Weighted GEEs for Longitudinal Data with Dropouts

Recall: GEEs

Assumptions:

1. Marginal mean & variance: $E(Y_{ij}) = \mu_{ij}$, $var(Y_{ij}) = \phi a_{ij}^{-1} v(\mu_{ij})$
2. Mean model: $g(\mu_{ij}) = \mathbf{X}_{ij}^T \beta$

GEEs:

$$\sum_{i=1}^{m} \mathbf{D_i^T V_i^{-1}} (\mathbf{Y_i} - \boldsymbol{\mu}_\mathbf{i}) = \mathbf{0}$$

where $\mathbf{D}_i = \frac{\partial \boldsymbol{\mu}_i}{\partial \beta^T}$ is $n_i \times p$, $(\mathbf{Y_i} - \boldsymbol{\mu}_\mathbf{i})$ is $n_i \times 1$, and $\mathbf{V}_i$ is an $n_i \times n_i$ working covariance matrix.

Characteristics of GEEs

- Study the relationship between the population-averaged mean response and a set of explanatory variables.

- $\hat{\beta}$ is consistent and asymptotically normal given the mean model $g(\mu_{ij}) = \mathbf{X}_{ij}^T\beta$ is correctly specified even when the correlation matrix $\mathbf{R_i(\alpha)}$ is misspecified.

- If the working correlation $\mathbf{R_i(\alpha)}$ is correctly specified, $\hat{\beta}$ is most efficient within the linear estimating function family.

GEEs with Missing Data (Dropouts)

- Ordinary GEEs only use the observed data and omit the missed occasions (Liang and Zeger, 1986).
- If the working correlation matrix is misspecified:
  - GEEs and sandwich estimators are consistent only under MCAR.
  - If the missingness is MAR or nonignorable, the GEE estimator is generally inconsistent.
- If the working correlation matrix is correctly specified: the GEE estimator is still consistent under MAR, but the sandwich estimator may not be consistent (Kenward and Molenberghs, 1998)

Notation

For each subject i:

$$\mathbf{Y}_i = \begin{pmatrix} \boldsymbol{Y}_{i,obs} \\ \boldsymbol{Y}_{i,mis} \end{pmatrix} = \begin{pmatrix} Y_{i1} \\ \vdots \\ Y_{iJ} \end{pmatrix}_{J\times 1}, \quad \mathbf{X}_i = \begin{pmatrix} X_{i1}^T \\ \vdots \\ X_{iJ}^T \end{pmatrix}_{J\times p},$$

$$M_{ij} = \begin{cases} 1, & Y_{ij} \quad observed \\ 0, & Y_{ij} \quad missing \end{cases}$$

$$R_i = 1 + \sum_{j=1}^{J} M_{ij}$$

Observed data of subject i:

$$(\boldsymbol{Y}_{i,obs}, R_i, \mathbf{X}_i)$$

Some Assumptions

- Monotone missingness, i.e. once a subject "drops out", no more measurements are obtained from this subject.
- The observation times are common for all subjects. (e.g. clinical visit, panel data)

Data Structure:

	pattern of $\mathbf{Y}_i$					pattern of $\mathbf{M}_i$					dropout time R_i
	t_1	t_2	t_3	$\cdots$	t_J	t_1	t_2	t_3	$\cdots$	t_J	
1	x	.	.	$\cdots$	.	1	0	0	$\cdots$	0	2
2	x	x	.	$\cdots$	.	1	1	0	$\cdots$	0	3
$\vdots$			$\cdots$					$\cdots$			$\vdots$
J	x	x	x	$\cdots$	x	1	1	1	$\cdots$	1	J+1

Un-weighted GEEs

$$
\begin{aligned}
\mathbf{U}(\beta) &= \sum_{i=1}^{m} \mathbf{D}_{\mathbf{i,obs}}^{\mathbf{T}} \mathbf{V}_{\mathbf{i,obs}}^{-1} (\boldsymbol{Y}_{i,obs} - \boldsymbol{\mu}_{i,obs}) \\
&= \sum_{i=1}^{m} \sum_{j=2}^{J+1} I(R_i = j) \mathbf{D}_i^T(j) \mathbf{V}_i(j)^{-1} [\mathbf{Y}_i(j) - \boldsymbol{\mu}_i(j)]
\end{aligned}
$$

where

- $\mathbf{Y}_i(j)$, $\boldsymbol{\mu}_i(j)$ denote the first $j-1$ elements of $\mathbf{Y}_i$ and $\boldsymbol{\mu}_i$
- $\mathbf{D}_i(j)$ and $\mathbf{V}_i(j)$ are defined in a similar way

Note: Only one term of $\sum_{j=2}^{J+1} I(R_i = j) \mathbf{D}_i^T(j) \mathbf{V}_i(j)^{-1} [\mathbf{Y}_i(j) - \boldsymbol{\mu}_i(j)]$ is non-zero.

