Bayesian methods for addressing missing data in health economic evaluation

Andrea Gabrio

(Thanks to Gianluca Baio and Alexina J. Mason)

http://www.ucl.ac.uk/statistics/research/statistics-health-economics/

ucakgab@ucl.ac.uk

DEPARTMENT OF STATISTICAL SCIENCE
UNIVERSITY COLLEGE LONDON

March 11, 2018

Outline

Outline

Outline

Outline

Outline

Outline

Outline

1. Health Economic Evaluation
   - What is it?
2. Missing Data
   - Missing Data Mechanism
3. Bayesian modelling for missing data
   - Modelling & advantages
   - Selection Models
4. Systematic Literature Review
   - Review of current approaches
5. Motivating example
   - Data & Models
   - Results
6. Model Extension
   - Results
7. Conclusions

Health Economic Evaluation

Objective: Combine costs & benefits of a given intervention into a rational scheme for allocating resources, increasingly often under a Bayesian framework

Health Economic Evaluation

Objective: Combine costs & benefits of a given intervention into a rational scheme for allocating resources, increasingly often under a Bayesian framework

Statistical model

- Estimates relevant **population** parameters $\boldsymbol{\theta}$
- Varies with the type of available data (& statistical approach!)

Health Economic Evaluation

Objective: Combine costs & benefits of a given intervention into a rational scheme for allocating resources, increasingly often under a Bayesian framework

$$\begin{aligned}\Delta_e &= f_e(\boldsymbol{\theta})\\ \Delta_c &= f_c(\boldsymbol{\theta})\\ &\dots\end{aligned}$$

Statistical model $\longrightarrow$ Economic model

- Estimates relevant **population** parameters $\boldsymbol{\theta}$
- Varies with the type of available data (& statistical approach!)

- Combines the parameters to obtain a population average measure for costs and clinical benefits
- Varies with the type of available data & statistical model used

Health Economic Evaluation

Objective: Combine costs & benefits of a given intervention into a rational scheme for allocating resources, increasingly often under a Bayesian framework

$$\Delta_e = f_e(\boldsymbol{\theta})$$
$$\Delta_c = f_c(\boldsymbol{\theta})$$
$$\ldots$$

$$\text{ICER} = \text{E}[\Delta_c]/\text{E}[\Delta_e]$$
$$\text{EIB} = k\text{E}[\Delta_e] - \text{E}[\Delta_c]$$
$$\text{CEAC} = \Pr(k\Delta_e - \Delta_c > 0)$$

Statistical model → **Economic model** → **Decision analysis**

Statistical model

- Estimates relevant **population** parameters $\boldsymbol{\theta}$
- Varies with the type of available data (& statistical approach!)

Economic model

- Combines the parameters to obtain a population average measure for costs and clinical benefits
- Varies with the type of available data & statistical model used

Decision analysis

- Summarises the economic model by computing suitable measures of "cost-effectiveness"
- Dictates the best course of actions, given current evidence
- Standardised process

Health Economic Evaluation

Objective: Combine costs & benefits of a given intervention into a rational scheme for allocating resources, increasingly often under a Bayesian framework

Uncertainty analysis

- Assesses the impact of uncertainty (eg in parameters or model structure) on the economic results
- Mandatory in many jurisdictions (including NICE)
- Fundamentally Bayesian

$$\Delta_e = f_e(\boldsymbol\theta)$$
$$\Delta_c = f_c(\boldsymbol\theta)$$
$$\ldots$$

ICER $= \mathrm{E}[\Delta_c]/\mathrm{E}[\Delta_e]$

EIB $= k\mathrm{E}[\Delta_e] - \mathrm{E}[\Delta_c]$

CEAC $= \Pr(k\Delta_e - \Delta_c > 0)$

Statistical model → **Economic model** → **Decision analysis**

Statistical model

- Estimates relevant **population** parameters $\boldsymbol\theta$
- Varies with the type of available data (& statistical approach!)

Economic model

- Combines the parameters to obtain a population average measure for costs and clinical benefits
- Varies with the type of available data & statistical model used

Decision analysis

- Summarises the economic model by computing suitable measures of "cost-effectiveness"
- Dictates the best course of actions, given current evidence
- Standardised process

- The available data usually look something like this:

		Demographics			HRQL data				Resource use data			
ID	**Trt**	Sex	Age	...	u_0	u_1	...	u_J	c_0	c_1	...	c_J
1	1	M	23	...	0.32	0.66	...	0.44	103	241	...	80
2	1	M	21	...	0.12	0.16	...	0.38	1 204	1 808	...	877
3	2	F	19	...	0.49	0.55	...	0.88	16	12	...	22
...	...	...	...	...	...	...	...	...	...	...	...	...

- The available data usually look something like this:

		Demographics			HRQL data				Resource use data			
ID	**Trt**	Sex	Age	...	u_0	u_1	...	u_J	c_0	c_1	...	c_J
1	1	M	23	...	0.32	0.66	...	0.44	103	241	...	80
2	1	M	21	...	0.12	0.16	...	0.38	1 204	1 808	...	877
3	2	F	19	...	0.49	0.55	...	0.88	16	12	...	22
...	...	...	...	...	...	...	...	...	...	...	...	...

and the **typical** analysis is based on the following steps:

1. Compute individual QALYs and total costs as

$$e_i = \sum_{j=1}^{J} (u_{ij} + u_{ij-1}) \frac{\delta_j}{2} \quad \text{and} \quad c_i = \sum_{j=0}^{J} c_{ij}, \qquad \left[\text{with: } \delta_j = \frac{\text{Time}_j - \text{Time}_{j-1}}{\text{Unit of time}}\right]$$

- The available data usually look something like this:

		Demographics			HRQL data				Resource use data			
ID	**Trt**	Sex	Age	...	u_0	u_1	...	u_J	c_0	c_1	...	c_J
1	1	M	23	...	0.32	0.66	...	0.44	103	241	...	80
2	1	M	21	...	0.12	0.16	...	0.38	1 204	1 808	...	877
3	2	F	19	...	0.49	0.55	...	0.88	16	12	...	22
...	...	...	...	...	...	...	...	...	...	...	...	...

and the **typical** analysis is based on the following steps:

1. Compute individual QALYs and total costs as

$$e_i = \sum_{j=1}^{J} (u_{ij} + u_{ij-1}) \frac{\delta_j}{2} \quad \text{and} \quad c_i = \sum_{j=0}^{J} c_{ij}, \qquad \left[\text{with: } \delta_j = \frac{\text{Time}_j - \text{Time}_{j-1}}{\text{Unit of time}}\right]$$

2. (Often implicitly) assume normality and linearity and model **independently** individual QALYs and total costs by controlling for baseline values

$$\begin{aligned} e_i &= \alpha_{e0} + \alpha_{e1} u_{0i} + \alpha_{e2}\text{Trt}_i + \varepsilon_{ei}\,[+\ldots], \qquad &\varepsilon_{ei} \sim \text{Normal}(0, \sigma_e) \\ c_i &= \alpha_{c0} + \alpha_{c1} c_{0i} + \alpha_{c2}\text{Trt}_i + \varepsilon_{ci}\,[+\ldots], \qquad &\varepsilon_{ci} \sim \text{Normal}(0, \sigma_c) \end{aligned}$$

What's wrong with this?

