WEB APPENDIX

Weight and survival estimation

The weight, $W(y_{(i)})$ is defined as,

$$W(y_{(i)}) = \begin{cases} \prod_{u=1}^{y_{(i)}} \dfrac{P[A(u)=0 \mid A(u-1)=0, B(u-1)=0, M(u-1)=0]}{P[A(u)=0 \mid A(u-1)=0, B(u-1)=0, M(u-1)=0, \bar{L}(u-1), V]}, & \text{if } A(y_{(i)})=0 \\ 0, & \text{if } A(y_{(i)})=1 \end{cases} \quad (1).$$

For observed visits prior to artificial censoring, the denominator of $W(y_{(i)})$ is a participant's probability of remaining not censored due to artificial censoring through visit $y_{(i)}$ given $\bar{L}(u-1)$, V, and that the participant remained event and censor-free prior to visit $y_{(i)}$. Similarly, the numerator is a participant's probability of remaining not artificially censored through visit $y_{(i)}$ given that the participant remained event and censor-free prior to visit $y_{(i)}$. For visits on or after artificial censoring, $W(y_{(i)}) = 0$.

To estimate the numerator and denominator of the weights for a given participant and visit, the following two pooled logistic regression models can be fit,

$$\text{logit } P[A(u)=0 \mid A(u-1)=0, B(u-1)=0, M(u-1)=0] = \alpha_{0u} \quad (2)$$

$$\text{logit } P[A(u)=0 \mid A(u-1)=0, B(u-1)=0, M(u-1)=0, \bar{L}(u-1), V] = \beta_{0u} + \beta_1' V + \beta_2' \bar{L}(u-1) \quad (3)$$

where logit $p = \ln(p/(1-p))$. The parameters α_{0u} and β_{0u} are the visit specific intercepts without and with inclusion of the time-fixed and time-varying common predictors in the pooled model, respectively. The parameter β_1' is the transpose of the column vector of log hazard ratios

for the time-fixed common predictors comprising V and β_2' is the transpose of the column vector of log hazard ratios for the time-varying common predictor histories that comprise $\bar{L}(u-1)$.

Assuming exchangeability and correct model specification, the following equations adapted from Robins and Finkelstein (3) may be used to estimate the survival function corrected for selection bias due to artificial censoring.

$$\hat{\lambda}(t_j) = \frac{\sum_{i \in D_j} \hat{W}_i(t_j)}{\sum_{i \in R_j} \hat{W}_i(t_j)} \tag{4}$$

$$\hat{S}(t) = \prod_{t_j \leq t} \left[1 - \hat{\lambda}(t_j) \right] \tag{5}$$

Let t_j be the time corresponding to the *jth* visit where an event was observed to occur during the visit, R_j be the subset of the baseline cohort where $y_i \geq t_j$, and D_j is the subset of R_j who develop the outcome at time t_j. Therefore, in equations (4) and (5), $\hat{W}_i(t_j)$ is the estimated weight for participant i at time t_j, $\hat{\lambda}(t_j)$ is the estimated hazard at time t_j, and $\hat{S}(t)$ is the estimated survival at time t. For the case when artificial censoring is non-informative, $\hat{W}_i(t_j) = 1$ for each i at all t_j's. Thus equation (4) reduces to the classical representation of the hazard at time t_j, d_j / r_j, where d_j and r_j are the number of observed events and risk set size at time t_j, respectively.

Simulated bias and mean squared error estimates of the survival function in various settings

Simulations were performed to demonstrate the bias in the IPCW survival function estimate that can occur in the context of each of the following: small sample size, strong selection bias, unmeasured common predictors, and model misspecification. For all examined scenarios, 500 simulations of sample size 50 or 500 were performed. Failure times were generated from a Weibull distribution (i.e., $S(t) = \exp\{-(t/\lambda)^{\sigma}\}$) where λ and σ for the baseline survival function was 9.0 and 2.5, respectively. Failure times were generated as a function of time-fixed binary covariates z_1 and z_2 where the relative hazard of failure was specified to be 12.2 for both covariates. The prevalence of z_1 and z_2 was 50% and the proportion of failure times that were censored was 60%.

