Supporting Information for "Causal comparative effectiveness analysis of dynamic continuous-time treatment initiation rules with sparsely measured outcomes and death" by Liangyuan Hu and Joseph W. Hogan

December 21, 2018

Appendix A

A.1 Model specification for imputations

The specific formulation used in our example models the CD4 trajectory in terms of 4 parameters: an intercept β_0, pre-treatment slope β_1, an instantaneous effect of ART at the time of initiation, captured in terms of a jump of size β_2, and a post-treatment slope β_4. Each of these temporal components is allowed to vary by individual, giving rise to subject-specific random effects $b_i = (b_{0i}, b_{1i}, b_{2i}, b_{3i})^\top$. The model further includes main effects of baseline covariates, plus an interaction between baseline covariates and the pre- and post-treatment slopes. The specific formulation is

$$\begin{aligned} m_i(t) &= (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})t + (\beta_2 + b_{2i})N_i^A(t) + (\beta_3 + b_{3i})(t - A_i)_+ \\ &\quad + L_i(0)\left\{\psi_0 + \psi_1 t + \psi_2 N_i^A(t) + \psi_3 (t - A_i)_+\right\}, \end{aligned} \tag{1}$$

where $a_+ = \max(0, a)$ is the positive part of a, and the ψ parameters represent coefficients for the main effect of $L_i(0)$ and its interaction with the trajectory terms. Hence the time trajectory differs by covariate profile, and within covariate profile further varies by individual, thus providing a rich structure for modeling true CD4 count. Components of $L_i(0)$ include the following: age at baseline (modeled with cubic spline), CDC symptom class (5-level categorical variable with levels mild (A), moderate (B), severe (C), asymptomatic (N) and missing); gender (1=male, 0=female); and CD4 cell count category (5 level categorical variable with categories 0-199, 200-349, 350-499, over 500, and missing). Time varying covariates include treatment initiation $N^A(t)$. Parameters for the fitted model appear in Table 1. Residual-versus-fitted plots (Figure 1) and examination of individual-specific fitted curves are used to assess fit of the CD4 submodel (see Figure 2 for a sample of 9 individuals).

The survival submodel specification is given in equation (11) in the main text. Baseline covariates include age at baseline (fitted using cubic spline), gender, CD4 categories as listed above and CDC class as listed above. Time varying covariates include treatment initiation $N^A(t)$ and current CD4 count $\widehat{m}(t)$. The fitted model appears in Table 3; the shape of the relationship between mortality and $m(t)$ appears in Figure 3, indicating a strong negative and nonlinear relationship (higher CD4 implies lower mortality).

We tested the proportional hazards assumption for each covariate using Schoenfeld residuals; test results appear in Table 4. There is no evidence that any of the covariates violate the proportional hazards assumpton.

A.2 Imputation algorithm

For those whose follow up is censored prior to t^*, we first impute a death indicator at t^*; for those whose imputed status is 'alive', we then impute a CD4 count. For those who are in follow-up at t^* but do not have an observed CD4 proximal to t^*, we impute a CD4 count at t^*. The specific imputation strategy is as follows

1. For those still in follow up at t^* but missing a proximal CD4 count, impute CD4 at time t^* by drawing $\widetilde{X} \sim N(\widehat{m}_i(t^*), \widehat{\sigma}^2(t^*))$, where $\widehat{m}_i(t^*)$ is the individual-specific prediction of true CD4 count.

2. For those whose follow up is censored at $C_i < t^*$,

   (a) Calculate $\widehat{S}_i^T(t^* \mid T > C_i)$, the estimated probability of survival at time t^* conditional on surviving to time C_i. Referring to equation (11) in the main text, note that

   $$\widehat{S}_i^T(t) \quad = \quad \exp\left[- \int_0^t \exp\left\{ g_1(\widehat{m}_i(s); \widehat{\gamma}_1) + g_2(L_i(0), N_i^A(s); \widehat{\gamma}_2) \right\} d\widehat{\Lambda}_0^T(s) \right],$$

   where $\widehat{\Lambda}_0^T(s)$ is the estimated cumulative baseline hazard function for mortality. Hence $\widehat{S}_i^T(t)$ can be estimated directly from the fitted hazard model, and $\widehat{S}_i^T(t^* \mid T > C_i) = \widehat{S}_i^T(t^*)/\widehat{S}_i^T(C_i)$.