- Potential correlation between costs & clinical benefits
    - Strong positive correlation — effective treatments are innovative and are associated with higher unit costs
    - Negative correlation — more effective treatments may reduce total care pathway costs e.g. by reducing hospitalisations, side effects, etc.

What's wrong with this?

- Potential correlation between costs & clinical benefits
  - Strong positive correlation — effective treatments are innovative and are associated with higher unit costs
  - Negative correlation — more effective treatments may reduce total care pathway costs e.g. by reducing hospitalisations, side effects, etc.
- Joint/marginal normality not realistic
  - Costs usually skewed and benefits may be bounded in $[0;1]$
  - Can use transformation (e.g. logs) — but care is needed when back transforming to the natural scale
  - Can use more suitable models (e.g. Gamma or log-Normal) — especially under the Bayesian approach

What's wrong with this?

- Potential correlation between costs & clinical benefits
  - Strong positive correlation — effective treatments are innovative and are associated with higher unit costs
  - Negative correlation — more effective treatments may reduce total care pathway costs e.g. by reducing hospitalisations, side effects, etc.
- Joint/marginal normality not realistic
  - Costs usually skewed and benefits may be bounded in $[0; 1]$
  - Can use transformation (e.g. logs) — but care is needed when back transforming to the natural scale
  - Can use more suitable models (e.g. Gamma or log-Normal) — especially under the Bayesian approach
- ... and of course **Partially Observed** data
  - Can have item and/or unit non-response
  - Missingness may occur in either or both benefits/costs
  - Focus in decision-making — not inference!

Missing Data

- Essential to investigate the possible reasons behind the missingness

Missing Data

- Essential to investigate the possible reasons behind the missingness

- Assumptions cannot be tested from the data but need to be formulated based on the available state of knowledge

Missing Data

- Essential to investigate the possible reasons behind the missingness
- Assumptions cannot be tested from the data but need to be formulated based on the available state of knowledge
- This formally translates into an assumed **missing data mechanism** (Rubin, 1987) that is linked to the data generating process
    - Missing Completely At Random (MCAR)
    - Missing At Random (MAR)
    - Missing Not At Random (MNAR)

Missing Data

- Essential to investigate the possible reasons behind the missingness
- Assumptions cannot be tested from the data but need to be formulated based on the available state of knowledge
- This formally translates into an assumed **missing data mechanism** (Rubin, 1987) that is linked to the data generating process
    - Missing Completely At Random (MCAR)
    - Missing At Random (MAR)
    - Missing Not At Random (MNAR)
- Well-defined statistical model for the complete data, and explicit assumptions about the missing value mechanism —"principled" approach to missingness

Missing Data Mechanism: MCAR

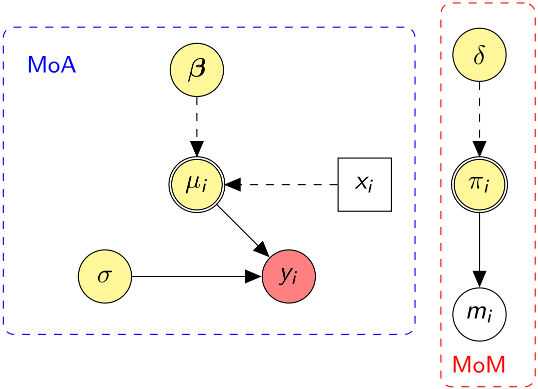


Figure: MoA=Model of Analysis, MoM=Model of Missingness

Missing Data Mechanism: MAR

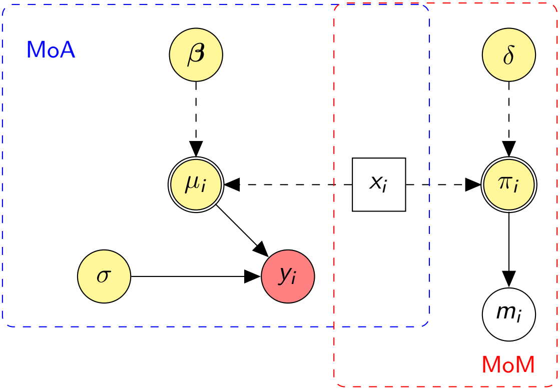


Figure: MoA=Model of Analysis, MoM=Model of Missingness

Missing Data Mechanism: MNAR

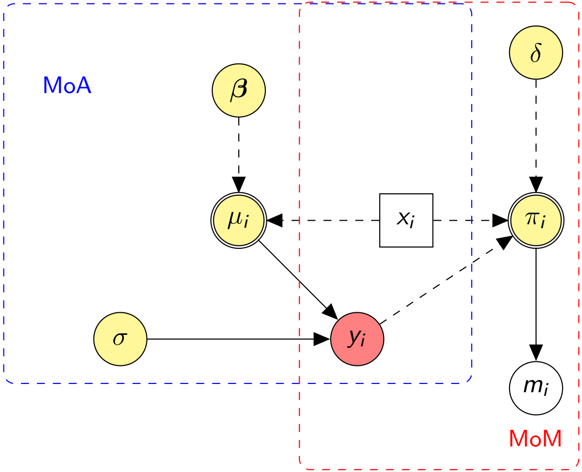


Figure: MoA=Model of Analysis, MoM=Model of Missingness

Missing Data Methods

- **Complete Case Analysis**
  - Elimination of partially observed cases
  - Simple but reduce efficiency and possibly bias parameter estimates
- **Single Imputation**
  - Imputation of missing data with a single value (mean, median, LVCF)
  - Does not account for the uncertainty in the imputation process
- **Multiple Imputation** (Rubin, 1987)
  - Missing data imputed H times to obtain H different imputed datasets
  - Each dataset is analysed and H sets of estimates are derived
  - Parameter estimates are combined into a single quantity
  - The uncertainty due to imputation is incorporated but the validity relies on the correct specification of the imputation model

Bayesian Modelling

- Parameters are given probability distributions that describe the uncertainty before (**prior**) and after (**posterior**) observing the data

$$p(\boldsymbol{\theta} \mid y) \propto p(y \mid \boldsymbol{\theta}) p(\boldsymbol{\theta})$$

- Incorporate both individual and parameter (missing data) uncertainty
- Naturally encode alternative missingness assumptions through the priors and assess the robustness of the results — **Sensitivity Analysis**
- Often not analytically tractable and iterative approximation methods, e.g. MCMC (Brooks et al., 2011), are required

A simple example

- For simplicity assume a Bivariate Normal for $\mathbf{MoA}(\boldsymbol{e}, \boldsymbol{c})$
- Can represent a joint distribution as a **conditional regression**

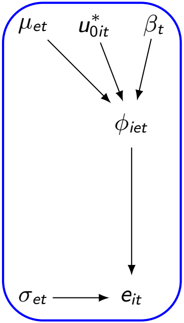
$$p(e, c) = p(e)p(c \mid e) = p(c)p(e \mid c)$$

