Modeling Longitudinal Binary Outcomes in a Small Matched-Pair Sample with Application to Cardiovascular Data: A Simulation Study

Jinyu Luo
University of Toronto
jinyu.luo@mail.utoronto.ca

Aya Mitani
University of Toronto
aya.mitani@utoronto.ca

Chun-Po Steve Fan
University Health Network
steve.fan@uhn.ca

Sudipta Saha
University Health Network
Sudipta.Saha@uhn.ca

Peiyu Li
University of Toronto
peiyu.li@mail.utoronto.ca

Abstract

Background: This study aimed to address the challenge of modelling small sample matched-pair longitudinal data in cardiovascular research. The independent working correlation structure in Generalized Estimating Equations (GEE), a robust method widely used to model endogenous follow-up data, relies on large-sample theory. Prior research noted significant constraints due to small sample sizes for continuous outcomes. Objectives: We evaluated the validity of the working independent correlation structure in GEE, specifically focusing on binary outcomes, through a simulation study. Methods: Initially, real hospital data were fitted, assuming a working exchangeable correlation structure to estimate the true values for simulation parameters. The simulations were designed to mimic the dropout process in the real-world scenario where previous survival outcomes and associated covariates influence the longitudinal outcomes. The simulated data involves hospital cohorts with longitudinal outcomes across two exposure groups, so cohorts were matched using propensity scores based on baseline characteristics to eliminate potential confounding effects in demographic and clinical characteristics. Due to the small sample size, standard errors were adjusted by degrees of freedom to prevent underestimation by the sandwich estimator. The simulated data was then analyzed using three different correlation structures in GEE: independent, exchangeable, and autoregressive (AR1). Results: The independent working correlation structure in GEEs consistently provides the

highest coverage probabilities for true parameter coefficients after adjusting for standard errors. Conclusion: Proper specification of the correlation structure is important for the robust analysis of small sample longitudinal data.

Introduction

Bicuspid aortic valve (BAV) is a congenital heart defect where the aortic valve has only two leaflets instead of three, affects approximately 1-2% of the general population and is frequently associated with aortic stenosis, regurgitation, and ascending aortic aneurysms (Hui et al., 2018). In contrast, tricuspid aortic valve (TAV) is the normal anatomical variant with three leaflets, can also develop pathological conditions, though it is not congenital. In both instances, regular monitoring is crucial because the majority of patients, especially those with aortic aneurysm, usually experience aortic dilation, which is associated with an increased risk of adverse aortic events and represents a potentially lethal condition, making it a considerable medical burden. However, the random dropout of patients from hospital monitoring programs poses a significant challenge, as it results in incomplete data, complicating statistical analysis and the development of effective prediction models.

It is evident that having aortic valve replacement can lead to better outcomes compared to medical therapy in both cases, as relief of severe stenosis is associated with improved patient results (Huntley et al., 2018).

Methods

First, please get familiar with the following resources:

- Creating Formats in Quarto as part of the Extensions mechanism.
- Journals Articles for Quarto.

