Joint longitudinal models for dealing with missing at random data in trial-based economic evaluations

Andrea Gabrio[1], **Rachael Hunter**[2], **Alexina J. Mason**[3], **Gianluca Baio**[4]

[1] Department of Methodology and Statistics, Faculty of Health Medicine and Life Sciences, UM, NL
[2] Department of Primary Care and Population Health, UCL, UK
[3] Department of Health Services Research and Policy, LSHTM, UK
[4] Department of Statistical Science, UCL, UK

email: a.gabrio@maastrichtuniversity.nl
GitHub page: https://github.com/AnGabrio

Session: Methodological aspects in health economic evaluation

Chair: Mathyn Vervaart

HEH1.Aud3 - EuHEA Conference, Friday 08 Jul 2022

Outline

Part 1

Introduction

Back to Table of content

Useful questions to ask about missingness

- **How much** missingness ?
  - If only few variables and small rates (e.g.$< 5\%$) unlikely to affect results

Useful questions to ask about missingness

- **How much** missingness ?
  - If only few variables and small rates (e.g.$< 5\%$) unlikely to affect results
- **Which** variables and patterns?
  - Outcomes vs predictors, dropout vs intermittent (different implications on inferences)

Useful questions to ask about missingness

- **How much** missingness ?
  - If only few variables and small rates (e.g.$< 5\%$) unlikely to affect results
- **Which** variables and patterns?
  - Outcomes vs predictors, dropout vs intermittent (different implications on inferences)
- **Why** missingness occurred?

Useful questions to ask about missingness

- **How much** missingness ?
  - If only few variables and small rates (e.g.$< 5\%$) unlikely to affect results
- **Which** variables and patterns?
  - Outcomes vs predictors, dropout vs intermittent (different implications on inferences)
- **Why** missingness occurred?
  - Random chance, individual characteristics observed/unobserved

Useful questions to ask about missingness

- **How much** missingness ?
  - If only few variables and small rates (e.g.$< 5\%$) unlikely to affect results
- **Which** variables and patterns?
  - Outcomes vs predictors, dropout vs intermittent (different implications on inferences)
- **Why** missingness occurred?
  - Random chance, individual characteristics observed/unobserved
- Different assumptions about the **mechanism** underlying missingness may have a strong impact on the validity of the analysis method
- **Rubin's taxonomy** (Rubin, 1986) groups the mechanisms into three classes
  - **Missing Completely At Random** - missingness does not depend on observed/unobserved data
  - **Missing At Random** - missingness does not depend on unobserved data given the observed data
  - **Missing Not At Random** - missingness depends on unobserved data given the observed data

- Standard analyses at the aggregated level (e.g. QALYs and total costs) require pre-processing the (longitudinal) data collected from the study (**ID**)

Intended data set (ID)

		Utilities				Costs			
i	t	u_0	u_1	$\cdots$	u_J	c_0	c_1	$\cdots$	c_J
1	1	0.32	0.66	$\cdots$	0.44	103	141	$\cdots$	180
2	2	0.33	0.54	$\cdots$	0.61	101	434	$\cdots$	511
3	1	0.12	0.16	$\cdots$	0.38	204	808	$\cdots$	877
4	2	0.41	0.47	$\cdots$	0.72	35	50	$\cdots$	90
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	2	0.49	0.55	$\cdots$	0.88	16	12	$\cdots$	22

How missingness is addressed in CEA ?

- In reality, utility/cost measurements for some individuals are missing at least at some occasions (**CD**)

Intended data set (ID)

		Utilities				Costs			
i	t	u_0	u_1	$\cdots$	u_J	c_0	c_1	$\cdots$	c_J
1	1	0.32	0.66	$\cdots$	0.44	103	141	$\cdots$	180
2	2	0.33	0.54	$\cdots$	0.61	101	434	$\cdots$	511
3	1	0.12	0.16	$\cdots$	0.38	204	808	$\cdots$	877
4	2	0.41	0.47	$\cdots$	0.72	35	50	$\cdots$	90
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	2	0.49	0.55	$\cdots$	0.88	16	12	$\cdots$	22