A simple example

- For simplicity assume a Bivariate Normal for **MoA($\boldsymbol{e, c}$)**
- Can represent a joint distribution as a **conditional regression**

$$p(e, c) = p(e)p(c \mid e) = p(c)p(e \mid c)$$

Marginal model for e



$$e_{it} \sim N(\phi_{iet}, \sigma_{et})$$

$$\phi_{iet} = \mu_{et} + \beta_t u^*_{0it}$$

$$\phi_{iet} = \text{marginal mean}$$
$$\sigma_{et} = \text{marginal sd}$$

A simple example

- For simplicity assume a Bivariate Normal for $\mathbf{MoA}(\boldsymbol{e}, \boldsymbol{c})$
- Can represent a joint distribution as a **conditional regression**

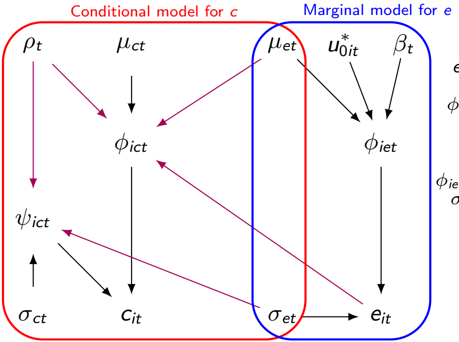
$$p(e, c) = p(e)p(c \mid e) = p(c)p(e \mid c)$$

$$c_{it} \mid e_{it} \sim N(\phi_{ict}, \psi_{ict})$$

$$\phi_{ict} = \mu_{ct} + \rho(e_{it} - \mu_{et})$$

$$\psi_{ict} = \sqrt{\sigma^2_{ct} - \sigma^2_{et}\rho^2}$$

ϕ_{ict} = conditional mean
ψ_{ict} = conditional sd



$$e_{it} \sim N(\phi_{iet}, \sigma_{et})$$

$$\phi_{iet} = \mu_{et} + \beta_t u^*_{0it}$$

ϕ_{iet} = marginal mean
σ_{et} = marginal sd

A simple example

- For simplicity assume a Bivariate Normal for **MoA(e, c)**
- Can represent a joint distribution as a **conditional regression**

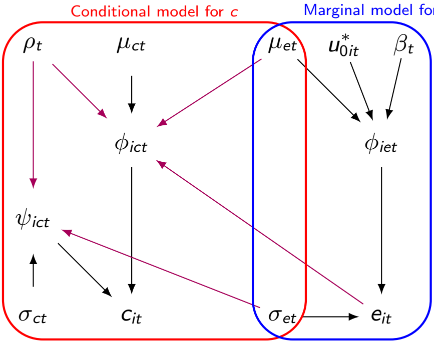
$$p(e, c) = p(e)p(c \mid e) = p(c)p(e \mid c)$$

$$c_{it} \mid e_{it} \sim N(\phi_{ict}, \psi_{ict})$$

$$\phi_{ict} = \mu_{ct} + \rho(e_{it} - \mu_{et})$$

$$\psi_{ict} = \sqrt{\sigma^2_{ct} - \sigma^2_{et}\rho^2}$$

ϕ_{ict} = conditional mean
ψ_{ict} = conditional sd



$$e_{it} \sim N(\phi_{iet}, \sigma_{et})$$

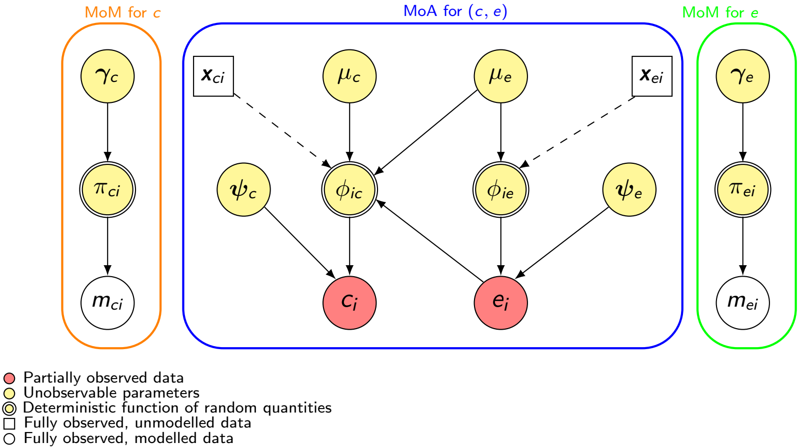
$$\phi_{iet} = \mu_{et} + \beta_t u^*_{0it}$$

ϕ_{iet} = marginal mean
σ_{et} = marginal sd

- Under MAR, no need to explicitly model the **MoM**

Nonignorable Missingness — Selection Models

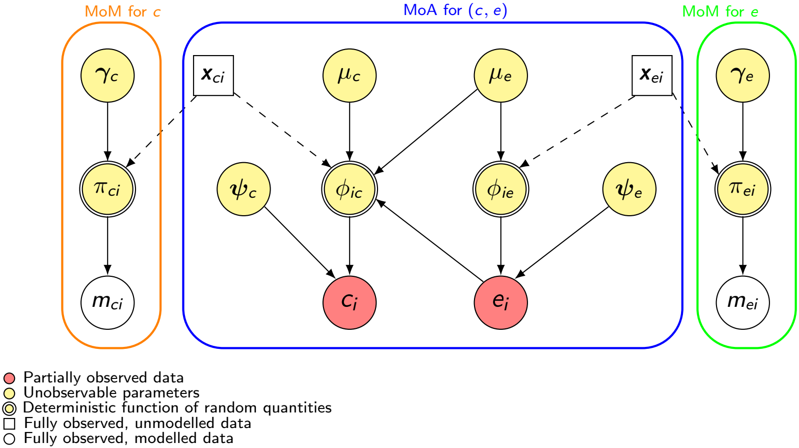
MCAR (e, c)



- $m_{ei} \sim \text{Bernoulli}(\pi_{ei});$ $\quad \text{logit}(\pi_{ei}) = \gamma_{e0}$
- $m_{ci} \sim \text{Bernoulli}(\pi_{ci});$ $\quad \text{logit}(\pi_{ci}) = \gamma_{c0}$

Nonignorable Missingness — Selection Models

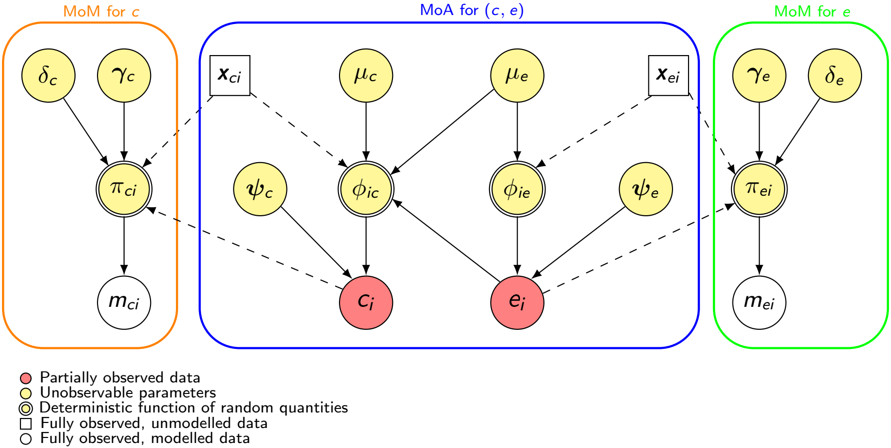
MAR (e, c)



