A Bayesian modelling framework for health care resource use and costs in trial-based HEE

Andrea Gabrio

Department of Methodology and Statistics, FHML (UM)

a.gabrio@maastrichtuniversity.nl

https://github.com/AnGabrio

https://angabrio.github.io

Bayes Conference, Leiden, 2025

Thursday, October 23, 2025

Part I

Introduction to statistical modelling in HTA

Individual-level data in HTA

- Typically collected from studies (e.g. RCTs) at multiple time points

		Demographics			HRQL data				Cost data				Clinical outcome			
ID	**Trt**	Sex	Age	…	u_0	u_1	…	u_J	c_0	c_1	…	c_J	y_0	y_1	…	y_J
1	1	M	23	…	0.32	0.66	…	0.44	103	241	…	80	y_{10}	y_{11}	…	y_{1J}
2	1	M	21	…	0.12	0.16	…	0.38	1 204	1 808	…	877	y_{20}	y_{21}	…	y_{2J}
3	2	F	19	…	0.49	0.55	…	0.88	16	12	…	22	y_{30}	y_{31}	…	y_{3J}
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…

y_{ij} = Survival time, event indicator (eg CVD), number of events, continuous measurement (eg blood pressure), …
u_{ij} = Utility-based score to value health (eg EQ-5D, SF-36, Hospital Anxiety & Depression Scale, …)
c_{ij} = Cost of healthcare resources/services (eg CSRI, iMCQ, …)

Individual-level data in HTA

- Typically collected from studies (e.g. RCTs) at multiple time points

		Demographics			HRQL data				Cost data				Clinical outcome			
ID	**Trt**	Sex	Age	…	u_0	u_1	…	u_J	c_0	c_1	…	c_J	y_0	y_1	…	y_J
1	1	M	23	…	0.32	0.66	…	0.44	103	241	…	80	y_{10}	y_{11}	…	y_{1J}
2	1	M	21	…	0.12	0.16	…	0.38	1 204	1 808	…	877	y_{20}	y_{21}	…	y_{2J}
3	2	F	19	…	0.49	0.55	…	0.88	16	12	…	22	y_{30}	y_{31}	…	y_{3J}
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…

y_{ij} = Survival time, event indicator (eg CVD), number of events, continuous measurement (eg blood pressure), …
u_{ij} = Utility-based score to value health (eg EQ-5D, SF-36, Hospital Anxiety & Depression Scale, …)
c_{ij} = Cost of healthcare resources/services (eg CSRI, iMCQ, …)

- Effectiveness/utility data at each time j and for each individual i often obtained after mapping country-specific value sets to the individual responses of L health domains assessed (profile):

$$u_{ij} = \text{mapping algorithm}[\text{HRQL}^1_{ij}, \ldots, \text{HRQL}^L_{ij}]$$

Individual-level data in HTA

- Typically collected from studies (e.g. RCTs) at multiple time points

		Demographics			HRQL data				Cost data				Clinical outcome			
ID	**Trt**	Sex	Age	...	u_0	u_1	...	u_J	c_0	c_1	...	c_J	y_0	y_1	...	y_J
1	1	M	23	...	0.32	0.66	...	0.44	103	241	...	80	y_{10}	y_{11}	...	y_{1J}
2	1	M	21	...	0.12	0.16	...	0.38	1 204	1 808	...	877	y_{20}	y_{21}	...	y_{2J}
3	2	F	19	...	0.49	0.55	...	0.88	16	12	...	22	y_{30}	y_{31}	...	y_{3J}
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

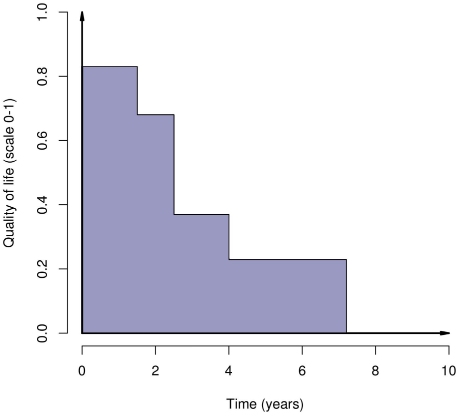
y_{ij} = Survival time, event indicator (eg CVD), number of events, continuous measurement (eg blood pressure), ...
u_{ij} = Utility-based score to value health (eg EQ-5D, SF-36, Hospital Anxiety & Depression Scale, ...)
c_{ij} = Cost of healthcare resources/services (eg CSRI, iMCQ, ...)

- Cost data at each time j and for each individual i obtained after applying service-specific unit prices p_k to each of the K resource use services collected and summing up their costs:

$$c_{ij} = \text{HRU}_{ij}^{1} \times p_1 + \ldots + \text{HRU}_{ij}^{K} \times p_K$$

Individual-level data in HTA

- Aggregate measures, e.g. QALYs and total costs, are then obtained by combining u_{ij} and c_{ij} data over the study follow-up time.



