

Supplementary Material for

“The Impact of Sparse Follow-up on Marginal Structural Models for Time-to-Event Data”

by Nassim Mojaverian, Erica E. M. Moodie, Alex Bliu, and Marina B. Klein

WEB APPENDIX

An Overview of Multiple Imputation by Chained Equations

There are a number of ways of performing the imputation. We adopted a chained equations approach. Let Y denote complete data which is a random sample from the distribution $P(Y/\theta)$. θ is a vector of unknown parameters that specifies the distribution of Y completely. The method of chained equations suggests that one can obtain a posterior distribution of θ by iteratively sampling from the conditional distributions of $P(Y_1/Y_{-1}, \theta_1), \dots, P(Y_p/Y_{-p}, \theta_p)$

where Y_{-j} is the predictor of a given Y_j . So in the l -th iteration, we have

$$\theta_1^{*(l)} \sim P(\theta_1 | Y_1^{obs}, Y_2^{(l-1)}, \dots, Y_p^{(l-1)})$$

$$Y_1^{*(l)} \sim P(Y_1 | Y_1^{obs}, Y_2^{(l-1)}, \dots, Y_p^{(l-1)}, \theta_1^{*(l)})$$

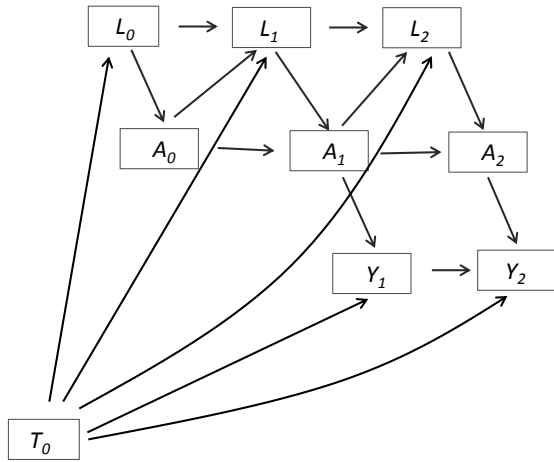
...

$$\theta_p^{*(l)} \sim P(\theta_p | Y_p^{obs}, Y_1^{(l)}, \dots, Y_{p-1}^{(l)})$$

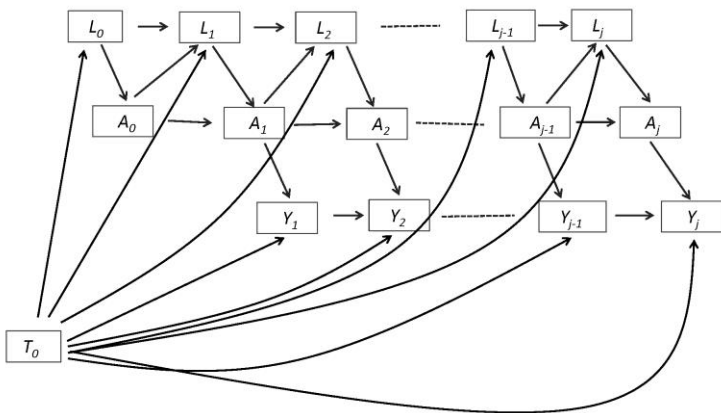
$$Y_p^{*(l)} \sim P(Y_p | Y_p^{obs}, Y_1^{(l)}, \dots, Y_{p-1}^{(l)}, \theta_p^{*(l)}),$$

where $Y_j^{(l)} = (Y_j^{obs}, Y_j^{*(l)})$ is the j -th imputed variable at the l -th iteration. This procedure continues until convergence. A suitable model for each variable containing missing measurements must be selected.

WEB FIGURES



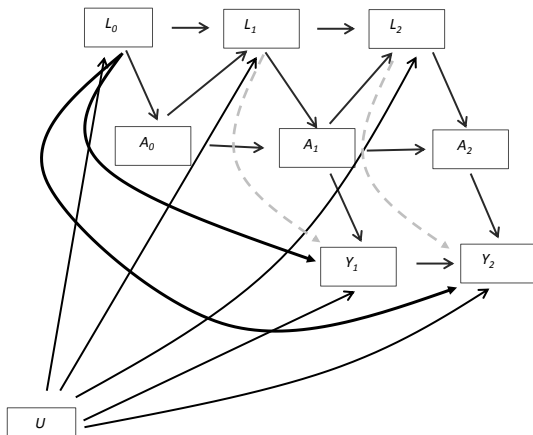
Web Figure 1. Causal diagram of Young et al.'s algorithm with two intervals.



