WEB MATERIAL

Evaluating Flexible Modeling of Continuous Covariates in Inverse-Weighted Estimators

Ryan P. Kyle, Erica E. M. Moodie, Marina B. Klein, and Michał Abrahamowicz

Contents:

Web Appendix 1. Brief survey of literature employing inverse weighting

We performed a brief survey of the literature to assess the extent to which the propensity score model used for inverse probability of treatment weighting of marginal structural models employed flexible modelling approaches. We selected three general epidemiology journals and three journals publishing HIV research: American Journal of Epidemiology, Epidemiology, International Journal of Epidemiology, Clinical Infectious Diseases, AIDS, and Lancet HIV. All original epidemiological (rather than methodological) research papers published in 2017 were included. In total, we identified 16 articles using inverse probability of treatment weighting. The majority of these focused on a point source treatment with a continuous outcome or, somewhat less often, a point-source treatment with a time-to-event outcome. Five out the 16 articles, or just over 30%, mentioned the use of some form of flexible modelling. See Web Table 1 for the included papers.

Web Table 1: Papers included in the literature survey of the publications of six journals in 2017

Reference	Authors	Citation	Title	PubMedID	Flexible Models
1	**Gardiner BJ et al.**	Clin Infect Dis. 2017 Nov 29;65(12):2000-2007.	Role of secondary prophylaxis with valganciclovir in the prevention of recurrent cytomegalovirus disease in solid organ transplant recipients.	29020220	Transformations
2	**Erlandson KM et al.**	Clin Infect Dis. 2017 Nov 29;65(12):2042-2049.	The impact of statin and angiotensin-converting enzyme inhibitor/angiotensin receptor blocker therapy on cognitive function in adults with Human Immunodeficiency Virus infection.	29020174	
3	**Mills JC et al.**	AIDS. 2017 Nov 28;31(18):2515-2524.	Comparative effectiveness of dual vs. single-action antidepressants on HIV clinical outcomes in HIV-infected people with depression.	28832409	
4	**Stoner MCD et al.**	AIDS. 2017 Sep 24;31(15):2127-2134.	The effect of school attendance and school dropout on incident HIV and HSV-2 among young women in rural South Africa enrolled in HPTN 068.	28692544	
5	**LaFleur J et al.**	AIDS. 2017 Sep 24;31(15):2095-2106.	Cardiovascular outcomes among HIV-infected veterans receiving atazanavir.	28692532	
6	**Lee JS et al.**	AIDS. 2017 Sep 10;31(14):1989-1997.	Incomplete viral suppression and mortality in HIV patients after antiretroviral therapy initiation.	28650383	Restricted quadratic splines
7	**Caniglia EC et al.**	Lancet HIV. 2017 Jun;4(6):e251-e259. Epub 2017 Apr 11.	Comparison of dynamic monitoring strategies based on CD4 cell counts in virally suppressed, HIV-positive individuals on combination antiretroviral therapy in high-income countries: a prospective, observational study.	28411091	Restricted cubic splines
8	**Bengtson AM et al.**	AIDS. 2017 Apr 24;31(7):1009-1016.	Relationship between ever reporting depressive symptoms and all-cause mortality in a cohort of HIV-infected adults in routine care.	28244956	Restricted cubic splines
9	**Marden JR et al.**	Am J Epidemiol. 2017 Oct 1;186(7):805-814.	Contribution of Socioeconomic Status at 3 Life-Course Periods to Late-Life Memory Function and Decline: Early and Late Predictors of Dementia Risk.	28541410	
10	**Melchior M et al.**	Int J Epidemiol. 2017 Oct 1;46(5):1641-1650.	Early cannabis initiation and educational attainment: is the association causal? Data from the French TEMPO study.	28520946	
11	**Lim S et al.**	Am J Epidemiol. 2017 Aug 1;186(3):297-304.	Impact of a supportive housing program on housing stability and sexually transmitted infections among young adults in New York City who were aging out of foster care.	28472264	
12	**Bédard A et al.**	Am J Epidemiol. 2017 Jul 1;186(1):21-28.	Time-dependent associations between body composition, physical activity, and current asthma in women: A marginal structural modeling analysis.	28453608	
13	**Todd JV et al.**	Am J Epidemiol. 2017 May 15;185(10):869-878.	Effects of antiretroviral therapy and depressive symptoms on all-cause mortality among HIV-infected women.	28430844	Restricted cubic splines
14	**Kramer MS et al.**	Am J Epidemiol. 2017 Apr 1;185(7):585-590.	Does fetal growth restriction cause later obesity? Pitfalls in analyzing causal mediators as confounders.	28338874	
15	**Fernandez CA et al.**	Int J Epidemiol. 2017 Apr 1;46(2):440-452.	Longitudinal course of disaster-related PTSD among a prospective sample of adult Chilean natural disaster survivors.	27283159	
16	**Ahrens KA et al.**	Am J Epidemiol. 2017 Mar 1;185(5):335-344.	Plurality of birth and infant mortality due to external causes in the United States, 2000-2010.	28180240	

Web Appendix 2. Additional simulation details: Data generation and results

Below, we provide tables detailing the data-generating parameters and results for the simulation studies discussed in the main paper. We also provide details of the computation time required by each of the modeling techniques we employed for the two-interval simulations. The high computational cost of fitting the covariate-balancing propensity score (Web Figure 1) necessitated the use of a small number of data sets, however we found simulation results and conclusions from 300, 500, and 1000 data sets to be very similar (Web Figure 2).

