Published in final edited form as:

Stat Med. 2012 August 30; 31(19): 2098–2109. doi:10.1002/sim.5317.

A simulation study of finite-sample properties of marginal structural Cox proportional hazards models

Daniel Westreich¹, Stephen R. Cole², Enrique F. Schisterman³, and Robert W. Platt⁴

¹Department of Obstetrics and Gynecology and Duke Global Health Institute, Duke University, Durham, NC

²Department of Epidemiology, UNC Gillings School of Global Public Health, Chapel Hill, NC

³Epidemiology Branch, Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, MD

⁴Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Quebec

Summary

Motivated by a previously published study of HIV treatment, we simulated data subject to time-varying confounding affected by prior treatment to examine some finite-sample properties of marginal structural Cox proportional hazards models. We compared (a) unadjusted, (b) regression-adjusted, (c) unstabilized and (d) stabilized marginal structural (inverse probability-of-treatment [IPT] weighted) model estimators of effect in terms of bias, standard error, root mean squared error (MSE) and 95% confidence limit coverage over a range of research scenarios, including relatively small sample sizes and ten study assessments. In the base-case scenario resembling the motivating example, where the true hazard ratio was 0.5, both IPT-weighted analyses were unbiased while crude and adjusted analyses showed substantial bias towards and across the null. Stabilized IPT-weighted analyses remained unbiased across a range of scenarios, including relatively small sample size; however, the standard error was generally smaller in crude and adjusted models. In many cases, unstabilized weighted analysis showed a substantial increase in standard error compared to other approaches. Root MSE was smallest in the IPT-weighted analyses for the base-case scenario. In situations where time-varying confounding affected by prior treatment was absent, IPT-weighted analyses were less precise and therefore had greater root MSE compared with adjusted analyses. The 95% confidence limit coverage was close to nominal for all stabilized IPT-weighted but poor in crude, adjusted, and unstabilized IPT-weighted analysis. Under realistic scenarios, marginal structural Cox proportional hazards models performed according to expectations based on large-sample theory and provided accurate estimates of the hazard ratio.

Keywords

Bias; Causal inference; Marginal structural models; Monte Carlo study

1. Introduction

Marginal structural Cox proportional hazards models were introduced in 1998 [1] and have rapidly disseminated into biomedical research over the last decade [2–19]. Marginal structural Cox proportional hazards models were developed to estimate causal effects in the presence of time-varying confounding affected by prior treatment. In such settings standard statistical methods, such as Cox proportional hazards models with time-varying covariates

[20] fail to consistently estimate the total effect of a time-varying treatment (or exposure) even in the absence of unmeasured confounding and unmeasured informative censoring [3].

The conditions necessary for a marginal structural Cox proportional hazards model to provide unbiased estimates of effects include: consistency, positivity, no unmeasured confounders, no unmeasured informative censoring, and correct model specification. Here we define consistency and positivity because these are less well-known than the remaining conditions. Consistency, not to be confused with statistical consistency (i.e., the property of convergence in probability to the true value as the sample size tends to infinity), is the assumption that an individual's potential outcome under her observed treatment is precisely her observed outcome: formally, $Y_i = Y_i^x$ if $X_i = x$ for participant i , where X is treatment, Y is the outcome and Y^x is the potential outcome under treatment x [21]. Positivity, also known as the experimental treatment assumption [22], requires that there are both treated and untreated individuals at every combination of the values of the confounders in the observed data: formally, for a discrete-valued treatment X and arbitrary confounder vector $\mathbf{Z}$, if $f(\mathbf{Z}_i)$

0 then $\Pr(X_i = x | \mathbf{Z}_i) > 0$ for all $x \in X$, where $f(\cdot)$ is the probability density function [23].

Given the above stated conditions, the large-sample properties of inverse probability-of-treatment (IPT) weighted estimators of a marginal structural Cox proportional hazards model have been described [1, 24]. In brief, the IPT-weighted estimator of a marginal structural Cox proportional hazards model provides a regular, asymptotically linear estimator of a contrast in potential outcomes [24]. Given data D_n , sample size n , and the influence function $\psi(D_i)$, a measurable random function with expectation 0, an estimator using sample size n , $\hat{\beta}_n$, of a parameter β is regular and asymptotically linear if $n^{1/2}(\hat{\beta}_n - \beta) = n^{-1/2} \sum_{i=1}^n \psi(D_i) + o_p(1)$ and if $n^{1/2}(\hat{\beta}_n - \beta)$ has a limiting distribution that does not depend on the local data generating process [25]. Again, under the above stated conditions, a hazard ratio from a marginal structural Cox proportional hazards model is asymptotically unbiased provided that the model for the probability of treatment (used to estimate the IPT weights) converges to the true probability of treatment at a rate of at least $n^{1/4}$.

Granting the (largely un-testable) conditions listed above and despite favorable large-sample properties, if marginal structural Cox proportional hazards models have poor finite sample properties then their general use will be limited. For instance, if IPT-weighted estimators are notably biased under realistic sample and effect sizes, then this finite-sample bias may well preclude their recommendation as a tool in applied research. To our knowledge, there exists no study of the finite-sample properties of the marginal structural Cox proportional hazards model under realistic research settings. Here we present results from such a study. In section 2, we review the marginal structural Cox proportional hazards model. In section 3, we describe a motivating example from HIV/AIDS. In section 4, we describe the simulation design. In section 5, we summarize the simulation results. We close in section 6 with a discussion.

2. Marginal structural Cox proportional hazards model

Let the data consist of n copies of the independent and identically distributed set $\{\mathbf{Z}(t), X(t), T_i\}$, where $i = 1$ to n , $\mathbf{Z}(t)$ is a vector of time-varying covariates, $X(t)$ is a time-varying treatment of interest, and T_i is the minimum of T_i^* and C_i where T_i^* is the time from study entry to the event of interest and C_i is the time from study entry to right-censoring. Uppercase letters represent random variables and lowercase letters represent possible realizations. We assume no unmeasured informative censoring, or

$f\left[T^*|Z(t), X(t)\right] = f\left[T^*|Z(t), X(t), C\right]$ The history of a time-varying variable is denoted using overbars, so that $\bar{X}_i = \bar{X}_i(t) = \{X_i(0), X_i(1), \dots, X_i(t)\}$ is the observed treatment history for participant i , and $\bar{x} = \bar{x}(t) = \{x(0), x(1), \dots, x(t)\}$ is a possible treatment history. Examples of possible histories $\bar{x}$ include always treated $\{1, 1, \dots, 1\}$ and never treated $\{0, 0, \dots, 0\}$.