   (b) Draw a binary death indicator $\widetilde{D}_i \sim \text{Bernoulli}(1 - \widehat{S}_i^T(t \mid T > C_i))$.

   (c) If $\widetilde{D}_i = 1$ then set $\widetilde{X}_i = 0$; else draw a missing CD4 count $\widetilde{Y}$ as in step 1 above.

3. Return $\widetilde{X}_i$.

A.3 Variance calculations

Let M denote the number of imputations for each missing observation, leading to M completed datasets. For each completed dataset, bootstrap resampling is used to compute the point estimator for the target parameters and its within-imputation variance. Let S denote the number of bootstrap samples on each of the M completed datasets, and for $s = 1, \ldots, S$ and $m = 1, \ldots, M$, let $\widehat{\theta}_{qj}^{(s,m)}$, $j = 1, 2, 3$, denote the point estimate of θ_{qj} derived from bootstrap sample s drawn from imputed dataset m. For imputed dataset m, the average of the point estimates across bootstrap samples is $\widehat{\theta}_{qj}^{(m)} = S^{-1} \sum_{s=1}^{S} \widehat{\theta}_{qj}^{(s,m)}$; hence the within-imputation variance estimator of $\mathrm{Var}(\widehat{\theta}^{(m)})$ is $\widehat{V}^{(m)} = (S-1)^{-1} \sum_{s=1}^{S} \left(\widehat{\theta}_{qj}^{(s,m)} - \widehat{\theta}_{qj}^{(m)}\right)^2$, and the estimate of within-imputation variance is

$$W = \frac{1}{M} \sum_{m=1}^{M} \widehat{V}^{(m)}.$$

The between-imputation variance estimator is

$$B = \frac{1}{M-1} \sum_{m=1}^{M} \left(\widehat{\theta}_{qj}^{(m)} - \widehat{\theta}_{qj}\right)^2,$$

where $\widehat{\theta}_{qj} = M^{-1} \sum_{m=1}^{M} \widehat{\theta}_{qj}^{(m)}$ is the mean over all imputation-specific estimates. Hence the estimator of total variance is $\widehat{\mathrm{Var}}(\widehat{\theta}_{qj}) = W + (1 + 1/M)B$.

To compute confidence intervals for point estimates, we assumed the (bootstrap) sampling distributions were well-approximated by a normal distribution, which was verified using q-q plots. Confidence intervals for mortality rates were based on a logit transformation of the sampling distribution.

Appendix B

B.1 Sensitivity analysis for weight truncation

To assess the impact of weight truncation on the estimation of the causal effects of DTRs, we conduct a sensitivity analysis describing estimated differences in mortality rate and median of X_q between dynamic regimes $q = \infty$ and $q = 500$ for $t^* = 1$ year and $t^* = 2$ years. We truncate the stabilized regime weights at the top and bottom 5%, 2.5% and 0% (no truncation). The results, shown in Table 1, suggest the point estimates and the confidence intervals for treatment effect on mortality were unchanged with different weighting schemes. Point estimates and variation associated with treatment effect on the composite outcome increased with less truncation; the confidence intervals indicated greater variability but no change in substantive conclusion about treatment effect. For the denominator weight model, we tested the proportional hazards assumption for each term included in the model and found no violations of the assumption. We summarize the distribution of the estimated weights for DTRs $q = \infty$ and $q = 500$ in Table 2.

Table 1: Sensitivity analysis depicting estimated differences in mortality rate and median of X_q between dynamic regimes $q = \infty$ and $q = 500$ for $t^* = 1$ year and $t^* = 2$ years. Comparison includes weight IPTW truncation at top and bottom 5%, top and bottom 2.5%, and no truncation. The parameter $\theta_{q1} = P(X_q = 0) = F_{X_q}(0)$ is the mortality rate, and $\theta_{q2} = F^{-1}_{X_q}(\frac{1}{2})$ is the median of X_q. The 95% confidence intervals are shown below the point estimates.