Web Figure 2. Causal diagram of Young et al.'s algorithm with j intervals.

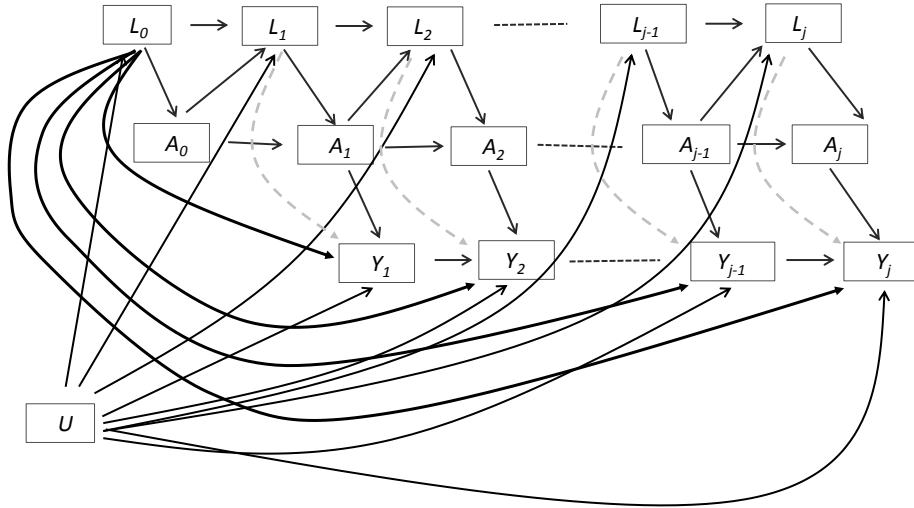
Web Figure 3. Causal diagram of data from the Canadian HIV–Hepatitis C Co-infection Cohort Study (CCC). In parts A and B, U denotes the latent individual-level health status (frailty) — an unmeasured quantity capturing health status that is not predictive of the exposure but affects measured health indicators such as CD4 cell count as well as the likelihood of developing liver fibrosis; this acts in a similar role to T_0 in Young et al.’s algorithm. L_0 denotes baseline covariates, including an indicator of injection drug use in the last six months, current CD4 cell count, current HIV viral load, age at baseline, duration of HCV infection at baseline, gender, indicator for Hepatitis B surface antigen, APRI score at baseline on the natural scale. Time-varying confounders L_j are an indicator of injection drug use in the last six months, current CD4 cell count, current HIV viral load, and current duration of HCV infection. The arrows from L_0 to the outcome are thicker to emphasize the link between baseline APRI and the definition of the outcome, liver fibrosis, which is taken to be an APRI score of 1.5 or greater. The gray dashed arrows from L_j to the outcome are meant to indicate a possible causal link between viral load and liver fibrosis, as there is some indication in the CCC data that poor viral suppression is associated with higher rates of fibrosis.

A)



Web Figure 3A. Causal diagram of CCC data with two intervals.

B)



Web Figure 3B. Causal diagram of CCC data with j intervals.

WEB TABLES

Web Table 1. Model coefficients for the censoring weighting in the Canadian HIV–Hepatitis C Co-infection Cohort Study. The coefficients are averaged over the 20 imputed data sets; confidence intervals were obtained by bootstrap.

Variable	<i>Censoring for Reasons Other Than Death or HCV Treatment</i>								
	Available Cases			LOCF			MI		
	OR	Lower 95% CL	Upper 95% CL	OR	Lower 95% CL	Upper 95% CL	OR	Lower 95% CL	Upper 95% CL
<i>Denominator</i>									
Age at baseline	0.99	0.99	1.03	1.00	0.99	1.00	1.01	0.99	1.09
Baseline lnAPRI	1.08	0.71	1.64	1.06	0.91	1.59	0.9	0.68	1.55
Female	1.01	0.86	1.25	0.97	0.93	1.08	0.96	0.82	1.45
Aboriginal	2.36	1.51	3.08	1.50	1.09	1.89	1.44	1.33	2.24
HCV RNA-positive at baseline	0.95	0.74	1.43	0.83	0.63	1.36	0.76	0.71	1.22
HCV duration	0.99	0.93	1.10	1.00	1.00	1.00	0.99	0.92	1.05
Hepatitis B surface antigen-positive at baseline	0.88	0.95	4.97	0.84	0.66	2.02	0.82	0.64	2.24
Previous alcohol	1.52	0.75	2.35	1.13	0.81	1.14	1.32	0.83	1.69
Previous CD4 (per 100 cells)	0.99	0.98	1.03	1.00	0.99	1.02	1.00	1.00	1.02
Previous HIV-RNA	0.99	0.92	1.09	0.99	0.95	1.05	1.00	0.92	1.07
Previous IDU	1.08	0.75	1.45	1.05	0.95	1.13	1.22	0.89	1.35
Previous ART	1.04	0.53	1.68	1.01	0.95	1.11	1.04	0.83	1.53
Previous smoking	0.97	0.55	1.16	0.91	0.93	1.03	0.65	0.33	1.02