QALY = "Area Under the Curve"

$$e_i = \sum_{j=1}^{J} \left(u_{ij} + u_{ij-1}\right) \frac{\delta_j}{2}$$

$$tc_i = \sum_{j=1}^{J} c_{ij}$$

$$\text{where } \delta_j = \frac{\text{Time}_j - \text{Time}_{j-1}}{\text{Unit of Time}}$$

Handling missing data in trial-based CEA

- Things to consider in the analysis (Faria et al. 2014):
  - Avoid biased methods (e.g. *case deletion*) in favour of those grounded on "plausible" **missingness assumptions** (e.g. *multiple imputation*)
  - Explore the impact of alternative assumptions on results in **sensitivity analysis**

Handling missing data in trial-based CEA

- Things to consider in the analysis (Faria et al. 2014):
  - Avoid biased methods (e.g. *case deletion*) in favour of those grounded on "plausible" **missingness assumptions** (e.g. *multiple imputation*)
  - Explore the impact of alternative assumptions on results in **sensitivity analysis**
- Analysis can be conducted at *different levels*:
  - HRQL and HRU responses at each time (**disaggregated**)
  - Utilities and costs at each time (**intermediate**)
  - QALYs and Total costs (**aggregated**)

Handling missing data in trial-based CEA

- Things to consider in the analysis (Faria et al. 2014):
  - Avoid biased methods (e.g. *case deletion*) in favour of those grounded on "plausible" **missingness assumptions** (e.g. *multiple imputation*)
  - Explore the impact of alternative assumptions on results in **sensitivity analysis**
- Analysis can be conducted at *different levels*:
  - HRQL and HRU responses at each time (**disaggregated**)
  - Utilities and costs at each time (**intermediate**)
  - QALYs and Total costs (**aggregated**)
- Need to choose the right **balance** between *modelling all evidence* and *fitting a stable model*

Current missing data practice

Ling at el. 2022 revealed a somewhat **unclear picture** about *missing data methods* in routine analyses (2016-2021):

- Different according to the level of data aggregation used in analysis
- **MI** often used at aggregated/intermediate level, while at disaggregated level missing values are often **"filled-in"** in some ways
- Typical of HRU (usually some answers skipped) compared to HRQL instruments (usually all answers skipped)
- In many cases HRUs are imputed based on *restrictive assumptions* (e.g. **assumed zero**) which, if not justified, can **misrepresent uncertainty** and **distort** the results

A Bayesian framework for trial-based CE data

- Why Bayesian (Spiegelhalter et al. 2004)?
  - **Modular structure** to handle multiple data features (e.g. skewness, clustering, missingness)
  - Imputation and estimation steps are **jointly** performed and always "congenial"
  - Direct implementation of **probabilistic sensitivity analysis** (PSA)
- **Aim** - apply framework to a real case study:
  - Assess impact of *different missing HRU assumptions* on posterior estimates
  - Fit separate models *at different levels of data aggregation* and compare the results

Part II

Case study

The PBS trial

- The *Positive Behaviour Support* (PBS) trial was an RCT involving individuals with mild to severe intellectual disabilities (Hassiotis et al. 2018)
  - Participants ($n = 244$) allocated to either **PBS** ($n_2 = 108$) or **TAU** ($n_1 = 136$)
  - HRQL and HRU data collected via EQ-5D-5L and CSRI questionnaires at baseline ($j = 0$), 6 and 12-months ($j = 1, 2$)
- Focus on the following individual-level CE data:
  - Utilities derived from EQ-5D answers and UK value sets (u_{ij})
  - HRUs for $K = 9$ types of healthcare services (hru_{ij}^{k})

Missingness rates

Outcome	Baseline ($j = 0$)	6-months ($j = 1$)	12-months ($j = 2$)	ic (outcome)
Utilities (u)	14 (5.7%)	23 (9.4%)	16 (6.6%)	40 (16.4%)
PSYDR (hru^1)	0	18 (7.4%)	12 (4.9%)	22 (9%)
PSYCH (hru^2)	2 (1.4%)	12 (4.9%)	9 (3.7%)	14 (5.7%)
PHYSI (hru^3)	1 (0.4%)	13 (5.3%)	10 (4.1%)	16 (6.6%)
DENT (hru^4)	1 (0.4%)	14 (5.7%)	12 (4.9%)	19 (7.8%)
SOCWORK (hru^5)	1 (0.4%)	14 (5.7%)	11 (4.5%)	18 (7.4%)
COMWORK (hru^6)	0	13 (5.3%)	11 (4.5%)	15 (6.1%)
GP (hru^7)	3 (2.1%)	15 (6.2%)	14 (5.7%)	22 (9%)
NURSE (hru^8)	2 (1.4%)	17 (7%)	13 (5.3%)	24 (9.8%)
THERAP (hru^9)	1 (0.4%)	15 (6.1%)	12 (4.9%)	19 (7.8%)
ic (time)	10 (4.2%)	22 (9%)	17 (7%)	**59 (24.2%)**

- EQ5D data only affected by *unit nonresponse*, while CSRI data affected by both *unit* and *item nonresponse*
- Overall proportion of incomplete cases is **substantial** ($\approx 24\%$)