	5%	2.5%	no truncation
$t^* = 1$			
$\widehat{\theta}_{q1}$	$-.008$	$-.008$	$-.008$
	$(-.015, -.001)$	$(-.014, -.001)$	$(-.014, -.001)$
$\widehat{\theta}_{q2}$	41	50	75
	(12, 70)	(14, 86)	(16, 134)
$t^* = 2$			
$\widehat{\theta}_{q1}$	$-.013$	$-.014$	$-.013$
	$(-.023, -.004)$	$(-.023, -.005)$	$(-.022, -.004)$
$\widehat{\theta}_{q2}$	51	56	81
	(14, 87)	(15, 98)	(11, 150)

Table 2: Distribution of inverse probability of treatment weights (IPTW) for dynamic treatment regimes $q = 500$ and $q = \infty$ at time points $t = 1$ and $t = 2$ years.

	min	2.5%	5%	median	95%	97.5%	max
$t^* = 1$							
$q = \infty$	0.86	3.57	5.46	18.01	66.10	83.20	112.00
$q = 500$	0.46	1.10	2.62	11.13	20.60	31.90	77.21
$t^* = 2$							
$q = \infty$	0.86	4.03	5.86	22.12	68.10	85.32	112.00
$q = 500$	0.68	1.80	3.04	16.14	26.60	33.50	85.12

Appendix C

C.1 Supplemental tables and figures

Table 3: Fitted CD4 submodel: Fixed effect estimates. See equation (1) for full model specification.

Number of Observations: 10036
Number of Groups: 1962

	Value	Std.Error	DF	t-value	p-value
(Intercept)	19.792310	0.5440822	8043	36.37743	0.0000
ns(age_c, df = 4)1	-0.826791	0.5297732	1948	-1.56065	0.1188
ns(age_c, df = 4)2	-1.313378	0.5219090	1948	-2.51649	0.0119
ns(age_c, df = 4)3	-4.221976	1.0584546	1948	-3.98881	0.0001
ns(age_c, df = 4)4	-1.298521	0.5451897	1948	-2.38178	0.0173
classA	-0.358940	0.4022471	1948	-0.89234	0.3723
classB	-0.281527	0.6411768	1948	-0.43908	0.6607
classC	-0.973655	0.6348106	1948	-1.53377	0.1252
classN	1.101933	0.4123147	1948	2.67255	0.0076
male	-0.696516	0.2411600	1948	-2.88819	0.0039
cd4.0	-9.822384	0.3967882	1948	-24.75473	0.0000
cd4.200	-1.266265	0.4349427	1948	-2.91134	0.0036
cd4.350	2.396883	0.4608261	1948	5.20127	0.0000
cd4.500	9.185741	0.4224343	1948	21.74478	0.0000
N^A(t)	3.498966	0.3951127	8043	8.85561	0.0000
N^A(t):classA	0.098775	0.4454290	8043	0.22175	0.8245
N^A(t):classB	-0.589170	0.6250478	8043	-0.94260	0.3459
N^A(t):classC	0.293096	0.6847882	8043	0.42801	0.6687
N^A(t):classN	-0.876555	0.5074726	8043	-1.72730	0.0842
N^A(t):male	0.505273	0.2750455	8043	1.83705	0.0662
N^A(t):cd4.0	4.769437	0.4305617	8043	11.07724	0.0000
N^A(t):cd4.200	2.250686	0.4743306	8043	4.74497	0.0000
N^A(t):cd4.350	-0.746502	0.5376364	8043	-1.38849	0.1650
N^A(t):cd4.500	-2.363437	0.5140559	8043	-4.59763	0.0000
N^A(t) : time to ARV	-0.001267	0.0003672	8043	-3.45096	0.0006
preArtMonths (t)	-0.054122	0.0206962	8043	-2.61508	0.0089
classA:preArtMonths	0.001707	0.0251862	8043	0.06778	0.9460
classB:preArtMonths	0.021122	0.0452437	8043	0.46685	0.6406
classC:preArtMonths	0.068495	0.0824450	8043	0.83079	0.4061
classN:preArtMonths	0.008884	0.0245975	8043	0.36116	0.7180
male:preArtMonths	0.030871	0.0174449	8043	1.76965	0.0768
cd4.0:preArtMonths	0.124723	0.0471944	8043	2.64274	0.0082
cd4.200:preArtMonths	-0.038744	0.0324292	8043	-1.19472	0.2322
cd4.350:preArtMonths	-0.016200	0.0273300	8043	-0.59277	0.5534
cd4.500:preArtMonths	-0.097350	0.0230936	8043	-4.21545	0.0000
postArtMonths (t-A)_+	0.067137	0.0246488	8043	2.72374	0.0065
classA:postArtMonths	-0.006954	0.0307434	8043	-0.22619	0.8211
classB:postArtMonths	-0.039469	0.0517607	8043	-0.76252	0.4458
classC:postArtMonths	-0.067821	0.0871196	8043	-0.77848	0.4363
classN:postArtMonths	-0.031397	0.0343065	8043	-0.91518	0.3601
male:postArtMonths	-0.071645	0.0211673	8043	-3.38470	0.0007
cd4.0:postArtMonths	-0.090715	0.0494893	8043	-1.83301	0.0668