Collect the data $\Big\downarrow$ Missingness occurs

Collected data set (CD)

		Utilities				Costs			
i	t	u_0	u_1	$\cdots$	u_J	c_0	c_1	$\cdots$	c_J
1	1	0.32	0.66	$\cdots$	0.44	103	141	$\cdots$	180
2	2	0.33	0.54	$\cdots$	0.61	101	434	$\cdots$	511
3	1	0.12	–	$\cdots$	–	204	–	$\cdots$	–
4	2	–	0.47	$\cdots$	0.72	35	–	$\cdots$	90
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	2	0.49	0.55	$\cdots$	–	–	12	$\cdots$	–

How missingness is addressed in CEA ?

- Computation of aggregated quantities causes removal of follow-up data for the partially-observed cases (**AD**)

Intended data set (ID)

		Utilities				Costs			
i	t	u_0	u_1	$\cdots$	u_J	c_0	c_1	$\cdots$	c_J
1	1	0.32	0.66	$\cdots$	0.44	103	141	$\cdots$	180
2	2	0.33	0.54	$\cdots$	0.61	101	434	$\cdots$	511
3	1	0.12	0.16	$\cdots$	0.38	204	808	$\cdots$	877
4	2	0.41	0.47	$\cdots$	0.72	35	50	$\cdots$	90
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	2	0.49	0.55	$\cdots$	0.88	16	12	$\cdots$	22

Collect the data

Missingness occurs

Collected data set (CD)

		Utilities				Costs			
i	t	u_0	u_1	$\cdots$	u_J	c_0	c_1	$\cdots$	c_J
1	1	0.32	0.66	$\cdots$	0.44	103	141	$\cdots$	180
2	2	0.33	0.54	$\cdots$	0.61	101	434	$\cdots$	511
3	1	0.12	–	$\cdots$	–	204	–	$\cdots$	–
4	2	–	0.47	$\cdots$	0.72	35	–	$\cdots$	90
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	2	0.49	0.55	$\cdots$	–	–	12	$\cdots$	–

Calculate QALYs/Total costs

Discard follow-up data

Aggregated data set (AD)

		Utilities		Costs	
i	t	u_0	e	c_0	c
1	1	0.32	0.47	103	424
2	2	0.33	0.49	101	1046
3	1	0.12	–	204	–
4	2	–	–	35	–
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	2	0.49	–	–	–

How missingness is addressed in CEA ?

- When focus is on complete cases, baseline data for partially-observe cases are also discarded (**ACD**)

Intended data set (ID)

		Utilities				Costs			
i	t	u_0	u_1	$\cdots$	u_J	c_0	c_1	$\cdots$	c_J
1	1	0.32	0.66	$\cdots$	0.44	103	141	$\cdots$	180
2	2	0.33	0.54	$\cdots$	0.61	101	434	$\cdots$	511
3	1	0.12	0.16	$\cdots$	0.38	204	808	$\cdots$	877
4	2	0.41	0.47	$\cdots$	0.72	35	50	$\cdots$	90
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	2	0.49	0.55	$\cdots$	0.88	16	12	$\cdots$	22

Collect the data

Missingness occurs

Collected data set (CD)

		Utilities				Costs			
i	t	u_0	u_1	$\cdots$	u_J	c_0	c_1	$\cdots$	c_J
1	1	0.32	0.66	$\cdots$	0.44	103	141	$\cdots$	180
2	2	0.33	0.54	$\cdots$	0.61	101	434	$\cdots$	511
3	1	0.12	–	$\cdots$	–	204	–	$\cdots$	–
4	2	–	0.47	$\cdots$	0.72	35	–	$\cdots$	90
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	2	0.49	0.55	$\cdots$	–	–	12	$\cdots$	–

Calculate QALYs/Total costs

Discard follow-up data

Aggregated data set (AD)

		Utilities		Costs	
i	t	u_0	e	c_0	c
1	1	0.32	0.47	103	424
2	2	0.33	0.49	101	1046
3	1	0.12	–	204	–
4	2	–	–	35	–
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	2	0.49	–	–	–

Identify the compelte cases

Remove the missing data

Aggregated complete data set (ACD)

		Utilities		Costs	
i	t	u_0	e	c_0	c
1	1	0.32	0.47	103	424
2	2	0.33	0.49	101	1046
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$n^{(cc)}$	1	0.29	0.30	116	1156

Some considerations ...

