**Online Appendix 1: Simulation Methods and Results for "SMART Longitudinal Analysis: A Tutorial for Using Repeated Outcome Measures from SMART Studies to Compare Adaptive Interventions" (Nahum-Shani et al)**

A simulation study was done to demonstrate the performance of the method proposed in the paper, and to compare the performance of variations of the method (known versus estimated weights, and different working correlation structures).

## Methods

A total of 5000 datasets were generated, each containing 500 simulated participants. Participants were responders or nonresponders with equal probability. There were assumed to be three measurement times: (1) immediately before the first-stage randomization; (2), after the first-stage randomization and treatment and immediately before the second-stage randomization; and (3) following the second-stage randomization and treatment. These were assumed to be equally spaced. Only non-responders were re-randomized.

There were a total of twelve possible mean parameters, representing the means of responders and nonresponders under each treatment at each time point, after removing means which would be redundant. For example, all adaptive interventions should have the same means at the first time point, because randomization had not yet occurred; however, future responders and future nonresponders could still have different means even before randomization, due to unobserved covariates. The values selected, loosely based on descriptive results from the ENGAGE dataset, are shown in Table A1. Conditional on response status (responder vs. non-responder) and on underlying adaptive intervention, outcomes at each time point were assumed to have an error variance of 5 units, and to have one of the following four conditional correlation structures:

- Independence: $\text{Corr}(Y_0,Y_1)=\text{Corr}(Y_0,Y_2)=\text{Corr}(Y_1,Y_2)=0.00$
- Autoregressive (AR-1): $\text{Corr}(Y_0,Y_1)=\text{Corr}(Y_1,Y_2)=\rho$; $\text{Corr}(Y_0,Y_2)=\rho^2$
- Exchangeable: $\text{Corr}(Y_0,Y_1)=\text{Corr}(Y_0,Y_2)=\text{Corr}(Y_1,Y_2)=\rho$
- Checkerboard: $\text{Corr}(Y_0,Y_1)=\text{Corr}(Y_1,Y_2)=0$; $\text{Corr}(Y_0,Y_2)=\rho$

where $\rho = 0.5$. The first three correlation structures are fairly well known to GEE users. The fourth was deliberately constructed to be unusual in an attempt to provide a more difficult test for the methods.

The method of Lu et al (2015) was applied to each dataset, under each of three working structures: independence, AR-1, or exchangeable. Also, weights were either treated as known or were re-estimated as described in the manuscript. Six parameters were assumed to be of interest, all of them contrasts between areas under the curve for different embedded adaptive interventions: (1,1) versus (1,-1), (1,1) versus (-1,1), (1,1) versus (-1,-1), (1,-1) versus (-1,+1), (1,-1) versus (-1,-1), and (-1,1) versus (-1,-1).

Four performance measures were of interest:

- Absolute bias, averaged across datasets and across the six parameters of interest
- Mean squared error MSE, averaged across datasets and across the six parameters of interest (this is expressed below as square root of MSE, because in theory this should consistently estimate the true standard error)
- Estimated standard error, averaged across datasets and across the six parameters of interest
- Simulated coverage for nominal 95% confidence intervals, averaged across datasets and across the six parameters of interest

Simulations were performed in R version 3.2.1. The 5000 simulations for the 9 covariance conditions and 2 weighting conditions, thus 90000 model fits, took about a day of computer running time total.

**Results**

Each of the methods gave an error variance estimate of approximately 5.76 on average across simulations. This would appear to be an overestimate of the data-generating variance parameter which was set to 5. However, the variance of 5 was intended to describe the variance of the responses conditional upon response status and embedded adaptive intervention (i.e., upon one of the six possible SMART design cells a-f in Figure 1 of the manuscript), while the fitted GEE model of interest was marginal across response status (i.e., compared only the four embedded adaptive interventions). It is reasonable to expect the marginal variance in this case to be higher than the conditional variance as it is taken over both responders and non-responders.

As shown in Table A2, when the correlation structures were misspecified, the correlation parameter $\rho$ tended to be underestimated relative to its true value of 0.5. When correlation structure was correctly specified, the estimated correlation parameter was on average near or slightly above the true value. The slightly higher correlation estimates may be due again to the difference between the data-generating model, which needed to be fully specified and therefore conditional upon response status, and the analysis model, which followed the proposed method of marginalizing over response status.

As shown in Table A3, under all simulated conditions, the method had practically no estimation bias for the parameters of interest. However, root mean squared error, shown in Table A4, is lower (i.e., precision is better, as well as statistical efficiency) when within-subject correlation is taken into account by using a non-independent working correlation. Root mean squared error (RMSE) is also lower when estimated weights are used compared to known weights. Ignoring within-subject correlation seemed to be less deleterious if estimated weights were being used.

The estimated standard errors, shown in Table A5, are very close to the true standard errors (the root mean squared error values in Table A4), as would be desirable. Average coverage for 95% confidence intervals was almost exactly nominal, as shown in Table A6.