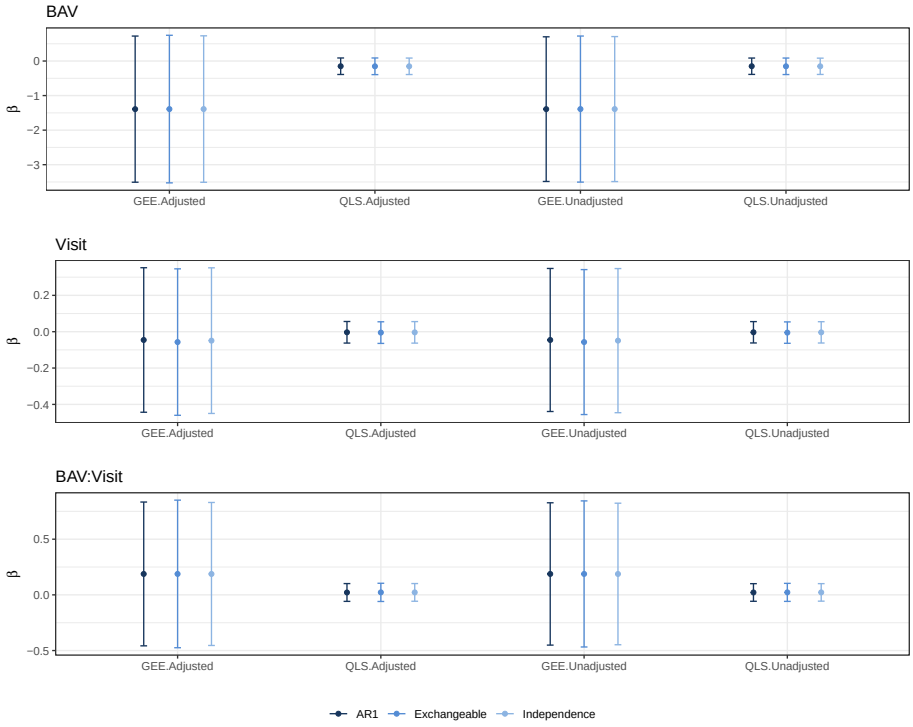
Structure of this repository

Everything for the extensions is in `_extensions`. See Quarto doc for details.

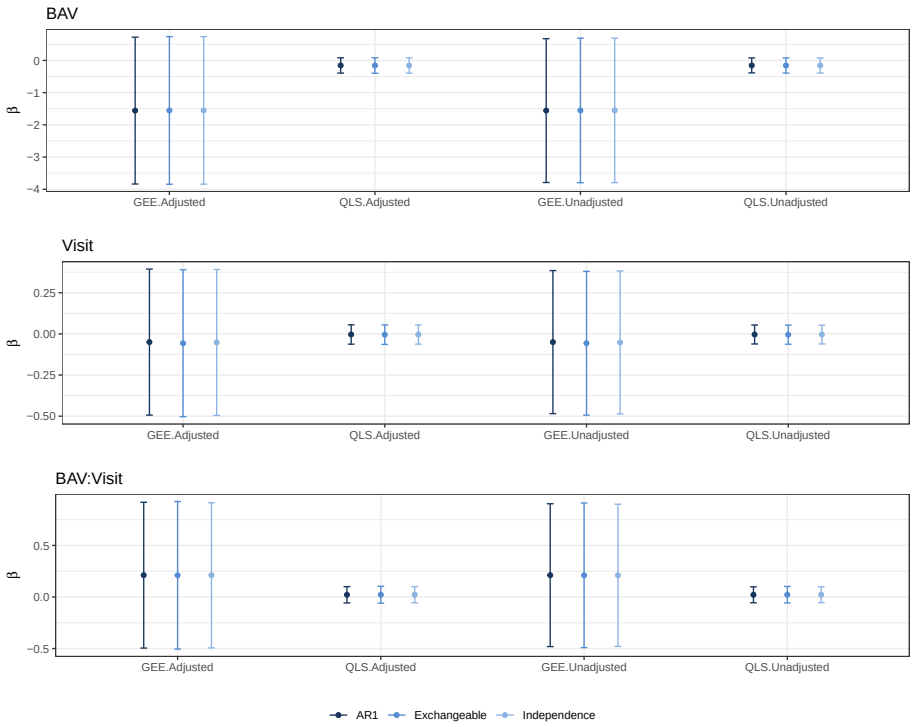
- In `partials`, you'll find the `.tex` partials that can be used and should be removed or tweaked,s
- Your extension can make shortcodes and lua filters available. This document shows the effect of the one provided in the `aft` format.
- `aft` format sets some defaults which are different from `pdf` or `html`, link setting links to URL in read inside PDF output.

