Μπευζιανης μεθοδων στην οικονομια της υγειας

(thanks/blame mostly to Google Translate)

Gianluca Baio

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

Research Seminars
Department of Statistics
Athens University of Economics and Business, Athens

Thursday 3 May 2018

Outline

1. **Health economic evaluation**
   - What is it?
   - How does it work?
2. Statistical modelling
   - Individual-level vs aggregated data
   - The importance of being a Bayesian
3. Some examples — you get to choose...
   - Individual level & partially observed data
   - Survival analysis in HTA
   - Value of information
4. Conclusions

Outline

1. Health economic evaluation
   - What is it?
   - How does it work?
2. **Statistical modelling**
   - Individual-level vs aggregated data
   - The importance of being a Bayesian
3. Some examples — you get to choose...
   - Individual level & partially observed data
   - Survival analysis in HTA
   - Value of information
4. Conclusions

Outline

1. Health economic evaluation
   - What is it?
   - How does it work?
2. Statistical modelling
   - Individual-level vs aggregated data
   - The importance of being a Bayesian
3. **Some examples — you get to choose...**
   - Individual level & partially observed data
   - Survival analysis in HTA
   - Value of information
4. Conclusions

Outline

1. Health economic evaluation
   - What is it?
   - How does it work?
2. Statistical modelling
   - Individual-level vs aggregated data
   - The importance of being a Bayesian
3. Some examples — you get to choose...
   - Individual level & partially observed data
   - Survival analysis in HTA
   - Value of information
4. **Conclusions**

Health technology assessment (HTA)

Objective: Combine costs & benefits of a given intervention into a rational scheme for allocating resources

Health technology assessment (HTA)

Objective: Combine costs & benefits of a given intervention into a rational scheme for allocating resources

Statistical model

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

Health technology assessment (HTA)

Objective: Combine costs & benefits of a given intervention into a rational scheme for allocating resources

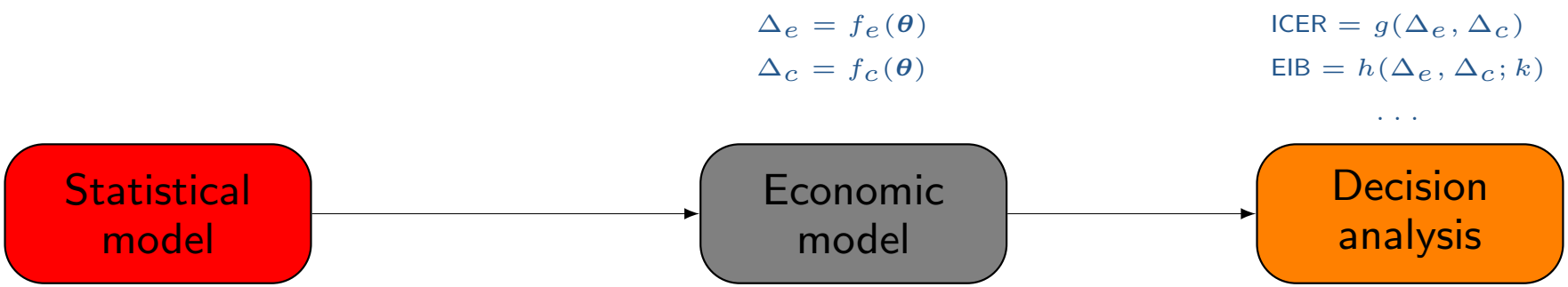
$$\Delta_e = f_e(\boldsymbol{\theta})$$
$$\Delta_c = f_c(\boldsymbol{\theta})$$

- 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 technology assessment (HTA)

Objective: Combine costs & benefits of a given intervention into a rational scheme for allocating resources