- $m_{ei} \sim \text{Bernoulli}(\pi_{ei})$; $\quad \text{logit}(\pi_{ei}) = \gamma_{e0} + \sum_{k=1}^{K} \gamma_{ek} x_{eik}$
- $m_{ci} \sim \text{Bernoulli}(\pi_{ci})$; $\quad \text{logit}(\pi_{ci}) = \gamma_{c0} + \sum_{h=1}^{H} \gamma_{ch} x_{cih}$

Nonignorable Missingness — Selection Models

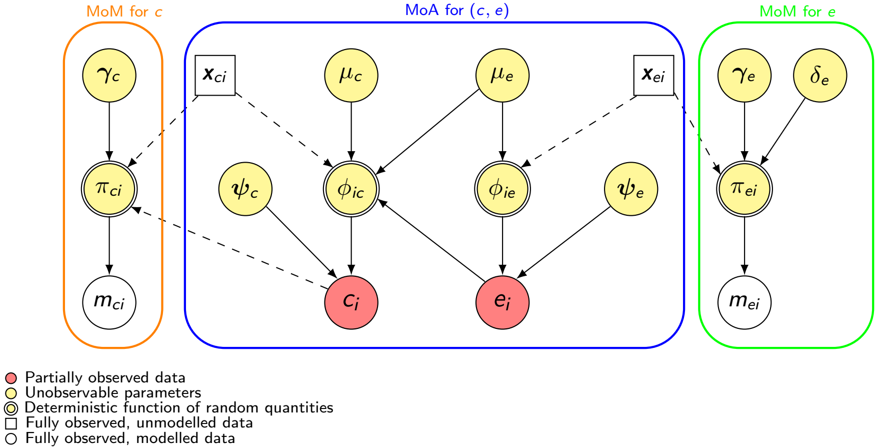
MNAR (e, c)



- $m_{ei} \sim \text{Bernoulli}(\pi_{ei})$; $\quad \text{logit}(\pi_{ei}) = \gamma_{e0} + \sum_{k=1}^{K} \gamma_{ek} x_{eik} + \delta_e e_i$
- $m_{ci} \sim \text{Bernoulli}(\pi_{ci})$; $\quad \text{logit}(\pi_{ci}) = \gamma_{c0} + \sum_{h=1}^{H} \gamma_{ch} x_{cih} + \delta_c c_i$

Nonignorable Missingness — Selection Models

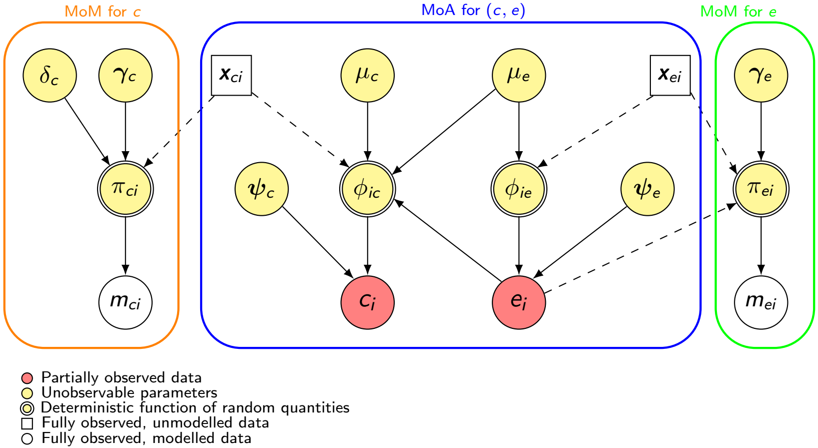
MNAR e; MAR c



- $m_{ei} \sim \text{Bernoulli}(\pi_{ei})$; $\quad \text{logit}(\pi_{ei}) = \gamma_{e0} + \sum_{k=1}^{K} \gamma_{ek} x_{eik} + \delta_e e_i$
- $m_{ci} \sim \text{Bernoulli}(\pi_{ci})$; $\quad \text{logit}(\pi_{ci}) = \gamma_{c0} + \sum_{h=1}^{H} \gamma_{ch} x_{cih}$

Nonignorable Missingness — Selection Models

MAR e; MNAR c



- $m_{ei} \sim \text{Bernoulli}(\pi_{ei})$; $\quad \text{logit}(\pi_{ei}) = \gamma_{e0} + \sum_{k=1}^{K} \gamma_{ek} x_{eik}$
- $m_{ci} \sim \text{Bernoulli}(\pi_{ci})$; $\quad \text{logit}(\pi_{ci}) = \gamma_{c0} + \sum_{h=1}^{H} \gamma_{ch} x_{cih} + \delta_c c_i$

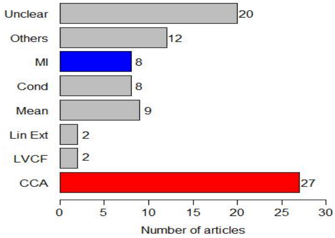
Systematic Literature Review

Two-fold purpose:

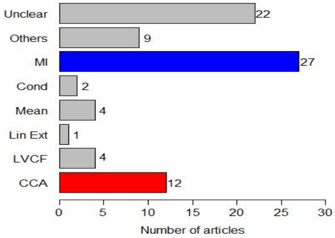
1. Provide some guidelines on the reporting and analysis of missingness
2. Review of the missing data methods in CEAs (2003-2015), updating the work of Noble et al. (2012)

- Original review focused only on missing costs in within-trial CEA studies
  - 88 articles between 2003-2009
- Include missing effects and update the review
  - 81 studies between 2009-2015

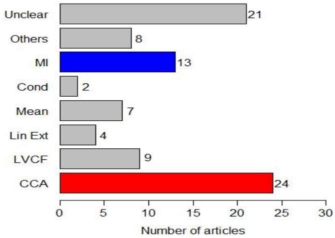
Current Methods



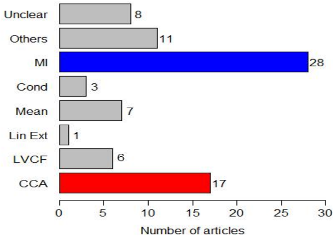
(a) Missing costs (2003-2009)



(b) Missing costs (2009-2015)



(c) Missing effects (2003-2009)



(d) Missing effects (2009-2015)

Conclusions

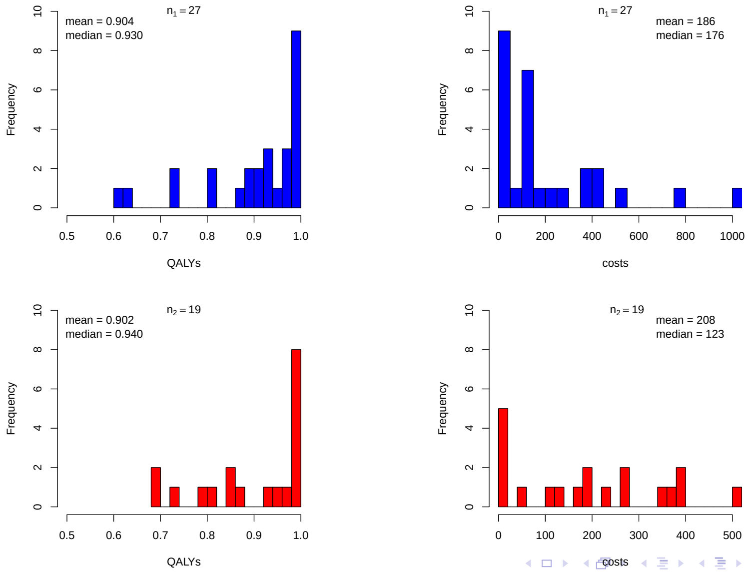
- High missing data proportions in within-trial CEAs may lead to imprecise economic evidences
- The review shows a movement towards more flexible methods in terms of missingness assumptions but:
  - Many studies do not provide transparent missing data information
  - Almost no study performs a sensitivity analysis
- Missing data handling can be improved by explicitly defining the assumptions and assess their impact on the conclusions.