Structural zeros

Outcome	Structural zeros
PSYDR (hru^1)	37 (15%)
PSYCH (hru^2)	159 (65%)
PHYSI (hru^3)	189 (77%)
DENT (hru^4)	35 (14%)
SOCWORK (hru^5)	47 (19%)
COMWORK (hru^6)	196 (80%)
GP (hru^7)	1 (0.4%)
NURSE (hru^8)	53 (22%)
THERAP (hru^9)	85 (35%)

- Many individuals have constant zero HRU values across all times (so called **"structural zeros"**)
- These values induce a *high degree of skewness* in the empirical distributions

Part III

Modelling framework

Fitting models at different levels of data aggregation

- Consider three alternative model specifications:
  - At the **aggregated** level (e_i, tc_i) – O'Hagan et al. 2001
  - At the **intermediate** level (u_{ij}, c_{ij}) – Gabrio et al. 2020
  - At the **disaggregated** level $(\text{HRQL}^{l}_{ij}, \text{HRU}^{k}_{ij})$

Fitting models at different levels of data aggregation

- Consider three alternative model specifications:
  - At the **aggregated** level (e_i, tc_i) – O'Hagan et al. 2001
  - At the **intermediate** level (u_{ij}, c_{ij}) – Gabrio et al. 2020
  - At the **disaggregated** level $(\text{HRQL}^{l}_{ij}, \text{HRU}^{k}_{ij})$
- Easier to specify model at **more aggregated** levels **BUT**:
  - when item nonresponse occurs, it is not possible to directly incorporate evidence from partially-observed cases into the analysis
  - need to either **discard** cases or **impute** them in some way *prior* to model fitting

Fitting models at different levels of data aggregation

- Consider three alternative model specifications:
  - At the **aggregated** level (e_i, tc_i) – O'Hagan et al. 2001
  - At the **intermediate** level (u_{ij}, c_{ij}) – Gabrio et al. 2020
  - At the **disaggregated** level $(\text{HRQL}_{ij}^{l}, \text{HRU}_{ij}^{k})$
- Easier to specify model at **more aggregated** levels **BUT**:
  - when item nonresponse occurs, it is not possible to directly incorporate evidence from partially-observed cases into the analysis
  - need to either **discard** cases or **impute** them in some way *prior* to model fitting
- Models fitted at **more disaggregated** level can overcome this limitation but are usually *more challenging* to fit

Aggregated-data model

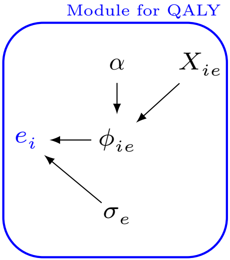
- Cross-sectional model for QALYs (e_i) and Total costs (tc_i)

$$p(e_i, tc_i) = p(tc_i | e_i) p(e_i)$$

Aggregated-data model

- Cross-sectional model for QALYs (e_i) and Total costs (tc_i)

$$p(e_i, tc_i) = p(tc_i | e_i) p(e_i)$$



Marginal model for e

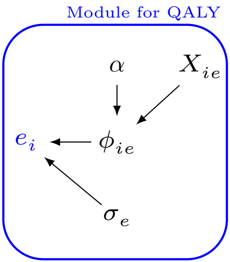
- Select a *distribution* $f(\cdot)$ for each variable (e.g. Normal, Gamma)

$$e_i \sim f(\phi_{ie}, \sigma_e)$$

Aggregated-data model

- Cross-sectional model for QALYs (e_i) and Total costs (tc_i)

$$p(e_i, tc_i) = p(tc_i | e_i) p(e_i)$$



Marginal model for e

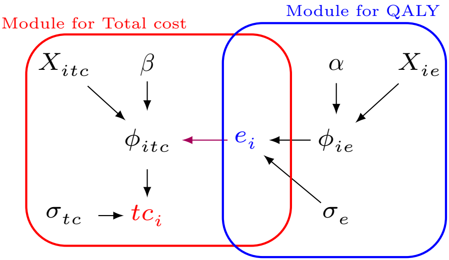
- Select a *distribution* $f(\cdot)$ for each variable (e.g. Normal, Gamma)
- Choose *structure* and *link function* $g(\cdot)$ to model dependence on other variables (e.g. $\log$)

$$e_i \sim f(\phi_{ie}, \sigma_e)$$
$$g(\phi_{ie}) = \alpha_0 + \alpha_1 X_{ie} + \ldots$$

Aggregated-data model

- Cross-sectional **joint** model for QALYs (e_i) and Total costs (tc_i)

$$p(e_i, tc_i) = p(tc_i|e_i)p(e_i)$$



Marginal model for e

$$e_i \sim f(\phi_{ie}, \sigma_e)$$
$$g(\phi_{ie}) = \alpha_0 + \alpha_1 X_{ie} + \ldots$$

Conditional model for tc

- Capture *correlation* between the outcomes

$$tc_i|e_i \sim f(\phi_{itc}, \sigma_{tc})$$
$$g(\phi_{itc}) = \beta_0 + \beta_1 e_i + \beta_2 X_{itc} + \ldots$$