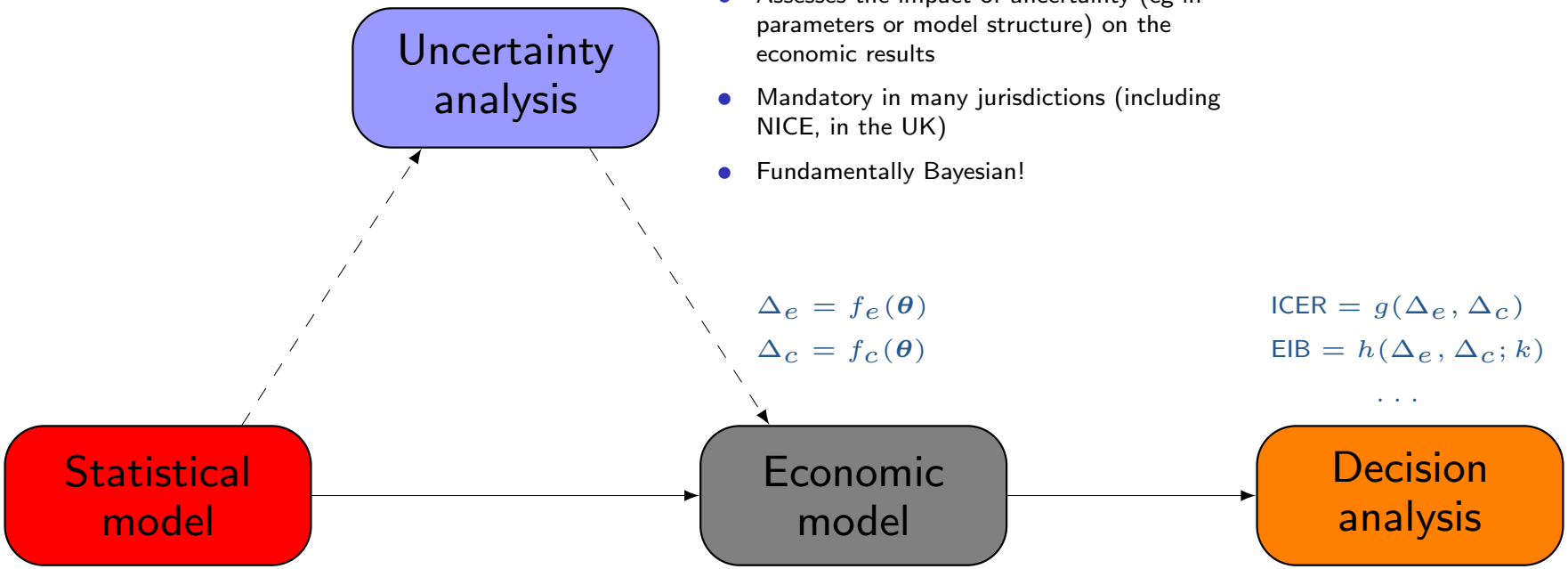
- Estimates relevant **population** parameters $\boldsymbol{\theta}$
- Varies with the type of available data (& statistical approach!)

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

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

Health technology assessment (HTA)

Objective: Combine costs & benefits of a given intervention into a rational scheme for allocating resources



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

- 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

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

1. ("Standard") Statistical modelling — **Individual level** data

		Demographics			HRQL data				Resource use 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} = Use of resources (drugs, hospital, GP appointments, ...)

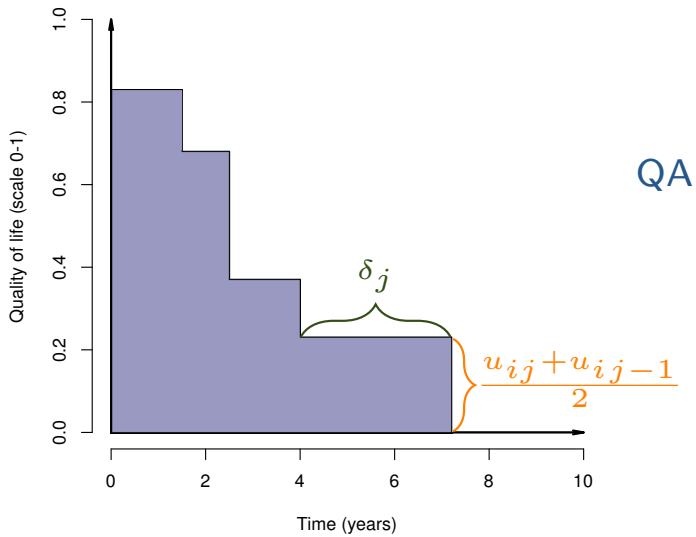
1. ("Standard") Statistical modelling — **Individual level** data

		Demographics			HRQL data				Resource use 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} = Use of resources (drugs, hospital, GP appointments, ...)

1. Compute individual QALYs and total costs as

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



QALY_i = "Area under the curve"

1. ("Standard") Statistical modelling — **Individual level** data

		Demographics			HRQL data				Resource use 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} = Use of resources (drugs, hospital, GP appointments, ...)

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

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

1. ("Standard") Statistical modelling — **Individual level** data

		Demographics			HRQL data				Resource use 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} = Use of resources (drugs, hospital, GP appointments, ...)