Scenario (I) corresponded to a censoring mechanism that does not induce selection bias. The sample size was 500 and censoring times were generated from an exponential distribution (i.e., $S(c) = \exp\{-(c/\mu)\}$) independently of z_1 and z_2. The μ for the baseline survival function was 2.7. Scenario (II) corresponded to a censoring mechanism that induces selection bias. The sample size was 500 and censoring times were generated from an exponential distribution as a function of z_1 and z_2. The μ for the baseline survival function was 12.2 and the relative hazard of censoring as a function of z_1 and z_2 was 4.5. Scenario (III) corresponded to a censoring mechanism that induces selection bias in the context of small sample size. The sample size was 50 and censoring times were generated from an exponential distribution as a function of z_1 and z_2. The μ for the baseline survival function was 12.2 and the relative hazard of censoring as a function of z_1 and z_2 was 4.5.

Scenario (IV) corresponded to a censoring mechanism that induces strong selection bias. The sample size was 500 and censoring times were generated from an exponential distribution as a function of z_1 and z_2. The μ for the baseline survival function was 33.1 and the relative hazard of censoring as a function of z_1 and z_2 was 12.2. Scenario (V) corresponded to a censoring mechanism that induces selection bias in the context of an unmeasured common predictor, z_2. The sample size was 500 and censoring times were generated from an exponential distribution as a function of z_1 and z_2. The μ for the baseline survival function was 12.2 and the relative hazard of censoring as a function of z_1 and z_2 was 4.5. Scenario (VI) corresponded to a censoring mechanism that induces selection bias in the context of a misspecified common predictor, z_2. The sample size was 500 and censoring times were generated from an exponential distribution as a function of z_1 and z_2. The μ for the baseline survival function was 12.2. The relative hazard of censoring as a function of z_1 and z_2 was 4.5.

Web Figures 1 and 2 show mean survival and mean squared error (MSE) for the standard KM and IPCW estimates for each of the above described simulation scenarios. Scenario (I) demonstrates that in the absence of selection bias due to censoring and a sufficiently large sample size the standard KM and IPCW estimators can be used to obtain unbiased estimates of survival with MSEs equal to the variance. However, as in scenario (II), in the presence of selection bias the standard KM estimator will likely be biased, while the IPCW estimator when necessary assumptions are met will yield an unbiased estimate of the survival function. When necessary assumptions are violated by small sample size (scenario (III)), strong selection bias (scenario (IV)), an unmeasured common predictor (scenario (V)), or model misspecification (scenario (VI)), the IPCW survival function may be biased with a substantial MSE. However like

in scenario (VI), violation of necessary assumptions does not always result in biased estimates. A similar pattern was observed for the median survival for the standard KM and IPCW estimates for each of the examined simulation scenarios.

Distribution of AIDS events by calendar time and HAART initiation

Web Figure 3 shows the distribution of AIDS events by calendar time and HAART initiation among 467 seroconverters in the MACS population. Nodes in dotted boxes correspond to participants who initiated HAART during follow-up.

For the administrative censoring at 1996 analysis, the node labeled "205 (AIDS ≤1996)" contributed the 205 AIDS events. The "2 (HAART <1996)" node, 26 from the "27 (HAART ≥1996)" node, the "13 (AIDS while HAART naïve)" node, 118 from the "138 (initiated HAART) node and 71 from the "82 (remained HAART naïve)" node contributed the 230 censored observations. Differences between the specified sample size for a given node and the actual number of participants contributing to the administrative censoring at 1996 analysis are due to the exclusion of participants who seroconverted at or after 1996 from the administrative censoring at 1996 analysis.

