Full Bayesian models to handle missing values in cost-effectiveness analysis from individual level data

Gianluca Baio

(Joint work with Andrea Gabrio and Alexina Mason)

University College London
Department of Statistical Science

g.baio@ucl.ac.uk

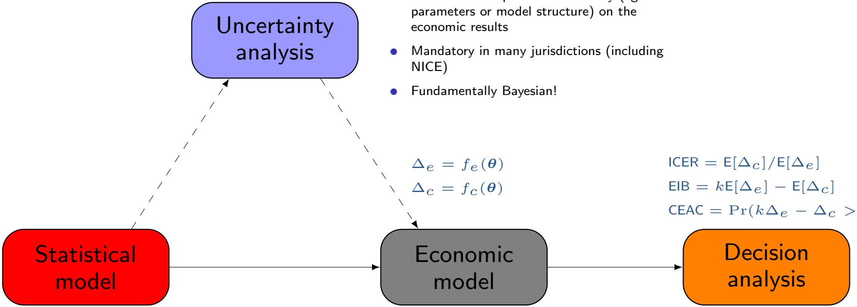
http://www.ucl.ac.uk/statistics/research/statistics-health-economics/
http://www.statistica.it/gianluca
https://github.com/giabaio

RSS 2017 International Conference
University of Strathclyde, Glasgow

Wednesday 6 September 2017

Health technology assessment (HTA)

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!

$\Delta_e = f_e(\boldsymbol{\theta})$

$\Delta_c = f_c(\boldsymbol{\theta})$

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

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

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

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
...	...	...	...	...	...	...	...	...	...	...	...	...

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

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

3. Estimate population average cost and effectiveness differentials and use bootstrap to quantify uncertainty

What's wrong with this?

- Potential correlation between costs & clinical benefits
  - Strong positive correlation — effective treatments are innovative and result from intensive and lengthy research ⇒ 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.
  - **NB**: In any case, the economic evaluation is based on both!

- 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
  - The missingness mechanisms may also be correlated

  - Focus in decision-making — not inference!

To be or not to be (Bayesians)?...

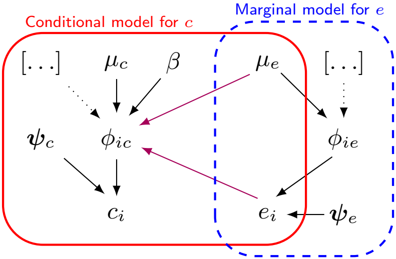
- In general, can represent a joint distribution as a **conditional regression**

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

$c_i \sim p(c \mid e, \phi_{ci}, \boldsymbol\psi_c)$

$g(\phi_{ci}) = \mu_c + \beta(e_i - \mu_e)\,[+\ldots]$

ϕ_{ci} = location
$\boldsymbol\psi_c$ = ancillary



$e_i \sim p(e \mid \phi_{ei}, \boldsymbol\psi_e)$

$g(\phi_{ei}) = \mu_e\,[+\ldots]$

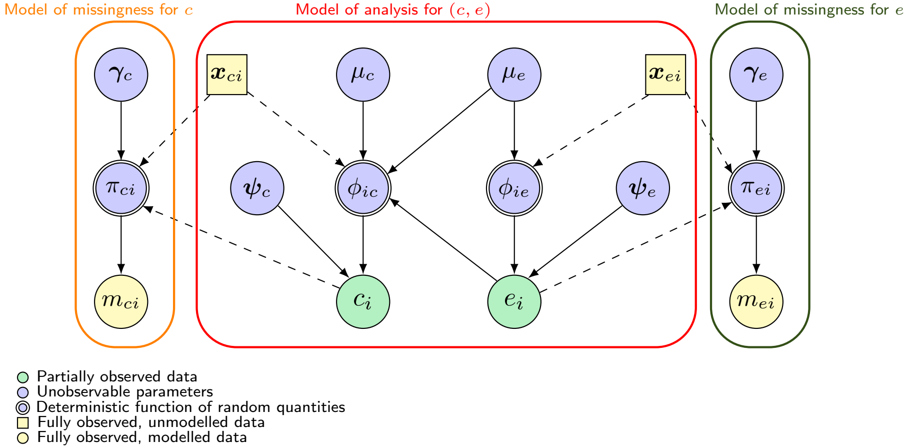
ϕ_{ei} = location
$\boldsymbol\psi_e$ = ancillary