1. Compute individual QALYs and total costs as

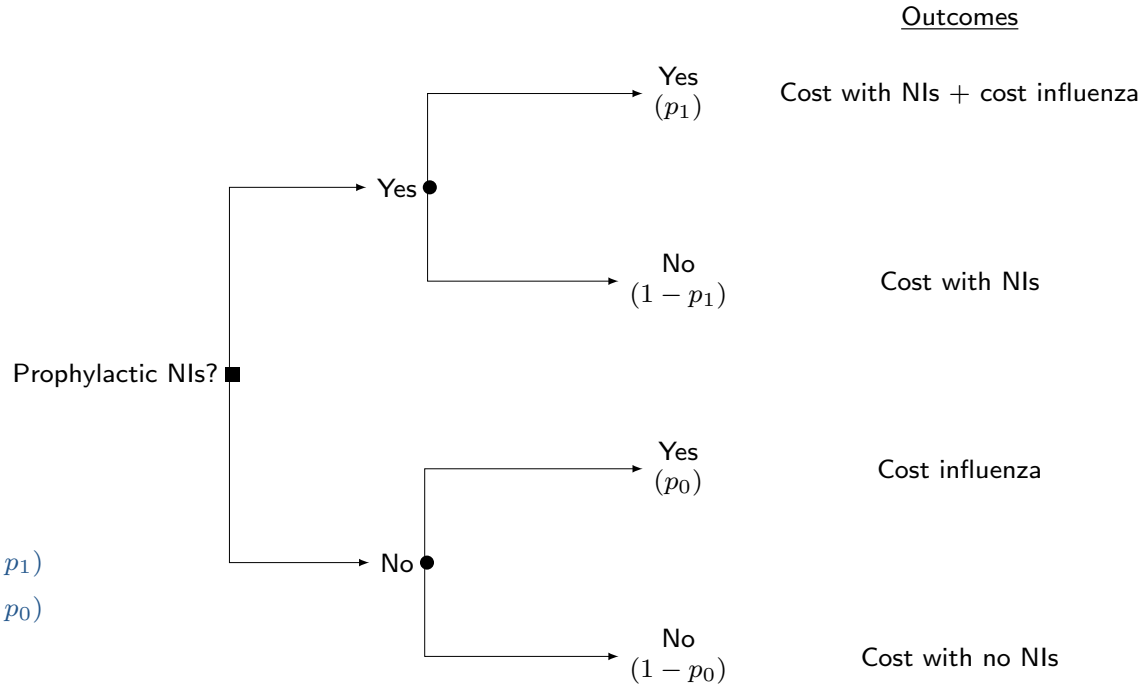
$$e_i = \sum_{j=1}^{J} (u_{ij} + u_{ij-1}) \frac{\delta_j}{2} \quad \text{and} \quad c_i = \sum_{j=0}^{J} c_{ij}, \qquad \left[\text{with: } \delta_j = \frac{\text{Time}_j - \text{Time}_{j-1}}{\text{Unit of time}}\right]$$

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

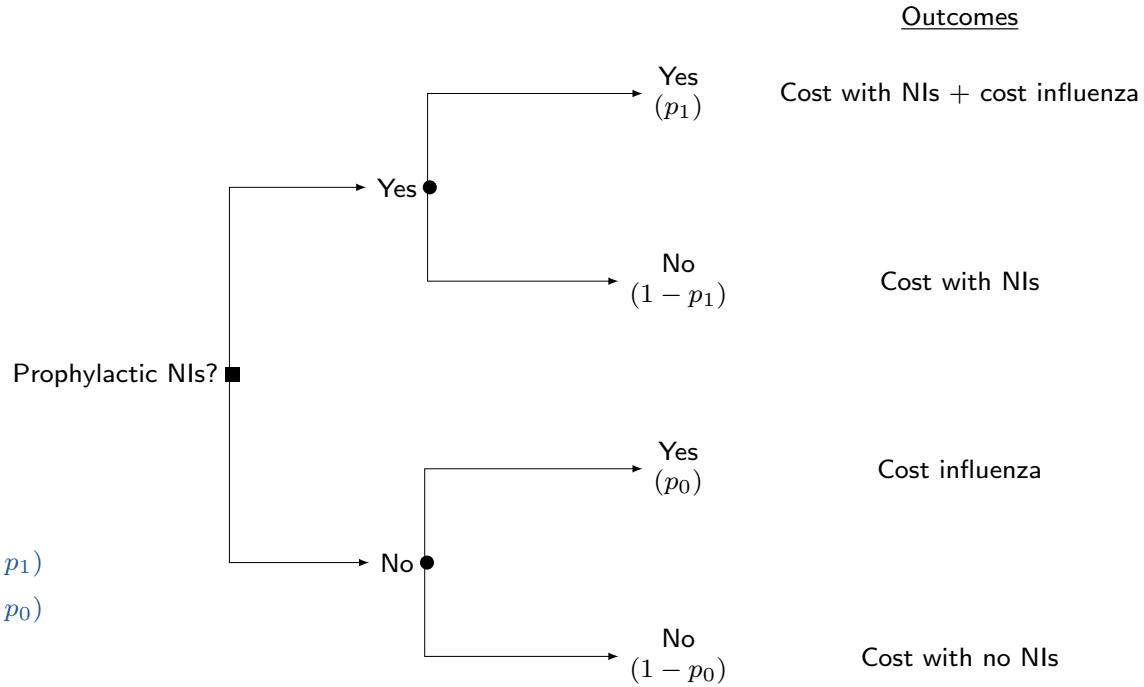
1. Build a **population level model** (eg decision tree/Markov model)



$$\mu_{e1} = -lp_1$$
$$\mu_{e0} = -lp_0$$
$$\mu_{c1} = \left(c^{NI} + c^{Inf}\right) p_1 + c^{NI}(1-p_1)$$
$$\mu_{c0} = \left(c^{NI} + c^{Inf}\right) p_0 + c^{NI}(1-p_0)$$