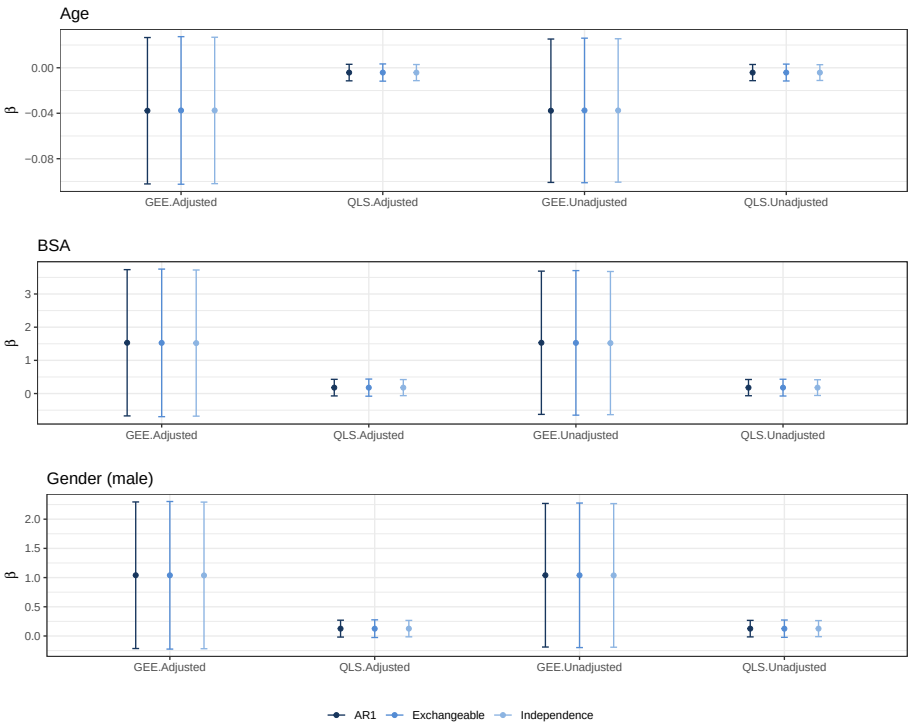
Source repository for this template format can found on Github

Code chunk


BAV
β
0
−1
−2
−3
GEE.Adjusted
QLS.Adjusted
GEE.Unadjusted
QLS.Unadjusted
Visit
β
0.2
0.0
−0.2
−0.4
GEE.Adjusted
QLS.Adjusted
GEE.Unadjusted
QLS.Unadjusted
BAV:Visit
β
0.5
0.0
−0.5
GEE.Adjusted
QLS.Adjusted
GEE.Unadjusted
QLS.Unadjusted
AR1
Exchangeable
Independence

add horizontal line representing the true value





	Correlation	Marginal Model	True Value	rho Estimate	SD	Bias	Relative Bias	MSE
Unadjusted	AR1	GEE	0.3	0.207	0.154	-0.093	-0.310	0.032
		QLS	0.3	0.249	0.180	-0.051	-0.171	0.035
	Exchangeable	GEE	0.3	0.098	0.116	-0.202	-0.675	0.055
		QLS	0.3	0.249	0.180	-0.051	-0.171	0.035
Adjusted	AR1	GEE	0.3	0.304	0.160	0.004	0.014	0.026
		QLS	0.3	0.306	0.178	0.006	0.020	0.032
	Exchangeable	GEE	0.3	0.179	0.124	-0.121	-0.404	0.030
		QLS	0.3	0.161	0.102	-0.139	-0.463	0.030

Correlation	Type	Term	True Value	Mean Estimate	SD	Bias	Relative Bias	MSE
Independence	GEE	Intercept	-2.412	-1.293	0.705	1.119	0.464	1.750
		BAV	-1.327	-1.390	1.853	-0.064	-0.048	3.438
		BAV:Visit	0.203	0.188	0.430	-0.015	-0.075	0.185
		Visit	-0.025	-0.049	0.249	-0.024	-0.985	0.063
	QLS	Intercept	-0.005	0.217	0.107	0.222	41.212	0.061
		BAV	0.010	-0.153	0.131	-0.162	-16.336	0.044
		BAV:Visit	0.016	0.022	0.046	0.006	0.388	0.002
		Visit	-0.124	-0.004	0.033	0.120	0.971	0.016