Un-weighted GEEs: Example

subject	Y_1	Y_2	Y_3	R
1	X	X	X	4
2	X	X		3
3	X			2

Then the estimating equation is

$$
\begin{aligned}
&\mathbf{D}_1^T(4)_{p\times 3}\mathbf{V}_1(4)^{-1}_{3\times 3}\left[\begin{pmatrix} Y_{11} \\ Y_{12} \\ Y_{13} \end{pmatrix} - \begin{pmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \end{pmatrix}\right] && \rightarrow subject1 \\
&+\mathbf{D}_2^T(3)_{p\times 2}\mathbf{V}_2(3)^{-1}_{2\times 2}\left[\begin{pmatrix} Y_{21} \\ Y_{22} \end{pmatrix} - \begin{pmatrix} \mu_{21} \\ \mu_{22} \end{pmatrix}\right] && \rightarrow subject2 \\
&+\mathbf{D}_3^T(2)_{p\times 1}\mathbf{V}_3(2)^{-1}_{1\times 1}[Y_{31} - \mu_{31}] && \rightarrow subject3
\end{aligned}
$$

Un-weighted GEEs Continued

For unweighted GEEs, we have

$$
\begin{aligned}
E\left[\mathbf{U}(\beta)\right] &= E_{\mathbf{Y}_i}\left\{E\left[\mathbf{U}(\beta)|\mathbf{Y}_i\right]\right\} \\
&= E_{\mathbf{Y}_i}\left\{\sum_{i=1}^{m}\sum_{j=2}^{J+1}\pi_{ij}\{\mathbf{Y}_i(j)\}\mathbf{D}_i^T(j)\mathbf{V}_i(j)^{-1}[\mathbf{Y}_i(j)-\boldsymbol{\mu}_i(j)]\right\} \\
&= \sum_{i=1}^{m}\sum_{j=2}^{J+1}\mathbf{D}_i^T(j)\mathbf{V}_i(j)^{-1}\left\{E_{\mathbf{Y}_i}[\pi_{ij}\{\mathbf{Y}_i(j)\}\mathbf{Y}_i(j)] - E_{\mathbf{Y}_i}[\pi_{ij}\{\mathbf{Y}_i(j\right. \\
&= \sum_{i=1}^{m}\sum_{j=2}^{J+1}\mathbf{D}_i^T(j)\mathbf{V}_i(j)^{-1} cov\left(\pi_{ij}\{\mathbf{Y}_i(j)\},\mathbf{Y}_i(j)\right)
\end{aligned}
$$

where $\pi_{ij}\{\mathbf{Y}_i(j)\} = Pr(R_i = j|\mathbf{Y}_i)$

- Under MCAR, π_{ij} is free of $\mathbf{Y}_i$. So (1) $=0$.
- Under MAR or NMAR, $\pi_{ij}\{\mathbf{Y}_i(j)\} = f(\mathbf{Y}_i(j))$. So (1) $\neq 0$, i.e, regular GEEs are biased.

Weighted GEEs

- First proposed by Robins, Rotnitzky and Zhao (1994, 1995)
- When the $\mathbf{Y}_{i,obs}$ are observed with probabilities π_{iR_i}, we can remove bias in estimating β by weighting the estimating equations with $w_{iR_i} = 1/\pi_{iR_i}$.
- Weighted GEE:

$$\sum_{i=1}^{m} \frac{1}{\pi_{iR_i}} \mathbf{D}^{\mathbf{T}}_{\mathbf{i,obs}} \mathbf{V}^{-1}_{\mathbf{i,obs}} (\boldsymbol{Y}_{\mathbf{i,obs}} - \boldsymbol{\mu}_{\mathbf{i,obs}}) = \mathbf{0}$$

where $\mathbf{D}_{i,obs} = \frac{\partial \boldsymbol{\mu}_{i,obs}}{\partial \beta^T}$ is $n_i \times p$, $(\mathbf{Y}_{\mathbf{i,obs}} - \boldsymbol{\mu}_{\mathbf{i,obs}})$ is $n_i \times 1$, $\mathbf{V}_{i,obs}$ is an $n_i \times n_i$ working covariance matrix, and $R_i = n_i + 1$.