A marginal structural Cox proportional hazards model is $\lambda_{\bar{x}}(t) = \lambda_0(t) \exp\{\delta_w g(\bar{x})\}$, where $\lambda_{\bar{x}}(t)$ is the hazard at time t had all participants (possibly contrary to fact) been assigned treatment history $\bar{x}$, $\lambda_0(t)$ is the hazard at time t had all participants been assigned treatment history $g(\bar{x}) = 0$. We assume the function of treatment history $g(\bar{x})$ is the cumulative average treatment $g(\bar{x}) = t^{-1} \sum_{k=0}^t x(k)$ but other specifications are possible. In the above model $\exp(\delta_w)$ is a hazard ratio for a unit increase in $g(\bar{x})$ and is a contrast of potential outcomes. Note that here $g(\bar{x})$ ranges from 0 to 1, so that a unit difference in the cumulative average treatment $g(\bar{x})$ corresponds to always versus never-treated, the contrast of interest in our setting.

The hazard ratio $\exp(\delta_w)$ may be estimated by fitting an *unadjusted* Cox proportional hazards model on the observed data *weighted* by a stabilized time-varying IPT weight, which is estimated from the observed data and defined as

$$W_i(t) = \prod_{k=0}^t f\left[X_i(k) | \bar{X}_i(k-1)\right] / f\left[X_i(k) | \bar{X}_i(k-1), \bar{Z}_i(k-1)\right]$$

where $f[\cdot|\cdot]$ is the conditional density function evaluated at the observed covariate values for a given participant. The unstabilized weights simply replace the numerator of $W_i(t)$ with 1. When $X(t)$ has a Bernoulli distribution, the denominator of the weight is equivalent to a time-varying generalization of the propensity score among the treated [26] and the complement of this generalization among the untreated. The denominator of this weight corresponds to each patient's probability of their entire treatment history up to each time point; these weights create a pseudopopulation in which treatment at each time is unrelated to the time-varying confounders. The numerator of the weight is a stabilizing factor which reduces the variance of the weighted estimator and which is a function of treatment that does not depend on time-varying confounders.

3. Motivating empirical example: Antiretroviral therapy for HIV and incidence of AIDS or death

Here we describe a motivating example from HIV/AIDS [6]. In this example, estimates from a marginal structural Cox proportional hazards model fit to observational data cohered well with substantive knowledge and existing randomized evidence. Moreover, these estimates differed from estimates from standard Cox proportional hazards models with time-varying covariates fit to the same data [6, 8].

The study population comprised participants from two ongoing US multicenter prospective studies of HIV namely, the Multicenter AIDS Cohort Study and the Women's Interagency HIV Study. The treatment was use of highly active antiretroviral therapy (henceforth therapy), and the outcome was time from study entry to the primary clinical AIDS defining illness or death from any cause. Participants ($n=1,498$) entered study for this investigation in 1995. Over a median (quartiles) follow-up of 5.4 (3.0, 5.9) years, approximately 47% of person-time was exposed to therapy and 382 participants had an incident diagnosis of

clinical AIDS or died (cumulative incidence 28%). The results of crude, adjusted, and marginal structural (IPT-weighted) Cox proportional hazards model analyses were as follows: 0.98 (95% confidence limits [CL]: 0.76, 1.26), 0.81 (95% CL: 0.61, 1.07), and 0.54 (95% CL: 0.38, 0.78), respectively [6]. Crude (panel A) and IPT-weighted (panel B) cumulative incidence curves are shown in Figure 1[27, 28]. Results were thus close to the null hazard ratio in both crude and adjusted analyses, but substantially different in the IPT-weighted analysis. Moreover, the estimate from the IPT-weighted analysis was closer than the crude or adjusted results to findings from randomized studies of therapy for the prevention of AIDS or death. For example, pooling results from two large randomized studies [29, 30] yields a summary hazard ratio for AIDS of 0.52 (95% CL: 0.43, 0.64) comparing treatment with therapy to standard care.

4. Simulation

Simulation Design

A simulated data record comprises values for $Z(k)$, $X(k)$, $T^{\bar{x}}$ and T , for $k = 0$ to 9 assessments. We first generated the potential survival time under no treatment, T^0 , from the Weibull distribution with hazard $\sigma \lambda t^{\sigma-1}$, where the Weibull shape parameter $\sigma = 0.75$ and rate parameter $\lambda = \exp(0)$. Next, we generated the potential survival times under the 10 possible non-zero treatment levels for cumulative average treatment with 10 assessments as $T^{\bar{x}} = T^0 \exp\{\varphi g(\bar{x})\}$ where $\varphi = 0.924$ for the base-case scenario. Therefore the hazard ratio $\exp(\delta_W)$ for always treated $g(\bar{x}) = 10/10 = 1$ compared to never treated $g(\bar{x}) = 0/10 = 0$ was $0.5 = \exp\{-0.924(0.75)\}$. Due to a dichotomous treatment at ten time points, there are 2^{10} possible treatment histories; however, we concern ourselves with the cumulative average treatment effect here. We note that these exposure patterns are similar but more general than the typical “once-treated, always-treated” intent-to-treat assumptions in applied work in HIV[6]; in most examples, the intent-to-treat assumption is only approximately true.

The measured time-varying confounder $Z(k)$ is determined in part by T^0 and by $X(k-1)$ by the models $\Pr[Z(k)=1] = \text{expit}[\beta_0 + \beta_1 X(k-1) + \beta_2 T^0]$ for $k = 1$ to 9, and

$\Pr[Z(0)=1] = \text{expit}[\beta_0 + \beta_2 T^0]$ for $k = 0$, where $\text{expit}(\cdot) = \exp(\cdot) / [1 + \exp(\cdot)]$. β_0 was chosen to maintain a marginal probability of 0.5, β_1 was set to $\log(2)$ in the base-case, and β_2 was set to $\log(0.5)$ for all scenarios. The time-varying treatment $X(k)$ is determined in part by $Z(k)$ by the model $\Pr[X(k)=1] = \text{expit}[\gamma_0 + \gamma_1 Z(k)]$, for $k = 0$ to 9, where γ_0 was chosen to maintain a marginal probability of 0.5 in the base-case, and γ_1 was set to $\log(2)$.

The observed event time T was assigned as the one of 11 potential event times that coincided with the observed treatment history, namely $T = T^{\bar{x}=g(\bar{x})}$, where

$g(\bar{x}) = 10^{-1} \sum_{k=0}^9 X(k)$. Observed event times were administratively censored to yield 20% events. Akin to prior illustrative examples [3], for simplicity we assume events always occur temporally posterior to all treatment assessments. Such an assumption is not required generally for marginal structural models [1, 31].