The MenSS Trial

- The MenSS pilot RCT (Bailey et al., 2016) 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%)

Complete Cases



Modelling strategy

model	**MoA (e, c)**	**MoM (e)**	**MoM (c)**
Base-Case	Independent Normal	MAR	MAR
MAR(e, c)	Joint Normal	MAR	MAR
MNAR(e)	Joint Normal	MNAR	MAR
MNAR(c)	Joint Normal	MAR	MNAR

- MNAR(e): $e^{mis} \approx (5 - 10\%)$ lower than $e^{obs} \rightarrow \delta_e \sim \text{N}(-2, 1)$
- MNAR(c): $c^{mis} \approx (60 - 70\%)$ higher than $c^{obs} \rightarrow \delta_c \sim \text{N}(0, 1)$

Results

Parameter	Base-Case		MAR (e,c)			MNAR (e)			MNAR (c)		
	Mean	95% CI	Mean	95% CI		Mean	95% CI		Mean	95% CI	
Control $(t=1)$											
mean QALY (μ^e_1)	0.886		0.874	0.840	0.907	0.855	0.807	0.893	0.863	0.826	0.899
mean cost (μ^c_1)	214		207.770	115.363	302.901	207.912	113.226	301.081	290.324	126.971	452.932
sd QALY (σ^e_1)			0.081	0.061	0.110	0.081	0.064	0.103	0.081	0.064	0.103
sd cost (σ^c_1)			257.964	197.201	341.123	259.517	191.160	344.420	267.924	197.633	356.626
Intervention $(t=2)$											
mean QALY (μ^e_2)	0.918		0.913	0.868	0.956	0.847	0.715	0.929	0.912	0.860	0.967
mean cost (μ^c_2)	189		189.170	110.778	267.963	188.497	108.829	267.280	316.032	106.835	516.946
sd QALY (σ^e_2)			0.092	0.066	0.130	0.094	0.070	0.124	0.092	0.069	0.122
sd cost (σ^c_2)			174.082	124.350	252.623	176.378	121.666	249.735	190.872	128.897	275.189
Incremental											
mean QALY increment (Δ_e)	0.032	-0.02 0.08	0.039	-0.016	0.095	-0.008	-0.122	0.072	0.049	-0.011	0.114
mean cost increment (Δ_c)	-25	-145 97	-18.600	-141.081	102.463	-19.415	-140.283	104.196	25.708	-121.593	194.326

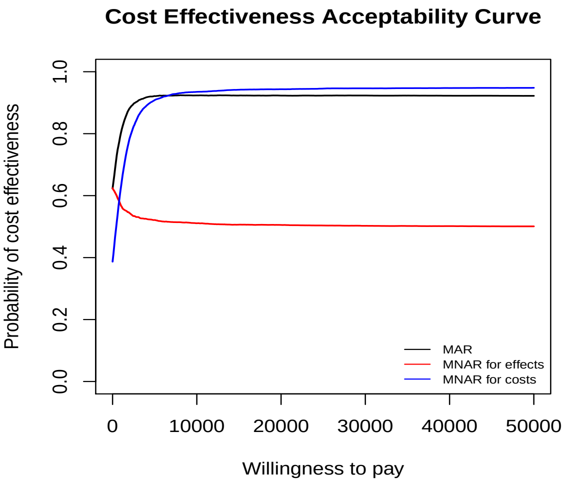
Results

	Base-Case		MAR (e,c)			MNAR (e)			MNAR (c)		
Parameter	**Mean**	**95% CI**	**Mean**	**95% CI**		**Mean**	**95% CI**		**Mean**	**95% CI**	
Control $(t=1)$											
mean QALY (μ^e_1)	0.886		0.874	0.840	0.907	0.855	0.807	0.893	0.863	0.826	0.899
mean cost (μ^c_1)	214		207.770	115.363	302.901	207.912	113.226	301.081	290.324	126.971	452.932
sd QALY (σ^e_1)			0.081	0.061	0.110	0.081	0.064	0.103	0.081	0.064	0.103
sd cost (σ^c_1)			257.964	197.201	341.123	259.517	191.160	344.420	267.924	197.633	356.626
Intervention $(t=2)$											
mean QALY (μ^e_2)	0.918		0.913	0.868	0.956	0.847	0.715	0.929	0.912	0.860	0.967
mean cost (μ^c_2)	189		189.170	110.778	267.963	188.497	108.829	267.280	316.032	106.835	516.946
sd QALY (σ^e_2)			0.092	0.066	0.130	0.094	0.070	0.124	0.092	0.069	0.122
sd cost (σ^c_2)			174.082	124.350	252.623	176.378	121.666	249.735	190.872	128.897	275.189
Incremental											
mean QALY increment (Δ_e)	0.032	-0.02 0.08	0.039	-0.016	0.095	-0.008	-0.122	0.072	0.049	-0.011	0.114
mean cost increment (Δ_c)	-25	-145 97	-18.600	-141.081	102.463	-19.415	-140.283	104.196	25.708	-121.593	194.326

Results

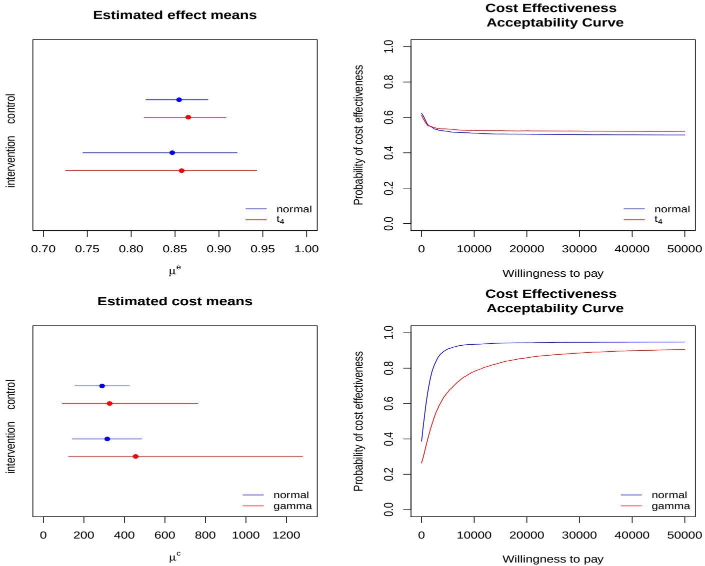
	Base-Case			MAR (e, c)			MNAR (e)			MNAR (c)		
Parameter	**Mean**	**95% CI**		**Mean**	**95% CI**		**Mean**	**95% CI**		**Mean**	**95% CI**	
Control $(t = 1)$												
mean QALY (μ_1^e)	0.886			0.874	0.840	0.907	0.855	0.807	0.893	0.863	0.826	0.899
mean cost (μ_1^c)	214			207.770	115.363	302.901	207.912	113.226	301.081	290.324	126.971	452.932
sd QALY (σ_1^e)				0.081	0.061	0.110	0.081	0.064	0.103	0.081	0.064	0.103
sd cost (σ_1^c)				257.964	197.201	341.123	259.517	191.160	344.420	267.924	197.633	356.626
Intervention $(t = 2)$												
mean QALY (μ_2^e)	0.918			0.913	0.868	0.956	0.847	0.715	0.929	0.912	0.860	0.967
mean cost (μ_2^c)	189			189.170	110.778	267.963	188.497	108.829	267.280	316.032	106.835	516.946
sd QALY (σ_2^e)				0.092	0.066	0.130	0.094	0.070	0.124	0.092	0.069	0.122
sd cost (σ_2^c)				174.082	124.350	252.623	176.378	121.666	249.735	190.872	128.897	275.189
Incremental												
mean QALY increment (Δ_e)	0.032	-0.02	0.08	0.039	-0.016	0.095	-0.008	-0.122	0.072	0.049	-0.011	0.114
mean cost increment (Δ_c)	-25	-145	97	-18.600	-141.081	102.463	-19.415	-140.283	104.196	25.708	-121.593	194.326