- Traditional trial-based CEAs are performed at the level of aggregated QALYs/Total costs despite the fact that collected data are **longitudinal** in nature
- This is not a problem with fully complete data as (e_i, c_i) are derived from (u_{ij}, c_{ij}) and therefore provide the same information in a "cross-sectional" setting
- This, however, is no longer true when missingness occurs since some longitudinal data are discarded and cannot be used in the analysis fitted at aggregated level
- Intuitively, addressing missing data at longitudinal level is more efficient since all observed data can be used to impute values/fit models

Some considerations ...

- Traditional trial-based CEAs are performed at the level of aggregated QALYs/Total costs despite the fact that collected data are **longitudinal** in nature
- This is not a problem with fully complete data as (e_i, c_i) are derived from (u_{ij}, c_{ij}) and therefore provide the same information in a "cross-sectional" setting
- This, however, is no longer true when missingness occurs since some longitudinal data are discarded and cannot be used in the analysis fitted at aggregated level
- Intuitively, addressing missing data at longitudinal level is more efficient since all observed data can be used to impute values/fit models

Question: How much does this matter?

Some considerations ...

- Traditional trial-based CEAs are performed at the level of aggregated QALYs/Total costs despite the fact that collected data are **longitudinal** in nature
- This is not a problem with fully complete data as (e_i, c_i) are derived from (u_{ij}, c_{ij}) and therefore provide the same information in a "cross-sectional" setting
- This, however, is no longer true when missingness occurs since some longitudinal data are discarded and cannot be used in the analysis fitted at aggregated level
- Intuitively, addressing missing data at longitudinal level is more efficient since all observed data can be used to impute values/fit models

Question: How much does this matter?

- **Aim:** Assess impact of addressing missing at different levels using alternative approaches in terms of bias/efficiency across a range of scenarios
- Focus on **MAR** assumption (standard assumption in CEAs) with only baseline outcome values as predictors (simplified setting)

Part 2

Methods

Back to Table of content

Review of popular approaches to handle missingness under standard linear regression analysis framework in trial-based CEA practice

1. Case deletion methods
2. Baseline imputation methods
3. Joint aggregated models
4. Joint longitudinal models

Case deletion methods

- Model fitted to the aggregated QALYs/Total costs (e_i, c_i) after processing utility/cost (u_{ij}, c_{ij}) data collected in the study (ACD or AD)

- Regression analysis is used to adjust for baseline variables, e.g. baseline utility/cost values (u_{i0}, c_{i0})

$$
\begin{aligned}
e_i | t_i, u_{i0} &\sim \text{Normal}(\alpha_0 + \alpha_1 t_i + \alpha_2 u_{i0}, \sigma^2_e) \\
c_i | t_i, c_{i0} &\sim \text{Normal}(\beta_0 + \beta_1 t_i + \beta_2 c_{i0}, \sigma^2_c)
\end{aligned}
$$

- Two alternatives:
  - **Complete Case Analysis** (CCA): use only complete cases (n_{cca}) to both fit the model and compute the baseline means $\bar{u}_0$ and $\bar{c}_0$
  - **Available Case Analysis** (ACA): use complete cases (n_{cca}) to fit the model but compute the baseline means $\bar{u}_0$ and $\bar{c}_0$ using all available observed cases $(n_{aca} \geq n_{cca})$

Missing data **analysis methods** in CEA

Review of popular approaches to handle missingness under standard linear regression analysis framework in trial-based CEA practice

1. Case deletion methods
2. Baseline imputation methods
3. Joint aggregated models
4. Joint longitudinal models

Baseline imputation methods

- Model fitted to the aggregated QALYs/Total costs (e_i, c_i) after processing utility/cost (u_{ij}, c_{ij}) data collected in the study (AD)
- Regression analysis is used to adjust for imputed baseline variables $(u^{\star}_{i0}, c^{\star}_{i0})$ using some value, e.g. **mean-imputed value** (MEAN)

$$
\begin{aligned}
e_i|t_i, u^{\star}_{i0} &\sim \text{Normal}(\alpha_0 + \alpha_1 t_i + \alpha_2 u^{\star}_{i0}, \sigma^2_e) \\
c_i|t_i, c^{\star}_{i0} &\sim \text{Normal}(\beta_0 + \beta_1 t_i + \beta_2 c^{\star}_{i0}, \sigma^2_c)
\end{aligned}
$$