- For example:

$$e_i \sim \text{Beta}(\phi_{ei}\psi_e, (1-\phi_{ei})\psi_e), \qquad \text{logit}(\phi_{ei}) = \mu_e\,[+\ldots]$$
$$c_i \mid e_i \sim \text{Gamma}(\psi_c\phi_{ci}, \psi_c), \qquad \log(\phi_{ci}) = \mu_c + \beta(e_i - \mu_e)\,[+\ldots]$$

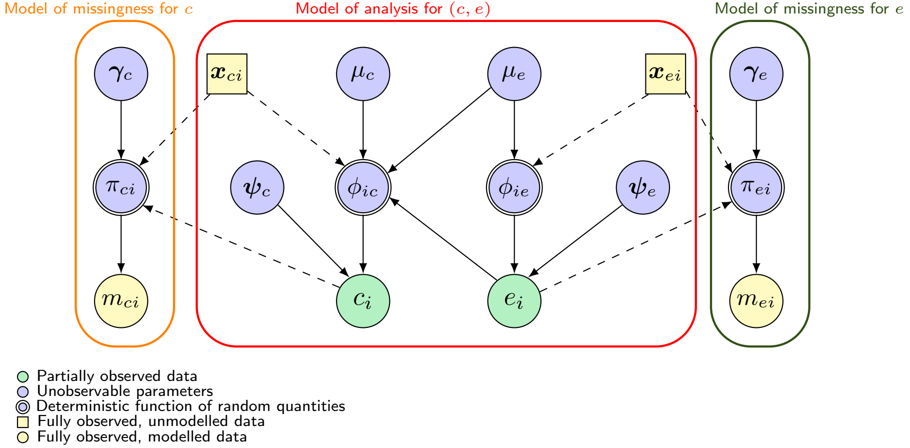
- Combining "modules" and fully characterising uncertainty about deterministic functions of random quantities is relatively straightforward using MCMC
- Prior information can help stabilise inference (especially with sparse data!), eg
  - Cancer patients are unlikely to survive as long as the general population
  - ORs are unlikely to be greater than ± 5

MCAR (e, c)



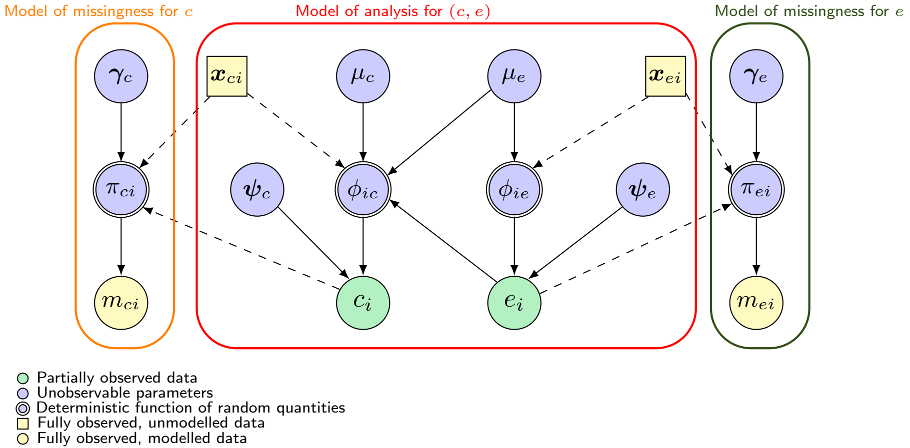
- $m_{ei} \sim \text{Bernoulli}(\pi_{ei})$; $\text{logit}(\pi_{ei}) = \gamma_{e0}$
- $m_{ci} \sim \text{Bernoulli}(\pi_{ci})$; $\text{logit}(\pi_{ci}) = \gamma_{c0}$

MAR (e, c)