Economic Evaluation

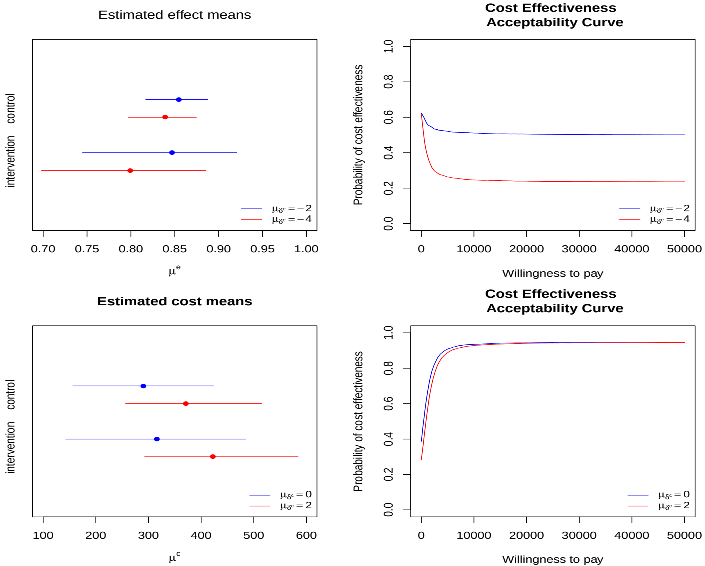


- Under $MNAR(e)$ the assessment radically changes, with the new interventions being not cost-effective compared with the control

Sensitivity Analysis (1)



Sensitivity Analysis (2)



Conclusions

- MAR is not likely to hold and the original study conclusions may overestimate the cost-effectiveness of the reference intervention
- The MNAR departures explored show how a relatively small variation in MoM(e) may substantially alter the decision output
- Lack of information about missingness may severely impair the analysis and force unrealistic assumptions
- Selection Models are a possible choice to handle nonignorable missingness but SA is necessary to assess the robustness of the results to alternative assumptions

Model Extension

- The models assume a Bivariate Normal for (e, c)
  - Simpler and closer to "standard" frequentist model
  - Account for baseline adjustment and correlation between outcomes

Model Extension

- The models assume a Bivariate Normal for (e, c)
  - Simpler and closer to “standard” frequentist model
  - Account for baseline adjustment and correlation between outcomes
- Capture skewness in the data with Bivariate Beta-Gamma for (e, c)
  - Model the relevant ranges: QALYs $\in (0, 1)$ and costs $\in (0, \infty)$
  - **But**: needs to rescale observed data to avoid 1 and 0 values

Model Extension

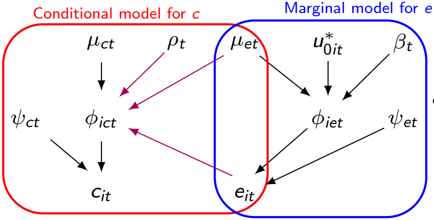
- The models assume a Bivariate Normal for (e, c)
  - Simpler and closer to "standard" frequentist model
  - Account for baseline adjustment and correlation between outcomes
- Capture skewness in the data with Bivariate Beta-Gamma for (e, c)
  - Model the relevant ranges: QALYs $\in (0, 1)$ and costs $\in (0, \infty)$
  - **But**: needs to rescale observed data to avoid 1 and 0 values
- Capture spike at 1 for QALYs with a Hurdle Model
  - Model e as a *mixture* of two components: $e^{<1}$ and e^{1}
  - e^{1} are considered as *structural values* (identically one)
  - May be extended to partially observed u_0

Beta-Gamma model

Conditional model for $c \mid e$

$$c_{it} \mid e_{it} \sim \text{Gamma}(\psi_{ct}\phi_{ict}, \psi_{ct})$$
$$\phi_{ict} = \mu_{ct} + \rho_t(e_{it} - \mu_{et})$$

$$\psi_{ct}\phi_{ict} = \text{shape}$$
$$\psi_{ct} = \text{rate}$$



Marginal model for e

$$e_{it} \sim \text{Beta}\left(\phi_{iet}\psi_{et}, (1-\phi_{iet})\psi_{et}\right)$$
$$\text{logit}(\phi_{iet}) = \mu_{et} + \beta_t u^*_{0it}$$

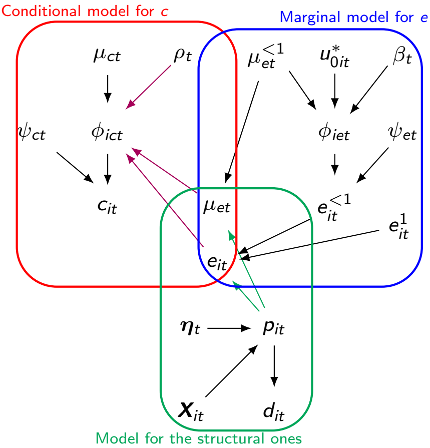
$$\phi_{iet}\psi_{et} = \text{shape}$$
$$(1-\phi_{iet})\psi_{et} = \text{shape}$$

Hurdle model for QALYs

Conditional model for $c \mid e$

$$c_{it} \mid e_{it} \sim \text{Gamma}(\psi_{ct}\phi_{ict}, \psi_{ct})$$
$$\phi_{ict} = \mu_{ct} + \rho_t(e_{it} - \mu_{et})$$

$$\psi_{ct}\phi_{ict} = \text{shape}$$
$$\psi_{ct} = \text{rate}$$



Marginal model for e

$$e_{it}^{1} : 1$$
$$e_{it}^{<1} \sim \text{Beta}\left(\phi_{iet}\psi_{et}, (1-\phi_{iet})\psi_{et}\right)$$
$$\text{logit}(\phi_{iet}) = \mu_{et} + \beta_t u_{0it}^{*}$$

$$e_{it} = p_{it}e_{it}^{1} + (1-p_{it})e_{it}^{<1}$$
$$\mu_{et} = p_t + (1-p_t)\mu_{et}^{<1}$$