Table 4: Fitted CD4 submodel: Random effects variance estimates and distribution of residuals.

```
Random effects:
 Formula: ~preArtMonths + postArtMonths | ptidno
 Structure: General positive-definite, Log-Cholesky parametrization
              StdDev     Corr
(Intercept)   3.91359199 (Intr) prArtM
preArtMonths  0.09076454  0.386
postArtMonths 0.16232370 -0.275 -0.454
Residual      3.96196571

Standardized Within-Group Residuals:
         Min           Q1          Med           Q3          Max
-4.710810881 -0.506312643 -0.004719634  0.484378098  6.595537128
```

Table 5: Fitted mortality model.

```
  n= 156368, number of events= 124

                                 coef exp(coef)  se(coef)      z Pr(>|z|)
ns(predicted.cd4, df = 4)1 -5.529860  0.003967  0.735769 -7.516 5.66e-14 ***
ns(predicted.cd4, df = 4)2 -4.577166  0.010284  1.052191 -4.350 1.36e-05 ***
ns(predicted.cd4, df = 4)3 -6.900041  0.001008  1.959493 -3.521 0.000429 ***
ns(predicted.cd4, df = 4)4 -1.715574  0.179861  2.560207 -0.670 0.502799

male                        0.009160  1.009202  0.183325  0.050 0.960150

ns(age_c, df = 3)1          0.850156  2.340011  0.377932  2.249 0.024481 *
ns(age_c, df = 3)2         -0.335923  0.714678  0.834747 -0.402 0.687372
ns(age_c, df = 3)3         -0.273305  0.760861  0.381530 -0.716 0.473782

cd4.0                       0.123630  1.131597  0.547881  0.226 0.821473
cd4.200                     0.900486  2.460800  0.590938  1.524 0.127552
cd4.350                     0.689812  1.993340  0.744415  0.927 0.354109
cd4.500                    -0.965751  0.380697  1.135172 -0.851 0.394906
classA                      0.275909  1.317728  0.271891  1.015 0.310211
classB                      0.182826  1.200605  0.396747  0.461 0.644934
classC                      0.414840  1.514128  0.374118  1.109 0.267496
classN                     -0.280110  0.755701  0.431519 -0.649 0.516258

postArtMonths (t-A)_+      -0.036337  0.964315  0.010400 -3.494 0.000476 ***

N^A(t)                      1.147999  3.151881  0.523526  2.193 0.028320 *
N^A(t):cd4.0               -0.077731  0.925214  0.601829 -0.129 0.897233
N^A(t):cd4.200             -0.587494  0.555718  0.697165 -0.843 0.399402
N^A(t):cd4.350             -0.737121  0.478489  0.918659 -0.802 0.422328
N^A(t):cd4.500              0.986362  2.681461  1.197227  0.824 0.410012
---
```

Table 6: Proportional hazards test on covariates for mortality submodel.