Aggregated-data model

- Mean incremental QALYs ($\Delta_e = \mu_{e2} - \mu_{e1}$) and Total costs ($\Delta_{tc} = \mu_{tc2} - \mu_{tc1}$) derived as *linear combination* of model parameters or through *simulation* (e.g. Monte Carlo methods)

Aggregated-data model

- Mean incremental QALYs ($\Delta_e = \mu_{e2} - \mu_{e1}$) and Total costs ($\Delta_{tc} = \mu_{tc2} - \mu_{tc1}$) derived as *linear combination* of model parameters or through *simulation* (e.g. Monte Carlo methods)
- **Advantages**
  - Need to model only two variables (e_i,tc_i)
  - Relatively easy to implement in most cases

Aggregated-data model

- Mean incremental QALYs ($\Delta_e = \mu_{e2} - \mu_{e1}$) and Total costs ($\Delta_{tc} = \mu_{tc2} - \mu_{tc1}$) derived as *linear combination* of model parameters or through *simulation* (e.g. Monte Carlo methods)
- **Advantages**
  - Need to model only two variables (e_i,tc_i)
  - Relatively easy to implement in most cases
- **Drawbacks**
  - HRQL/HRU item nonreponses are either *discarded* or *imputed* a priori
  - Incomplete utilities / costs are either *discarded* or *imputed* a priori
  - Often partially-missing HRUs/ costs are imputed as **zero**, and used to calculate tc_i to ensure higher completion rates

Intermediate-data model

- Longitudinal **joint** model for utilities (u_{ij}) and costs (c_{ij})

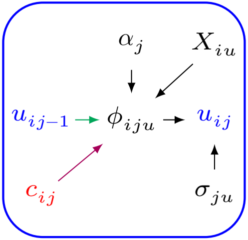
$$p(u_{ij}, c_{ij}) = p(u_{ij} | u_{ij-1}, c_{ij}) p(c_{ij} | u_{ij-1}, c_{ij-1})$$

Intermediate-data model

- Longitudinal **joint** model for utilities (u_{ij}) and costs (c_{ij})

$$p(u_{ij},c_{ij}) = p(u_{ij}|u_{ij-1},c_{ij})p(c_{ij}|u_{ij-1},c_{ij-1})$$

Module for utility at j



Conditional models for u

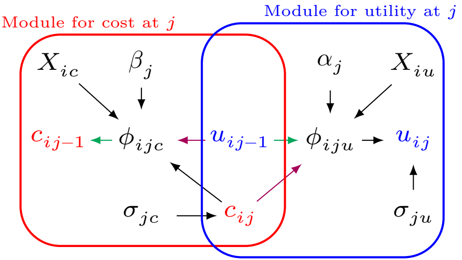
- Select a *distribution*, *structure* and *link function* for each variable
- Capture *correlation* between outcomes and *dependence* over time

$$\begin{aligned} u_{ij}|u_{ij-1},c_{ij} &\sim f(\phi_{iju},\sigma_{ju}) \\ g(\phi_{iju}) &= \alpha_{0j} + \alpha_{1j}X_{iu} + \\ &\quad \alpha_{2j}u_{ij-1} + \alpha_{3j}c_{ij} + \ldots \end{aligned}$$

Intermediate-data model

- Longitudinal **joint** model for utilities (u_{ij}) and costs (c_{ij})

$$p(u_{ij}, c_{ij}) = p(u_{ij} | u_{ij-1}, c_{ij}) p(c_{ij} | u_{ij-1}, c_{ij-1})$$



Conditional models for u

$$\begin{aligned} u_{ij} | u_{ij-1}, c_{ij} &\sim f(\phi_{iju}, \sigma_{ju}) \\ g(\phi_{iju}) &= \alpha_{0j} + \alpha_{1j} X_{iu} + \\ &\quad \alpha_{2j} u_{ij-1} + \alpha_{3j} c_{ij} + \ldots \end{aligned}$$

Conditional models for c

$$\begin{aligned} c_{ij} | u_{ij-1}, c_{ij-1} &\sim f(\phi_{ijc}, \sigma_{jc}) \\ g(\phi_{ijc}) &= \beta_{0j} + \beta_{1j} X_{ic} + \\ &\quad \beta_{2j} u_{ij-1} + \beta_{3j} c_{ij-1} + \ldots \end{aligned}$$

Intermediate-data model

- Time-specific mean utilities (μ_{ju}) and costs (μ_{jc}) derived as *linear combination* of model parameters or through *simulation* (e.g. Monte Carlo methods), and then used to calculate Δ_e and Δ_{tc}

Intermediate-data model

- Time-specific mean utilities (μ_{ju}) and costs (μ_{jc}) derived as *linear combination* of model parameters or through *simulation* (e.g. Monte Carlo methods), and then used to calculate Δ_e and Δ_{tc}
- **Advantages**
  - Directly capture the *longitudinal structure* of the data
  - Missingness in (u_{ij},c_{ij}) is directly handled