(Table continues)

Numerator

Age at baseline	1.00	0.98	1.03	1.00	0.97	1.05	1.00	0.97	1.09
Baseline lnAPRI	0.99	0.57	1.97	1.01	0.98	2.03	0.95	0.64	1.48
Female	0.99	0.90	1.40	0.97	0.90	1.31	0.95	0.65	1.49
Aboriginal	1.68	1.31	2.67	1.50	0.95	1.94	1.50	1.15	2.26
HCV RNA-positive at baseline	0.97	0.74	1.39	0.85	0.35	1.18	0.73	0.42	1.17
HCV duration	0.99	0.93	1.06	0.99	0.96	1.05	0.99	0.93	1.03
Hepatitis B surface antigen-positive at baseline	0.88	0.62	3.52	0.87	0.60	3.21	0.92	0.62	2.17

Censoring Due to Death

Variable	Available Cases			LOCF			MI		
	OR	Lower 95% CL	Upper 95% CL	OR	Lower 95% CL	Upper 95% CL	OR	Lower 95% CL	Upper 95% CL
	<i>Denominator</i>								
Age at baseline	1.09	0.98	1.12	1.05	0.98	1.07	1.08	1.01	1.11
Baseline lnAPRI	0.92	0.89	1.01	0.9	0.86	0.99	0.96	0.92	1.02
Female	0.96	0.92	1.06	0.96	0.93	1.04	1.05	0.94	1.09
Aboriginal	0.83	0.64	1.15	0.7	0.56	1.09	0.97	0.66	1.13
HCV RNA-positive at baseline	0.92	0.82	1.09	0.85	0.66	1.05	1.02	0.93	1.07
HCV duration	0.99	0.98	1.04	0.99	0.98	1.05	0.99	0.97	1.02
Hepatitis B surface antigen-positive at baseline	1.12	0.92	1.38	1.68	0.97	2.09	1.49	0.78	2.86
Previous alcohol	1.22	0.99	1.47	1.34	0.98	1.62	1.36	0.98	1.64
Previous CD4 (per 100 cells)	0.99	0.99	1.00	0.99	0.99	1.01	0.99	0.99	1.00
Previous HIV-RNA	1.06	0.94	1.08	1.00	0.96	1.02	1.03	0.96	1.06
Previous IDU	1.04	0.95	1.06	1.00	0.92	1.06	1.03	0.94	1.23

Previous ART	1.12	0.84	1.19	1.12	0.86	1.16	1.08	0.91	1.17
Previous smoking	1.25	0.97	1.64	1.36	1.04	1.62	1.07	1.02	1.19

Numerator

Age at baseline	1.05	0.99	1.07	1.05	0.99	1.07	1.07	1.00	1.08
Baseline lnAPRI	0.99	0.95	1.02	0.98	0.86	1.01	1.02	0.91	1.34
Female	0.99	0.99	1.03	0.96	0.93	1.02	1.09	0.97	1.41
Aboriginal	0.87	0.63	1.11	0.74	0.58	1.05	0.94	0.62	1.07
HCV RNA-positive at baseline	0.93	0.52	1.05	0.9	0.73	1.12	1.05	0.85	1.19
HCV duration	0.99	0.99	1.02	1.00	0.98	1.08	0.99	0.99	1.07
Hepatitis B surface antigen-positive at baseline	1.75	0.91	2.65	1.89	0.98	2.16	1.71	1.07	3.15

Censoring Due to Start of HCV Treatment

Variable	Available Cases			LOCF			MI		
	OR	Lower 95% CL	Upper 95% CL	OR	Lower 95% CL	Upper 95% CL	OR	Lower 95% CL	Upper 95% CL

Denominator

Age at baseline	1.00	0.99	1.01	1.00	0.99	1.00	1.00	0.99	1.00
Baseline lnAPRI	2.95	1.35	4.57	2.27	1.24	4.41	2.5	1.23	3.75
Female	0.69	0.48	1.16	0.64	0.45	1.13	0.56	0.31	1.10
Aboriginal	0.69	0.52	1.42	0.69	0.52	1.44	0.64	0.41	1.35
HCV RNA-positive at baseline	8.05	5.63	14.21	6.02	4.33	13.86	6.45	3.62	11.78
HCV duration	0.99	0.97	1.02	0.98	0.97	1.01	0.97	0.93	1.01
Hepatitis B surface antigen-positive at baseline	0.61	0.54	1.28	0.43	0.36	1.12	0.45	0.31	1.49
Previous alcohol	0.75	0.42	1.33	0.47	0.28	1.30	0.60	0.23	1.54