SD and MSE: * Standard Deviation and Mean Squared Error

But you can set `echo` option to `true` locally in the chunk

include mean SD

Correlation	Type	Term	True Value	Mean Estimate	SD	Bias	Relative Bias	MSE
AR1	GEE	Intercept	-2.412	-1.309	0.698	1.104	0.458	1.706
		BAV	-1.327	-1.392	1.833	-0.066	-0.050	3.365
		BAV:Visit	0.203	0.188	0.420	-0.015	-0.073	0.177
		Visit	-0.025	-0.046	0.243	-0.021	-0.839	0.060
	QLS	Intercept	-0.005	0.215	0.105	0.220	40.889	0.060
		BAV	0.010	-0.150	0.128	-0.160	-16.181	0.042
		BAV:Visit	0.016	0.021	0.045	0.005	0.343	0.002
		Visit	-0.124	-0.003	0.032	0.121	0.975	0.016

SD and MSE: * Standard Deviation and Mean Squared Error

Correlation	Type	Term	True Value	Mean Estimate	SD	Bias	Relative Bias	MSE
Exchangeable	GEE	Intercept	-2.412	-1.286	0.707	1.127	0.467	1.768
		BAV	-1.327	-1.391	1.862	-0.064	-0.048	3.471
		BAV:Visit	0.203	0.188	0.439	-0.015	-0.073	0.193
		Visit	-0.025	-0.057	0.247	-0.032	-1.304	0.062
	QLS	Intercept	-0.005	0.218	0.107	0.223	41.441	0.061
		BAV	0.010	-0.153	0.133	-0.162	-16.388	0.044
		BAV:Visit	0.016	0.022	0.047	0.006	0.401	0.002
		Visit	-0.124	-0.005	0.033	0.119	0.960	0.015

SD and MSE: * Standard Deviation and Mean Squared Error

Corstr	Type	Term	True Value	Mean Estimate	SD	Bias	Relative Bias	MSE
Independence	GEE	Intercept	-2.412	-2.603	3.160	-0.190	-0.079	10.020
		Age	-0.033	-0.038	0.040	-0.004	-0.120	0.002
		BAV	-1.327	-1.552	1.963	-0.225	-0.170	3.904
		BAV:Visit	0.203	0.210	0.473	0.007	0.037	0.224
		BSA	1.376	1.521	1.385	0.145	0.105	1.940
		Male	0.987	1.037	1.628	0.050	0.051	2.652
		Visit	-0.025	-0.052	0.279	-0.027	-1.100	0.078
	QLS	Intercept	-0.005	0.079	0.340	0.085	15.701	0.123
		Age	-0.004	-0.004	0.004	0.000	-0.018	0.000
		BAV	0.010	-0.157	0.133	-0.167	-16.798	0.045
		BAV:Visit	0.016	0.022	0.046	0.006	0.367	0.002
		BSA	0.211	0.179	0.151	-0.032	-0.151	0.024
		Male	0.104	0.126	0.086	0.022	0.216	0.008
		Visit	-0.124	-0.004	0.033	0.120	0.968	0.016

Corstr, SD, MSE: * Correlation Structure, Standard Deviation and Mean Squared Error

Text color

Our format makes applying color on inline text possible using the `[content]{color=<name>}` syntax. Let's see an example.

Here we are using a special feature of our format which is the coloring because pink is a **nice** color.