Intermediate-data model

- Time-specific mean utilities (μ_{ju}) and costs (μ_{jc}) derived as *linear combination* of model parameters or through *simulation* (e.g. Monte Carlo methods), and then used to calculate Δ_e and Δ_{tc}
- **Advantages**
  - Directly capture the *longitudinal structure* of the data
  - Missingness in (u_{ij},c_{ij}) is directly handled
- **Drawbacks**
  - More challenging to fit (i.e. $J \times 2$ variables)
  - HRQL / HRU item nonreponses are either *discarded* or *imputed* a priori
  - Often partially-missing HRUs are imputed as **zero**, and used to calculate c_{ij} to ensure higher completion rates

Disaggregated-data model

- Longitudinal **joint** model for utilities (u_{ij}) and HRUs (hru^k_{ij})

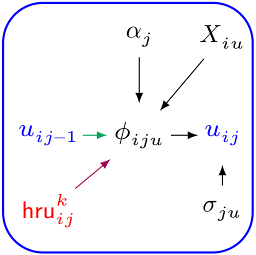
$$p(u_{ij}, \mathsf{hru}^k_{ij}) = p(u_{ij} | u_{ij-1}, \mathsf{hru}^k_{ij}) p(\mathsf{hru}^k_{ij} | u_{ij-1}, \mathsf{hru}^k_{ij-1})$$

Disaggregated-data model

- Longitudinal **joint** model for utilities (u_{ij}) and HRUs (hru_{ij}^{k})

$$p(u_{ij},\text{hru}_{ij}^{k}) = p(u_{ij}|u_{ij-1},\text{hru}_{ij}^{k})p(\text{hru}_{ij}^{k}|u_{ij-1},\text{hru}_{ij-1}^{k})$$

Module for utility at j



Conditional models for u

- Select a *distribution*, *structure* and *link function* for each variable
- Capture *correlation* between outcomes and *dependence* over time

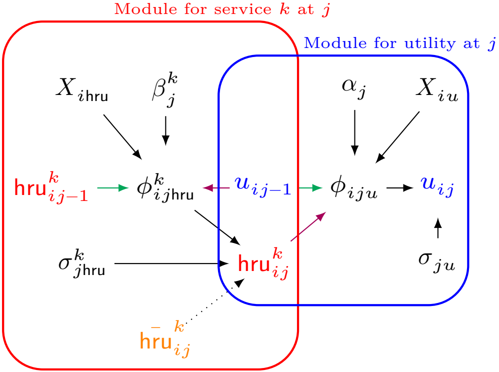
$$u_{ij}|u_{ij-1},\text{hru}_{ij}^{k} \sim f(\phi_{iju},\sigma_{ju})$$

$$g(\phi_{iju}) = \alpha_{0j} + \alpha_{1j}X_{iu} + \alpha_{2j}u_{ij-1} + \alpha_{3j}\text{hru}_{ij}^{k} + \ldots$$

Disaggregated-data model

- Longitudinal **joint** model for utilities (u_{ij}) and HRUs (hru_{ij}^{k})

$$p(u_{ij}, \text{hru}_{ij}^{k}) = p(u_{ij} | u_{ij-1}, \text{hru}_{ij}^{k}) p(\text{hru}_{ij}^{k} | u_{ij-1}, \text{hru}_{ij-1}^{k})$$



Conditional models for hru^k

- Specify outcome models for only *non-structural* values for service k

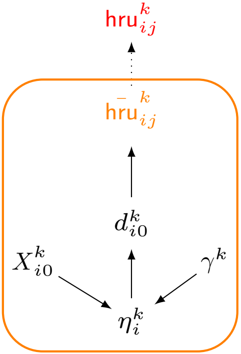
$$\text{hru}_{ij}^{k} | u_{ij-1}, \text{hru}_{ij-1}^{k} \sim f(\phi_{ijc}, \sigma_{jc})$$

$$g(\phi_{ijc}) = \beta_{0j} + \beta_{1j} X_{ic} + \beta_{2j} u_{ij-1} + \beta_{3j} \text{hru}_{ij-1}^{k} + \ldots$$

Disaggregated-data model

- Longitudinal **joint** model for utilities (u_{ij}) and HRUs (hru^k_{ij})

$$p(u_{ij},\text{hru}^k_{ij}) = p(u_{ij}|u_{ij-1},\text{hru}^k_{ij})p(\text{hru}^k_{ij}|u_{ij-1},\text{hru}^k_{ij-1})$$



Module zeros for serive k

Logistic models for $\overline{\text{hru}}^k$

- Handle **structural zeros** in service k using an **hurdle approach**

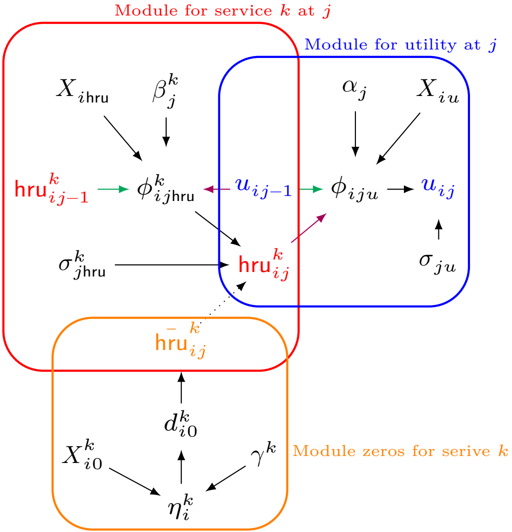
$$d^k_{i0} := \mathbb{I}(\text{hru}_{i\forall j} = 0) \sim \text{Bernoulli}(\eta^k_i)$$
$$\text{logit}(\eta^k_i) = \gamma^k_0 + \gamma^k_1 X_{i0} + \ldots$$