NB: in this case, the "data" are typically represented by summary statistics for the parameters of interest $\boldsymbol{\theta} = (p_0, p_1, l, \ldots)$, but may also have access to a combination of ILD and summaries

1. Build a **population level model** (eg decision tree/Markov model)

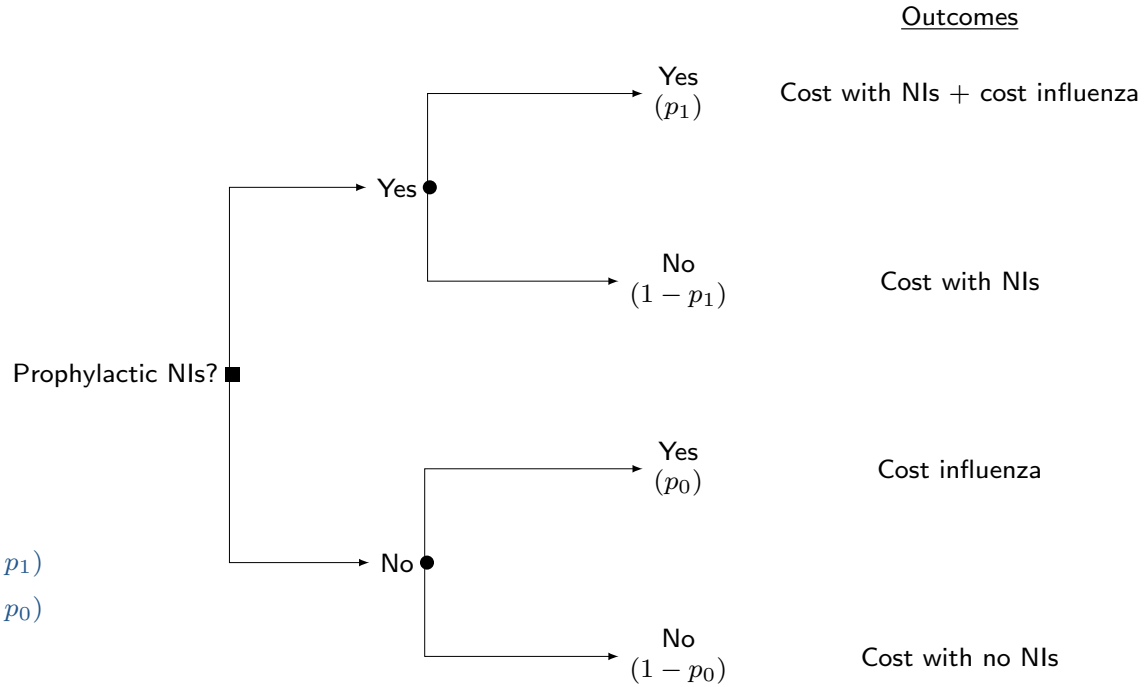


$$\mu_{e1} = -lp_1$$
$$\mu_{e0} = -lp_0$$
$$\mu_{c1} = \left(c^{NI} + c^{Inf}\right) p_1 + c^{NI}(1-p_1)$$
$$\mu_{c0} = \left(c^{NI} + c^{Inf}\right) p_0 + c^{NI}(1-p_0)$$

NB: in this case, the "data" are typically represented by summary statistics for the parameters of interest $\boldsymbol{\theta} = (p_0, p_1, l, \ldots)$, but may also have access to a combination of ILD and summaries