- Missing outcome values (e_i, c_i) are imputed through the linear model either:
  - replacing missing data with point predictions
  - replacing missing data with point predictions and an error term

Review of popular approaches to handle missingness under standard linear regression analysis framework in trial-based CEA practice

1. Case deletion methods
2. Baseline imputation methods
3. Joint aggregated models
4. Joint longitudinal models

Joint aggregated models

- Model fitted to the aggregated QALYs/Total costs (e_i, c_i) after processing utility/cost (u_{ij}, c_{ij}) data collected in the study (AD)

- A joint distribution is simultaneously fitted to aggregated and baseline variables, which can also be conditionally specified as

$$
\begin{aligned}
u_{i0} \sim \text{Normal}(\mu_{u0}, \sigma^2_{u0}), \quad e_i | t_i, u_{i0} &\sim \text{Normal}(\alpha_0 + \alpha_1 t_i + \alpha_2 u_{i0}, \sigma^2_e) \\
c_{i0} \sim \text{Normal}(\mu_{c0}, \sigma^2_{c0}), \quad c_i | t_i, c_{i0} &\sim \text{Normal}(\beta_0 + \beta_1 t_i + \beta_2 c_{i0}, \sigma^2_c)
\end{aligned}
$$

- Two alternative methods can be used to directly impute missing values within this bivariate modelling framework
  - **Multiple Imputation** (MI): separate imputation and analysis steps
  - **Full Bayesian** (FB): jointly perform imputation and analysis steps

Review of popular approaches to handle missingness under standard linear regression analysis framework in trial-based CEA practice

1. Case deletion methods
2. Baseline imputation methods
3. Joint aggregated models
4. Joint longitudinal models

Joint longitudinal models

- Model fitted to the utility/cost (u_{ij}, c_{ij}) data at each time point collected in the study (CD)

- A joint distribution is simultaneously fitted to all variables, which can also be conditionally specified (under a lag-1 dependence assumption) as

$$
\begin{aligned}
u_{i0} \sim \text{Normal}(\mu_{u0}, \sigma^2_{u0}), \quad u_{ij} | t_i, u_{ij-1} &\sim \text{Normal}(\alpha_{0t} + \alpha_{1t} t_i + \alpha_{2t} u_{ij-1}, \sigma^2_e) \\
c_{i0} \sim \text{Normal}(\mu_{c0}, \sigma^2_{c0}), \quad c_{ij} | t_i, c_{ij-1} &\sim \text{Normal}(\beta_{0t} + \beta_{1t} t_i + \beta_{2t} c_{ij-1}, \sigma^2_c)
\end{aligned}
$$

- The joint multivariate distributions of (u_{i0}, u_{ij}) and (c_{i0}, c_{ij}) can be
  - approximated via MI (L-MI): first multiply-impute all variables and then fit the model and combine estimates across the imputed datasets
  - directly fitted via FB (L-FB): fit model to partially-observed data using weakly informative priors