- Original hru^k_{ij} are expressed as:

$$\text{hru}^k_{ij} = \text{hru}^k_{ij} \times (1-\pi^k_0) + \overline{\text{hru}}^k_{ij} \times \pi^k_0$$

Disaggregated-data model

- Longitudinal **joint** model for utilities (u_{ij}) and HRUs (hru^k_{ij})



Disaggregated-data model

- Estimates for the overall mean HRUs for each service k and time j are then obtained as:
  - the linear combination $\mu^k_{j\text{hru}} = (1 - \pi^k_0)\mu^k_{j\text{hru}}$
  - which are then combined with unit prices p^k to generate mean costs μ^k_{jc} and summed across all services and times to obtain μ_{tc} and Δ_{tc}

Disaggregated-data model

- Estimates for the overall mean HRUs for each service k and time j are then obtained as:
  - the linear combination $\mu^k_{j\text{hru}} = (1 - \pi^k_0)\mu^k_{j\text{hru}}$
  - which are then combined with unit prices p^k to generate mean costs μ^k_{jc} and summed across all services and times to obtain μ_{tc} and Δ_{tc}
- **Advantages**
  - Directly capture the *longitudinal structure* of the data
  - Missingness in (u_{ij},hru^k_{ij}) directly handled **without** the need of prior imputation

Disaggregated-data model

- Estimates for the overall mean HRUs for each service k and time j are then obtained as:
  - the linear combination $\mu^k_{j\text{hru}} = (1-\pi^k_0)\mu^k_{j\text{hru}}$
  - which are then combined with unit prices p^k to generate mean costs μ^k_{jc} and summed across all services and times to obtain μ_{tc} and Δ_{tc}
- **Advantages**
  - Directly capture the *longitudinal structure* of the data
  - Missingness in $(u_{ij},\text{hru}^k_{ij})$ directly handled **without** the need of prior imputation
- **Drawbacks**
  - More challenging to fit (i.e. $J \times (K+1)$ variables)
  - Need a compromise between **model complexity and feasibility** (e.g. in small samples)

Implementation

- All models fitted to the PBS data via *JAGS*, called from *R*
  - Total of 20000 iterations per chain, with a burn-in of 10000
  - Convergence assessed via standard measures (e.g. *Rhat*, *Trace plots*)
  - Fit to observed data assessed through **graphical posterior predictive checks** and Bayesian **information measures** (e.g. *WAIC*)
- For each model, the distribution for each outcome was selected after comparing the fit of alternative choices:
  - **Normal** for utilities / QALYs
  - **Gamma** for costs / Total costs (compared to Normal and LogNormal)
  - **Normal + hurdle for structural zeros** for HRUs (compared to Poisson and Negative Binomial)

Part IV

Results

Posterior mean QALY/Total cost estimates

- Compare alternative *models* and *missingness strategies*:
  - **S1**) **Aggregated data**: all cases (**ALL**); zero-imputed HRU (**IMP-H**); zero-imputed HRU and costs (**IMP-HC**)
  - **S2**) **Intermediate data**: all cases (**ALL**); zero-imputed HRU (**IMP-H**)
  - **S3**) **Disaggregated data**: all cases (**ALL**)

Posterior mean QALY/Total cost estimates

- Compare alternative *models* and *missingness strategies*:
  - **S1**) **Aggregated data**: all cases (**ALL**); zero-imputed HRU (**IMP-H**); zero-imputed HRU and costs (**IMP-HC**)
  - **S2**) **Intermediate data**: all cases (**ALL**); zero-imputed HRU (**IMP-H**)
  - **S3**) **Disaggregated data**: all cases (**ALL**)

Approach	$\mu_{tc(t=1)}$ mean	$\mu_{tc(t=1)}$ 95% CI	$\mu_{tc(t=2)}$ mean	$\mu_{tc(t=2)}$ 95% CI	$\mu_{e(t=1)}$ mean	$\mu_{e(t=1)}$ 95% CI	$\mu_{e(t=2)}$ mean	$\mu_{e(t=2)}$ 95% CI
Strategy 1 - total cost & QALY								
ALL	2543	(2157;2938)	2754	(2249;3310)	**0.487**	(0.449;0.575)	**0.609**	(0.569;0.649)
IMP-H	2888	(2411;3393)	2379	(1939;2872)	**0.488**	(0.45;0.526)	**0.61**	(0.571;0.651)
IMP-HC	2395	(1897;2899)	2237	(1175;2749)	**0.486**	(0.449;0.523)	**0.61**	(0.572;0.649)
Strategy 2 - cost & utility at each time								
ALL	2607	(2253;2971)	2701	(2278;3145)	**0.494**	(0.463;0.527)	**0.6**	(0.566;0.635)
IMP-H	2453	(2087;2843)	2273	(1874;2664)	**0.494**	(0.462;0.526)	**0.6**	(0.565;0.633)
Strategy 3 - HRU category & utility at each time								
ALL	2687	(2173;3194)	2587	(1995;3206)	**0.513**	(0.475;0.55)	**0.599**	(0.565;0.634)