2. Use point estimates for the parameters to build the "base-case" (average) evaluation

1. Build a **population level model** (eg decision tree/Markov model)

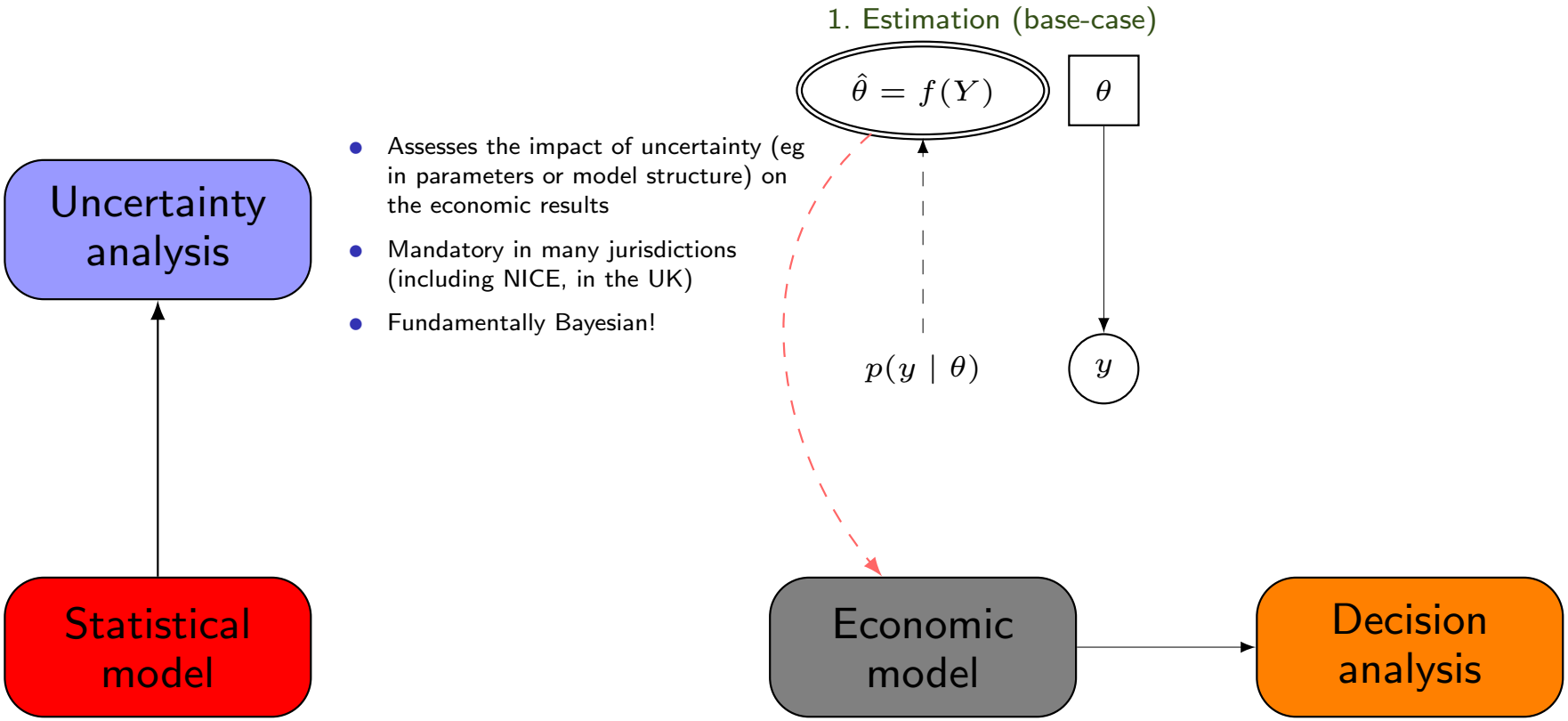


$$\mu_{e1} = -lp_1$$
$$\mu_{e0} = -lp_0$$
$$\mu_{c1} = \left(c^{NI} + c^{Inf}\right) p_1 + c^{NI}(1-p_1)$$
$$\mu_{c0} = \left(c^{NI} + c^{Inf}\right) p_0 + c^{NI}(1-p_0)$$

NB: in this case, the "data" are typically represented by summary statistics for the parameters of interest $\boldsymbol{\theta} = (p_0, p_1, l, \ldots)$, but may also have access to a combination of ILD and summaries

2. Use point estimates for the parameters to build the "base-case" (average) evaluation
3. Use resampling methods (eg bootstrap) to propage uncertainty in the point estimates and perform uncertainty analysis

"Standard" approach to HTA — "Two-stage"