In preliminary analyses, using the default selection probability of p=0.05 for the fractional polynomial algorithm (12), we found very poor identification of the true functional form, resulting from insufficient power for the test of non-linearity. Using the default settings for the FP algorithm when the true covariate-exposure relationship was quadratic led to the selection of the (incorrect) linear model in 39% and 9% of simulated data sets with sample sizes of N=250 and 500, respectively. Thus, in the final analyses reported below and in the main manuscript, we used the *best-fitting* FP model (corresponding to a model selection probability of 1.0) to overcome the low power.

Web Table 2: Data generating model for single-time-interval simulation studies

	Exposure model$^{\pi}$ parameters					Outcome model† parameters				
	α_0	α_1	α_2	α_3	α_4	β_0	β_1	β_2	β_3	β_4
Linear	5.57	0.58	−2.78	0.15	−0.02					
Quadratic	2.75	1.5	−2.25	0.4	−0.04	−0.46	−0.32	−0.05	0.01	5
Exponential	5.5	0.55	−0.30	0.19	−0.06					

$^{\pi}logit\big(P(X=1|L)\big) = \alpha_0 + \alpha_1(ARV status) + \alpha_2\big(log(GGT)\big) + \alpha_3(Female) + \alpha_4(Age)$

$^{\dagger}\ Y = \beta_0 + \beta_1(X) + \beta_2(Female) + \beta_3(Age) + \beta_4\big(log(GGT)\big) + \varepsilon \sim N(0, 1)$

Web Table 3: Data generating model for two-time-interval simulation studies

	Exposure model$^{\pi}$ parameters			Outcome model† parameters				
	α_0	α_1	α_2	β_0	β_1	β_2	β_3	β_4
Linear	−0.75	0.50	−1.50	0	1.00	1.50	0.10	0.90
Quadratic	−2.50	0.25	1.00	0	3.00	2.50	1.00	0.75
Exponential	−1.00	2.00	0.30	0	1.00	1.50	0.10	0.60

$^{\pi}logit\big(P(X_1=1|L_1)\big) = \alpha_0 + \alpha_1(L_1), logit\big(P(X_2=1|X_1, L_2)\big) = \alpha_0 + \alpha_1(L_2) + \alpha_2(X_1)$

$^{\dagger}Y = \beta_0 + \beta_1 X_1 + \beta_1 X_2 + \beta_3 L_1 + \beta_4 L_2 + \varepsilon\, N(0,1)$

Web Table 4: Results of single-time-interval simulation studies

95% Boot Cov = 95% bootstrap coverage interval, SMD = standardized mean difference for log_{10}(γ-glutamyltransferase), FP = Fractional polynomials, CBPS = Covariate-balancing propensity score, MCSE = Monte Carlo standard error, MSE = mean squared error. Bootstrap coverage was computed using the nonparametric percentiles approach.

	Linear Scenario					**Quadratic Scenario**					**Exponential Scenario**				
***N* and Model Specification**	**Bias**	**Root MSE**	**MCSE**	**95% Boot Cov**	**SMD**	**Bias**	**Root MSE**	**MCSE**	**95% Boot Cov**	**SMD**	**Bias**	**Root MSE**	**MCSE**	**95% Boot Cov**	**SMD**
250															
Linear	−0.05	0.27	0.27	0.87	0.07	0.28	0.41	0.30	0.92	0.15	−1.77	1.88	0.63	0.14	0.84
Quadratic	–	–	–	–	–	0.03	0.25	0.25	0.95	0.05	–	–	–	–	–
Exponential	–	–	–	–	–	–	–	–	–	–	−0.07	0.25	0.24	0.93	0.07
Splines	−0.13	0.25	0.22	0.84	0.06	−0.02	0.25	0.25	0.97	0.05	−0.12	0.27	0.24	0.90	0.06
FP	−0.11	0.25	0.22	0.85	0.06	0.01	0.24	0.24	0.96	0.04	−0.10	0.25	0.23	0.90	0.05
CBPS	−0.23	0.32	0.23	0.67	0.11	−0.11	0.27	0.25	0.98	0.06	−0.15	0.28	0.23	0.84	0.06
500															
Linear	−0.02	0.17	0.17	0.92	0.05	0.33	0.40	0.23	0.73	0.15	−1.78	1.83	0.44	0.003	0.82
Quadratic	–	–	–	–	–	0.05	0.18	0.18	0.93	0.03	–	–	–	–	–
Exponential	–	–	–	–	–	–	–	–	–	–	−0.05	0.18	0.17	0.95	0.05
Splines	−0.06	0.15	0.13	0.91	0.04	0.02	0.18	0.18	0.96	0.04	−0.09	0.18	0.16	0.91	0.03
FP	−0.05	0.15	0.14	0.90	0.04	0.03	0.18	0.18	0.93	0.03	−0.08	0.18	0.16	0.91	0.04
CBPS	−0.15	0.22	0.16	0.72	0.08	−0.07	0.20	0.18	0.96	0.05	−0.11	0.19	0.16	0.850	0.04