	rho	chisq	p
male	0.07305	6.79e-01	0.410
ns(age_c, df = 3)1	-0.09634	1.09e+00	0.297
ns(age_c, df = 3)2	-0.08321	8.62e-01	0.353
ns(age_c, df = 3)3	-0.05758	3.17e-01	0.573
cd4.0	-0.12566	1.62e+00	0.202
cd4.200	0.00635	4.86e-03	0.944
cd4.350	-0.02277	6.34e-02	0.801
cd4.500	0.03056	1.11e-01	0.739
classA	0.09080	1.06e+00	0.303
classB	0.10337	1.33e+00	0.248
classC	-0.13256	2.20e+00	0.138
classN	0.10103	1.31e+00	0.252
ns(predicted.cd4, df = 4)1	-0.15634	2.25e+00	0.133
ns(predicted.cd4, df = 4)2	-0.12531	1.31e+00	0.252
ns(predicted.cd4, df = 4)3	-0.05161	2.45e-01	0.621
ns(predicted.cd4, df = 4)4	0.06332	3.11e-01	0.577
postArtMonths	0.05998	3.66e-01	0.545
N^A(t)	-0.00216	6.11e-04	0.980
N^A(t):cd4.0	0.04550	2.25e-01	0.635
N^A(t):cd4.200	-0.06772	5.54e-01	0.457
N^A(t):cd4.350	-0.09623	1.12e+00	0.290
N^A(t):cd4.500	-0.02389	7.02e-02	0.791
GLOBAL	NA	2.35e+01	0.375

Table 7: Proportional hazards test on covariates for weight model.

	rho	chisq	p
as.factor(male)1	-0.024698	7.76e-01	0.37846
cdcclass_bA	-0.016814	3.64e-01	0.54604
cdcclass_bB	-0.010932	1.55e-01	0.69383
cdcclass_bC	-0.024645	2.28e+00	0.13105
cdcclass_bN	-0.028996	2.31e+00	0.12854
ns(age_b_c, df = 3)1	-0.043891	2.47e+00	0.11630
ns(age_b_c, df = 3)2	0.000578	4.33e-04	0.98340
ns(age_b_c, df = 3)3	-0.028656	1.06e+00	0.30257
ns(cd4mr, df = 2)1	-0.016403	1.39e-01	0.70927
ns(cd4mr, df = 2)2	-0.026378	8.96e-01	0.34385
ns(wazmr, df = 3)1	0.041298	1.52e+00	0.21744
ns(wazmr, df = 3)2	0.052773	2.30e+00	0.12943
ns(wazmr, df = 3)3	0.012629	1.52e-01	0.69683
ns(hazmr, df = 3)1	-0.046355	2.74e+00	0.09774
ns(hazmr, df = 3)2	-0.011046	1.40e-01	0.70797
ns(hazmr, df = 3)3	0.004891	2.83e-02	0.86651
GLOBAL	NA	1.89e+01	0.27388

To empirically check why results from the weighted and unweighted analyses differ, we compare the distributions of baseline covariates (mean (SD) or count(%)) between 'immediate initiation' and 'never treat', shown in the table below. Significant differences are observed in the distributions of CD4, WAZ and HAZ, between the two groups of patients, and ignoring covariate imbalance (unweighted analysis) would lead to biased conclusion.

Table 8: Comparing baseline covariates between 'immediate initiation' and 'never treat'.

	Immediate initiation $n = 885$	Never treat $n = 616$	p value
CD4	249.54 (285.01)	512.42 (328.65)	<.001
WAZ	−2.73 (1.80)	−2.25 (1.65)	<.001
HAZ	−2.16 (1.42)	−1.85 (1.52)	.001
Age	12.26 (1.42)	12.20 (1.42)	.452
CDC class			.009
mild	89 (10.1%)	67 (10.9%)	
moderate	35 (4.0%)	15 (2.4%)	
severe	50 (5.6%)	18 (2.9%)	
asymptomatic	77 (8.7%)	76 (12.3%)	
Male	414 (46.8%)	246 (39.9%)	.010