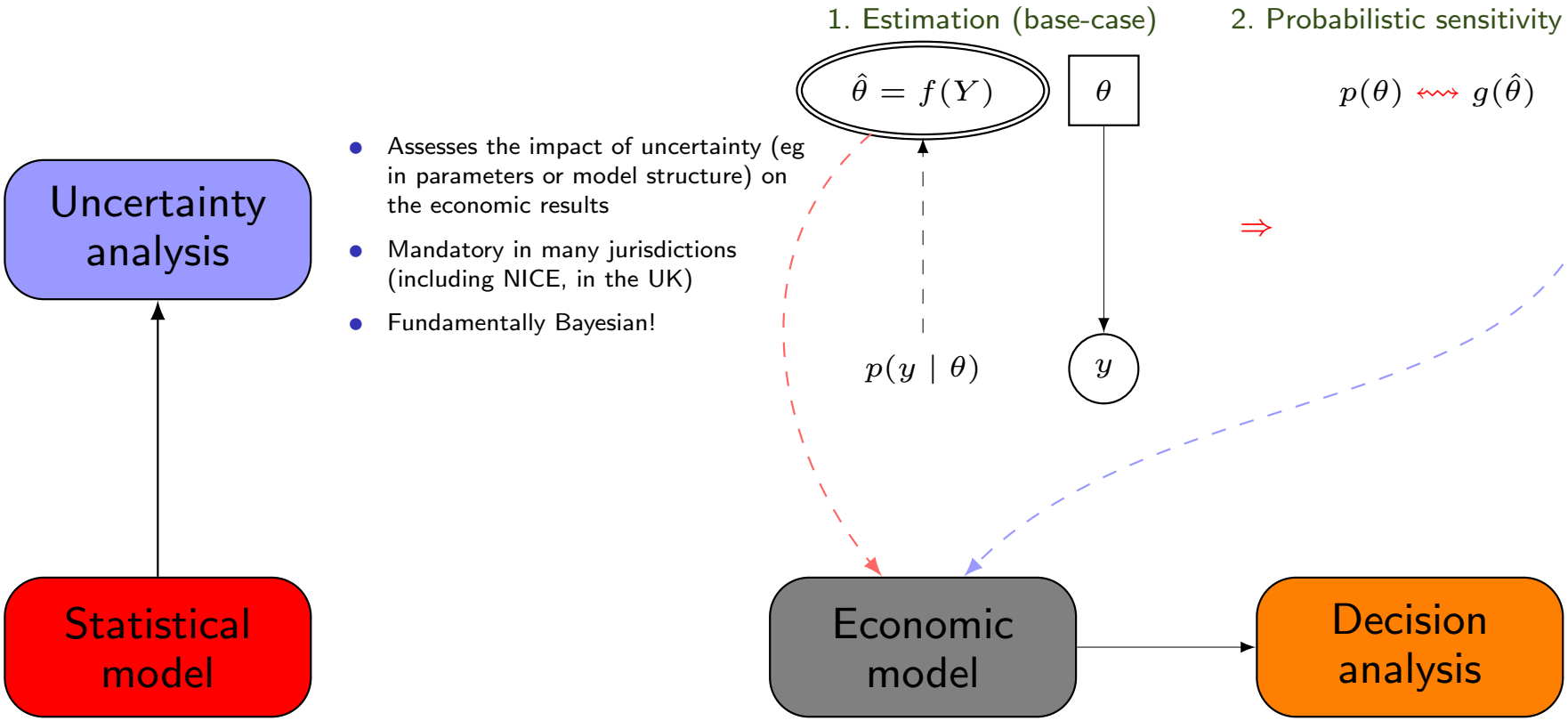
- Assesses the impact of uncertainty (eg in parameters or model structure) on the economic results
- Mandatory in many jurisdictions (including NICE, in the UK)
- Fundamentally Bayesian!

- 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

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

"Standard" approach to HTA — "Two-stage"



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

- 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

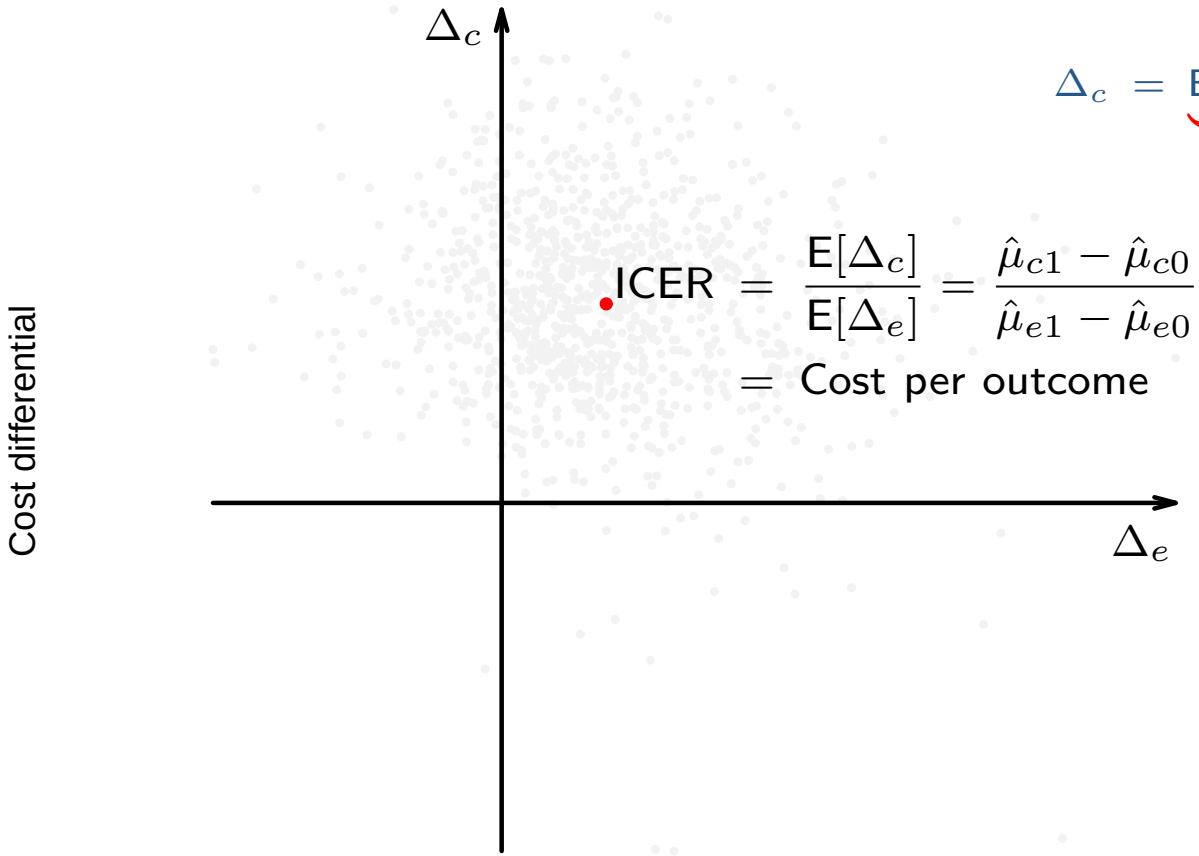
- Summarises the economic model by computing suitable measures of "cost-effectiveness"
- Dictates the best course of actions, given current evidence
- Standardised process

