Received <day> <Month>, <year>; | Revised <day> <Month>, <year>; | Accepted <day> <Month>, <year>

DOI: xxx/xxxx

ARTICLE TYPE

Model-assisted analyses of longitudinal, ordinal outcomes with absorbing states: Online Supplement

Jonathan S. Schildcrout[1] | Frank E. Harrell, Jr.[1] | Patrick J. Heagerty[2] | Sebastien Haneuse[3] | Chiara Di Gravio[1] | Shawn Garbett[1] | Paul J. Rathouz[4] | Bryan E. Shepherd[2]

[1]Department of Biostatistics, Vanderbilt University Medical Center, Nashville, Tennessee 37232, U.S.A.
[2]Department of Biostatistics, University of Washington School of Public Health, Seattle, WA U.S.A.
[3]Department of Biostatistics, T.H. Chan School of Public Health, Harvard University, Boston, MA, U.S.A.
[4]Department of Population Health, Dell Medical Center, University of Texas, Austin Texas, U.S.A.

Correspondence
*Jonathan S. Schildcrout, Department of Biostatistics, Vanderbilt University Medical Center, Nashville, Tennessee 37232, U.S.A.
Email: jonny.schild@vanderbilt.edu

Online Supplement

A1 | ADDITIONAL SIMULATION STUDY SCENARIO: LOW RESPONSE DEPENDENCE

We expand the simulation study shown in Section 5 to include a lower response dependence scenario. Specifically we set the dependence model parameter matrix, $\mathbf{\Gamma}$ to be $\begin{pmatrix} 20 & 0 & 0 \\ -20 & 2 & 0 \\ -20 & 0 & 2 \end{pmatrix}$ which induced a day to day state transition matrix shown in Table A1. We show the resulting average estimates, empirical standard errors and coverages in Table A2. We observe that in this setting, when the proportional odds assumption is violated, β_{xt}^{ppo1} may not be a weighted average among dichotomization-specific interaction estimates from $ppo2$. For example, when $(\beta_{xt,2}^{ppo2}, \beta_{xt,3}^{ppo2}) = (-0.2, -0.4)$, $\beta_{xt}^{ppo1} = -0.297$ is not a weighted average of the three estimated interactions from $ppo2$, i.e., 0.206 (for $k \leq 1$), $0.005 = 0.206 - 0.201$ (for $k \leq 2$), and $-0.194 = 0.206 - 0.400$ (for $k \leq 3$). This is also shown in the top right panel of Figure A1 where, under the severe PO assumption violation, the $MB_{ppo1}(t_{ij})$ estimators capture a distinct quantity as compared to the other estimators. We highlight that this is a very severe form of PO assumption violation that does not resemble the VIOLET data.

A2 | LONGITUDINAL VERSUS CROSS-SECTIONAL MODELING COMPARISONS

In this simulation study, to examine the relative benefit of using longitudinal modeling approaches over a cross-sectional approach, we generated data under the $(\beta_{xt,2}^{ppo2}, \beta_{xt,3}^{ppo2}) = (0, 0)$ setting which corresponds to $ppo1$ or Model 1. We studied the

TABLE A1: Transition matrix for the low response dependence setting

$Y_i(t_{ij})$	$Y_i(t_{ij-1})$			
	1	2	3	4
1	1.000	0.008	0.008	0.013
2	0.000	0.418	0.084	0.125
3	0.000	0.161	0.612	0.193
4	0.000	0.413	0.296	0.668

TABLE A2: Average parameter estimates, empirical standard errors, and coverage probabilities across 1000 replicates for model fits from *ppo*1 and *ppo*2. In all cases, data were generated from a *ppo*2 model and fit with *ppo*2 and *ppo*1 models. In settings where $(\beta^{ppo2}_{xt,2}, \beta^{ppo2}_{xt,3}) = (0,0)$, the *ppo*1 model is correctly specified. When this is not the case, it is misspecified and estimated coefficients differ from *ppo*2 model. In this table we only show the low response dependence scenario