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

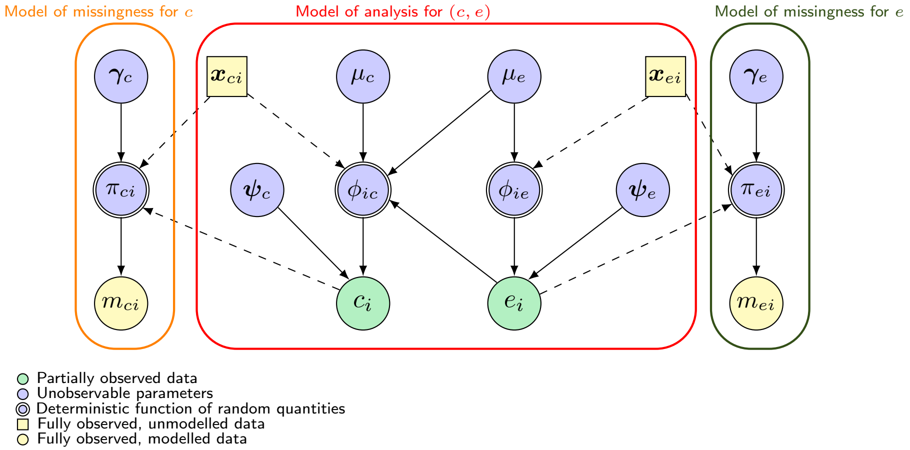
MNAR (e, c)



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

Dealing with missing data — selection models

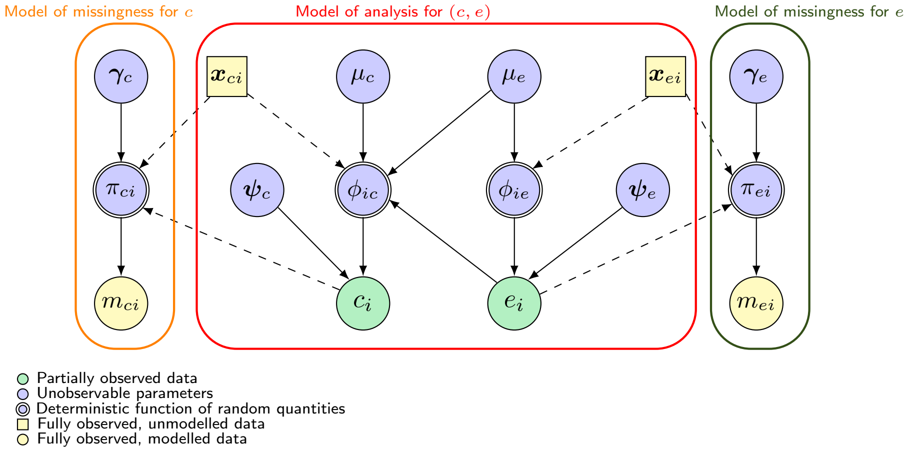
MNAR e; MAR c



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

Dealing with missing data — selection models

MAR e; MNAR c

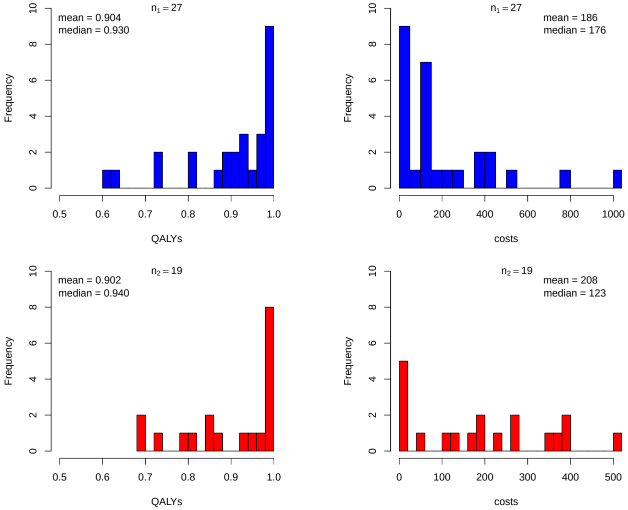


- $m_{ei} \sim \text{Bernoulli}(\pi_{ei})$; $\quad \text{logit}(\pi_{ei}) = \gamma_{e0} + \sum_{k=1}^{K} \gamma_{ek} x_{eik}$
- $m_{ci} \sim \text{Bernoulli}(\pi_{ci})$; $\quad \text{logit}(\pi_{ci}) = \gamma_{c0} + \sum_{h=1}^{H} \gamma_{ch} x_{cih} + \gamma_{cH+1} c_i$