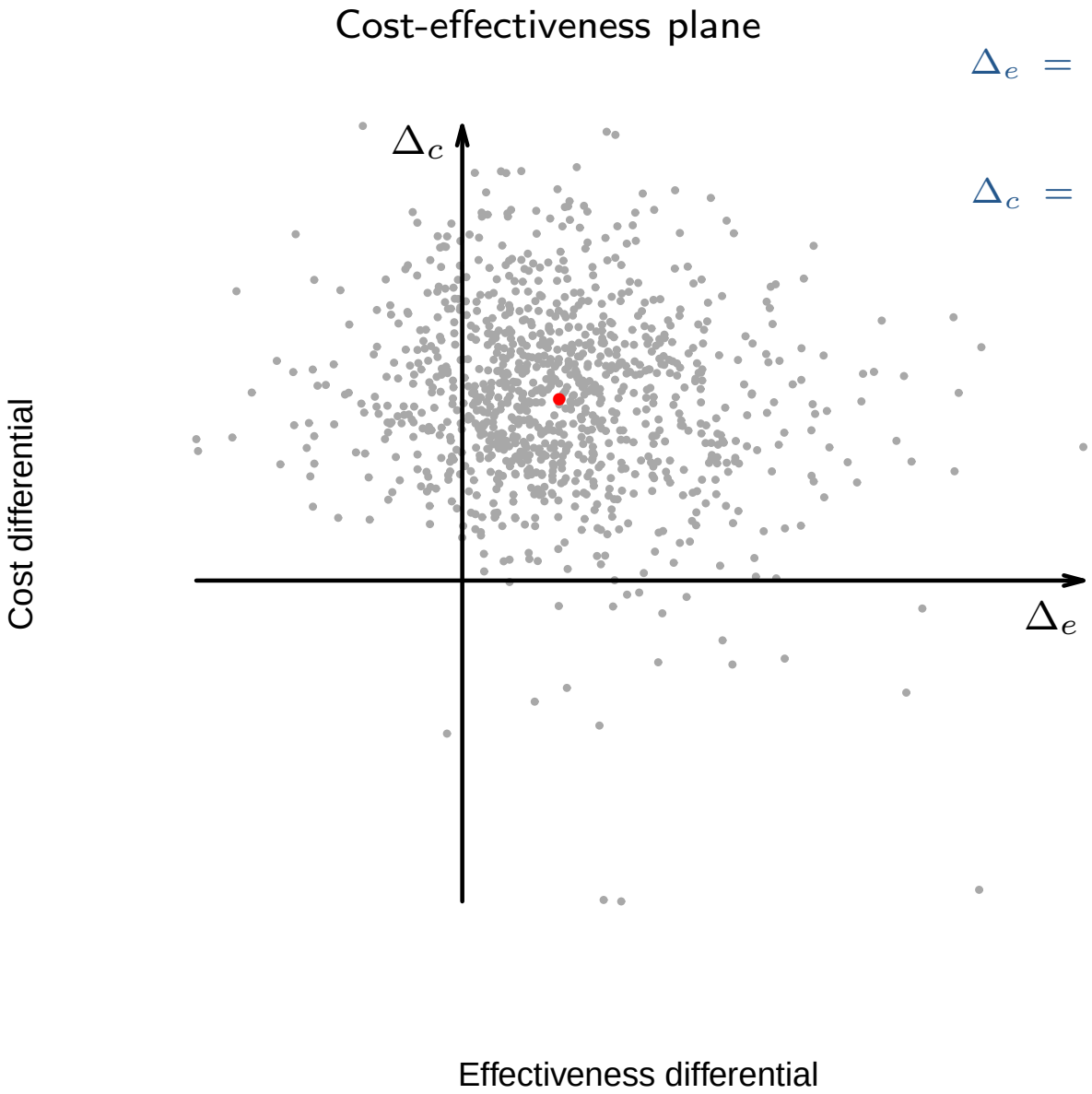
"Two-stage approach" (Spiegelhalter, Abrams & Myles, 2004)