			*ppo*2 estimates				*ppo*1 estimates	
Dependence	$(\beta^{ppo2}_{xt,2}, \beta^{ppo2}_{xt,3})$		$\beta^{ppo2}_{x,0}$	β^{ppo2}_{xt}	$\beta^{ppo2}_{xt,2}$	$\beta^{ppo2}_{xt,3}$	$\beta^{ppo1}_{x,0}$	β^{ppo1}_{xt}
Low	(0,0)	True value	-0.1	0.2	0	0	-0.1	0.2
		Average Estimate	-0.098	0.188	0.005	0.007	-0.098	0.197
		Standard Error	[0.093]	[0.180]	[0.075]	[0.102]	[0.093]	[0.131]
		Coverage	(0.941)	(0.944)	(0.955)	(0.952)	(0.941)	(0.946)
	(-0.1,-0.2)	True value	-0.1	0.2	-0.1	-0.2	NA	NA
		Average Estimate	-0.094	0.189	-0.098	-0.196	-0.089	-0.049
		Standard Error	[0.087]	[0.171]	[0.069]	[0.096]	[0.087]	[0.125]
		Coverage	(0.954)	(0.961)	(0.948)	(0.960)	NA	NA
	(-0.2,-0.4)	True value	-0.1	0.2	-0.2	-0.4	NA	NA
		Average Estimate	-0.104	0.206	-0.201	-0.400	-0.089	-0.297
		Standard Error	[0.091]	[0.182]	[0.065]	[0.094]	[0.091]	[0.134]
		Coverage	(0.943)	(0.948)	(0.948)	(0.949)	NA	NA

high and moderate response dependence and with a) complete follow-up (similar to VIOLET), 2) dropout completely at random (DCAR) with 50% of subjects at risk for uniform dropout from $j = 5, \ldots, 19$, and 3) dropout at random (DAR) with a probability of 0.075 of being censored after each observation in the highest ($k = 4$; at-home) outcome state. We then compared the averaged estimates and empirical standard errors for three estimators of $MB_{ppo1}(t_{ij} = 10)$: 1) OMTM1 summarized at $t_{ij} = 10$, 2) GEE with an independence working covariance structure (GEE-I) summarized at $t_{ij} = 10$, and 3) a cross-sectional analysis at $t_{ij} = 10$ that ignores the longitudinal structure. We expect that by using longitudinal methods, e.g., OMTM1 and GEE-I, we can improve efficiency of the estimators by exploiting the linear (on the log_{10} scale) structure of the intervention effect estimates over time. Among the longitudinal modeling approaches, by using OMTM1 instead of GEE-I, we can improve efficiency further by incorporating an accurate model for higher moments in addition to the univariate CPM captured by GEE-I. The likelihood-based OMTM1 estimator, which is based on a properly specified model, is also anticipated to be more robust to the DAR mechanism we study here. We summarize results across 750 replicates.

The data generating model yielded a treatment effect estimate log odds ratio at $t_{ij} = 10$ to be equal to 0.1. As shown in Table A3, with complete follow-up data and in the high response data setting (most similar to VIOLET), the average parameter estimates (empirical standard errors) for the OMTM1, GEE-I, and cross-sectional analyses were 0.100 (0.102), 0.100 (0.113) and 0.099 (0.120), respectively. Compared to the cross-sectional estimator, the OMTM1 and GEE-I