Figure 2 is a causal diagram [32] representing the simulated data. In Figure 2, arrows represent direct causal effects, or contrasts in potential outcomes [33, 34]. In Figure 2, $Z(k+1)$ is time-varying confounder affected by prior treatment; $Z(k+1)$ is related to outcome through T^0 . Therefore, failing to adjust for $Z(k+1)$ will bias the estimate of treatment effect on T through the open bias path $X(k+1) - Z(k+1) - T^0 - T$. Adjusting for $Z(k+1)$ blocks this aforementioned bias path, but opens a collider-stratification bias path $X(k) - Z(k+1) - T^0 - T$ and again results in a biased estimate of the treatment effect on T .

Simulation base-case values were chosen so that weighted and adjusted estimates coincided with results from previously published examples [6, 35, 36]. We explored 40 scenarios defined by varying the sample size n as 1500 for the base-case and 200; the true effect of treatment as a hazard ratio of 0.5 for the base-case and 1.0; a treatment prevalence of 0.5 for the base-case and 0.2; and the magnitude of treatment $X(k)$ on confounder $Z(k+1)$ as $\beta_1 = \log(2)$ for the base-case and $\{0.5, 0.75, 1, 1.5\}$. We conducted 5000 simulations for each scenario. The simulation standard error for 5000 simulations is approximately 1/3% for estimates of the 95% CL coverage [37].

Simulation Analysis

We conducted four analyses estimating the hazard ratio for the effect of treatment on outcome. First, we estimated the unadjusted association obtained from a standard Cox proportional hazards model with the sole covariate $g(\bar{X})$ as $h_t(t) = h_0(t) \exp[\delta_c g(\bar{X})]$, where $h_0(t)$ is an unspecified hazard at time t for the group with $g(\bar{X})=0$ and δ_c is the estimated crude log hazard ratio for a unit difference in $g(\bar{X})$. This analysis estimates the unadjusted hazard ratio for a unit change in the cumulative average treatment.

Second, we estimated the adjusted association obtained from a standard Cox proportional hazards model with the two covariates, $g(\bar{X})$ (defined as above) and the vector of covariates $\mathbf{Z}$ as $h_t(t) = h_0(t) \exp\{\delta_A g(\bar{X}) + \eta \mathbf{Z}\}$, where $h_0(t)$ is an unspecified hazard at time t for the group with $g(\bar{X})=\mathbf{Z}=0$, δ_A is the estimated log hazard ratio for a unit difference in $g(\bar{X})$ holding the vector of covariate indicators $\mathbf{Z}$ constant at any level. This analysis estimated the $\mathbf{Z}$ -adjusted hazard ratio for a unit change in the same cumulative average treatment. We conducted a sensitivity analysis using cumulative average Z instead of a vector of indicators for each $Z(k)$; results were similar as in main analysis.

Third and fourth we estimated the association obtained from a marginal structural Cox proportional hazards model estimated using IPT weights as $h_t(t) = h_0(t) \exp[\delta_w g(\bar{X})]$, where $h_0(t)$ is an unspecified hazard at time t for the group with $g(\bar{X})=0$, δ_w is the estimated log hazard ratio for a unit difference in $g(\bar{X})$.

For each of these four analyses (i.e., crude, adjusted, unstabilized weighted, stabilized weighted) we estimated simulated bias, simulated standard error, simulated 95% CL coverage, and simulated root mean squared error (MSE). Bias was computed as the average difference between an estimator and the true log hazard ratio. The average standard error was based on the robust variance. The Monte Carlo standard error was defined as the simulated standard deviation of an estimator. Confidence limit coverage was computed as the proportion of times the CL contained the true hazard ratio. Root mean squared error was calculated by taking the square root of the sum of the bias and the square of the Monte Carlo standard error. Convergence was obtained in all simulations.

Additional simulations were performed using a different causal structure with only two time points but a wider range of simulation scenarios (432 scenarios, including use of a lognormal failure time distribution). These alternate simulations provided the same qualitative conclusions as the main results, and so are not described further; details are available from the authors.

5. Simulation results

Simulations for the base-case reflected key characteristics of the motivating example [6]. For example, the IPT-weighted hazard ratio was 0.54 in the motivating example [6], and 0.51 in the simulated data. Likewise, the adjusted hazard ratio in the motivating example was 0.81, and 0.85 in the simulated data. Tables 1 through 4 show estimated bias, standard error, root mean squared error, and 95% CL coverage, respectively, for the 40 scenarios explored.

Table 1 shows bias; in the base-case scenario (first row), bias was notably large and toward the null in both the unadjusted and adjusted analyses, severely attenuating and reversing the true log-hazard ratio, respectively. In contrast, the stabilized IPT-weighted analysis was essentially unbiased. These findings were robust to changes in the sample size, true hazard ratio, and the effect of treatment on the time-varying confounder. However, analysis with unstabilized weights was notably biased with the fairly rare treatment (i.e. 20%). This bias may reflect non-positivity due to the relatively small probability of treatment [23]. Crude and adjusted analyses were generally biased. However, when $\beta = 0$ (i.e., there was no effect of treatment on the time-varying confounder), confounding was unaffected by prior treatment and thus a simple adjusted analysis was unbiased. For example, when taking the base-case scenario but with $\beta = 0$, bias for the adjusted analysis was -0.055 , and for the weighted analyses was 0.010 (Table 1, row 5).

Table 2 shows estimated Monte Carlo standard error for all scenarios and analyses. In the base-case scenario (first row) and in general, the standard error was smallest for crude analyses, higher in the adjusted analyses, higher still in the stabilized weighted analyses, but all were of the same magnitude. However, the unstabilized weighted standard errors were notably larger, as expected. The effect of stabilization on reduction in standard error was dramatic when treatment was rare (i.e., 20%), and modest when treatment was common (i.e., 50%). When $\beta = 0$ and thus an adjusted analysis was unbiased (Table 1), the pattern and magnitude of standard errors remained essentially the same.

Table 3 shows root MSE. Here, both weighted analyses (unstabilized root MSE = 0.583, stabilized root MSE = 0.582) were clearly superior to either crude or adjusted analyses in the main scenario (crude root MSE = 1.293, adjusted root MSE = 0.702). In general, the adjusted and stabilized weighted analyses were superior to the crude analyses. However, comparing adjusted to weighted analyses, there were few clear patterns of results across scenarios. It is of note that when weights are unnecessary (i.e. $\beta = 0$), the adjusted analysis has a superior root MSE to either IPT weighted analysis.