Web Table 5: Results of two-time-interval simulation studies

N and Model	Linear Scenario										Quadratic Scenario									
	Bias		Root MSE		MCSE		95% Boot Coverage		SMD		Bias		Root MSE		MCSE		95% Boot Coverage		SMD	
	β_1	β_2	β_1	β_2	β_1	β_2	β_1	β_2	β_1	β_2	β_1	β_2	β_1	β_2	β_1	β_2	β_1	β_2	β_1	β_2
250																				
Linear	0.04	0.04	0.31	0.34	0.31	0.34	0.91	0.89	0.02	0.12	0.15	−1.41	0.28	1.44	0.24	0.29	0.94	0.01	0.07	1.12
Quad.	–	–	–	–	–	–	–	–	–	–	−0.07	−0.09	0.36	0.35	0.36	0.34	0.92	0.91	0.03	0.12
Exp.	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Splines	0.04	0.08	0.32	0.35	0.31	0.34	0.91	0.87	0.01	0.11	−0.08	−0.12	0.34	0.33	0.33	0.31	0.94	0.91	0.03	0.10
FP	0.05	0.07	0.32	0.34	0.31	0.34	0.91	0.88	0.02	0.10	−0.08	−0.12	0.34	0.33	0.33	0.30	0.92	0.90	0.03	0.11
CBPS	0.23	0.23	0.41	0.39	0.35	0.32	0.76	0.67	0.07	0.09	−0.05	−0.12	0.42	0.39	0.41	0.37	0.87	0.84	0.06	0.12
500																				
Linear	0.02	0.02	0.25	0.26	0.25	0.26	0.92	0.88	0.01	0.12	0.16	0.16	0.23	1.44	0.16	0.21	0.86	0.00	0.06	1.12
Quad.	–	–	–	–	–	–	–	–	–	–	−0.02	−0.02	0.28	0.27	0.28	0.27	0.93	0.85	0.02	0.09
Exp.	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Splines	0.01	0.01	0.25	0.26	0.25	0.26	0.91	0.89	0.01	0.11	−0.04	−0.04	0.28	0.26	0.28	0.25	0.94	0.88	0.01	0.07
FP	0.02	0.02	0.24	0.26	0.24	0.26	0.91	0.88	0.01	0.11	−0.04	−0.04	0.28	0.26	0.28	0.25	0.93	0.88	0.02	0.08
CBPS	0.15	0.15	0.33	0.30	0.30	0.26	0.76	0.76	0.05	0.05	0.03	0.03	0.36	0.34	0.36	0.34	0.88	0.83	0.04	0.08

Quad. = Quadratic, Exp. = Exponential, FP = Fractional polynomials, CBPS = Covariate-balancing propensity score, MCSE = Monte Carlo standard error, MSE = mean squared error, SMD = standardized mean difference. Bootstrap coverage was computed via the nonparametric percentiles approach.

N and **Model**	**Exponential Scenario**									
	Bias		**Root MSE**		**MCSE**		**95% Boot Coverage**		**SMD**	
	β_1	β_2	β_1	β_2	β_1	β_2	β_1	β_2	β_1	β_2
250										
Linear	−0.18	−0.03	0.33	0.26	0.27	0.26	0.90	0.95	0.18	0.14
Quad.	–	–	–	–	–	–	–	–	–	–
Exp.	0.00	0.03	0.27	0.32	0.27	0.32	0.94	0.91	0.12	0.16
Splines	0.00	0.06	0.25	0.30	0.25	0.29	0.95	0.94	0.03	0.10
FP	0.00	0.05	0.24	0.29	0.24	0.28	0.94	0.95	0.04	0.10
CBPS	0.30	0.34	0.57	0.58	0.49	0.47	0.87	0.86	0.05	0.07
500										
Linear	−0.16	−0.16	0.25	0.21	0.20	0.20	0.87	0.94	0.18	0.15
Quad.	–	–	–	–	–	–	–	–	–	–
Exp.	0.01	0.01	0.19	0.22	0.22	0.22	0.93	0.94	0.08	0.12
Splines	0.01	0.01	0.16	0.20	0.20	0.20	0.96	0.95	0.02	0.08
FP	0.01	0.01	0.16	0.19	0.19	0.19	0.94	0.96	0.03	0.08
CBPS	0.23	0.23	0.53	0.52	0.45	0.45	0.71	0.70	0.05	0.04

Quad. = Quadratic, Exp. = Exponential, FP = Fractional polynomials, CBPS = Covariate-balancing propensity score, MCSE = Monte Carlo standard error, MSE = mean squared error, SMD = standardized mean difference. Bootstrap coverage was computed via the nonparametric percentiles approach.

