMed: 13424810]
46. Maller RA, Zhou X. Analysis of parametric models for competing risks. *Statistica Sinica*. 2002; 12:725–750.
47. Efron, B. *The Jackknife, the Bootstrap, and Other Resampling Plans*. Society for Industrial and Applied Mathematics; Philadelphia: 1982.
48. Efron, B.; Tibshirani, R. *An introduction to the Bootstrap*. Chapman Hall; London: 1993.
49. Fieller EC. A Fundamental Formula in the Statistics of Biological Assay, and Some Applications. *Quarterly Journal of Pharmacy and Pharmacology*. 1944; 17:117–123.
50. Zerbe GO. On Fieller's Theorem and the General Linear Model. *The American Statistician*. 1978; 32:103–105.

Figure 1.

Mixture model estimates for CIF for HIV RNA suppression (solid thick line) and HAART switching or discontinuation (dashed thick line) and non-parametric estimations (thin solid and dotted lines) for the (a) <1998, (b) ≥ 1998 -2001, (c) ≥ 2001 -2003, and (d) ≥ 2003 HAART periods.

Figure 2.

Mixture model estimates for unadjusted Adverse-Benefit Ratio (solid thick line; HAART switching or discontinuation CIF / HIV RNA CIF) with corresponding pointwise 95% CI (long-dash thick line), adjusted for HIV RNA at HAART initiation (numbered dotted lines: 1- 25th, 2- 50th, 3 -75th percentile of HIV RNA levels), and the non-parametric estimation (thin solid line) for (a) <1998, (b) ≥1998-2001, (c) ≥2001-2003, and (d) ≥2003 HAART periods.

Figure 3.

Mixture model estimates for the subhazard (left column) for HIV RNA suppression (solid line) and HAART switching or discontinuation (dashed line) and Adverse-Benefit Subhazard Ratio (right column; pointwise 95% CI dashed lines) for <1998 (a and b), ≥1998-2001 (c and d), ≥2001-2003 (e and f), and ≥2003 (g and h) HAART periods.

Table 1
Parameter estimates from the application of mixture model methods to HIV data.

Parameter	Mixture* Probability		Generalized Gamma HIV RNA Suppression		Generalized Gamma HAART Switch/ Discontinuation	
	Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.
Location						
Intercept	-0.5200	0.0867	4.3089	0.0628	4.3306	0.1190
Calendar Period						
<1998	0 (ref)		0 (ref)		0 (ref)	
≥ 1998 to < 2001	0.9539	0.1286	-0.1378	0.0746	0.4111	0.1357
≥ 2001 to < 2003	1.1789	0.1563	-0.2967	0.0837	0.1953	0.1777
≥ 2003	1.3779	0.1834	-0.1682	0.0914	-0.3638	0.2223
Log ₁₀ HIV RNA**	-0.1551	0.0739	0.3100	0.0405	0.3351	0.0786
Scale			0.8265	0.0209	1.4685	0.0484
Shape			-0.2589	0.0703	0.5444	0.1068

$$* \text{ Mixture Probability } \pi_i = \frac{e^{\sum \text{Parameter Estimate}}}{1 + e^{\sum \text{Parameter Estimate}}}$$

** Centered at Log10(70000)