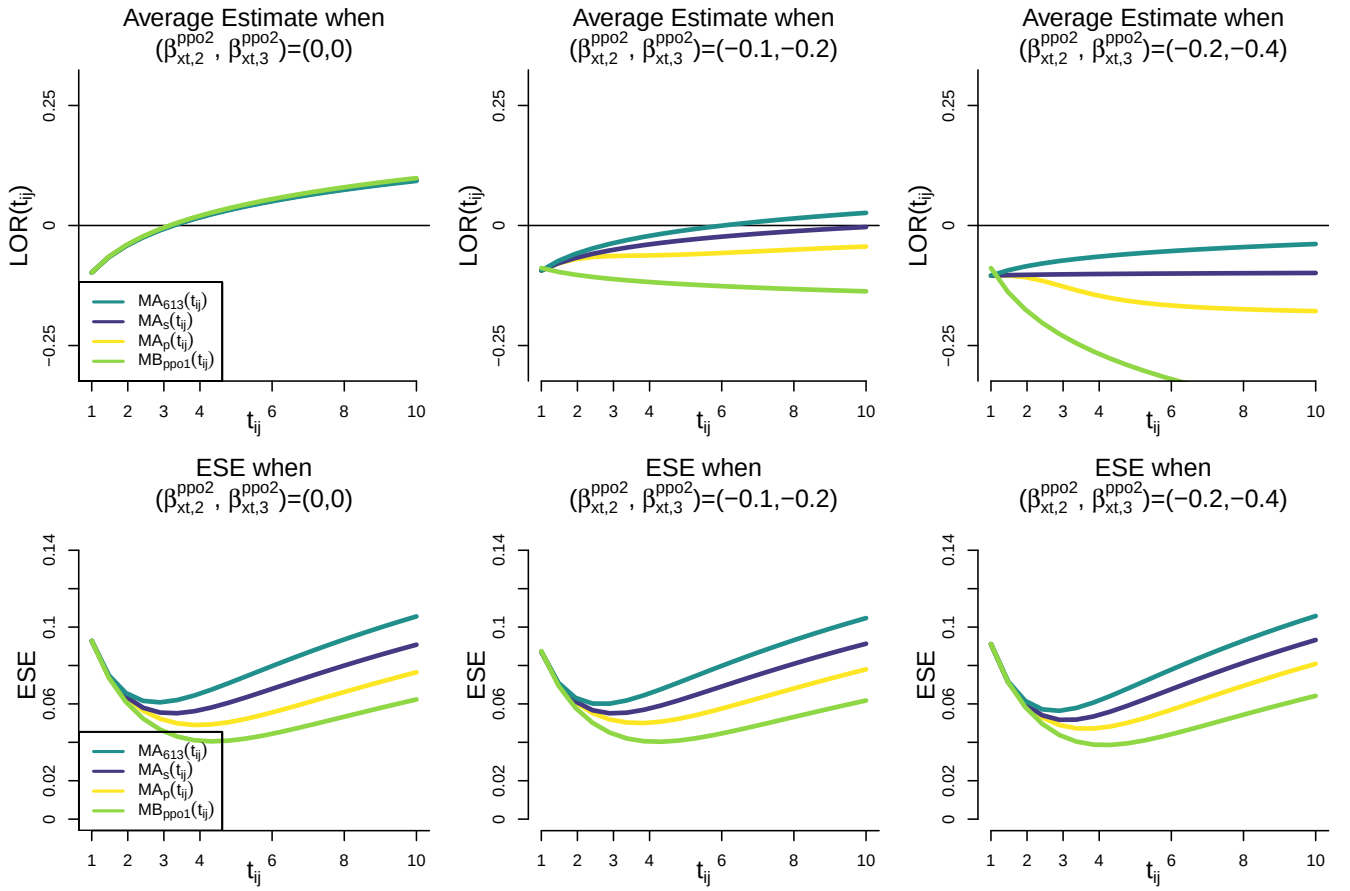

FIGURE A1: Model-based and model-assisted summaries: For low response dependence settings, we report averages of the $t_{ij}-$specific estimates of intervention effects using a global log odds ratio $[LOR(t_{ij})]$ and the empirical standard errors (ESE) across 1000 replicates. We includes scenarios with no violation $[(\beta^{ppo2}_{xt,2}, \beta^{ppo2}_{xt,3}) = (0,0)]$, moderate violation $[(\beta^{ppo2}_{xt,2}, \beta^{ppo2}_{xt,3}) = (-0.1,-0.2)]$, and severe violation $[(\beta^{ppo2}_{xt,2}, \beta^{ppo2}_{xt,3}) = (-0.2,-0.4)]$ of the proportional odds assumption for the intervention effect.

estimators had relative efficiencies of $(0.120/0.102)^2 = 1.38$ and $(0.113/0.102)^2 = 1.23$, respectively. With moderate response dependence and/or DCAR, relative efficiency for the longitudinal approaches over cross-sectional analyses were even more dramatic and were as high as $(0.148/0.092)^2 = 2.59$ under DCAR and moderate response dependence. We can also see, particularly with high response dependence, the improved accuracy of the OMTM1 estimators under DAR.

□

TABLE A3: Longitudinal versus cross-sectional modeling: Average Estimates (Empirical Standard Errors) for the $MB_{ppo1}(t_{ij} = 10)$ intervention effect.

		Response Dependence	
	Estimation	High	Moderate
Complete Follow-up	OMTM1	0.100 (0.102)	0.098 (0.087)
	GEE-I	0.100 (0.113)	0.100 (0.104)
	Cross-sectional	0.099 (0.120)	0.100 (0.129)
DCAR	OMTM1	0.102 (0.109)	0.100 (0.092)
	GEE-I	0.100 (0.127)	0.100 (0.113)
	Cross-sectional	0.099 (0.144)	0.100 (0.148)
DAR	OMTM1	0.101 (0.112)	0.098 (0.100)
	GEE-I	0.095 (0.124)	0.097 (0.121)
	Cross-sectional	0.087 (0.142)	0.094 (0.161)