The process for generating the data used in the simulation studies with a time-to-event outcome was largely identical to that used for all other single-time-interval simulations. For these simulations, the distribution of the response was forced to be positive by adding a constant (specified here as 10) to all generated values.

We then simulated 20% censoring based on computing quantiles of the resulting event times. All observations with event times greater than the 80th percentile of the distribution were censored. The data were analyzed using an accelerated failure time model.

Due to the reduction in effective sample size as a result of the censoring, we specified a smaller variance for the additive random error term (0.25 instead of 1.0).

Web Table 6: Data generating model for single-time-interval simulation studies with time-to-event outcome

	Exposure model$^{\pi}$ parameters					*Outcome model† parameters*				
	α_0	α_1	α_2	α_3	α_4	β_0	β_1	β_2	β_3	β_4
Linear	5.57	0.58	−2.78	0.15	−0.02					
Quadratic	2.75	1.5	−2.25	0.4	−0.04	−0.46	−0.32	−0.05	0.01	5
Exponential	5.5	0.55	−0.30	0.19	−0.06					

$^{\pi}logit\big(P(X=1|L)\big) = \alpha_0 + \alpha_1(ARVstatus) + \alpha_2\big(log(GGT)\big) + \alpha_3(Female) + \alpha_4(Age)$

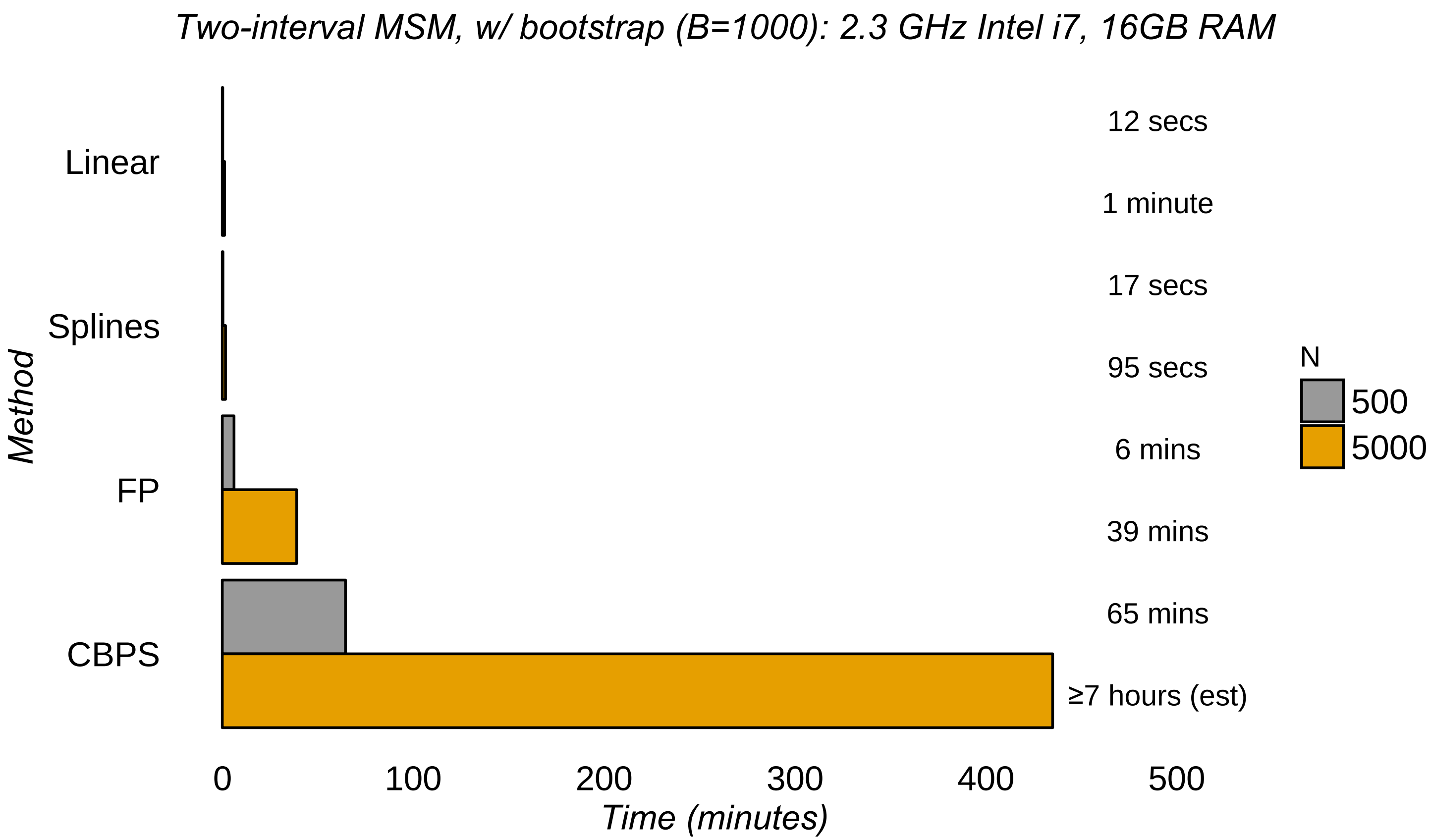
$^{\dagger}\ Y = \beta_0 + \beta_1(X) + \beta_2(Female) + \beta_3(Age) + \beta_4\big(log(GGT)\big) + \varepsilon \sim N(0, 0.25) + 10$