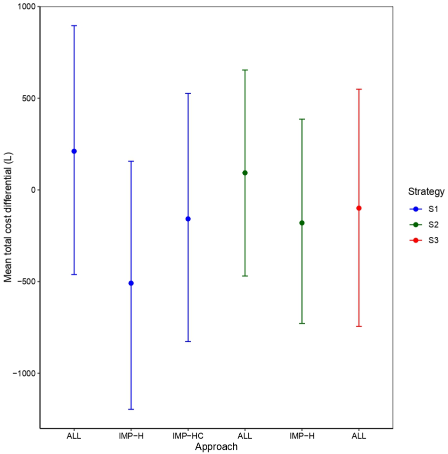
Posterior mean QALY/Total cost esitmates

- Compare alternative *models* and *missingness strategies*:
  - **S1**) **Aggregated data**: all cases (**ALL**); zero-imputed HRU (**IMP-H**); zero-imputed HRU and costs (**IMP-HC**)
  - **S2**) **Intermediate data**: all cases (**ALL**); zero-imputed HRU (**IMP-H**)
  - **S3**) **Disaggregated data**: all cases (**ALL**)

Approach	$\mu_{tc(t=1)}$ mean	$\mu_{tc(t=1)}$ 95% CI	$\mu_{tc(t=2)}$ mean	$\mu_{tc(t=2)}$ 95% CI	$\mu_{e(t=1)}$ mean	$\mu_{e(t=1)}$ 95% CI	$\mu_{e(t=2)}$ mean	$\mu_{e(t=2)}$ 95% CI
Strategy 1 - total cost & QALY								
ALL	**2543**	(2157;2938)	**2754**	(2249;3310)	0.487	(0.449;0.575)	0.609	(0.569;0.649)
IMP-H	**2888**	(2411;3393)	**2379**	(1939;2872)	0.488	(0.45;0.526)	0.61	(0.571;0.651)
IMP-HC	**2395**	(1897;2899)	**2237**	(1175;2749)	0.486	(0.449;0.523)	0.61	(0.572;0.649)
Strategy 2 - cost & utility at each time								
ALL	**2607**	(2253;2971)	**2701**	(2278;3145)	0.494	(0.463;0.527)	0.6	(0.566;0.635)
IMP-H	**2453**	(2087;2843)	**2273**	(1874;2664)	0.494	(0.462;0.526)	0.6	(0.565;0.633)
Strategy 3 - HRU category & utility at each time								
ALL	**2687**	(2173;3194)	**2587**	(1995;3206)	0.513	(0.475;0.55)	0.599	(0.565;0.634)

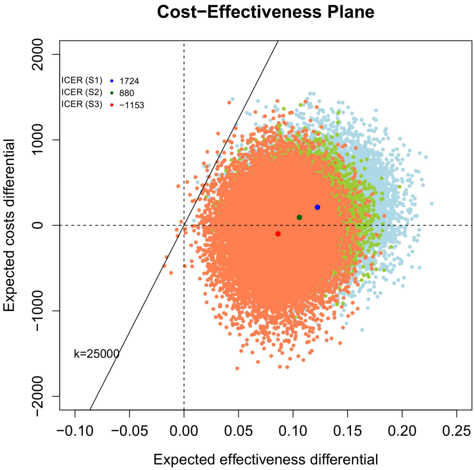
Posterior mean incremental TC esitmates

- **Key results**:
  - **S1**/**S2** Δ_{tc} based on zero-imputed data are systematically lower
  - **S3** Δ_{tc} located between **S1**/**S2** alternative missingness strategies



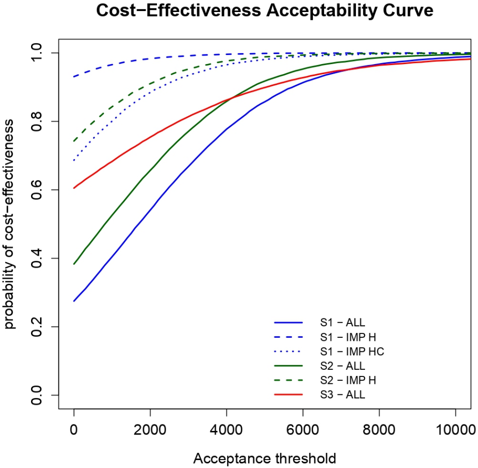
Economic evaluation

- Assess cost-effectiveness based on standard tools
  - **CE plane**: representation of the joint distribution of (Δ_e, Δ_{tc})