Part 3

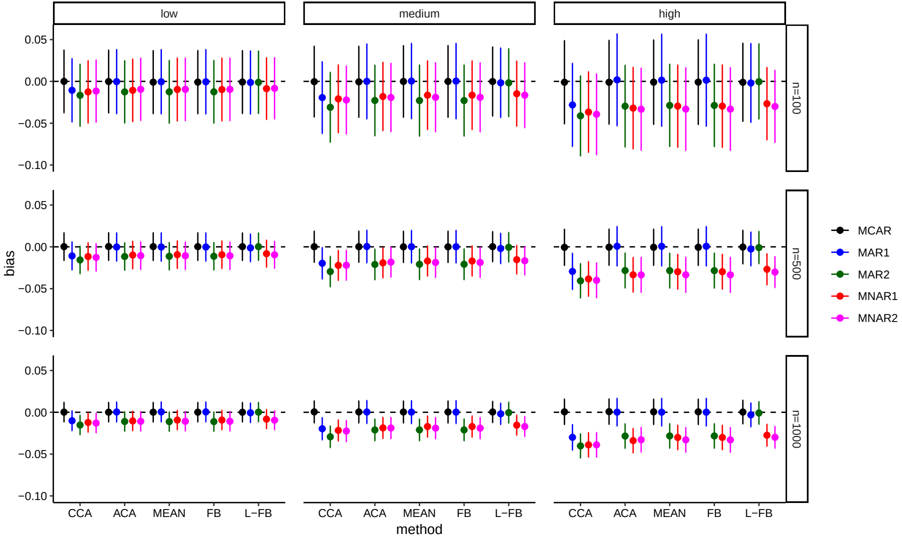
Simulation Study

Back to Table of content

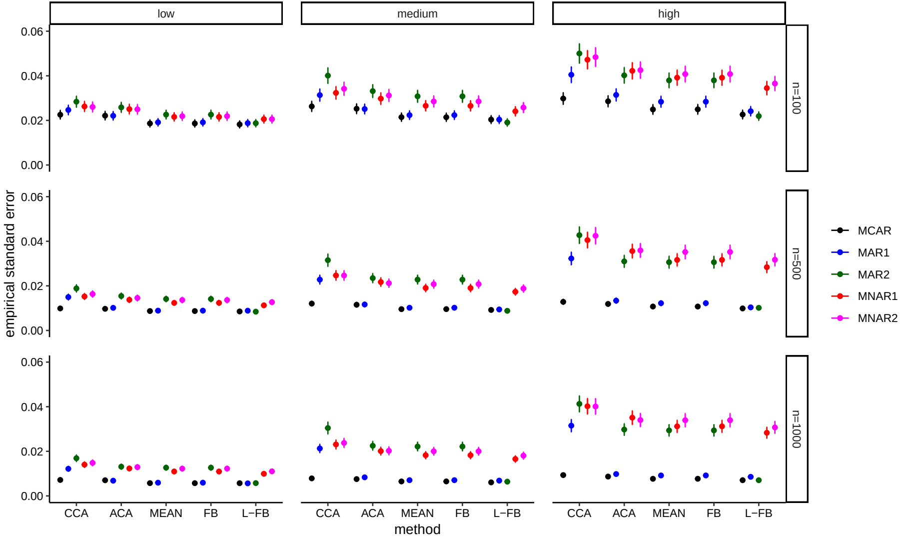
Design

- **Setting**: normally-distributed outcome collected at 3 time points in a RCT
- **Objective**: assess performance of methods under different assumptions about data generating and missing data processes
- **Scenarios**: total of 45 different scenarios generated by varying
  - sample size: 100, 500, 1000
  - missing data proportion: low $= 0.15$, medium $= 0.3$, high $= 0.5$
  - dropout across arms introduced by modelling dropout probabilities as function of outcomes at previous times via logistic regression
- **Missingness scenarios**:
  - Dropout is totally random (**MCAR**)
  - Dropout random at $j = 0$ but more likely at $j = 1, 2$ for individuals with higher utilities at $j = 0$ (**MAR1**) or $j = 1$ (**MAR2**)
  - Dropout more likely at $j = 1, 2$ for individuals with higher utilities at same time with dropout at $j = 0$ being random (**MNAR1**) or more likely for individuals with higher utilities at same time (**MNAR2**)