Last, Table 4 shows simulated 95% CL coverage from all scenarios. In the base-case scenario and all other scenarios, unstabilized and stabilized weighted analyses had essentially nominal coverage with a range from 93% to 96%. Crude and adjusted analyses often had less than nominal coverage, with ranges of 7% to 84% for the crude and 60% to 95% for the adjusted.

One additional simulation was performed to examine the convergence of the marginal structural model as a function of sample size. With other parameters held at base-case values, we generated 5000 samples varying sample size n in $\{200, 400, 800, 1500, 2500, 5000\}$, and performed a linear regression of the estimated Monte Carlo standard error on n^k , where k varied between -2 and 1 . We found the k that minimized the residual error to be -0.524 . Therefore, we estimated that the standard error approached zero at a rate of approximately $n^{-1/2}$, in line with the theoretical expectation.

6. Discussion

This study describes properties of marginal structural Cox proportional hazards models in the presence and absence of time-varying confounding affected by prior treatment. We found that IPT-weighted marginal structural models performed as well as or better than unadjusted or adjusted Cox proportional hazards regression models in terms of bias and 95% CL coverage across a wide range of scenarios motivated by an empirical example. As expected, we also found that unadjusted and adjusted Cox models often perform better than weighted models in terms of standard error.

This divergence in performance between metrics of bias and standard error leads to predictably mixed results when considering root MSE: the advantage of a weighted model depends on whether the increase in variance associated with weighting is outweighed by the reduction in bias that the weighted model provides. For instance, in the fourth scenario the adjusted analysis was biased yet precise yielding the same root MSE of 0.591 as the analogous stabilized weighted analysis, which was unbiased yet less precise. However, this adjusted estimate, which was both precise and biased, had relatively poor CL coverage (85%) compared to the weighted analysis (95%). Of course, when a time-varying confounder is not affected by prior treatment (i.e., $\beta = 0$), the adjusted model is unbiased and thus has a clear advantage in both standard error and root MSE over marginal structural models.

Previously, it was unknown whether discrepancies in crude, adjusted, and weighted results in the original analysis [6] could be attributed to poor finite-sample performance of marginal structural models. These simulation results suggest a better explanation for observed discrepancies is a failure of the crude and adjusted analyses to produce unbiased estimates of the effect, and the comparison of root MSE in the base-case scenario further suggests that the increased variance associated with weighting versus adjusting was not too high a price to pay in the original analysis. On a related note, while adjusted results are not generally expected to precisely coincide with marginal results in non-collapsible effect estimates (such as the hazard ratio and rate ratio [38]), non-collapsibility does not play a role if the null hypothesis is true. Because contrasts in performance of adjusted and marginal models was similar in both null and non-null scenarios, the differences seen in non-null scenarios are unlikely to be explained by non-collapsibility.

This work has limitations. First, we only explored a limited region of the possible joint parameter space. We defined 40 scenarios by combinations of values of n , the marginal probability of treatment, the true hazard ratio, and the association between the treatment and time-varying confounder. However, as noted above, we simulated an additional 432 scenarios in fewer time points; those results support the findings presented. Second, we included only a single, dichotomous time-varying confounder; real world data and examples [6] often include higher dimension covariate vectors. Thus, despite the fact that our simulation was able to approximate results from an empirical example, this work may not generalize to all real world situations of time-varying confounding affected by prior treatment. Future work might extend these results by expanding on the dimension of confounders and allowing the outcome to occur at any point during follow up: methods to simulate such data are under development [39]. We assumed that models, including the model for weights, were correctly specified.

Last, in applied settings, the existence of poorly or unmeasured confounders, informative censoring, and model mis-specification may all sabotage inferences from observational data. Nonetheless, without adequate finite-sample behavior when these assumptions do hold, the use of marginal structural models could not be recommended. That finite-sample behavior

appears to be good: marginal structural Cox models fit with stabilized inverse probability of weights appear to give unbiased results with approximately nominal coverage down to sample size of 200 and marginal prevalence of exposure of 20% (e.g., about 40 exposed individuals), making them a potentially appropriate approach to even small data sets. Of course, the need to control for additional confounders may hamper the use of these methods in such small data sets.

We note that in this simulation, outcomes could only occur after the 10 time points (e.g., a subject could not die at time $t=4$). This situation is analogous to numerous real world scenarios; e.g., intrapartum exposures (say, to iron supplementation) and outcomes measured at birth [40] or in the infant thereafter; or early life exposures such as education before age 30 or parity and later-life outcomes such as development of dementia. Alternate ways of simulating data appropriate for marginal structural models may have more flexibility in this regard: for example, an alternative approach was published recently by Young et al. that addresses this issue [41]. Another possible approach is to reverse the algorithm for the parametric g-formula [42]: assume parametric distributions for interrelations among exposure, covariates, and outcome; simulate data; then estimate using the parametric g-formula to estimate parameters of a marginal structural model. While the parametric g-formula requires substantial assumptions, these assumptions can be met in a simulation context, and thus allows the unbiased estimation of the marginal structural model using a method independent of inverse probability weights. Both these methods may be explored in further work.

Those caveats notwithstanding, the present results provide, to our knowledge, the first simulation study lending support to the use of marginal structural Cox proportional hazards models in a setting representative of real-world data. Key strengths of this study include a large number of study assessments, a relatively large number of scenarios explored based on a motivating empirical example, small Monte Carlo simulation error, and coherence with theoretical expectations. In conclusion, in the presence of time-varying confounding affected by prior treatment, standard Cox proportional hazard models provide biased effect estimates, while IPT-weighted estimation of marginal structural Cox proportional hazards models provides approximately unbiased and relatively precise effect estimates.

Acknowledgments

FUNDING

This work was supported by the National Institutes of Health, National Institute of Allergy and Infectious Disease [grant number T32-AI-07001 to D.W.; grant numbers R01-AA-01759 to S.R.C.]; the National Institutes of Health the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development [Intramural Research Program to E.F.S.; grant numbers K99-HD-06-3961 and 4R00-HD-06-3961-02 to D.W.]; and the Chercheur-boursier award, and by core support to the Montreal Children's Hospital Research Institute, from the Fonds de la Recherche en Santé' du Quebec to [R.W.P.].