Corstr	Type	Term	True Value	Mean Estimate	SD	Bias	Relative Bias	MSE
AR1	GEE	Intercept	-2.412	-2.618	3.152	-0.206	-0.085	9.978
		Age	-0.033	-0.038	0.040	-0.004	-0.128	0.002
		BAV	-1.327	-1.557	1.949	-0.231	-0.174	3.854
		BAV:Visit	0.203	0.212	0.467	0.009	0.043	0.218
		BSA	1.376	1.530	1.383	0.154	0.112	1.936
		Male	0.987	1.040	1.623	0.053	0.054	2.638
		Visit	-0.025	-0.050	0.276	-0.025	-1.011	0.077
	QLS	Intercept	-2.412	0.078	0.336	2.490	1.032	6.313
		Age	-0.033	-0.004	0.004	0.029	0.874	0.001
		BAV	-1.327	-0.155	0.129	1.172	0.883	1.389
		BAV:Visit	0.203	0.021	0.045	-0.182	-0.895	0.035
		BSA	1.376	0.179	0.150	-1.197	-0.870	1.456
		Male	0.987	0.126	0.085	-0.861	-0.872	0.749
		Visit	-0.025	-0.004	0.032	0.021	0.856	0.001

Corstr, SD, MSE: * Correlation Structure, Standard Deviation and Mean Squared Error

Corstr	Type	Term	True Value	Mean Estimate	SD	Bias	Relative Bias	MSE
Exchangeable	GEE	Intercept	-2.412	-2.612	3.152	-0.200	-0.083	9.973
		Age	-0.033	-0.038	0.040	-0.004	-0.120	0.002
		BAV	-1.327	-1.553	1.970	-0.226	-0.170	3.933
		BAV:Visit	0.203	0.210	0.482	0.007	0.037	0.232
		BSA	1.376	1.527	1.402	0.151	0.110	1.989
		Male	0.987	1.039	1.630	0.052	0.052	2.658
		Visit	-0.025	-0.057	0.280	-0.032	-1.288	0.079
	QLS	Intercept	-2.412	0.079	0.337	2.491	1.033	6.320
		Age	-0.033	-0.004	0.004	0.029	0.875	0.001
		BAV	-1.327	-0.157	0.133	1.170	0.882	1.387
		BAV:Visit	0.203	0.022	0.047	-0.181	-0.891	0.035
		BSA	1.376	0.179	0.149	-1.198	-0.870	1.456
		Male	0.987	0.125	0.085	-0.862	-0.873	0.750
		Visit	-0.025	-0.005	0.033	0.020	0.810	0.002

Corstr, SD, MSE: * Correlation Structure, Standard Deviation and Mean Squared Error

This is possible thanks to the Lua Filter included in the custom extension format.

Using references

I did not read this book (Cameron and Trivedi, 2013) but it must be interesting.

Differences between `aft-html` and `aft-pdf`:

- For the HTML format, we are using Pandoc citeproc to include the bibliog-

raphy. Here `reference-section-title` controls the title for the chapter that will be used.
- For the PDF format, `natbib` is used by default and the bibliography is included with a title by the LaTeX template.

References

Cameron, A. Colin and Pravin K. Trivedi (2013), *Regression Analysis of Count Data*, 2nd edition. Cambridge University Press, Cambridge.

Hui, Sonya K., Chun-Po Steve Fan, Shakira Christie, Christopher M. Feindel, Tirone E. David, and Maral Ouzounian (2018), "The aortic root does not dilate over time after replacement of the aortic valve and ascending aorta in patients with bicuspid or tricuspid aortic valves." *The Journal of Thoracic and Cardiovascular Surgery*, 156, 5–13.e1, URL https://doi.org/10.1016/j.jtcvs.2018.02.094.

Huntley, Geoffrey D, Jeremy J Thaden, Said Alsidawi, Hector I Michelena, Joseph J Maleszewski, William D Edwards, Christopher G Scott, Sorin V Pislaru, Patricia A Pellikka, Kevin L Greason, Naser M Ammash, Joseph F Malouf, Maurice Enriquez-Sarano, and Vuyisile T Nkomo (2018), "Comparative study of bicuspid vs. tricuspid aortic valve stenosis." *European Heart Journal - Cardiovascular Imaging*, 19, 3–8, URL https://doi.org/10.1093/ehjci/jex211.