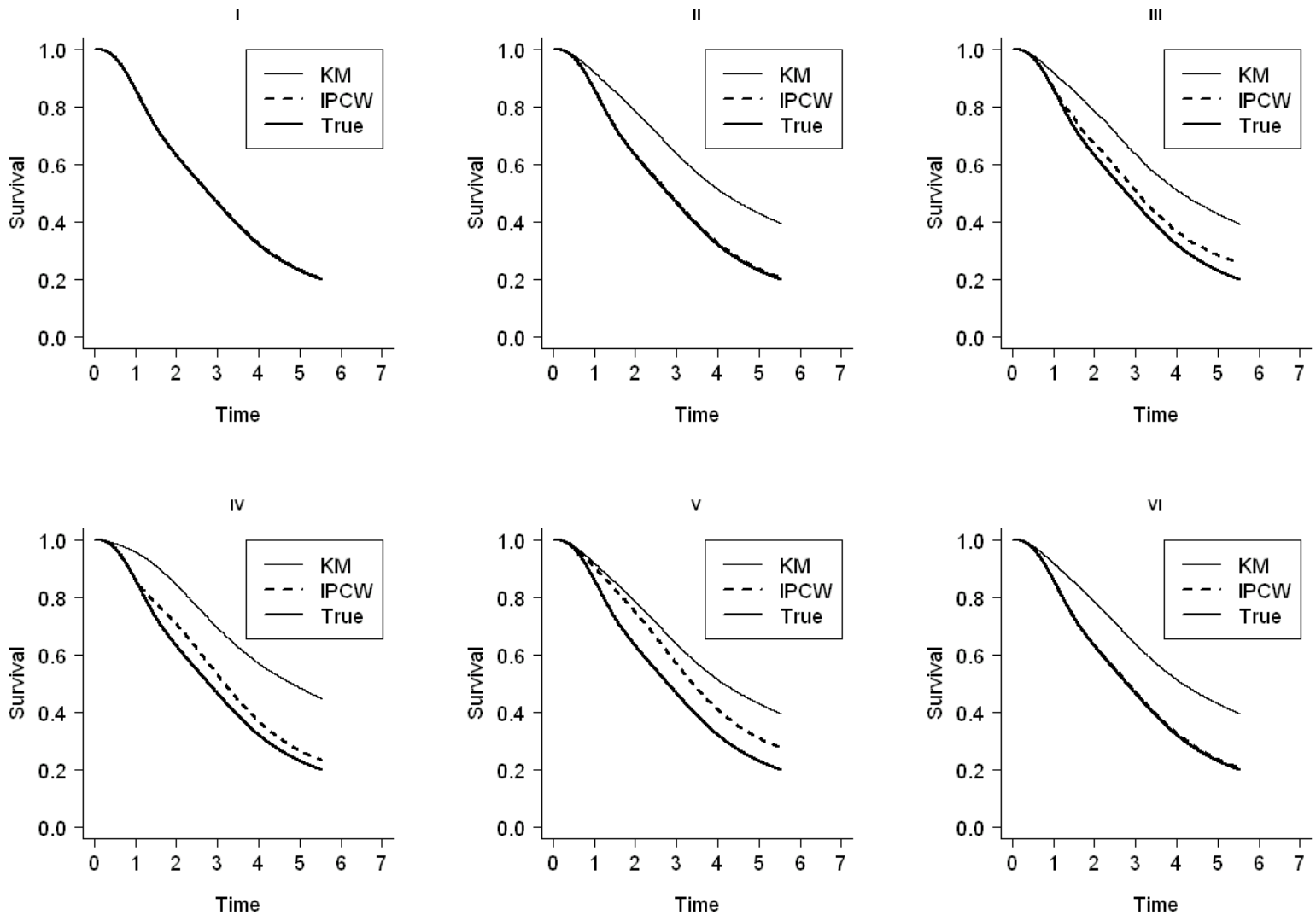
For the artificial censoring at HAART initiation analysis, the nodes labeled "205 (AIDS ≤1996)" and "13 (AIDS while HAART naïve)" contributed the 218 AIDS events. The nodes labeled "220 AIDS-free" and "29 Post-HAART initiation" contributed the 249 censored observations.