- 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%)

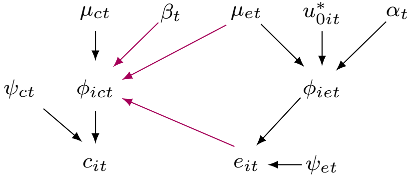
- The MenSS pilot RCT evaluates the cost-effectiveness of a new digital intervention to reduce the incidence of STI in young men with respect to the SOC
  - QALYs calculated from utilities (EQ-5D 3L)
  - Total costs calculated from different components (no baseline)



1. **Bivariate Normal**
   - Simpler and closer to "standard" frequentist model
   - Account for correlation between QALYs and costs

Conditional model for $c \mid e$

$$c_{it} \mid e_{it} \sim \text{Normal}(\phi_{cit}, \psi_{ct})$$
$$\phi_{cit} = \mu_{ct} + \beta_t(e_{it} - \mu_{et})$$



Marginal model for e

$$e_{it} \sim \text{Normal}(\phi_{eit}, \psi_{et})$$
$$\phi_{eit} = \mu_{et} + \alpha_t(u_{0it} - \bar{u}_{0t})$$
$$= \mu_{et} + \alpha_t u^*_{0it}$$

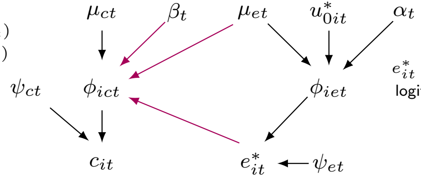
Modelling

1. **Bivariate Normal**
   - Simpler and closer to "standard" frequentist model
   - Account for correlation between QALYs and costs
2. **Beta-Gamma**
   - Account for correlation between outcomes
   - Model the relevant ranges: QALYs $\in (0,1)$ and costs $\in (0,\infty)$
   - **But**: needs to rescale observed data $e^*_{it} = (e_{it} - \epsilon)$ to avoid spikes at 1

Conditional model for $c \mid e^*$

$c_{it} \mid e^*_{it} \sim \text{Gamma}(\psi_{ct}\phi_{cit}, \psi_{ct})$

$\log(\phi_{cit}) = \mu_{ct} + \beta_t(e^*_{it} - \mu_{et})$



Marginal model for e^*

$e^*_{it} \sim \text{Beta}\left(\phi_{eit}\psi_{et}, (1-\phi_{eit})\psi_{et}\right)$

$\text{logit}(\phi_{eit}) = \mu_{et} + \alpha_t(u_{0it} - \bar{u}_{0t})$

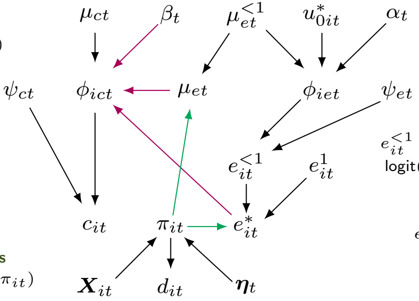
$= \mu_{et} + \alpha_t u^*_{0it}$

Modelling

1. **Bivariate Normal**
   - Simpler and closer to "standard" frequentist model
   - Account for correlation between QALYs and costs
2. **Beta-Gamma**
   - Account for
   - Model the relevant ranges: QALYs $\in (0,1)$ and costs $\in (0,\infty)$
   - **But**: needs to rescale observed data $e^*_{it} = (e_{it} - \epsilon)$ to avoid spikes at 1
3. **Hurdle model**
   - Model e_{it} as a **mixture** to account for correlation between outcomes, model the relevant ranges and account for structural values
   - May expand to account for partially observed baseline utility u_{0it}