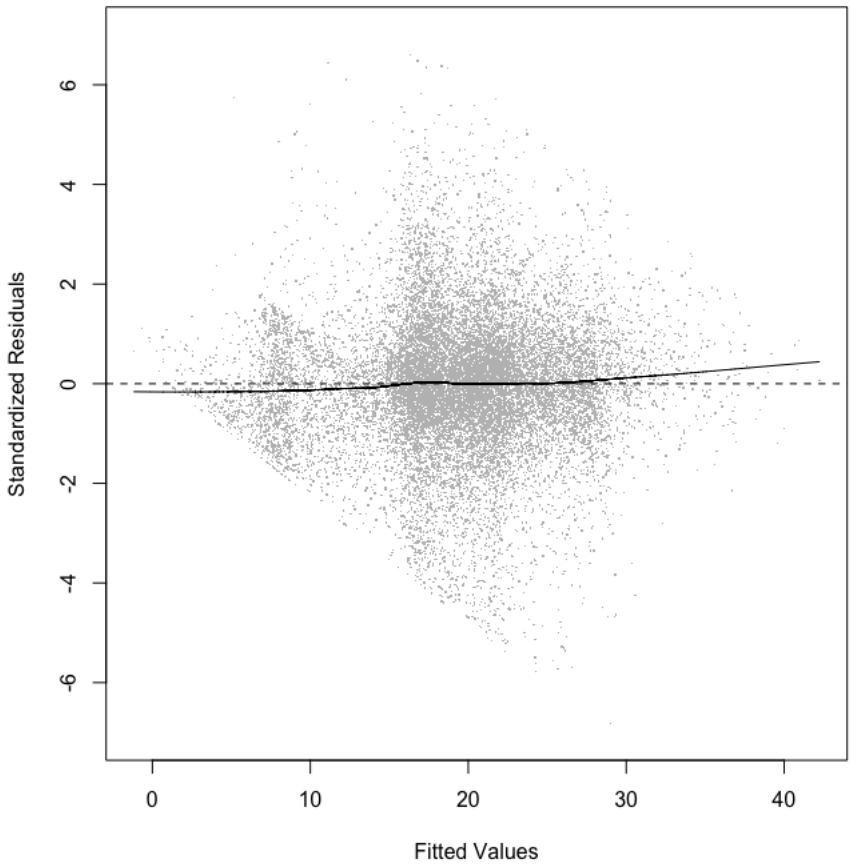

Figure 1: Residual-versus-fitted plot for CD4 submodel, with lowess curve.