Web Table 7: Results of single-time-interval simulation studies for accelerated failure time (AFT) outcome model with 20% censoring

N and Model Specification	Linear Scenario					Quadratic Scenario					Exponential Scenario				
	Bias	**Root MSE**	**MCSE**	**95% Boot Cov**	**SMD**	**Bias**	**Root MSE**	**MCSE**	**95% Boot Cov**	**SMD**	**Bias**	**Root MSE**	**MCSE**	**95% Boot Cov**	**SMD**
250															
Linear	−0.07	0.15	0.13	0.91	0.07	0.41	0.49	0.26	0.79	0.14	−1.53	1.63	0.57	0.18	0.83
Quadratic	-	-	-	-	-	0.04	0.19	0.19	0.98	0.04	-	-	-	-	-
Exponential	-	-	-	-	-	-	-	-	-	-	−0.06	0.24	0.24	0.91	0.08
Splines	−0.13	0.18	0.14	0.79	0.06	−0.01	0.17	0.17	0.98	0.05	−0.13	0.22	0.17	0.81	0.05
FP	−0.10	0.16	0.12	0.84	0.06	0.03	0.17	0.17	0.98	0.04	−0.09	0.18	0.16	0.89	0.05
CBPS	−0.07	0.15	0.13	0.91	0.12	0.11	0.24	0.22	0.97	0.06	−0.11	0.22	0.19	0.85	0.06
500															
Linear	−0.06	0.11	0.09	0.90	0.05	0.45	0.48	0.18	0.37	0.15	−1.48	1.54	0.41	0.02	0.81
Quadratic	-	-	-	-	-	0.06	0.14	0.13	0.94	0.03	-	-	-	-	-
Exponential	-	-	-	-	-	-	-	-	-	-	−0.06	0.17	0.16	0.91	0.05
Splines	−0.09	0.13	0.10	0.80	0.04	0.01	0.12	0.12	0.97	0.03	−0.09	0.16	0.13	0.81	0.03
FP	−0.07	0.11	0.09	0.85	0.04	0.04	0.12	0.12	0.96	0.03	−0.07	0.13	0.11	0.89	0.04
CBPS	−0.06	0.11	0.09	0.90	0.08	0.12	0.19	0.15	0.91	0.05	−0.06	0.15	0.14	0.89	0.05