Conditional model for $c \mid e^*$

$$c_{it} \mid e^*_{it} \sim \text{Gamma}(\psi_{ct}\phi_{cit}, \psi_{ct})$$
$$\log(\phi_{cit}) = \mu_{ct} + \beta_t(e^*_{it} - \mu_{et})$$



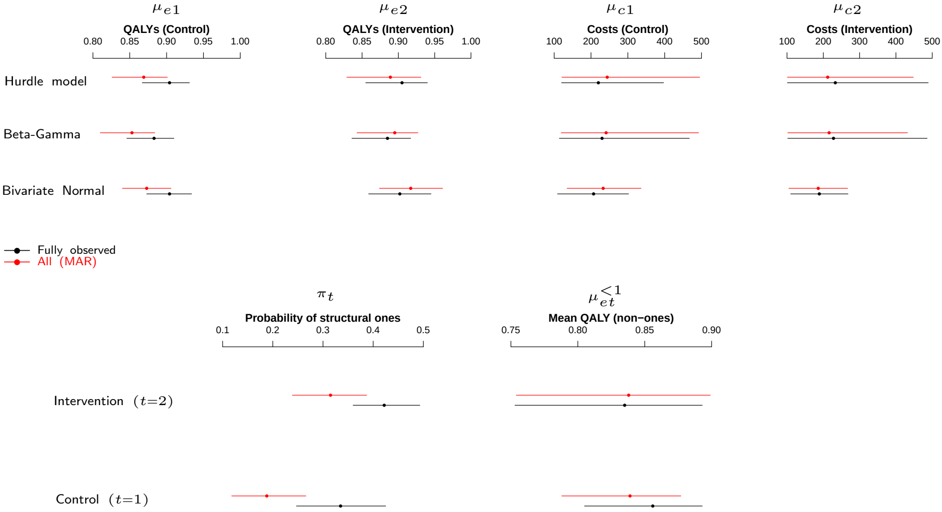
Mixture model for e

$$e^1_{it} := 1$$
$$e^{<1}_{it} \sim \text{Beta}(\phi_{eit}\psi_{et}, (1-\phi_{eit})\psi_{et})$$
$$\text{logit}(\phi_{eit}) = \mu^{<1}_{et} + \alpha_t(u_{0it} - \bar{u}_{0t}) = \mu^{<1}_{et} + \alpha_t u^*_{0it}$$
$$e^*_{it} = \pi_{it} e^1_{it} + (1-\pi_{it}) e^{<1}_{it}$$
$$\mu_{et} = (1-\bar{\pi}_t)\mu^{<1}_{et} + \bar{\pi}_t$$

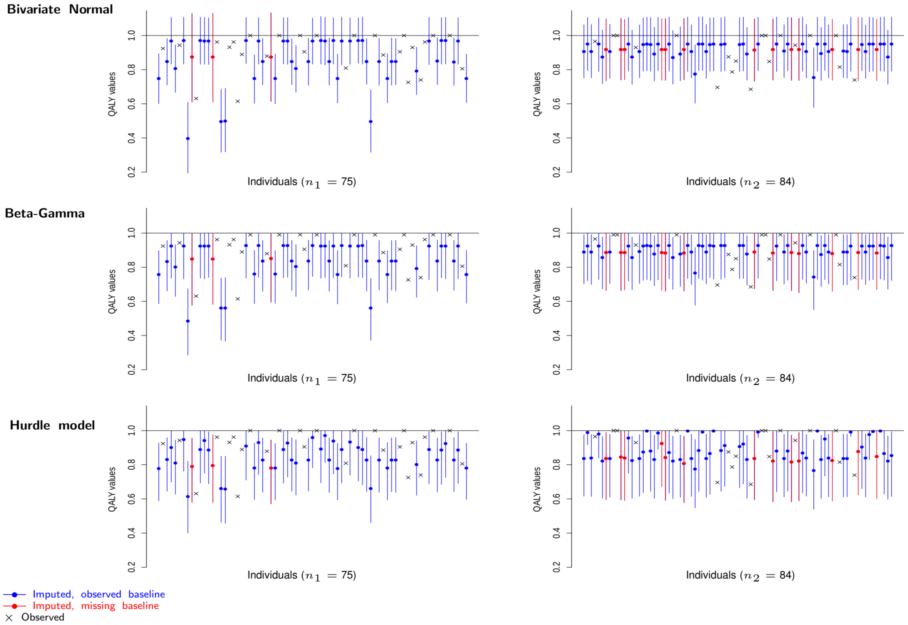
Model for the structural ones

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

Results — estimation of the main parameters (CCA + MAR)