Previous CD4 (per 100 cells)	1.00	0.99	1.00	0.99	0.99	1.00	0.99	0.99	1.00
Previous HIV-RNA	0.97	0.94	1.04	0.99	0.96	1.02	0.88	0.72	1.14
Previous IDU	0.59	0.35	1.12	0.55	0.34	1.08	0.70	0.56	1.20
Previous ART	1.02	0.55	1.62	0.35	0.22	1.27	1.35	0.67	2.92
Previous smoking	0.91	0.65	1.58	0.38	0.13	1.39	0.44	0.15	1.22

Numerator

Age at baseline	1.01	0.98	1.02	1.02	0.99	1.03	1.01	0.99	1.02
Baseline lnAPRI	3.21	1.58	5.62	2.55	1.16	5.14	1.95	1.18	4.67
Female	0.85	0.79	1.17	0.62	0.47	1.13	0.95	0.65	1.09
Aboriginal	0.88	0.61	1.23	0.62	0.60	1.17	0.50	0.31	1.12
HCV RNA-positive at baseline	7.42	3.87	15.29	5.56	3.29	12.73	5.35	3.27	11.42
HCV duration	0.99	0.99	1.01	0.97	0.94	1.00	0.99	0.99	1.00
Hepatitis B surface antigen-positive at baseline	0.69	0.45	1.32	0.38	0.17	1.15	0.91	0.25	1.65

Abbreviations: ART, antiretroviral therapy; CL, confidence limit; HCV, hepatitis C virus; IDU, injection drug use; MI, multiple imputation; lnAPRI, natural logarithm of the AST-to-platelet ratio index; LOCF, last observation carried forward; OR, odds ratio.

Web Table 2. Model coefficients for the treatment weighting in the available case analysis of the Canadian HIV–Hepatitis C Co-infection Cohort Study. The coefficients are averaged over the 20 imputed data sets; confidence intervals were obtained by bootstrap.

Variable	Treatment Model: Denominator			Treatment Model: Numerator		
	OR	Lower 95% CL	Upper 95% CL	OR	Lower 95% CL	Upper 95% CL
Age at baseline	0.92	0.84	0.99	0.91	0.84	0.99
Baseline lnAPRI	1.13	0.72	1.79	1.16	0.76	1.84
Female	0.8	0.58	1.07	0.81	0.60	1.08
Aboriginal	1.11	0.76	1.59	1.15	0.79	1.65
HCV RNA-positive at baseline	0.95	0.68	1.31	0.92	0.68	1.29
HCV duration	0.99	0.93	1.06	1.01	0.95	1.08
Hepatitis B surface antigen-positive at baseline	1.36	0.65	2.53	1.44	0.69	2.57
Previous alcohol abuse	10.82	7.95	13.88	0.91	0.84	0.99
Previous CD4 (100 cells/ μ L)	0.96	0.92	1.00			
Previous HIV RNA (log copies/ml)	0.96	0.91	1.02			
Previous IDU	0.88	0.68	1.14			
Previous ART	1.20	0.76	1.81			
Previous smoking	1.71	1.28	2.49			

Age and HCV duration were scaled so that the coefficient is associated with a change of 5 years.

Abbreviations: ART, antiretroviral therapy; CL, confidence limit; HCV, hepatitis C virus; IDU, injection drug use; lnAPRI, natural logarithm of the AST-to-platelet ratio index; OR, odds ratio.

Web Table 3. Model coefficients for the treatment weighting in the LOCF analysis of the Canadian HIV–Hepatitis C Co-infection Cohort Study. The coefficients are averaged over the 20 imputed data sets; confidence intervals were obtained by bootstrap.