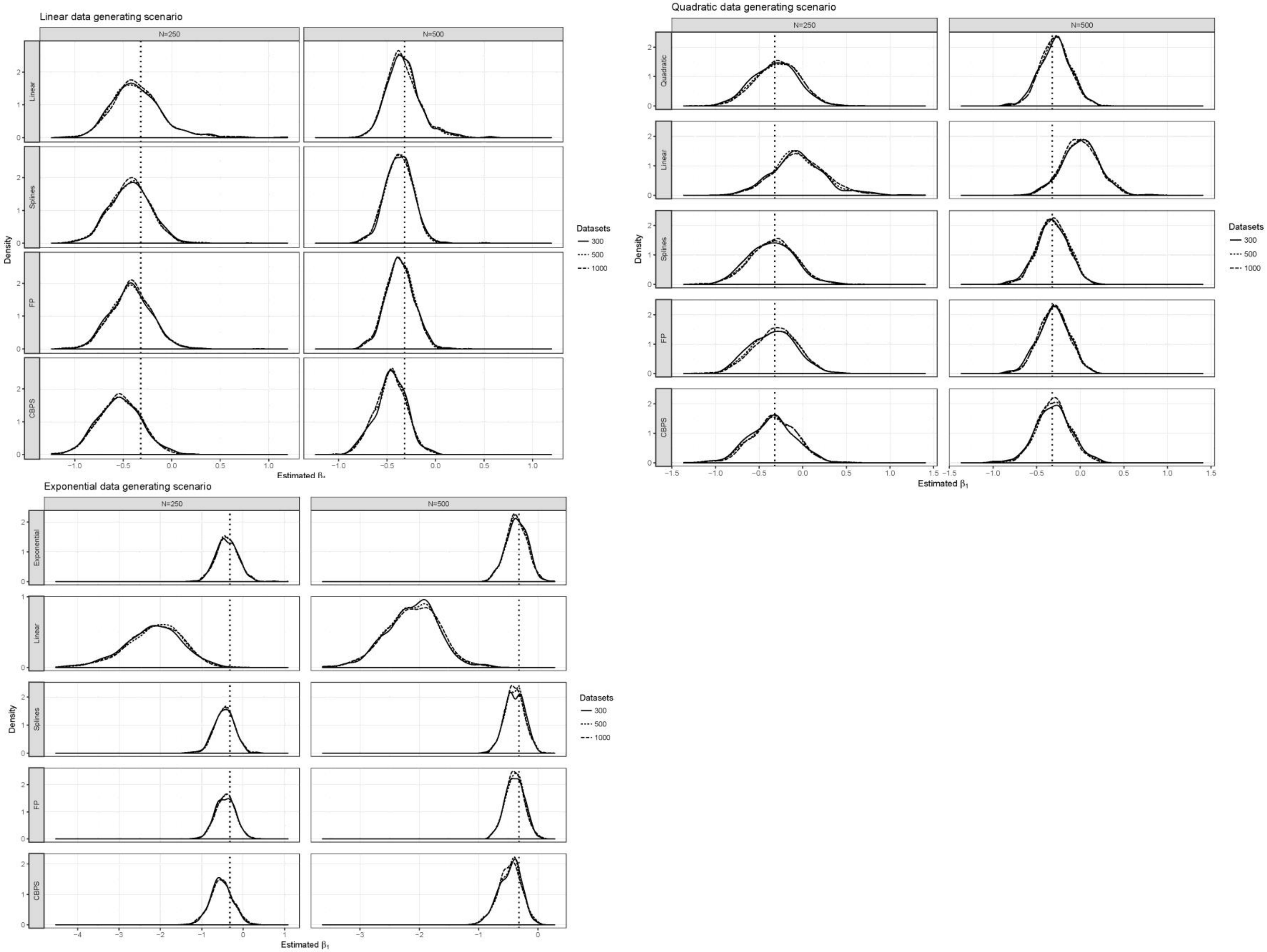
95% Boot Cov = 95% bootstrap coverage interval, SMD = standardized mean difference for $\log_{10}$(γ-glutamyltransferase), FP = Fractional polynomials, CBPS = Covariate-balancing propensity score, MCSE = Monte Carlo standard error, MSE = mean squared error. Bootstrap coverage was computed using the nonparametric percentiles approach.

Web Figure 1: Computational time comparisons for conventional and alternative exposure model specifications



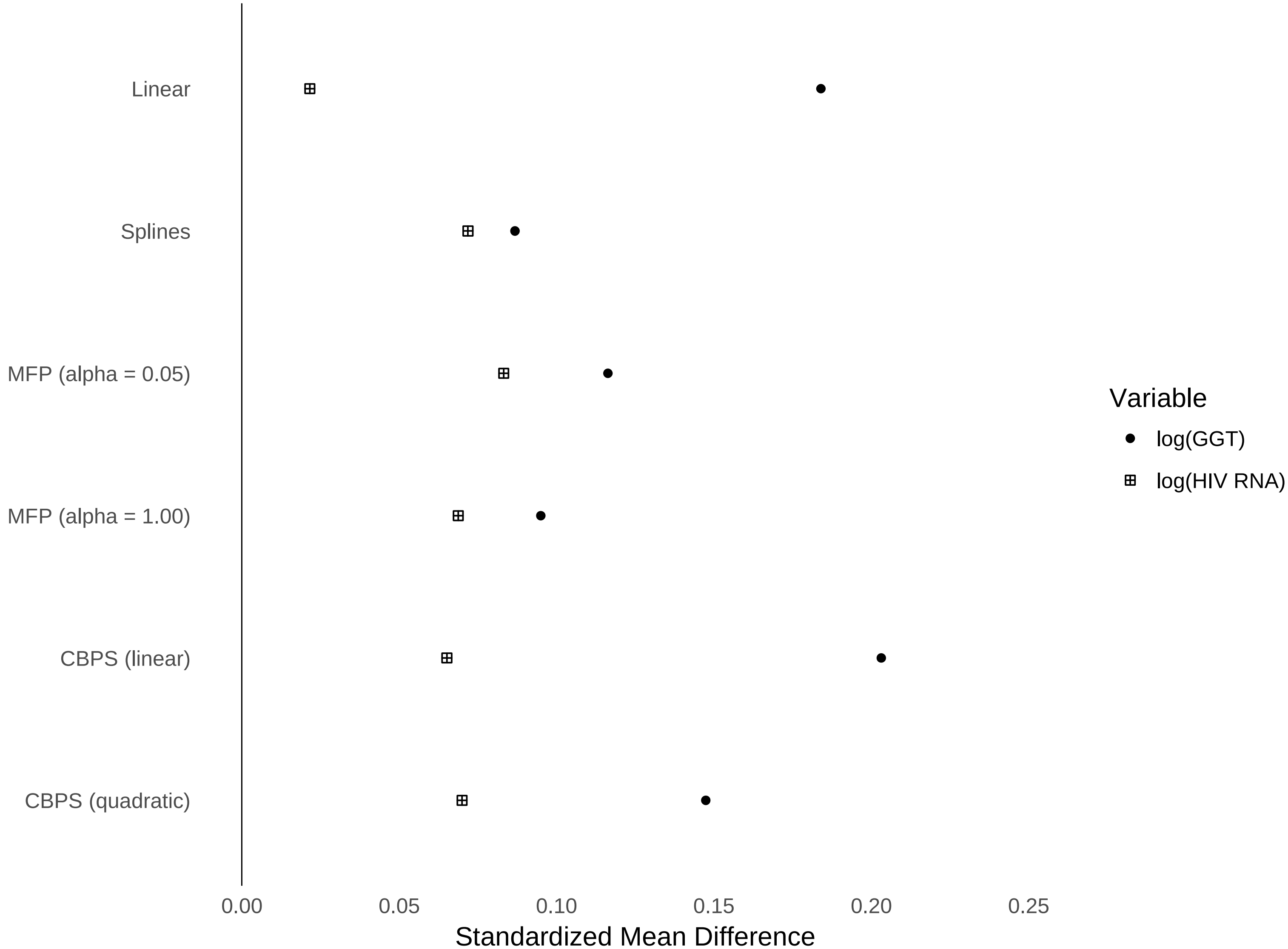
FP = Fractional polynomials, CBPS = Covariate-balancing propensity score