Bayesian multiple imputation (under MAR)



MNAR

- We observe $n_{01} = 13$ and $n_{02} = 22$ individuals with $u_{0it} = 1$ and $u_{jit} = \texttt{NA}$, for $j > 1$
- 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

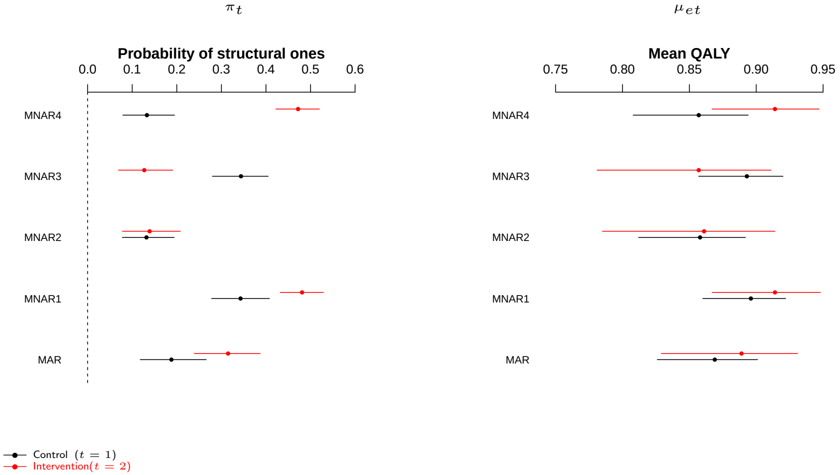
MNAR1. Set $d_{it} = 1$ for all individuals with unit observed baseline utility

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

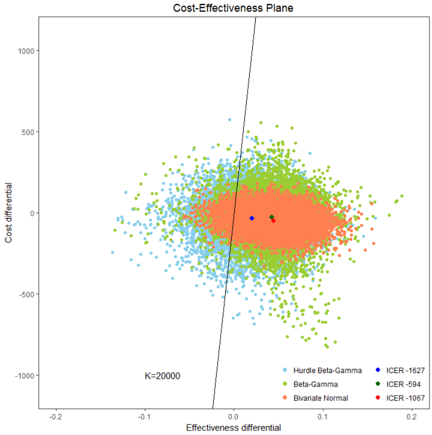
MNAR3. Set $d_{it} = 1$ for the $n_{01} = 13$ individuals with $u_{0i1} = 1$ and $d_{it} = 0$ for the $n_{02} = 22$ individuals with $u_{0i2} = 1$

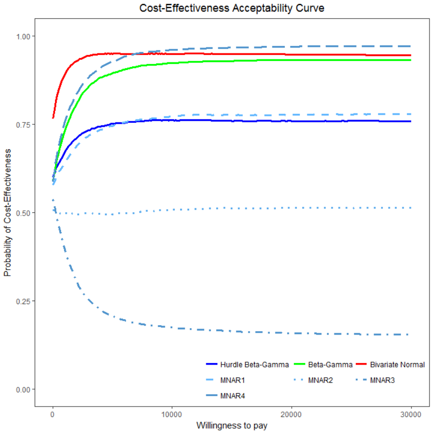
MNAR4. Set $d_{it} = 0$ for the $n_{01} = 13$ individuals with $u_{0i1} = 1$ and $d_{it} = 1$ for the $n_{02} = 22$ individuals with $u_{0i2} = 1$

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

- 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)

- Need specialised software + coding skills
  - R package **missingHE** under development to implement a set of general models
  - Preliminary work available at https://github.com/giabaio/missingHE
  - Eventually, will be able to combine with existing packages (eg **BCEA**: http://www.statistica.it/gianluca/BCEA; https://github.com/giabaio/BCEA) to perform the whole economic analysis

Thank you!