Distribution of weights for HAART initiation

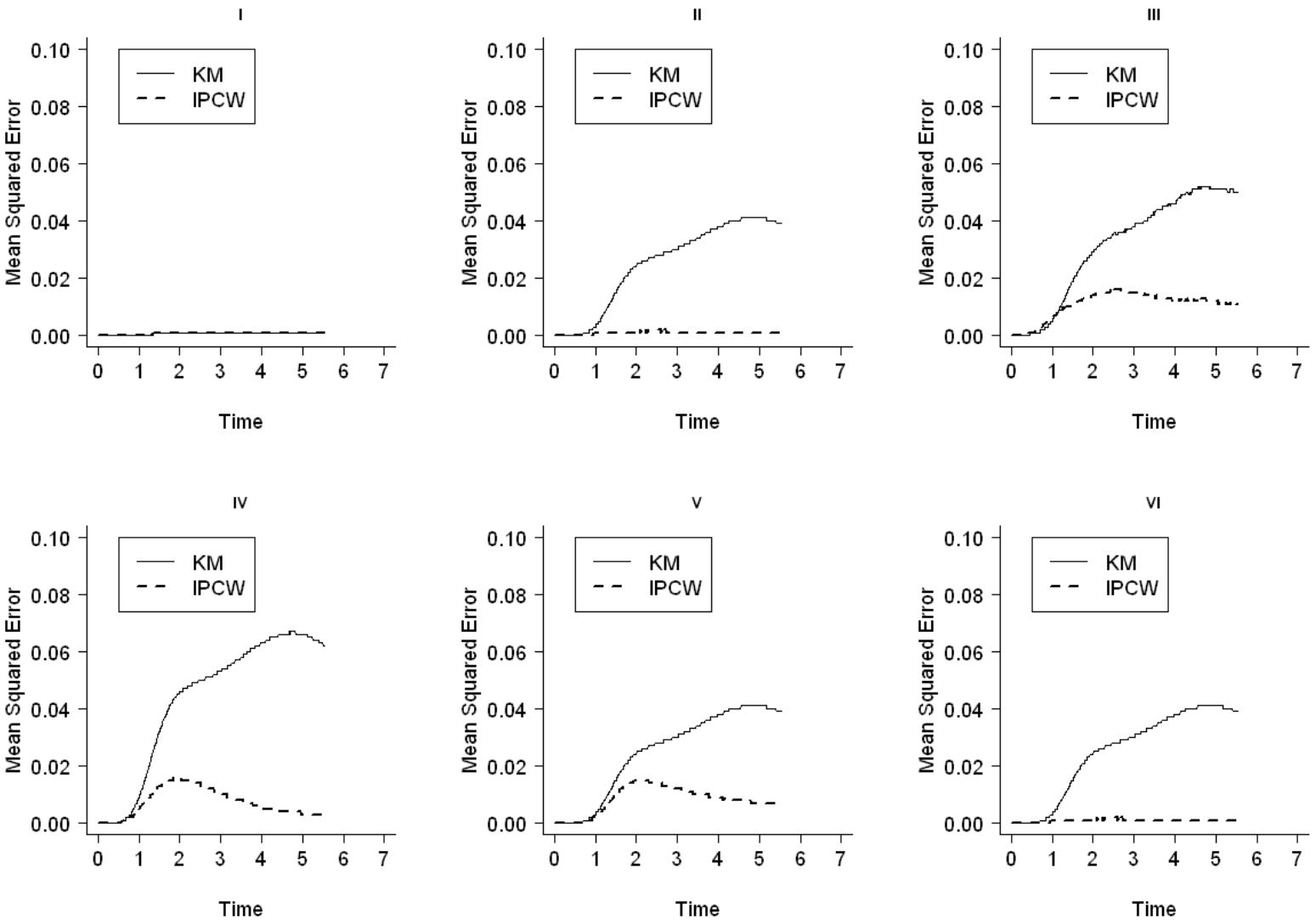
Web Figure 4 shows the distribution of the weights for HAART initiation among the 467 seroconverters used in the IPCW analysis by years since infection. Lines at the 1st and 99th percentile are included in Web Figure 4. For all times, the minimum, 1st percentile, 25th percentile, median, 75th percentile, 99th percentile, and maximum weights were, 0.10, 0.30, 1.00, 1.00, 1.00, 1.80, and 7.47, respectively. The mean weight was 0.98. The standard deviation of the weights was 0.26. Although the range of the weights increased with years since infection, no weights were extreme.

Web Figure 1. Simulated Bias of Estimates of the Survival Function in Various Settings.