Web Figure 2: Comparison of the distribution of parameter estimates given three data-generating scenarios for 300, 500, and 1000 simulated data sets.



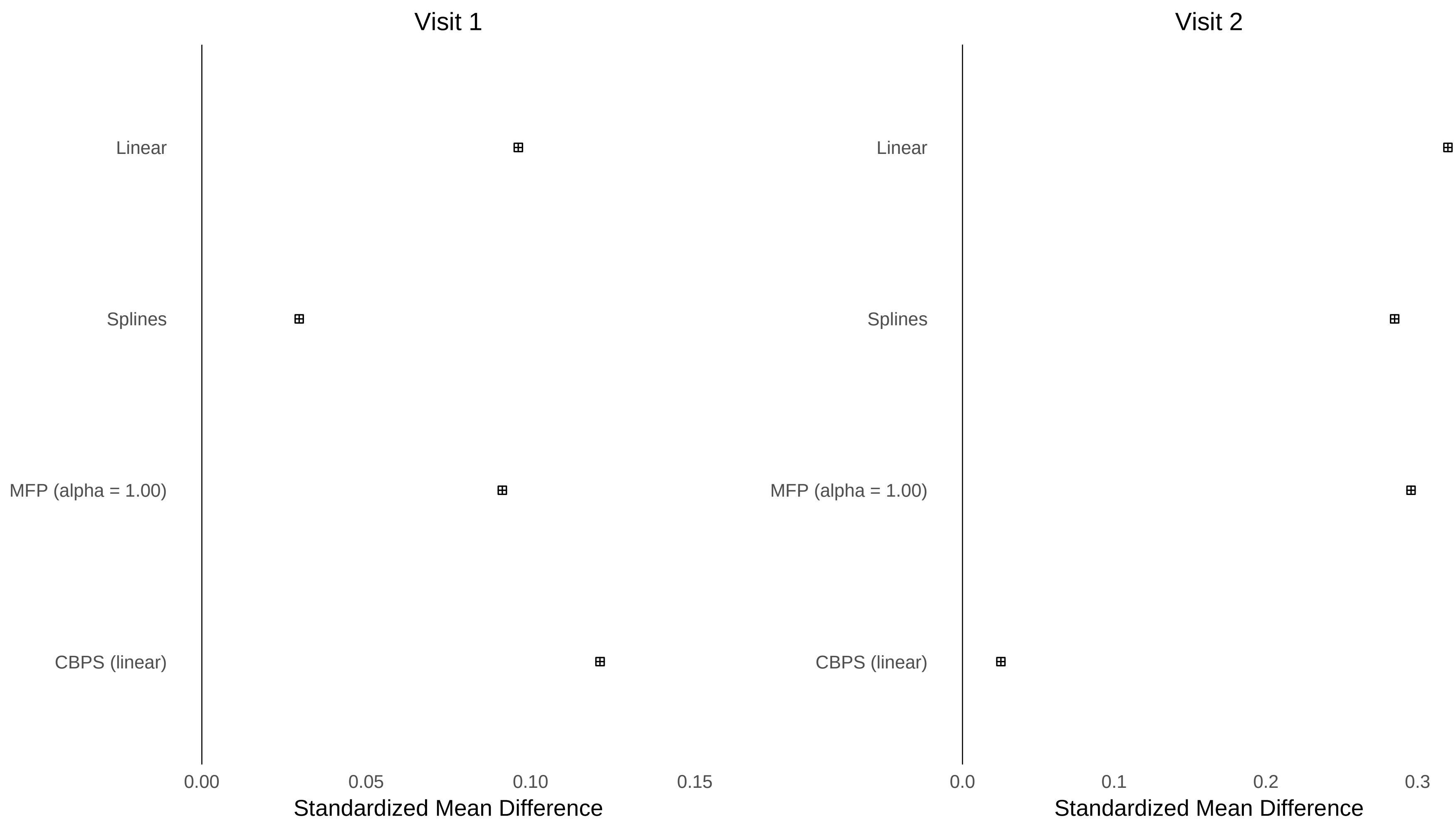
Web Appendix 3. Additional details of the Canadian Co-infection Cohort analysis