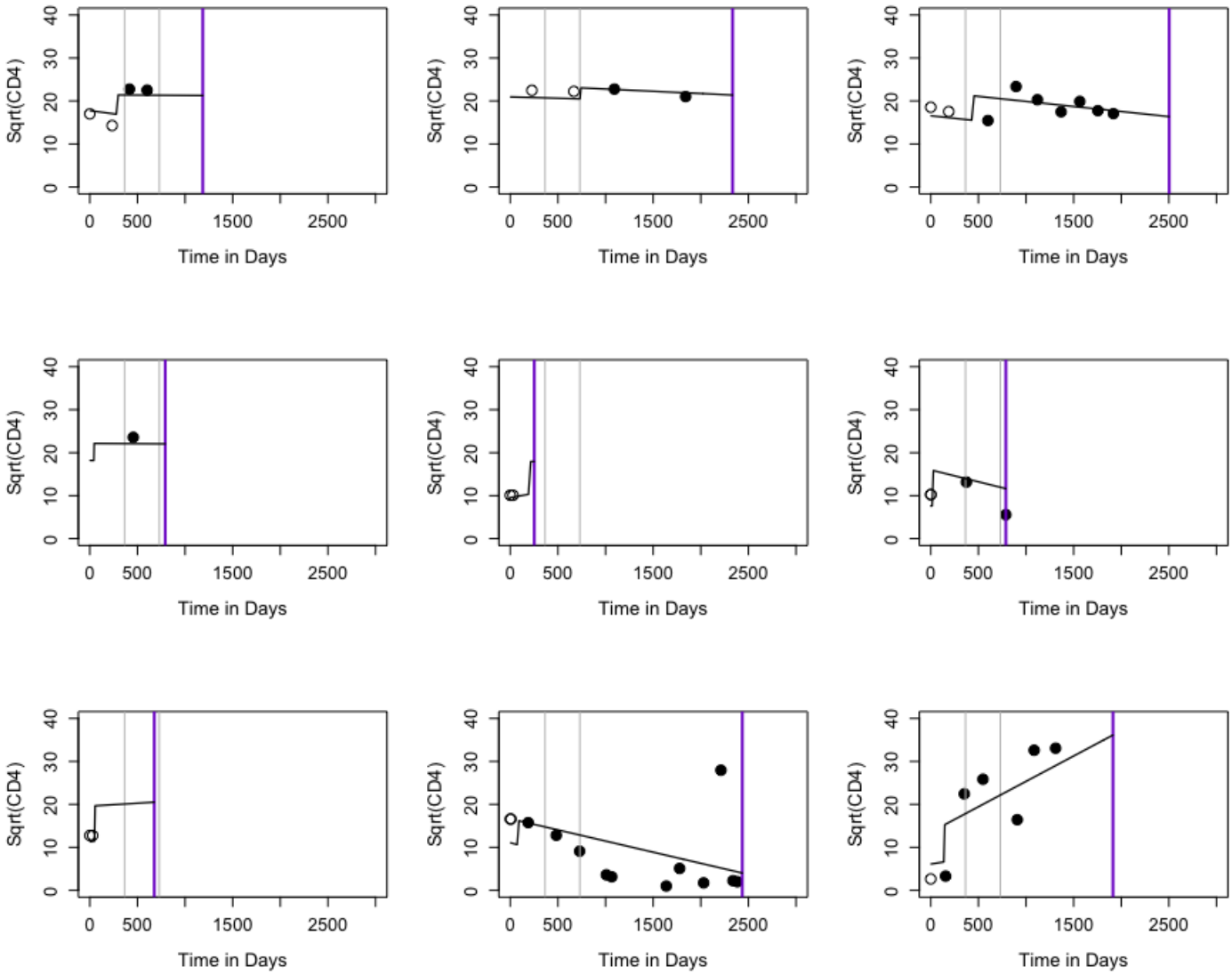

Figure 2: Fitted lines from imputation model for CD4 and ART initiation status during follow up for the 9 randomly selected individuals in Figure 1 from the main text. Empty circles indicate no ART and filled circles represent on ART. Two gray lines denote one year and two years post diagnosis. Purple line indicates end of follow up.

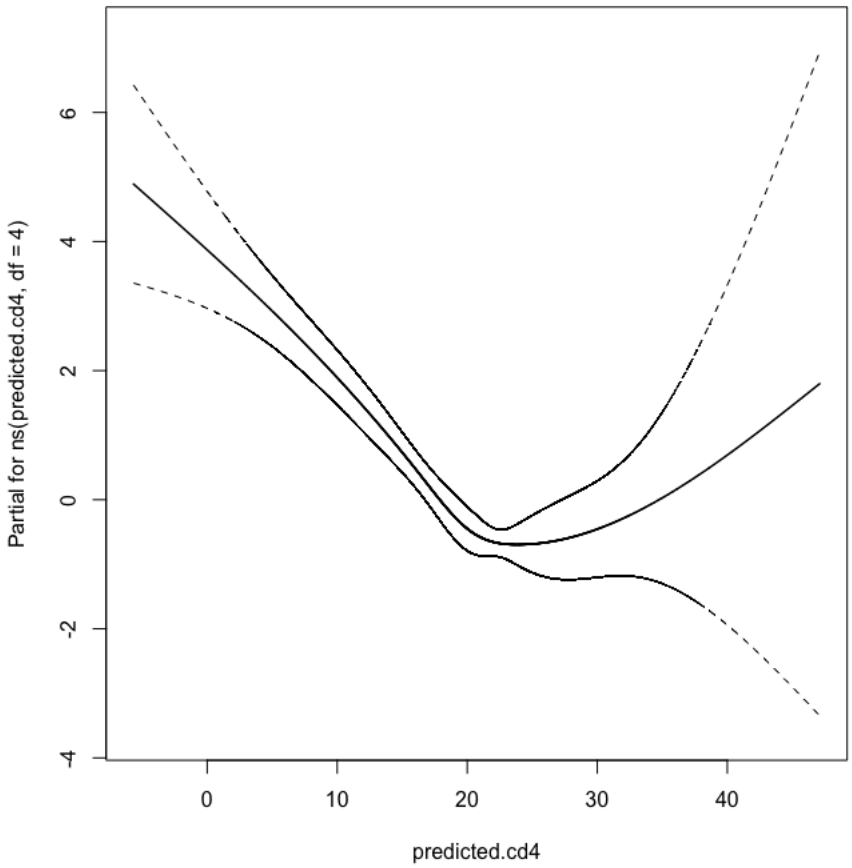

Figure 3: Effect of fitted CD4 $\widehat{m}(t)$ on hazard of death at t in mortality submodel.