Data in this manuscript were collected by the Women's Interagency HIV Study (WIHS) Collaborative Study Group with centers (Principal Investigators) at New York City/Bronx Consortium (Kathryn Anastos); Brooklyn, NY (Howard Minkoff); Washington, DC, Metropolitan Consortium (Mary Young); The Connie Wofsy Study Consortium of Northern California (Ruth Greenblatt); Los Angeles County/Southern California Consortium (Alexandra Levine); Chicago Consortium (Mardge Cohen); Data Coordinating Center (Stephen Gange). The WIHS is funded by the National Institute of Allergy and Infectious Diseases (U01-AI-35004, U01-AI-31834, U01-AI-34994, U01-AI-34989, U01-AI-34993, and U01-AI-42590) and by the National Institute of Child Health and Human Development (U01-HD-32632). The study is co-funded by the National Cancer Institute, the National Institute on Drug Abuse, and the National Institute on Deafness and Other Communication Disorders. Funding is also provided by the National Center for Research Resources (UCSF-CTSI Grant Number UL1 RR024131). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

Data in this manuscript were collected by the Multicenter AIDS Cohort Study (MACS) with centers (Principal Investigators) at The Johns Hopkins Bloomberg School of Public Health (Joseph B. Margolick, Lisa P. Jacobson), Howard Brown Health Center, Feinberg School of Medicine, Northwestern University, and Cook County Bureau of Health Services (John P. Phair, Steven M. Wolinsky), University of California, Los Angeles (Roger Detels), and University of Pittsburgh (Charles R. Rinaldo). The MACS is funded by the National Institute of Allergy and Infectious Diseases, with additional supplemental funding from the National Cancer Institute. U01-AI-35042, UL1-RR025005 (GCRC), U01-AI-35043, U01-AI-35039, U01-AI-35040, U01-AI-35041. Website located at <http://www.statepi.jhsph.edu/macs/mac.html>.

References

1. Robins, JM. 1997 Proceedings of the American Statistical Association. 1998. Marginal Structural Models; p. 1-10.
2. Hernán MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*. 2000; 11:561–570. [PubMed: 10955409]
3. Hernán MA, Brumback BA, Robins JM. Marginal Structural Models to Estimate the Joint Causal Effect of Nonrandomized Treatments. *JASA*. 2001; 96:440–448.
4. Choi HK, Hernan MA, Seeger JD, Robins JM, Wolfe F. Methotrexate and mortality in patients with rheumatoid arthritis: a prospective study. *Lancet*. 2002; 359:1173–1177. [PubMed: 11955534]
5. Cook NR, Cole SR, Hennekens CH. Use of a marginal structural model to determine the effect of aspirin on cardiovascular mortality in the Physicians' Health Study. *Am J Epidemiol*. 2002; 155:1045–1053. [PubMed: 12034583]
6. Cole SR, Hernán MA, Robins JM, Anastos K, Chmiel J, Detels R, et al. Effect of highly active antiretroviral therapy on time to acquired immunodeficiency syndrome or death using marginal structural models. *Am J Epidemiol*. 2003; 158:687–694. [PubMed: 14507605]
7. Barrón Y, Cole SR, Greenblatt RM, Cohen MH, Anastos K, DeHovitz JA, et al. Effect of discontinuing antiretroviral therapy on survival of women initiated on highly active antiretroviral therapy. *Aids*. 2004; 18:1579–1584. [PubMed: 15238776]
8. Sterne JA, Hernan MA, Ledergerber B, Tilling K, Weber R, Sendi P, et al. Long-term effectiveness of potent antiretroviral therapy in preventing AIDS and death: a prospective cohort study. *Lancet*. 2005; 366:378–384. [PubMed: 16054937]
9. Wang C, Vlahov D, Galai N, Cole SR, Bareta J, Pollini R, et al. The effect of HIV infection on overdose mortality. *Aids*. 2005; 19:935–942. [PubMed: 15905674]
10. Eisner MD, Wang Y, Haight TJ, Balmes J, Hammond SK, Tager IB. Secondhand smoke exposure, pulmonary function, and cardiovascular mortality. *Ann Epidemiol*. 2007; 17:364–373. [PubMed: 17300955]
11. López-Gatell H, Cole SR, Hessol NA, French AL, Greenblatt RM, Landesman S, et al. Effect of tuberculosis on the survival of women infected with human immunodeficiency virus. *Am J Epidemiol*. 2007; 165:1134–1142. [PubMed: 17339383]
12. Delaney JA, Daskalopoulou SS, Suissa S. Traditional versus marginal structural models to estimate the effectiveness of beta-blocker use on mortality after myocardial infarction. *Pharmacoepidemiol Drug Saf*. 2009; 18:1–6. [PubMed: 18949804]
13. López-Gatell H, Cole SR, Margolick JB, Witt MD, Martinson J, Phair JP, et al. Effect of tuberculosis on the survival of HIV-infected men in a country with low tuberculosis incidence. *Aids*. 2008; 22:1869–1873. [PubMed: 18753866]
14. Petersen ML, van der Laan MJ, Napravnik S, Eron JJ, Moore RD, Deeks SG. Long-term consequences of the delay between virologic failure of highly active antiretroviral therapy and regimen modification. *Aids*. 2008; 22:2097–2106. [PubMed: 18832873]
15. Tentori F, Albert JM, Young EW, Blayney MJ, Robinson BM, Pisoni RL, et al. The survival advantage for haemodialysis patients taking vitamin D is questioned: findings from the Dialysis Outcomes and Practice Patterns Study. *Nephrol Dial Transplant*. 2009; 24:963–972. [PubMed: 19028748]
16. Wiesbauer F, Heinze G, Mitterbauer C, Harnoncourt F, Horl WH, Oberbauer R. Statin use is associated with prolonged survival of renal transplant recipients. *J Am Soc Nephrol*. 2008; 19:2211–2218. [PubMed: 18650477]