Weighted GEEs (2)

Example:

Suppose for subject 1, we observe

x x . .

then we weights $\mathbf{D}_{1,obs}^{T}\mathbf{V}_{1,obs}^{-1}(\mathbf{Y}_{1,obs} - \boldsymbol{\mu}_{1,obs})$ by $\frac{1}{\pi_{13}}$.

Here π_{13} is the probability of dropout at time 3 for subject 1.

If $\pi_{13} = 0.1$, then subject 1 will represent 10 subjects who are similar to himself.

Weighted GEEs (3)

Example of weighted GEE:

subject	Y_1	Y_2	Y_3	R
1	X	X	X	4
2	X	X		3
3	X			2

Then the weighted estimating equation is

$$\frac{1}{\pi_{14}} \mathbf{D}_{1,obs}^T \mathbf{V}_{1,obs}^{-1} \left[\begin{pmatrix} Y_{11} \\ Y_{12} \\ Y_{13} \end{pmatrix} - \begin{pmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{13} \end{pmatrix} \right] \quad \rightarrow subject1$$

$$+ \frac{1}{\pi_{13}} \mathbf{D}_{2,obs}^T \mathbf{V}_{2,obs}^{-1} \left[\begin{pmatrix} Y_{21} \\ Y_{22} \end{pmatrix} - \begin{pmatrix} \mu_{21} \\ \mu_{22} \end{pmatrix} \right] \quad \rightarrow subject2$$

$$+ \frac{1}{\pi_{12}} \mathbf{D}_{3,obs}^T \mathbf{V}_{3,obs}^{-1} [Y_{31} - \mu_{31}] \quad \rightarrow subject3$$

Weighted GEEs (4)

Recall the Weighted GEEs is

$$\sum_{i=1}^{m} \frac{1}{\pi_{iR_i}} \mathbf{D}_{\mathbf{i,obs}}^{\mathbf{T}} \mathbf{V}_{\mathbf{i,obs}}^{-1} (\boldsymbol{Y}_{\mathbf{i,obs}} - \boldsymbol{\mu}_{\mathbf{i,obs}}) = \mathbf{0}$$

Denote the left side of the above equation as $\mathbf{U}_w(\beta)$.

$\mathbf{U}_w(\beta)$ weights the standard GEEs using the inverse probability of dropping out at the observed dropout time.

Weighted GEE provides a consistent estimator under MAR since $E[\mathbf{U}_w(\beta)] = 0$. why?

Weighted GEEs (5)

Note

$$\begin{aligned}
\mathbf{U}_w(\beta) &= \sum_{i=1}^{m} \frac{1}{\pi_{iR_i}} \mathbf{D}_{\mathbf{i,obs}}^{\mathbf{T}} \mathbf{V}_{\mathbf{i,obs}}^{-1} (\boldsymbol{Y}_{\mathbf{i,obs}} - \boldsymbol{\mu}_{\mathbf{i,obs}}) \\
&= \sum_{i=1}^{m} \sum_{j=2}^{J+1} \frac{I[R_i = j]}{\pi_{ij}} \mathbf{D}_i^T(j) \mathbf{V}_i(j)^{-1} [\mathbf{Y}_i(j) - \boldsymbol{\mu}_i(j)]
\end{aligned}$$