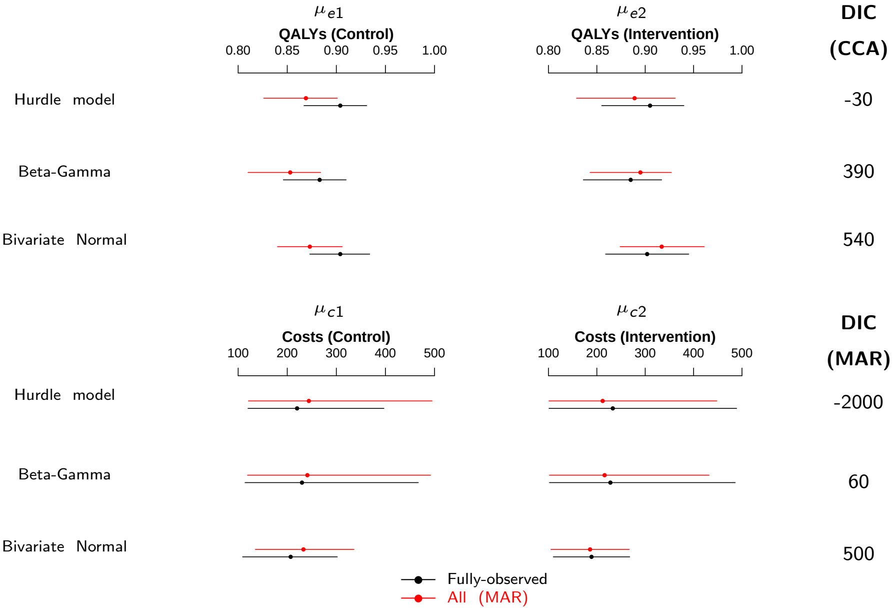
$$\phi_{iet}\psi_{et} = \text{shape}$$
$$(1-\phi_{iet})\psi_{et} = \text{shape}$$

Model for the structural ones

$$d_{it} : \mathbb{I}(e_{it}) = 1 \sim \text{Bernoulli}(p_{it})$$
$$\text{logit}(p_{it}) = \boldsymbol{X}_{it}\boldsymbol{\eta}_t$$

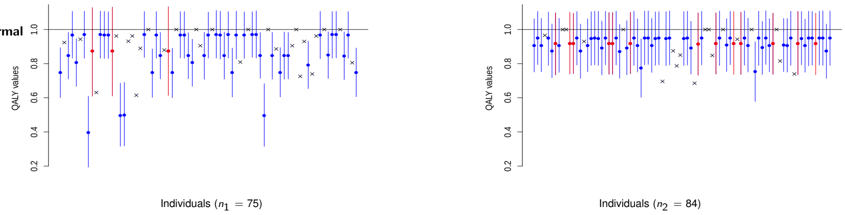
$$p_{it} = \text{probability of } e_{it}^{1}$$

Mean estimates (CCA + MAR)



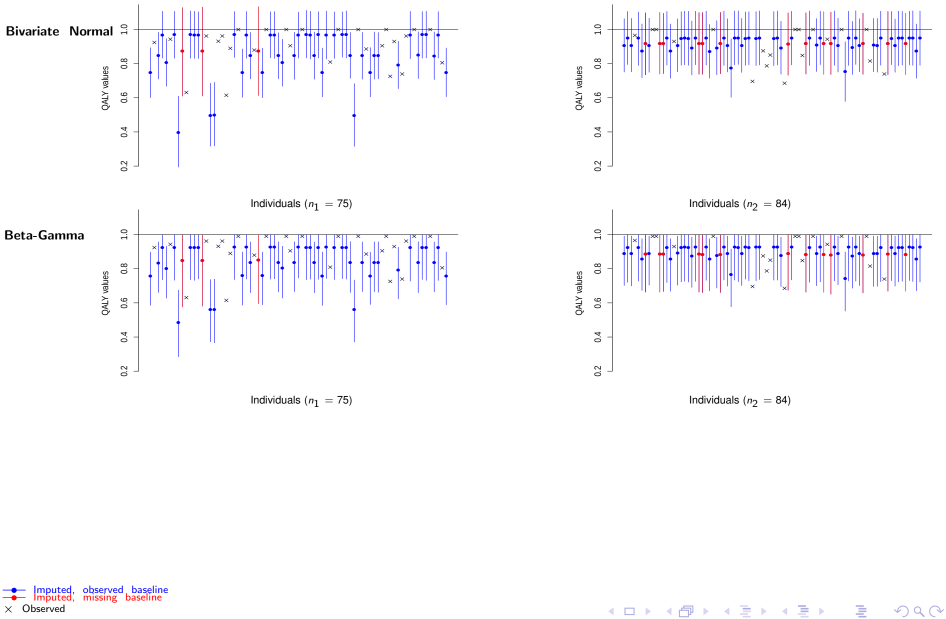
Imputations (under MAR)

Bivariate Normal

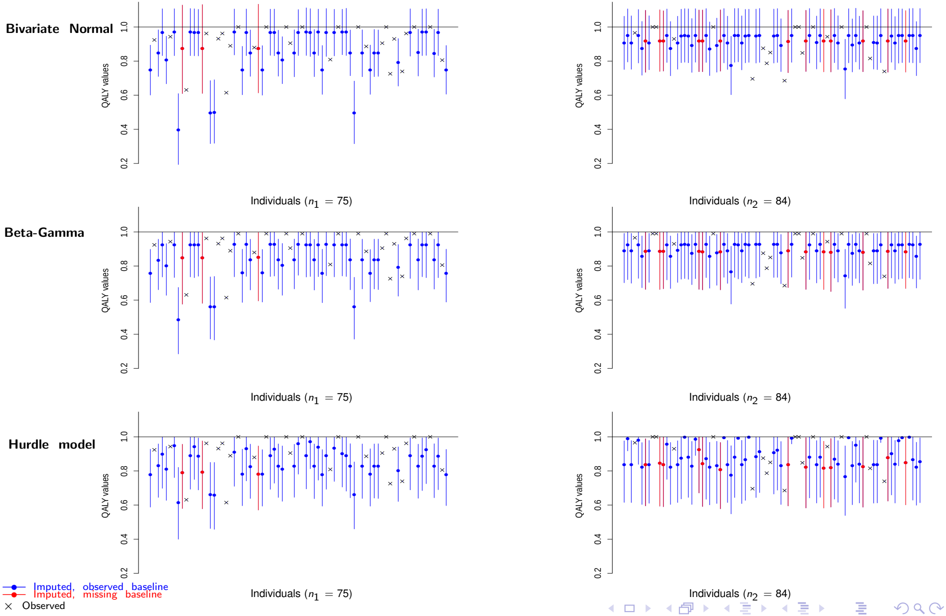


Imputed, observed baseline
Imputed, missing baseline
× Observed

Imputations (under MAR)



Imputations (under MAR)