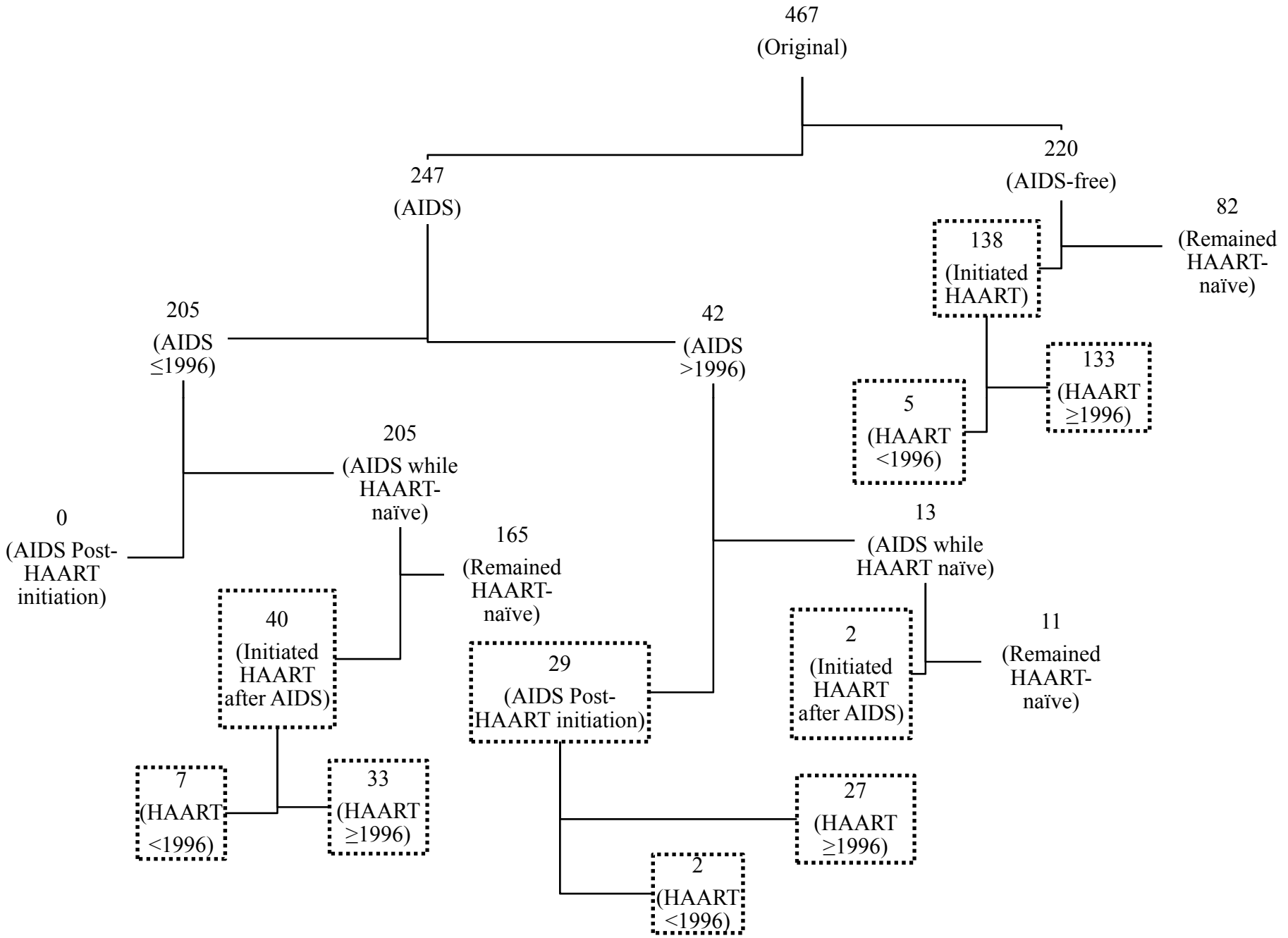
Abbreviations: IPCW, inverse probability-of-censoring weights; KM, Kaplan-Meier.


I
II
III
IV
V
VI
KM
IPCW
True
Survival
Time
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5
6
7