Therefore,

$$\begin{aligned}
E\left[\mathbf{U}_w(\beta)\right] &= E_{\mathbf{Y}_i, \mathbf{X}_i} \left\{ E\left[\mathbf{U}_w(\beta) | \mathbf{Y}_i, \mathbf{X}_i\right] \right\} \\
&= \sum_{i=1}^{m} E_{\mathbf{Y}_i, \mathbf{X}_i} \left\{ \sum_{j=2}^{J+1} \frac{E\left[I[R_i = j] | \mathbf{Y}_i, \mathbf{X}_i\right]}{\pi_{ij}} \mathbf{D}_i^T(j) \mathbf{V}_i(j)^{-1} [\mathbf{Y}_i(j) - \boldsymbol{\mu}_i(j)] \right\} \\
&= \sum_{i=1}^{m} E_{\mathbf{Y}_i, \mathbf{X}_i} \left\{ \sum_{j=2}^{J+1} \mathbf{D}_i^T(j) \mathbf{V}_i(j)^{-1} \left\{ \mathbf{Y}_i(j) - \boldsymbol{\mu}_i(j) \right\} \right\} \\
&= 0 \qquad (2)
\end{aligned}$$

Weighted GEEs (6)

- Weighted GEEs requires known π_{ij} or consistently estimated π_{ij} for each subject given their observed measurement history and any relevant covariates.
- Regular logistic regression for R:

$$logit(Pr(R_i = j)) = \alpha_{j0} + \alpha_1 Y_{i,j-1} + \boldsymbol{\alpha}_2^T \mathbf{X}_i$$

where $j = 2, \cdots J - 1$.

$\Rightarrow$

If $R_i \leq J$, then

$$\begin{aligned} \pi_{i,R_i} &= Pr(i\text{th subject drops out at } R_i) \\ &= \frac{exp^{\alpha_{R_i,0} + \alpha_1 Y_{i,R_i-1} + \boldsymbol{\alpha}_2^T \mathbf{X}_i}}{1 + exp^{\alpha_{R_i,0} + \alpha_1 Y_{i,R_i-1} + \boldsymbol{\alpha}_2^T \mathbf{X}_i}} \end{aligned}$$

If $R_i = J + 1$ (complete case), then

$$\begin{aligned} \pi_{i,R_i} &= Pr(i\text{th subject completes the study}) \\ &= 1 - \sum_{j=2}^{J} \pi_{ij} \end{aligned}$$

Weighted GEEs (7)

- Discrete survival model for R:

$$logit\{Pr(R_i = j | R_i \geq j)\} = \alpha_{j0} + \alpha_1 Y_{i,j-1} + \boldsymbol{\alpha}_2^T \mathbf{X}_i$$

$\Rightarrow$

If $R_i < J + 1$, then

$$\begin{aligned} \pi_{iR_i} &= Pr(i\text{th subject drop outs at } R_i) \\ &= \frac{exp^{\alpha_{R_i,0} + \alpha_1 Y_{i,R_i-1} + \boldsymbol{\alpha}_2^T \mathbf{X}_i}}{\prod_{j=2}^{R_i} \{1 + exp^{\alpha_{j0} + \alpha_1 Y_{i,j-1} + \boldsymbol{\alpha}_2^T \mathbf{X}_i}\}} \end{aligned}$$

If $R_i = J + 1$ (complete case), then

$$\begin{aligned} \pi_{iR_i} &= Pr(i\text{th subject completes the study}) \\ &= \prod_{j=2}^{J} (1 - P_{ij}) = \frac{1}{\prod_{j=2}^{J} \{1 + exp^{\alpha_{j0} + \alpha_1 Y_{i,j-1} + \boldsymbol{\alpha}_2^T \mathbf{X}_i}\}} \end{aligned}$$

Simulations

Fitzmaurice et. al, 1995.

TABLE 1

Percentage relative bias of the linear effect β_L, for the marginal logistic model with parameters $(\beta_0, \beta_G, \beta_L, \beta_Q) = (-0.125, 0.25, 0.2, -0.1)$ and serial odds ratio association parameter $\alpha = 2.5$