Economic evaluation

- Assess cost-effectiveness based on standard tools
  - **CE acceptability curve**: proportion of dots lying below the straight line (*sustainability area*) in CE plane upon varying the threshold k



Part V

Discussion

Summary

- CEAs often conducted on quantities derived from *questionnaire data* affected by **missingness**
  - Lack of a **"gold standard"** to handle missing at item level (e.g. HRUs)
  - Current practice often relies on *ad-hoc methods*(e.g. **zero-imputing**)
  - These can **distort** the data and lead to *incorrect* inferences

Summary

- CEAs often conducted on quantities derived from *questionnaire data* affected by **missingness**
  - Lack of a **"gold standard"** to handle missing at item level (e.g. HRUs)
  - Current practice often relies on *ad-hoc methods*(e.g. **zero-imputing**)
  - These can **distort** the data and lead to *incorrect* inferences
- Propose Bayesian framework for CEA data to handle *item nonresponse* in HRUs
  - Allow to fit models at **different levels** of data-aggregation
  - Can handle different types of **data features** (correlation, skewness, structural zeros)
  - Directly quantify impact of **missingness uncertainty** on results
  - Implemented using *freely-available* software (e.g. JAGS)

Application

- Application to PBS data shows that:
  - Model estimates and CE results are **affected** by the approach used
  - Aggregated models require **implicit assumptions**(e.g. *zero*) about disaggregated missingness
  - Possibly lead to a substantial **loss of information** and **distort** the results

Application

- Application to PBS data shows that:
  - Model estimates and CE results are **affected** by the approach used
  - Aggregated models require **implicit assumptions**(e.g. *zero*) about disaggregated missingness
  - Possibly lead to a substantial **loss of information** and **distort** the results
- **Advantages**:
  - Quantify impact of missingness and parameter uncertainty **coherently** with the analysis model
  - **Flexible** model specification at *different aggregation levels*

Application

- Application to PBS data shows that:
  - Model estimates and CE results are **affected** by the approach used
  - Aggregated models require **implicit assumptions**(e.g. *zero*) about disaggregated missingness
  - Possibly lead to a substantial **loss of information** and **distort** the results
- **Advantages**:
  - Quantify impact of missingness and parameter uncertainty **coherently** with the analysis model
  - **Flexible** model specification at *different aggregation levels*
- **Limitations**:
  - Same assumption for cases with different **missingness patterns**
  - **Challenging** to fit as number of variables and parameters increases

Recommendations & Conclusions

- **Design**: **minimise** item missingness (e.g. shorter follow-ups or fewer questionnaire items)
- **Data processing**: **ad-hoc** imputation should be **avoided** unless clearly motivated
- **Analysis**: account for *data features*, including **missingness**:
  - Fully-missing HRQL/HRU & partially-missing utilities/costs $\rightarrow$ **Intermediate model**
  - Partially-missing HRU & costs $\rightarrow$ **Disaggregated model**

Recommendations & Conclusions

- **Design**: **minimise** item missingness (e.g. shorter follow-ups or fewer questionnaire items)
- **Data processing**: **ad-hoc** imputation should be **avoided** unless clearly motivated
- **Analysis**: account for *data features*, including **missingness**:
  - Fully-missing HRQL/HRU & partially-missing utilities/costs $\rightarrow$ **Intermediate model**
  - Partially-missing HRU & costs $\rightarrow$ **Disaggregated model**
- Always need to find a **balance** between **ideal approach** and its **feasibility** based on the available data:
  - **Simplify** model (i.e. dependence among variables) to reduce the number of parameters
  - Use *more* **informative priors** to handle sparse data
  - **Aggregate** some types of HRUs to reduce number of variables

References

Faria, R., Gomes, M., Epstein, D., & White, I. R. (2014). A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. *Pharmacoeconomics*, *32*(12), 1157–1170.

Gabrio, A., Daniels, M. J., & Baio, G. (2020). A bayesian parametric approach to handle missing longitudinal outcome data in trial-based health economic evaluations. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *183*(2), 607–629.

Hassiotis, A., Poppe, M., Strydom, A., Vickerstaff, V., Hall, I., Crabtree, J., Omar, R., King, M., Hunter, R., Bosco, A., Biswas, A., Ratti, V., Blickwedel, J., Cooper, V., Howie, W., & Crawford, M. (2018). Positive behaviour support training for staff for treating challenging behaviour in people with intellectual disabilities: A cluster RCT. *Health Technology Assessment*, *22*.

Ling, X., Gabrio, A., Mason, A., & Baio, G. (2022). A scoping review of item-level missing data in within-trial cost-effectiveness analysis. *Value in Health*.

O'Hagan, A., Stevens, J. W., & Montmartin, J. (2001). Bayesian cost-effectiveness analysis from clinical trial data. *Statistics in Medicine*, *20*(5), 733–753.

Plummer, M. et al. (2003). JAGS: A program for analysis of bayesian graphical models using gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, *124*, 1–10.

Spiegelhalter, D. J., Abrams, K. R., & Myles, J. P. (2004). *Bayesian approaches to clinical trials and health-care evaluation*. John Wiley & Sons.