MNAR

- We observe $n_{01} = 13$ and $n_{02} = 22$ individuals with $u_{0it} = 1$ and $u_{jit} = \texttt{NA}$, for $j = 1, 2, 3$
- For those individuals, we cannot compute directly the structural one indicator d_{it} and so need to make assumptions/model this
    - Sensitivity analysis to alternative MNAR departures from MAR

MNAR

- We observe $n_{01} = 13$ and $n_{02} = 22$ individuals with $u_{0it} = 1$ and $u_{jit} = \mathtt{NA}$, for $j = 1, 2, 3$
- For those individuals, we cannot compute directly the structural one indicator d_{it} and so need to make assumptions/model this
  - Sensitivity analysis to alternative MNAR departures from MAR

MNAR1. Set $d_{it} = 1$ for all individuals with unit observed baseline utility

MNAR

- We observe $n_{01} = 13$ and $n_{02} = 22$ individuals with $u_{0it} = 1$ and $u_{jit} = \texttt{NA}$, for $j = 1, 2, 3$
- For those individuals, we cannot compute directly the structural one indicator d_{it} and so need to make assumptions/model this
  - Sensitivity analysis to alternative MNAR departures from MAR

MNAR1. Set $d_{it} = 1$ for all individuals with unit observed baseline utility

MNAR2. Set $d_{it} = 0$ for all individuals with unit observed baseline utility

MNAR

- We observe $n_{01} = 13$ and $n_{02} = 22$ individuals with $u_{0it} = 1$ and $u_{jit} = \mathtt{NA}$, for $j = 1, 2, 3$
- For those individuals, we cannot compute directly the structural one indicator d_{it} and so need to make assumptions/model this
  - Sensitivity analysis to alternative MNAR departures from MAR

MNAR1. Set $d_{it} = 1$ for all individuals with unit observed baseline utility

MNAR2. Set $d_{it} = 0$ for all individuals with unit observed baseline utility

MNAR3. Set $d_{it} = 1$ for the $n_{01} = 13$ individuals with $u_{0i1} = 1$ and $d_{it} = 0$ for the $n_{02} = 22$ individuals with $u_{0i2} = 1$

MNAR

- We observe $n_{01} = 13$ and $n_{02} = 22$ individuals with $u_{0it} = 1$ and $u_{jit} = \texttt{NA}$, for $j = 1, 2, 3$
- For those individuals, we cannot compute directly the structural one indicator d_{it} and so need to make assumptions/model this
  - Sensitivity analysis to alternative MNAR departures from MAR

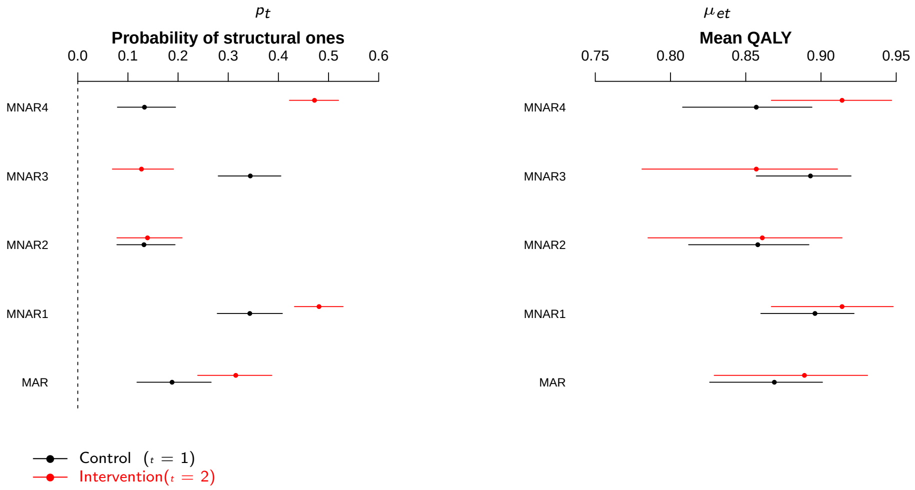
MNAR1. Set $d_{it} = 1$ for all individuals with unit observed baseline utility

MNAR2. Set $d_{it} = 0$ for all individuals with unit observed baseline utility

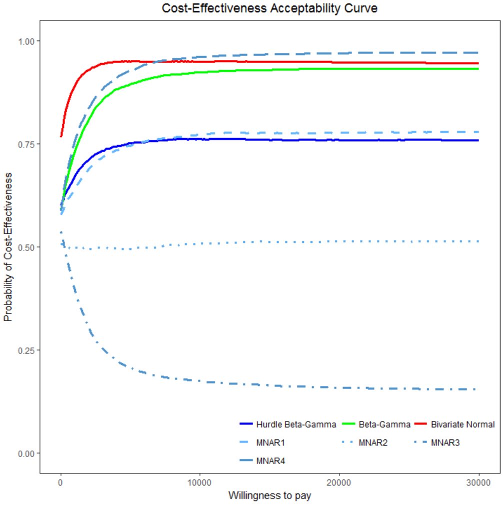
MNAR3. Set $d_{it} = 1$ for the $n_{01} = 13$ individuals with $u_{0i1} = 1$ and $d_{it} = 0$ for the $n_{02} = 22$ individuals with $u_{0i2} = 1$

MNAR4. Set $d_{it} = 0$ for the $n_{01} = 13$ individuals with $u_{0i1} = 1$ and $d_{it} = 1$ for the $n_{02} = 22$ individuals with $u_{0i2} = 1$

Results — MNAR



Cost-effectiveness analysis



Conclusions

- A full Bayesian approach to handling missing data extends standard "imputation methods"
  - Can consider MAR and MNAR with relatively little expansion to the basic model

Conclusions

- A full Bayesian approach to handling missing data extends standard "imputation methods"
    - Can consider MAR and MNAR with relatively little expansion to the basic model
- Particularly helpful in cost-effectiveness analysis, to account for
    - Asymmetrical distributions for the main outcomes
    - Correlation between costs & benefits
    - Structural values (eg spikes at 1 for utilities or spikes at 0 for costs)

Conclusions

- A full Bayesian approach to handling missing data extends standard "imputation methods"
  - Can consider MAR and MNAR with relatively little expansion to the basic model
- Particularly helpful in cost-effectiveness analysis, to account for
  - Asymmetrical distributions for the main outcomes
  - Correlation between costs & benefits
  - Structural values (eg spikes at 1 for utilities or spikes at 0 for costs)
- MNAR can never be ruled out
  - Necessary to explore plausible MNAR departures
  - Assess and quantify impact of uncertainty on inferences and (more importantly) on the decision process

References

Bailey, J., Webster, R., Hunter, R., Griffin, M., N., F., Rait, G., Estcourt, C., Michie, S., Anderson, J., Stephenson, J., Gerressu, M., Sinag Ang, C., and Murray, E. (2016). The men's safer sex project: intervention development and feasibility randomised controlled trial of an interactive digital intervention to increase condom use in men. *Health Technology Assessment*, 20.

Brooks, S., Gelman, A., Jones, G., and Meng, X. (2011). *Handbook of Markov Chain Monte Carlo*. CRC press.

Noble, S., Hollingworth, W., and Tilling, K. (2012). Missing data in trial-based cost-effectiveness analysis: the current state of play. *Health Economics*, 21:187–200.

Rubin, D. (1987). *Multiple Imputation for Nonresponse in Surveys*. John Wiley and Sons, New York,USA.