17. Fox MP, Brooks DR, Kuhn L, Aldrovandi G, Sinkala M, Kankasa C, et al. Role of breastfeeding cessation in mediating the relationship between maternal HIV disease stage and increased child mortality among HIV-exposed uninfected children. *Int J Epidemiol.* 2009; 38:569–576. [PubMed: 19047077]
18. Westreich D, MacPhail P, Van Rie A, Malope-Kgokong B, Ive P, Rubel D, et al. Effect of pulmonary tuberculosis on mortality in patients receiving HAART. *AIDS.* 2009; 23:707–715. [PubMed: 19279444]
19. Westreich DJ, Sanne I, Maskew M, Malope-Kgokong B, Conradie F, Majuba P, et al. Tuberculosis treatment and risk of stavudine substitution in first-line antiretroviral therapy. *Clin Infect Dis.* 2009; 48:1617–1623. [PubMed: 19385733]
20. Cox DR. Regression models and life-tables. *Journal of the Royal Statistical Society Series B (Methodological).* 1972; 34:187–220.
21. Cole SR, Frangakis CE. The consistency statement in causal inference: a definition or an assumption? *Epidemiology.* 2009; 20:3–5. [PubMed: 19234395]
22. Mortimer KM, Neugebauer R, van der Laan M, Tager IB. An application of model-fitting procedures for marginal structural models. *Am J Epidemiol.* 2005; 162:382–388. [PubMed: 16014771]
23. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol.* 2008; 168:656–664. [PubMed: 18682488]
24. Robins JM. Marginal Structural Models versus Structural Nested Models as Tools for Causal Inference. In: Willard, Miller J., editor. *Statistical Models in Epidemiology, the Environment, and Clinical Trials The IMA Volumes in Mathematics and its Applications.* NY: Springer-Verlag; 1999. p. 95-134.
25. Tsiatis, AA. *Semiparametric Theory and Missing Data.* New York: Springer; 2006.
26. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika.* 1983; 70:41–55.
27. Cole SR, Hernán MA. Adjusted survival curves with inverse probability weights. *Comput Methods Programs Biomed.* 2004; 75:45–49. [PubMed: 15158046]
28. Westreich D, Cole SR, Tien PC, Chmiel JS, Kingsley L, Funk MJ, et al. Time scale and adjusted survival curves for marginal structural Cox models. In review at *American Journal of Epidemiology.* 2009
29. Cameron DW, Heath-Chiozzi M, Danner S, Cohen C, Kravcik S, Maurath C, et al. Randomised placebo-controlled trial of ritonavir in advanced HIV-1 disease. The Advanced HIV Disease Ritonavir Study Group. *Lancet.* 1998; 351:543–549. [PubMed: 9492772]
30. Hammer SM, Squires KE, Hughes MD, Grimes JM, Demeter LM, Currier JS, et al. A controlled trial of two nucleoside analogues plus didanosine in persons with human immunodeficiency virus infection and CD4 cell counts of 200 per cubic millimeter or less. AIDS Clinical Trials Group 320 Study Team. *N Engl J Med.* 1997; 337:725–733. [PubMed: 9287227]
31. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology.* 2000; 11:550–560. [PubMed: 10955408]
32. Pearl J. Causal Diagrams for Empirical Research. *Biometrika.* 1995; 82:669–688.
33. Hernán MA. A definition of causal effect for epidemiological research. *J Epidemiol Community Health.* 2004; 58:265–271. [PubMed: 15026432]
34. Rubin DB. Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. *Journal of Educational Psychology.* 1974; 66
35. Albrecht MA, Bosch RJ, Hammer SM, Liou SH, Kessler H, Para MF, et al. Nelfinavir, efavirenz, or both after the failure of nucleoside treatment of HIV infection. *N Engl J Med.* 2001; 345:398–407. [PubMed: 11496850]
36. Mellors JW, Munoz A, Giorgi JV, Margolick JB, Tassoni CJ, Gupta P, et al. Plasma viral load and CD4+ lymphocytes as prognostic markers of HIV-1 infection. *Ann Intern Med.* 1997; 126:946–954. [PubMed: 9182471]
37. Burton A, Altman DG, Royston P, Holder RL. The design of simulation studies in medical statistics. *Stat Med.* 2006; 25:4279–4292. [PubMed: 16947139]

38. Greenland S. Absence of confounding does not correspond to collapsibility of the rate ratio or rate difference. *Epidemiology*. 1996; 7:498–501. [PubMed: 8862980]
39. Young, JG.; Hernán, MA.; Picciotto, S.; Robins, JM. Simulation from Structural Survival Models under Complex Time-Varying Data Structures; Joint Statistical Meeting; Denver, CO. 2008.
40. Bodnar LM, Davidian M, Siega-Riz AM, Tsiatis AA. Marginal structural models for analyzing causal effects of time-dependent treatments: an application in perinatal epidemiology. *Am J Epidemiol*. 2004; 159:926–934. [PubMed: 15128604]
41. Young JG, Hernan MA, Picciotto S, Robins JM. Relation between three classes of structural models for the effect of a time-varying exposure on survival. *Lifetime Data Anal*. 2010; 16:71–84. [PubMed: 19894116]
42. Taubman SL, Robins JM, Mittleman MA, Hernan MA. Intervening on risk factors for coronary heart disease: an application of the parametric g-formula. *Int J Epidemiol*. 2009; 38:1599–1611. [PubMed: 19389875]

Figure 1. Effect of highly active antiretroviral therapy (compared to no therapy) on time to AIDS or death under (A) crude and (B) weighted cumulative incidence curves.

Figure 2.

Causal diagram: T^0 represents the unmeasured event time under no treatment, $X(k)$ represents the time-varying treatment at assessment k , $Z(k)$ represents a time-varying confounder at assessment k affected by prior time-varying treatment, and T represents the observed event time.

Table 1

Simulated bias for unadjusted, covariate-adjusted and stabilized and unstabilized inverse probability-of-treatment weighted hazard ratios across 40 scenarios, 5,000 trials per scenario.

	Scenario				Crude	Adjusted	Weighted	
	N	Pr(X(t)=1)	HR	exp(β_1)			unstabilized	stabilized
1	1500	0.5	0.5	2.0	1.234	0.530	0.004	0.005
2	1500	0.5	1.0	2.0	1.217	0.583	0.005	0.004
3	1500	0.5	0.5	1.5	1.240	0.287	0.005	0.006
4	1500	0.5	1.0	1.5	1.218	0.340	0.009	0.008
5	1500	0.5	0.5	1.0	1.227	-0.055	0.010	0.010
6	1500	0.5	1.0	1.0	1.205	-0.003	0.005	0.005
7	1500	0.5	0.5	0.75	1.207	-0.305	0.006	0.005
8	1500	0.5	1.0	0.75	1.189	-0.247	0.002	0.002
9	1500	0.5	0.5	0.5	1.169	-0.652	0.003	0.002
10	1500	0.5	1.0	0.5	1.159	-0.587	0.009	0.010
11	1500	0.2	0.5	2.0	1.201	0.554	-0.337	0.006
12	1500	0.2	1.0	2.0	1.184	0.608	-0.151	0.003
13	1500	0.2	0.5	1.5	1.200	0.297	-0.248	0.013
14	1500	0.2	1.0	1.5	1.186	0.348	-0.115	0.002
15	1500	0.2	0.5	1.0	1.172	-0.071	-0.278	-0.011
16	1500	0.2	1.0	1.0	1.164	-0.012	-0.178	-0.010
17	1500	0.2	0.5	0.75	1.161	-0.313	-0.266	-0.014
18	1500	0.2	1.0	0.75	1.157	-0.242	-0.171	-0.002
19	1500	0.2	0.5	0.5	1.126	-0.641	-0.288	-0.022
20	1500	0.2	1.0	0.5	1.144	-0.547	-0.030	0.003
21	200	0.5	0.5	2.0	1.268	0.543	0.078	0.072
22	200	0.5	1.0	2.0	1.231	0.620	0.079	0.070
23	200	0.5	0.5	1.5	1.234	0.251	0.030	0.027
24	200	0.5	1.0	1.5	1.223	0.352	0.052	0.047
25	200	0.5	0.5	1.0	1.222	-0.114	0.036	0.035
26	200	0.5	1.0	1.0	1.216	0.008	0.044	0.048
27	200	0.5	0.5	0.75	1.182	-0.389	0.015	0.015