Variable	Treatment Model: Denominator			Treatment Model: Numerator		
	OR	Lower 95% CL	Upper 95% CL	OR	Lower 95% CL	Upper 95% CL
Age at baseline	1.00	0.87	1.10	1.00	0.87	1.09
Baseline lnAPRI	1.29	0.74	2.82	1.30	0.73	2.63
Female	0.78	0.51	1.22	0.80	0.52	1.24
Aboriginal	1.38	0.89	2.44	1.33	0.91	2.51
HCV RNA-positive at baseline	1.05	0.52	1.54	1.03	0.53	1.50
HCV duration	0.99	0.92	1.11	0.99	0.92	1.08
Hepatitis B surface antigen-positive at baseline	1.22	0.35	2.42	1.22	0.35	2.43
Previous alcohol abuse	8.02	4.59	11.03			
Previous CD4 (100 cells/ μ L)	1.00	0.91	1.05			
Previous HIV RNA (log copies/ml)	1.03	0.86	1.04			
Previous IDU	0.88	0.62	1.48			
Previous ART	1.15	0.84	3.46			
Previous smoking	0.90	0.69	1.73			

Age and HCV duration were scaled so that the coefficient is associated with a change of 5 years.

Abbreviations: ART, antiretroviral therapy; CL, confidence limit; HCV, hepatitis C virus; IDU, injection drug use; lnAPRI, natural logarithm of the AST-to-platelet ratio index; OR, odds ratio.

Web Table 4. Model coefficients for the treatment weighting in the multiple imputation analysis of the Canadian HIV–Hepatitis C Co-infection Cohort Study. The coefficients are averaged over the 20 imputed data sets; confidence intervals were obtained by bootstrap.

Variable	Treatment Model: Denominator			Treatment Model: Numerator		
	OR	Lower 95% CL	Upper 95% CL	OR	Lower 95% CL	Upper 95% CL
Age at baseline	0.99	0.85	1.00	0.99	0.88	0.99
Baseline lnAPRI	1.14	0.80	1.66	1.18	0.82	1.69
Female	0.86	0.65	1.02	0.85	0.66	1.04
Aboriginal	1.15	0.86	1.54	1.16	0.85	1.55
HCV RNA-positive at baseline	0.99	0.76	1.32	0.96	0.75	1.30
HCV duration	1.02	0.95	1.06	0.99	0.95	1.06
Hepatitis B surface antigen-positive at baseline	1.42	0.86	2.37	1.43	0.83	2.42
Previous alcohol abuse	5.60	4.72	7.10			
Previous CD4 (100 cells/ μ L)	0.99	0.94	1.00			
Previous HIV RNA (log copies/ml)	0.97	0.97	1.04			
Previous IDU	0.87	0.73	1.05			
Previous ART	1.09	0.87	1.37			
Previous smoking	1.35	1.18	1.71			

Age and HCV duration were scaled so that the coefficient is associated with a change of 5 years.

Abbreviations: ART, antiretroviral therapy; CL, confidence limit; HCV, hepatitis C virus; IDU, injection drug use; lnAPRI, natural logarithm of the AST-to-platelet ratio index; OR, odds ratio.

Web Table 5. Risk Associated with Alcohol Abuse in the Canadian HIV–Hepatitis C Co-infection Cohort Study. Inverse probability weights were truncated at the 99th percentile.

Method of Analysis	Variable	OR	Lower 95% CL	Upper 95% CL
Available case	<i>Alcohol abuse</i>	1.57	0.91	2.37
	Age at baseline	1.06	0.94	1.18
	Baseline lnAPRI	8.24	4.21	18.66
	Female	1.65	0.95	2.67
	Aboriginal	0.52	0.19	1.05
	HCV RNA-positive at baseline	1.56	0.86	4.51
	HCV duration	1.00	0.89	1.21
	Hepatitis B surface antigen-positive at baseline	1.00	0.00	1.92
LOCF	<i>Alcohol abuse</i>	1.76	0.65	2.63
	Age at baseline	1.00	0.86	1.15
	Baseline lnAPRI	6.45	2.04	11.94
	Female	1.35	0.91	2.77
	Aboriginal	0.86	0.33	1.34
	HCV RNA-positive at baseline	1.55	0.73	2.22
	HCV duration	1.00	0.92	1.16
	Hepatitis B surface antigen-positive at baseline	0.86	0.00	1.18

(Table continues)

MI	<i>Alcohol abuse</i>	<i>1.34</i>	<i>1.11</i>	<i>1.58</i>
	Age at baseline	0.99	0.85	1.09
	Baseline lnAPRI	4.95	3.14	7.24
	Female	1.26	0.89	1.53
	Aboriginal	1.02	0.63	1.47
	HCV RNA-positive at baseline	1.12	0.94	2.02
	HCV duration	1.00	0.97	1.08
	Hepatitis B surface antigen-positive at baseline	1.13	0.64	2.01

Abbreviations: ART, antiretroviral therapy; CL, confidence limit; HCV, hepatitis C virus; IDU, injection drug use; MI, multiple imputation; lnAPRI, natural logarithm of the AST-to-platelet ratio index; LOCF, last observation carried forward; OR, odds ratio.