True model†			*Results for the working model*								
Mechanism	*% drop-out*	$(\gamma_0, \gamma_P, \gamma_C)$	*GEE (unweighted), MCAR*			*GEE (weighted), MAR*			*ML, MAR*		*ML, NI, exchangeable*
			Independence	*Exchangeable*	*Serial*	*Independence*	*Exchangeable*	*Serial*	*Exchangeable*	*Serial*	
Random drop-out	33	(−1.7, 0.5, 0)	24.10	1.53	6.78	0.00	0.00	0.00	−2.23	0.00	–
	50	(−0.9, 0.5, 0)	39.58	7.23	15.38	0.00	0.00	0.00	−8.91	0.00	–
	67	(0.3, 0.5, 0)	61.60	13.81	29.59	0.00	0.00	0.00	−22.07	0.00	–
Informative drop-out	33	(−1.7, 0.0, 0.5)	42.66	28.29	31.81	29.59	29.58	29.76	31.60	29.62	11.72
	50	(−0.9, 0.0, 0.5)	70.23	47.83	53.90	48.68	48.60	48.93	47.91	48.80	9.08
	67	(0.3, 0.0, 0.5)	110.38	74.92	86.83	77.74	77.58	78.11	70.06	76.09	5.52
Informative drop-out	33	(−1.0, 0.5, −0.5)	−18.26	−30.16	−27.59	−29.16	−29.13	−29.34	−28.37	−29.13	1.02
	50	(−0.2, 0.5, −0.5)	−29.91	−48.72	−44.00	−47.67	−47.55	−47.93	−51.02	−47.67	−9.67
	67	(1.0, 0.5, −0.5)	−47.38	−77.84	−67.65	−76.11	−75.38	−76.47	−86.64	−75.02	−17.09
Informative drop-out	33	(−2.5, 0.5, 0.5)	63.99	35.86	42.23	29.40	29.40	29.59	24.46	29.02	1.18
	50	(−1.7, 0.5, 0.5)	103.69	67.43	75.95	48.20	48.14	48.48	34.23	47.46	−1.28
	67	(−0.5, 0.5, 0.5)	160.35	112.59	127.97	77.09	76.92	77.46	43.42	74.79	−10.80

†The true drop-out model is $\text{logit}\{\text{pr}(R_{it} = 0 \mid R_{i1} = \ldots = R_{i,t-1} = 1, y_{i1}, \ldots, y_{it}, \gamma)\} = \gamma_0 + \gamma_P y_{i,t-1} + \gamma_C y_{it}$.

Weighted GEEs: Limitations

- Weighted GEEs provide an unbiased estimator under MCAR and MAR.
- π_{iR_i}, the observed probability, need to be bounded away from 0. If π_{iR_i} is too small, say 0.0001, then this observation will represent 10000 subjects. For finite samples, this may overweight and lead to biased estimator.
- Weighted GEEs can be inefficient since information from incompletely observed data is ignored.

Weighted GEEs: Loss of Efficiency

- Weighted GEEs throw away the subset of subject-occasions that are difficult to use because of missing responses and/or covariates, and re-weight the rest to make them more representative.

- For example, Robins, Rotnitzky and Zhao (1995) discard observations with missing responses:
  - Suppose for subject 1, we observe

    x x . .

    then in the weighted GEE, we only use $\begin{pmatrix} Y_{11} \\ Y_{12} \end{pmatrix}$ and $\begin{pmatrix} \mathbf{X}_{11}^T \\ \mathbf{X}_{12}^T \end{pmatrix}$, although we also observed $\begin{pmatrix} \mathbf{X}_{13}^T \\ \mathbf{X}_{14}^T \end{pmatrix}$.

Schizophrenia Study

- 312 patients received drug therapy for schizophrenia; 101 patients received a placebo

- Variables:
  - subject ID number
  - Outcome = IMPS79: overall severity of illness (continuous)
  - WEEK: 0,1,3,6
  - DRUG: 0=placebo 1=drug (chlorpromazine, fluphenazine, or thioridazine)
  - SEX: 0=female 1=male

Schizophrenia Study: Results from GEE Analysis

Unweighted GEE estimators:

Parameter	Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept	5.3617	0.0904	5.1846	5.5388	59.34	<.0001
SWEEK	-0.3762	0.0856	-0.5439	-0.2085	-4.40	<.0001
DRUG	0.0483	0.1050	-0.1575	0.2542	0.46	0.6455
SWEEK*DRUG	-0.6595	0.0980	-0.8516	-0.4674	-6.73	<.0001

Weighted GEE estimators:

Parameter	Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept	5.3865	0.1965	5.0014	5.7717	27.41	<.0001
SWEEK	-0.2807	0.1055	-0.4875	-0.0738	-2.66	0.0078
DRUG	-0.1408	0.2162	-0.5645	0.2829	-0.65	0.5149
SWEEK*DRUG	-0.7700	0.1261	-1.0172	-0.5228	-6.11	<.0001

	drug=0	drug=1
unweighted GEE	$E(imps79\|\mathbf{X}) = 5.36 - 0.38\sqrt{t}$	$E(imps79\|\mathbf{X}) = 5.40 - 1.03\sqrt{t}$
weighted GEE	$E(imps79\|\mathbf{X}) = 5.39 - 0.28\sqrt{t}$	$E(imps79\|\mathbf{X}) = 5.25 - 1.05\sqrt{t}$

Augmented Weighted GEE

- Weighted GEE may lose efficiency since we throw away the subset of subject-occasions that are difficult to use because of missing responses and/or covariates.