	Scenario			Crude	Adjusted	Weighted	
	N	Pr(X(0)=1)	HR			unadjusted	stabilized
28	200	0.5	1.0	0.75	1.192	-0.257	0.058
29	200	0.5	0.5	0.5	1.187	-0.733	0.048
30	200	0.5	1.0	0.5	1.186	-0.595	0.074
31	200	0.2	0.5	2.0	1.154	0.491	-0.969
32	200	0.2	1.0	2.0	1.132	0.604	-0.694
33	200	0.2	0.5	1.5	1.142	0.207	-0.908
34	200	0.2	1.0	1.5	1.153	0.334	-0.633
35	200	0.2	0.5	1.0	1.189	-0.085	-0.923
36	200	0.2	1.0	1.0	1.157	-0.014	-0.602
37	200	0.2	0.5	0.75	1.158	-0.373	-0.778
38	200	0.2	1.0	0.75	1.154	-0.257	-0.734
39	200	0.2	0.5	0.5	1.106	-0.750	-0.889
40	200	0.2	1.0	0.5	1.131	-0.598	-0.617
							0.006

Table 2

Simulated standard error for unadjusted, covariate-adjusted and stabilized and unstabilized inverse probability-of-treatment weighted hazard ratios across 40 scenarios, 5,000 trials per scenario.

	Scenario				Crude	Adjusted	Weighted	
	N	Pr(X(t)=1)	HR	exp(β_1)			unstabilized	stabilized
1	1500	0.5	0.5	2.0	0.368	0.376	0.434	0.433
2	1500	0.5	1.0	2.0	0.369	0.377	0.440	0.439
3	1500	0.5	0.5	1.5	0.365	0.375	0.432	0.431
4	1500	0.5	1.0	1.5	0.366	0.375	0.435	0.435
5	1500	0.5	0.5	1.0	0.359	0.374	0.425	0.424
6	1500	0.5	1.0	1.0	0.361	0.374	0.430	0.429
7	1500	0.5	0.5	0.75	0.356	0.375	0.421	0.420
8	1500	0.5	1.0	0.75	0.357	0.374	0.426	0.425
9	1500	0.5	0.5	0.5	0.352	0.377	0.415	0.414
10	1500	0.5	1.0	0.5	0.353	0.377	0.420	0.419
11	1500	0.2	0.5	2.0	0.449	0.456	1.778	0.508
12	1500	0.2	1.0	2.0	0.442	0.449	1.789	0.505
13	1500	0.2	0.5	1.5	0.446	0.455	1.773	0.505
14	1500	0.2	1.0	1.5	0.439	0.447	1.791	0.501
15	1500	0.2	0.5	1.0	0.442	0.454	1.739	0.501
16	1500	0.2	1.0	1.0	0.435	0.446	1.743	0.496
17	1500	0.2	0.5	0.75	0.439	0.455	1.720	0.498
18	1500	0.2	1.0	0.75	0.432	0.446	1.732	0.493
19	1500	0.2	0.5	0.5	0.436	0.457	1.693	0.494
20	1500	0.2	1.0	0.5	0.428	0.447	1.706	0.489
21	200	0.5	0.5	2.0	1.022	1.100	1.177	1.158
22	200	0.5	1.0	2.0	1.026	1.102	1.187	1.168
23	200	0.5	0.5	1.5	1.011	1.093	1.165	1.145
24	200	0.5	1.0	1.5	1.017	1.096	1.184	1.165
25	200	0.5	0.5	1.0	0.999	1.097	1.152	1.133
26	200	0.5	1.0	1.0	1.001	1.092	1.167	1.147
27	200	0.5	0.5	0.75	0.988	1.096	1.144	1.124

	Scenario			Crude	Adjusted	Weighted	
	N	Pr(X(0)=1)	HR			unadjusted	stabilized
28	200	0.5	1.0	0.75	0.992	1.092	1.133
29	200	0.5	0.5	0.5	0.978	1.109	1.115
30	200	0.5	1.0	0.5	0.979	1.102	1.125
31	200	0.2	0.5	2.0	1.248	1.336	1.379
32	200	0.2	1.0	2.0	1.233	1.318	1.368
33	200	0.2	0.5	1.5	1.242	1.331	1.370
34	200	0.2	1.0	1.5	1.223	1.331	1.361
35	200	0.2	0.5	1.0	1.229	1.329	1.353
36	200	0.2	1.0	1.0	1.211	1.309	1.351
37	200	0.2	0.5	0.75	1.221	1.331	1.351
38	200	0.2	1.0	0.75	1.201	1.305	1.336
39	200	0.2	0.5	0.5	1.212	1.336	1.342
40	200	0.2	1.0	0.5	1.193	1.313	1.335

Table 3

Simulated root mean squared error for unadjusted, covariate-adjusted and stabilized and unstabilized inverse probability-of-treatment weighted hazard ratios across 40 scenarios, 5,000 trials per scenario.