Results - **Bias**



Results - **Empirical SE**



Part 4

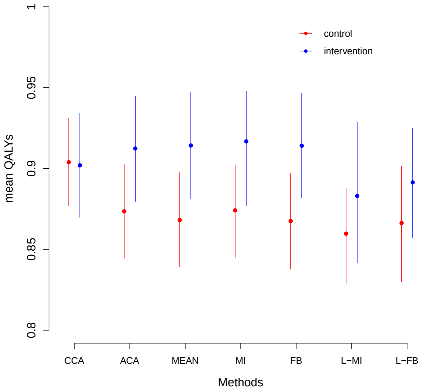
Application: the MenSS trial

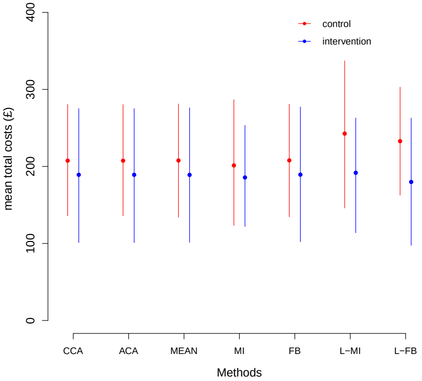
Back to Table of content

- The MenSS pilot RCT evaluates the cost-effectiveness of a new digital intervention to reduce the incidence of STI in young men with respect to the SOC
  - QALYs calculated from utilities (EQ-5D 3L)
  - Total costs calculated from different components (no baseline)

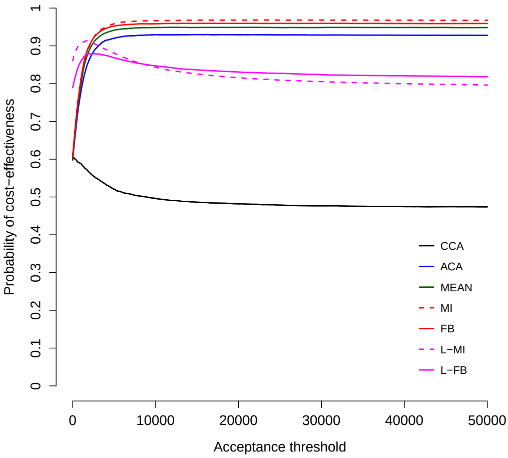
Time	**Type of outcome**	**observed (%)**	**observed (%)**
		Control (n_1=75)	Intervention (n_2=84)
Baseline	utilities	72 (96%)	72 (86%)
3 months	utilities and costs	34 (45%)	23 (27%)
6 months	utilities and costs	35 (47%)	23 (27%)
12 months	utilities and costs	43 (57%)	36 (43%)
Complete cases	utilities and costs	27 (44%)	19 (23%)

- Partially-observed cases are mostly associated with lower utilities and higher costs in the control arm while no clear pattern emerges in the intervention arm





- With the exception of CCA, all methods show similar estimates for mean QALYs/Costs with small differences observed for the joint longitudinal models (L-MI,L-FB)



- With the exception of CCA, all methods show a relatively high chance of cost-effectiveness, with milder conclusions drawn by L-MI and L-FB

Part 5

Discussion

Back to Table of content

Conclusions

- The main objective was to assess impact on trial-based CEA results of alternative missingness approaches with focus on:
  - Differences between aggregated and longitudinal methods
  - Use and extension of standard modelling frameworks used by practitioners
- Results based on the scenarios explored in the simulation study and analyses of case studies indicate that:
  - Depending on the specification of the missingness mechanism, simple methods (CCA,ACA,MEAN) and even joint aggregate models (MI,FB) may lead to biased results under MAR
  - Joint longitudinal models are the most robust approach to MAR assumptions as they incorporate all information from partially-observed cases
  - The magnitude of the differences between methods changes depending on sample size and missingness proportion
  - Potential benefit of L-MI/L-FB is likely to increase when the number complete cases is small (MenSS)

Recommendations and limitations

- Looking at observed distributions of utilities/costs by time and arm could provide some insights on potential benefit of using longitudinal vs aggregate models
- Presence of substantial observed differences at any time between complete cases and partially-observed cases suggests longitudinal models provide more robust inferences under MAR
- Both MI and FB are valid approaches to implement longitudinal models (MLE also possible)
- Exploration of the impact of additional data complexities which are typical of CEA data have not been explored and could be considered in the future:
  - Correlation between utilities/costs
  - Skewness and presence of "structural values" in both outcomes' distributions
- **Conclusions**: models that take into account the longitudinal nature of utility/cost data provide reliable estimates under a wider range of MAR assumptions compared to standard models fitted to aggregate QALYs/Total costs.