**Conclusions**

Regardless of the choice of weights or working structure, the results are unbiased and have unbiased standard errors, and as a consequence they also have nominal coverage. Statistical efficiency is poorer (hence, RMSE and standard errors are higher) when correlation is not modeled well: that is, when a strong true correlation structure is present (exchangeable or AR-1) but a method which ignores correlation (working independence structure with known weights) is used. Efficiency can be improved by using estimated weights. Correlation apparently does not have to be modeled if estimated weights are used, at least in a context like that of the current simulation. The performance differences between working structures would presumably be larger if there were a larger number of measurement occasions rather than only three.

**Table A1**

**Mean Parameters for Data-Generating Model**

| Parameter | Assumed Value |
|---|---|
| $E(Y_0 \mid R=1)$ | 33.52 |
| $E(Y_0 \mid R=0)$ | 32.49 |
| $E(Y_1 \mid A_1 = +1, R=1)$ | 31.83 |
| $E(Y_1 \mid A_1 = +1, R=0)$ | 30.75 |
| $E(Y_1 \mid A_1 = -1, R=1)$ | 33.33 |
| $E(Y_1 \mid A_1 = -1, R=0)$ | 31.92 |
| $E(Y_2 \mid A_1 = +1, R=1)$ | 32.57 |
| $E(Y_2 \mid A_1 = +1, R=0, A_2 = +1)$ | 30.43 |
| $E(Y_2 \mid A_1 = +1, R=0, A_2 = -1)$ | 27.88 |
| $E(Y_2 \mid A_1 = -1, R=1)$ | 31.78 |
| $E(Y_2 \mid A_1 = -1, R=0, A_2 = +1)$ | 31.71 |
| $E(Y_2 \mid A_1 = -1, R=0, A_2 = -1)$ | 31.29 |

**Table A2**

**Mean estimated correlation parameter (true value 0.50 except for true independence)**

| | | Using Known Weights | | | Using Estimated Weights | | |
|---|---|---|---|---|---|---|---|
| Working Structure: | | Indep. | Exch. | AR-1 | Indep. | Exch. | AR-1 |
| | Indep. | 0.000 | 0.070 | 0.068 | 0.000 | 0.070 | 0.068 |
| True | Exch. | 0.000 | 0.507 | 0.505 | 0.000 | 0.508 | 0.506 |
| Structure: | AR-1 | 0.000 | 0.434 | 0.506 | 0.000 | 0.435 | 0.506 |
| | Check. | 0.000 | 0.216 | 0.068 | 0.000 | 0.217 | 0.069 |

**Table A3**

**Mean absolute bias**

| | | Using Known Weights | | | Using Estimated Weights | | |
|---|---|---|---|---|---|---|---|
| Working Structure: | | Indep. | Exch. | AR-1 | Indep. | Exch. | AR-1 |
| | Indep. | 0.002 | 0.002 | 0.002 | 0.002 | 0.002 | 0.002 |
| True | Exch. | 0.007 | 0.007 | 0.007 | 0.006 | 0.006 | 0.006 |
| Structure: | AR-1 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| | Check. | 0.003 | 0.003 | 0.003 | 0.002 | 0.002 | 0.002 |

**Table A4**

**Square root of mean squared error**

| | | Using Known Weights | | | Using Estimated Weights | | |
|---|---|---|---|---|---|---|---|
| Working Structure: | | Indep. | Exch. | AR-1 | Indep. | Exch. | AR-1 |
| | Indep. | 0.212 | 0.211 | 0.211 | 0.209 | 0.209 | 0.209 |
| True | Exch. | 0.232 | 0.187 | 0.192 | 0.190 | 0.185 | 0.187 |
| Structure: | AR-1 | 0.235 | 0.205 | 0.203 | 0.205 | 0.202 | 0.202 |
| | Check. | 0.210 | 0.200 | 0.208 | 0.199 | 0.197 | 0.199 |

**Table A5**

**Mean standard error estimate**

| | | Using Known Weights | | | Using Estimated Weights | | |
|---|---|---|---|---|---|---|---|
| Working Structure: | | Indep. | Exch. | AR-1 | Indep. | Exch. | AR-1 |
| | Indep. | 0.211 | 0.210 | 0.210 | 0.208 | 0.208 | 0.208 |
| True | Exch. | 0.235 | 0.191 | 0.195 | 0.193 | 0.188 | 0.190 |
| Structure: | AR-1 | 0.235 | 0.205 | 0.203 | 0.204 | 0.201 | 0.201 |
| | Check. | 0.211 | 0.201 | 0.208 | 0.199 | 0.197 | 0.199 |

**Table A6**

**Average empirical coverage of 95% confidence intervals**

| | | Using Known Weights | | | Using Estimated Weights | | |
|---|---|---|---|---|---|---|---|
| Working Structure: | | Indep. | Exch. | AR-1 | Indep. | Exch. | AR-1 |
| | Indep. | 0.950 | 0.949 | 0.950 | 0.949 | 0.949 | 0.949 |
| True | Exch. | 0.951 | 0.954 | 0.953 | 0.953 | 0.953 | 0.953 |
| Structure: | AR-1 | 0.949 | 0.949 | 0.950 | 0.948 | 0.949 | 0.949 |
| | Check. | 0.948 | 0.948 | 0.948 | 0.949 | 0.948 | 0.949 |