	Scenario				Crude	Adjusted	Weighted	
	N	Pr(X(t)=1)	HR	exp(β_1)			unstabilized	stabilized
1	1500	0.5	0.5	2.0	1.293	0.702	0.583	0.582
2	1500	0.5	1.0	2.0	1.278	0.740	0.592	0.591
3	1500	0.5	0.5	1.5	1.298	0.570	0.583	0.581
4	1500	0.5	1.0	1.5	1.277	0.595	0.592	0.591
5	1500	0.5	0.5	1.0	1.284	0.510	0.574	0.573
6	1500	0.5	1.0	1.0	1.263	0.509	0.579	0.577
7	1500	0.5	0.5	0.75	1.264	0.580	0.568	0.566
8	1500	0.5	1.0	0.75	1.247	0.553	0.576	0.574
9	1500	0.5	0.5	0.5	1.226	0.789	0.560	0.559
10	1500	0.5	1.0	0.5	1.217	0.743	0.570	0.569
11	1500	0.2	0.5	2.0	1.297	0.798	3.146	0.685
12	1500	0.2	1.0	2.0	1.278	0.824	3.190	0.682
13	1500	0.2	0.5	1.5	1.295	0.677	3.093	0.679
14	1500	0.2	1.0	1.5	1.279	0.685	3.109	0.678
15	1500	0.2	0.5	1.0	1.268	0.621	3.015	0.677
16	1500	0.2	1.0	1.0	1.257	0.610	3.012	0.678
17	1500	0.2	0.5	0.75	1.257	0.681	2.954	0.675
18	1500	0.2	1.0	0.75	1.249	0.647	2.993	0.673
19	1500	0.2	0.5	0.5	1.225	0.852	2.897	0.672
20	1500	0.2	1.0	0.5	1.235	0.782	2.942	0.658
21	200	0.5	0.5	2.0	1.791	1.602	1.628	1.596
22	200	0.5	1.0	2.0	1.776	1.626	1.646	1.616
23	200	0.5	0.5	1.5	1.763	1.537	1.612	1.582
24	200	0.5	1.0	1.5	1.753	1.549	1.623	1.590
25	200	0.5	0.5	1.0	1.735	1.521	1.576	1.547
26	200	0.5	1.0	1.0	1.744	1.521	1.607	1.578
27	200	0.5	0.5	0.75	1.709	1.559	1.585	1.550

	Scenario			Crude	Adjusted	Weighted	
	N	Pr(X(0)=1)	HR exp(β_1)			unstablized	stabilized
28	200	0.5	1.0	0.75	1.712	1.528	1.590
29	200	0.5	0.5	0.5	1.705	1.678	1.570
30	200	0.5	1.0	0.5	1.700	1.619	1.570
31	200	0.2	0.5	2.0	1.985	1.898	5.347
32	200	0.2	1.0	2.0	1.958	1.908	5.459
33	200	0.2	0.5	1.5	1.978	1.861	5.372
34	200	0.2	1.0	1.5	1.958	1.851	5.414
35	200	0.2	0.5	1.0	1.999	1.862	5.243
36	200	0.2	1.0	1.0	1.938	1.817	5.246
37	200	0.2	0.5	0.75	1.968	1.883	5.233
38	200	0.2	1.0	0.75	1.936	1.833	5.214
39	200	0.2	0.5	0.5	1.921	1.956	5.231
40	200	0.2	1.0	0.5	1.907	1.888	5.167
							1.823

Table 4

Simulated 95% confidence limit coverage probability for unadjusted, covariate-adjusted and stabilized and unstabilized inverse probability-of-treatment weighted hazard ratios across 40 scenarios, 5,000 trials per scenario.

	Scenario				Crude	Adjusted	Weighted	
	N	Pr(X(t)=1)	HR	exp(β_1)			unstabilized	stabilized
1	1500	0.5	0.5	2.0	0.075	0.705	0.955	0.955
2	1500	0.5	1.0	2.0	0.080	0.660	0.953	0.953
3	1500	0.5	0.5	1.5	0.077	0.880	0.949	0.949
4	1500	0.5	1.0	1.5	0.080	0.846	0.948	0.949
5	1500	0.5	0.5	1.0	0.068	0.943	0.947	0.947
6	1500	0.5	1.0	1.0	0.085	0.947	0.958	0.959
7	1500	0.5	0.5	0.75	0.069	0.870	0.951	0.950
8	1500	0.5	1.0	0.75	0.082	0.901	0.950	0.950
9	1500	0.5	0.5	0.5	0.081	0.601	0.950	0.949
10	1500	0.5	1.0	0.5	0.088	0.650	0.945	0.946
11	1500	0.2	0.5	2.0	0.233	0.771	0.714	0.950
12	1500	0.2	1.0	2.0	0.233	0.724	0.718	0.947
13	1500	0.2	0.5	1.5	0.233	0.890	0.728	0.951
14	1500	0.2	1.0	1.5	0.231	0.870	0.737	0.949
15	1500	0.2	0.5	1.0	0.242	0.941	0.732	0.947
16	1500	0.2	1.0	1.0	0.240	0.945	0.739	0.945
17	1500	0.2	0.5	0.75	0.249	0.896	0.741	0.950
18	1500	0.2	1.0	0.75	0.244	0.909	0.733	0.947
19	1500	0.2	0.5	0.5	0.270	0.714	0.738	0.947
20	1500	0.2	1.0	0.5	0.234	0.773	0.735	0.954
21	200	0.5	0.5	2.0	0.761	0.905	0.930	0.934
22	200	0.5	1.0	2.0	0.780	0.899	0.932	0.934
23	200	0.5	0.5	1.5	0.770	0.934	0.930	0.932
24	200	0.5	1.0	1.5	0.780	0.926	0.942	0.943
25	200	0.5	0.5	1.0	0.780	0.939	0.938	0.938
26	200	0.5	1.0	1.0	0.771	0.938	0.937	0.935
27	200	0.5	0.5	0.75	0.782	0.924	0.933	0.938

	Scenario			Crude	Adjusted	Weighted	
	N	Pr(X(0)=1)	HR			unadjusted	stabilized
28	200	0.5	1.0	0.75	0.777	0.934	0.938
29	200	0.5	0.5	0.5	0.772	0.880	0.933
30	200	0.5	1.0	0.5	0.770	0.907	0.938
31	200	0.2	0.5	2.0	0.836	0.922	0.945
32	200	0.2	1.0	2.0	0.842	0.906	0.939
33	200	0.2	0.5	1.5	0.836	0.928	0.942
34	200	0.2	1.0	1.5	0.829	0.921	0.941
35	200	0.2	0.5	1.0	0.818	0.933	0.939
36	200	0.2	1.0	1.0	0.830	0.939	0.939
37	200	0.2	0.5	0.75	0.819	0.931	0.933
38	200	0.2	1.0	0.75	0.826	0.928	0.937
39	200	0.2	0.5	0.5	0.835	0.908	0.943
40	200	0.2	1.0	0.5	0.834	0.921	0.942