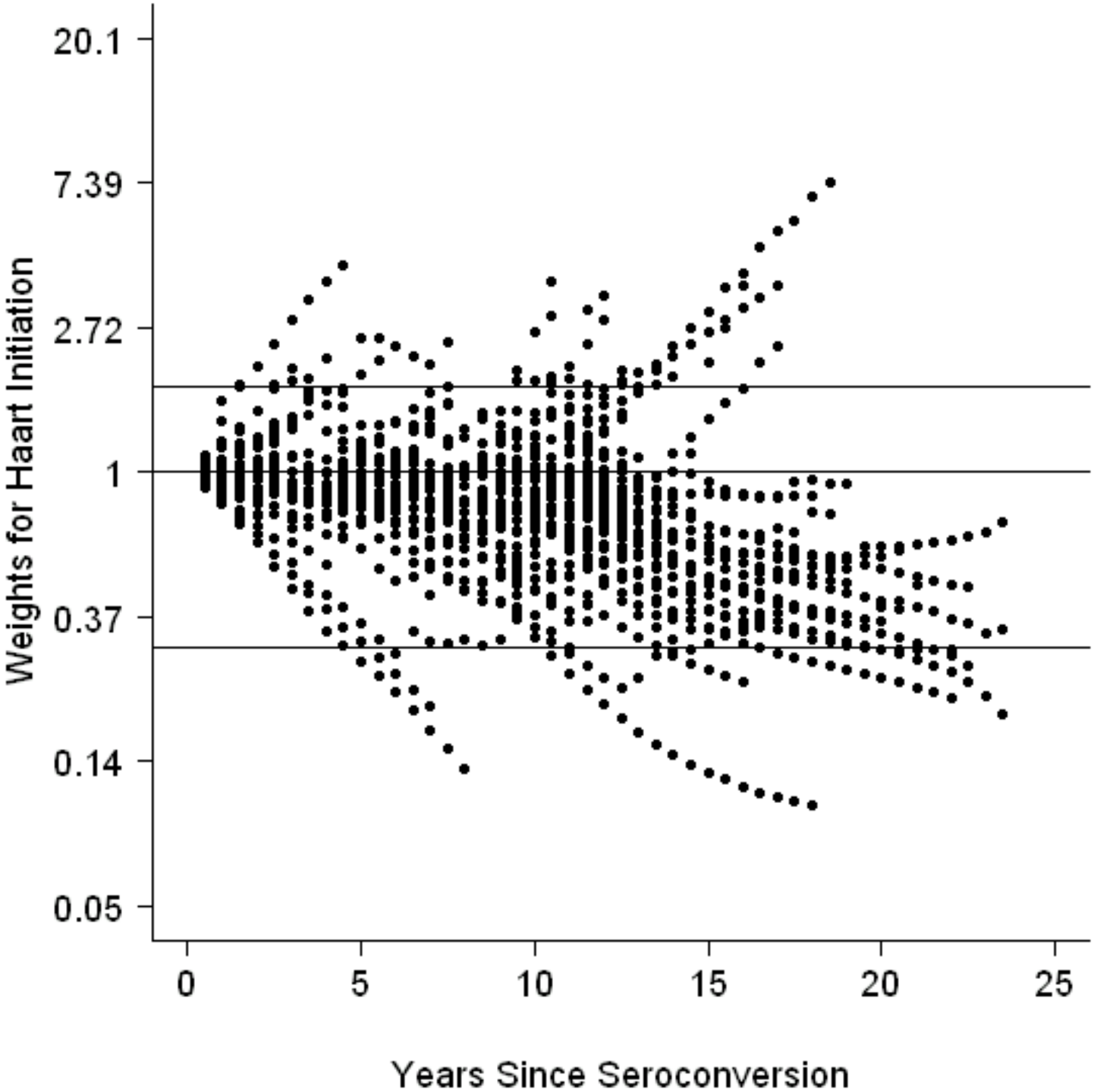
Web Figure 2. Simulated Mean Squared Error of Estimates of the Survival Function in Various Settings. Abbreviations: IPCW, inverse probability-of-censoring weights; KM, Kaplan-Meier.


I
II
III
IV
V
VI
KM
IPCW
Mean Squared Error
Time
0.00
0.02
0.04
0.06
0.08
0.10
0 1 2 3 4 5 6 7

Web Figure 3. Distribution of AIDS Events by Calendar time and HAART Initiation among 467 Seroconverters in the Multicenter AIDS cohort study, 1984-2008. Nodes in Dotted Boxes Correspond to Participants who Initiated HAART during Follow-up. Abbreviations: HAART, Highly Active Antiretroviral Therapy.


467
(Original)
247
(AIDS)
220
(AIDS-free)
82
(Remained HAART-naïve)
138
(Initiated HAART)
205
(AIDS ≤1996)
42
(AIDS >1996)
133
(HAART ≥1996)
5
(HAART <1996)
205
(AIDS while HAART-naïve)
0
(AIDS Post-HAART initiation)
165
(Remained HAART-naïve)
13
(AIDS while HAART naïve)
40
(Initiated HAART after AIDS)
29
(AIDS Post-HAART initiation)
2
(Initiated HAART after AIDS)
11
(Remained HAART-naïve)
7
(HAART <1996)
33
(HAART ≥1996)
27
(HAART ≥1996)
2
(HAART <1996)

Web Figure 4. Distribution of Weights for HAART Initiation by Years since Seroconversion among 467 Seroconverters in the Multicenter AIDS Cohort Study, 1984-2008.


20.1
7.39
2.72
1
0.37
0.14
0.05
Weights for Haart Initiation
0
5
10
15
20
25
Years Since Seroconversion