Web Figure 3: Standardized mean differences for $\log_{10}$(γ-glutamyltransferase) and $\log_{10}$(HIV viral load) under conventional and flexible exposure model specifications



MFP = Fractional polynomials, CBPS = Covariate-balancing propensity score

Web Figure 4: Standardized mean differences for $log_{10}(CD4^{+})$ under conventional and alternative exposure model specifications



MFP = Fractional polynomials, CBPS = Covariate-balancing propensity score

Web References

1. Gardiner BJ, Chow JK, Price LL et al. Role of secondary prophylaxis with valganciclovir in the prevention of recurrent cytomegalovirus disease in solid organ transplant recipients. *Clin Infect Dis.* 2017;65(12):2000–2007.
2. Erlandson KM, Kitch D, Wester CW et al. The impact of statin and angiotensin-converting enzyme inhibitor/angiotensin receptor blocker therapy on cognitive function in adults with human immunodeficiency virus infection. *Clin Infect Dis.* 2017;65(12):2042–2049.
3. Mills JC, Harman JS, Cook RL et al. Comparative effectiveness of dual vs. single-action antidepressants on HIV clinical outcomes in HIV-infected people with depression. *AIDS*. 2017;31(18):2515–2524.
4. Stoner MCD, Pettifor A, Edwards JK et al. The effect of school attendance and school dropout on incident HIV and HSV-2 among young women in rural South Africa enrolled in HPTN 068. *AIDS.* 2017;31(15):2127–2134.
5. LaFleur J, Bress AP, Rosenblatt L et al. Cardiovascular outcomes among HIV-infected veterans receiving atazanavir. *AIDS*. 2017;31(15):2095–2106.
6. Lee JS, Cole SR, Richardson DB et al. Incomplete viral suppression and mortality in HIV patients after antiretroviral therapy initiation. *AIDS.* 2017;31(14):1989–1997.
7. Caniglia EC, Cain LE, Sabin CA et al. Comparison of dynamic monitoring strategies based on CD4 cell counts in virally suppressed, HIV-positive individuals on combination antiretroviral therapy in high-income countries: a prospective, observational study. *Lancet HIV.* 2017;4(6):e251–e259.
8. Bengtson AM, Pence BW, Moore R et al. Relationship between ever reporting depressive symptoms and all-cause mortality in a cohort of HIV-infected adults in routine care. *AIDS.* 2017;31(7):1009–1016.

9. Marden JR, Tchetgen Tchetgen EJ, Kawachi I, Glymour MM. Contribution of Socioeconomic Status at 3 Life-Course Periods to Late-Life Memory Function and Decline: Early and Late Predictors of Dementia Risk. *Am J Epidemiol*. 2017;186(7):805–814.

10. Melchior M, Bolze C, Fombonne E et al. Early cannabis initiation and educational attainment: is the association causal? Data from the French TEMPO study. *Int J Epidemiol*. 2017;46(5):1641–1650.

11. Lim S, Singh TP, Gwynn RC. Impact of a supportive housing program on housing stability and sexually transmitted infections among young adults in New York City who were aging out of foster care. *Am J Epidemiol*. 2017;186(3):297–304.

12. Bédard A, Serra I, Dumas O, et al. Time-dependent associations between body composition, physical activity, and current asthma in women: a marginal structural modeling analysis. *Am J Epidemiol.* 2017;186(1):21–28.

13. Todd JV, Cole SR, Pence BW, et al. Effects of antiretroviral therapy and depressive symptoms on all-cause mortality among HIV-infected women. *Am J Epidemiol.* 2017;185(10):869–878.

14. Kramer MS, Zhang X, Dahhou M, et al. Does fetal growth restriction cause later obesity? Pitfalls in analyzing causal mediators as confounders. *Am J Epidemiol.* 2017;185(7):585–590.

15. Fernandez CA, Vicente B, Marshall BD, et al. Longitudinal course of disaster-related PTSD among a prospective sample of adult Chilean natural disaster survivors. *Int J Epidemiol.* 2017;46(2):440–452.

16. Ahrens KA, Thoma ME, Rossen LM, Warner M, Simon AE. Plurality of birth and infant mortality due to external causes in the United States, 2000–2010. *Am J Epidemiol.* 2017;185(5):335–344.