- Augmented weighted GEE (Scharfstein, Rotnitzky & Robins, 1999; Davidian, Tsiatis & Leon, 2005) were proposed to improve the efficiency of the estimators by incorporating additional information to predict the missing responses
  - Double robustness
  - More efficiency

Augmented Weighted GEE (2)

- Augmented weighted GEE:

$$\sum_{i=1}^{m} \left[\frac{I(R_i = J+1)}{\pi_{i,J+1}} \mathbf{D}_i \mathbf{V}_i^{-1} (\mathbf{Y}_i - \boldsymbol{\mu}_i) + \mathbf{A}_i \right] = 0$$

where

$$\mathbf{A}_i = \sum_{j=2}^{J} \left\{ I(R_i = j) - \frac{I(R_i = J+1)}{\pi_{i,J+1}} \pi_{i,j} \right\} \delta_j(\boldsymbol{Y}_{i,obs}, \mathbf{X}_i)$$

$\delta_j(\mathbf{X}_i)$ is an arbitrary $p \times 1$ vector function of the observed data when $R_i = j$.

- Augmented weighted GEE estimator is still consistent since the augmented term $\mathbf{A}_i$ has mean 0.

$$\begin{aligned} E(\mathbf{A}_i) &= E[E(\mathbf{A}_i | \boldsymbol{Y}_{i,obs}, \mathbf{X}_i)] \\ &= E[\sum_{j=2}^{J} \left\{ \pi_{ij} - \frac{\pi_{i,J+1}}{\pi_{i,J+1}} \pi_{i,j} \right\} \delta_j(\boldsymbol{Y}_{i,obs}, \mathbf{X}_i)] \\ &= 0 \end{aligned}$$

Augmented Weighted GEE (3)

- In augmented weighted GEE, complete cases contribute

$$\frac{1}{\pi_{i,J+1}} d(\mathbf{X}_i)(\mathbf{Y}_i - \boldsymbol{\mu}_i) - \sum_{j=2}^{J} \frac{\pi_{i,j}}{\pi_{i,J+1}} \delta_j(\boldsymbol{Y}_{i,obs}, \mathbf{X}_i)$$

  while incomplete cases contribute

$$\begin{aligned} \mathbf{A}_i &= \sum_{j=2}^{J} \{I(R_i = j)\} \, \delta_j(\boldsymbol{Y}_{i,obs}, \mathbf{X}_i) \\ &= \delta_{R_i}(\boldsymbol{Y}_{i,obs}, \mathbf{X}_i) \end{aligned}$$

- If $\delta_{R_i}(\boldsymbol{Y}_{i,obs}, \mathbf{X}_i) = \mathbf{D}_{\mathbf{i,obs}}^{\mathbf{T}} \mathbf{V}_{\mathbf{i,obs}}^{-\mathbf{1}}(\boldsymbol{Y}_{\mathbf{i,obs}} - \boldsymbol{\mu}_{\mathbf{i,obs}})/\pi_{ij}$,then we recover the weighted GEEs.
- All available data are exploited in augmented weighted GEEs.
- Augmented weighted GEEs have potentials to increase efficiency in $\hat{\beta}$, given the wisely chosen $d(\mathbf{X}_i)$ and $\delta_j(\boldsymbol{Y}_{i,obs}, \mathbf{X}_i)$.

Summary

- GEEs yield a consistent estimator of β under MCAR but yield a biased estimator of β under MAR unless the true correlation is used.

- Weighted GEEs yield an unbiased estimator of βunder MAR given the drop-out model is correctly specified. It allows the working correlation to be misspecified.

- Under NMAR, both unweighted and weighted GEEs usually yield a biased βestimator even when the correlation structure is correctly specified.

- Weighted GEE estimator is often not the most efficient estimator. Augmented GEEs have potentials to increase the efficiency.