Cost-effectiveness plane

$$\Delta_e = \underbrace{\mathsf{E}[e \mid \hat{\boldsymbol{\theta}}_1]}_{\hat{\mu}_{e1}} - \underbrace{\mathsf{E}[e \mid \hat{\boldsymbol{\theta}}_0]}_{\hat{\mu}_{e0}}$$

$$\Delta_c = \underbrace{\mathsf{E}[c \mid \hat{\boldsymbol{\theta}}_1]}_{\hat{\mu}_{c1}} - \underbrace{\mathsf{E}[c \mid \hat{\boldsymbol{\theta}}_0]}_{\hat{\mu}_{c0}}$$

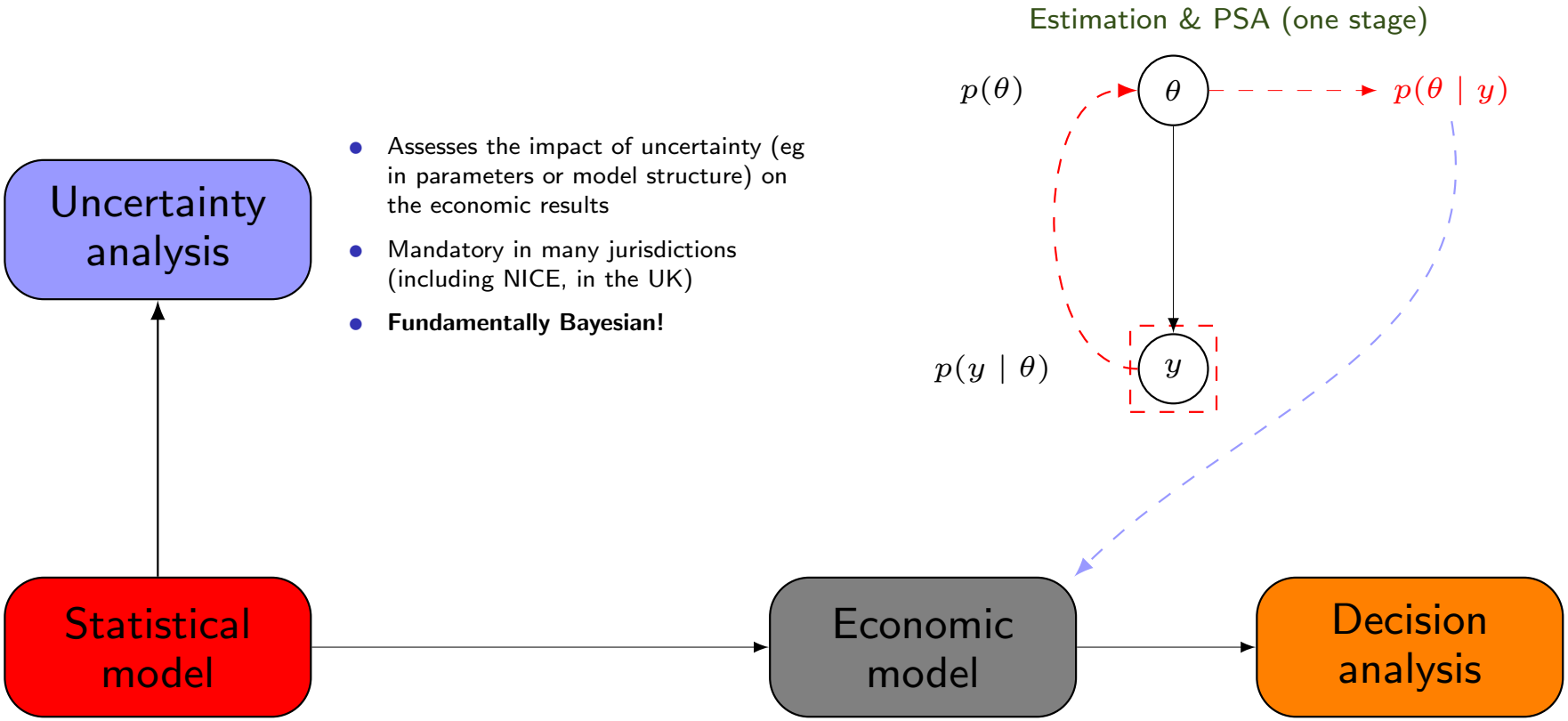




$$\Delta_e = \underbrace{\mathrm{E}[e \mid \boldsymbol{\theta}_1]}_{\mu_{e1}} - \underbrace{\mathrm{E}[e \mid \boldsymbol{\theta}_0]}_{\mu_{e0}}$$

$$\Delta_c = \underbrace{\mathrm{E}[c \mid \boldsymbol{\theta}_1]}_{\mu_{c1}} - \underbrace{\mathrm{E}[c \mid \boldsymbol{\theta}_0]}_{\mu_{c0}}$$

Bayesian approach to HTA



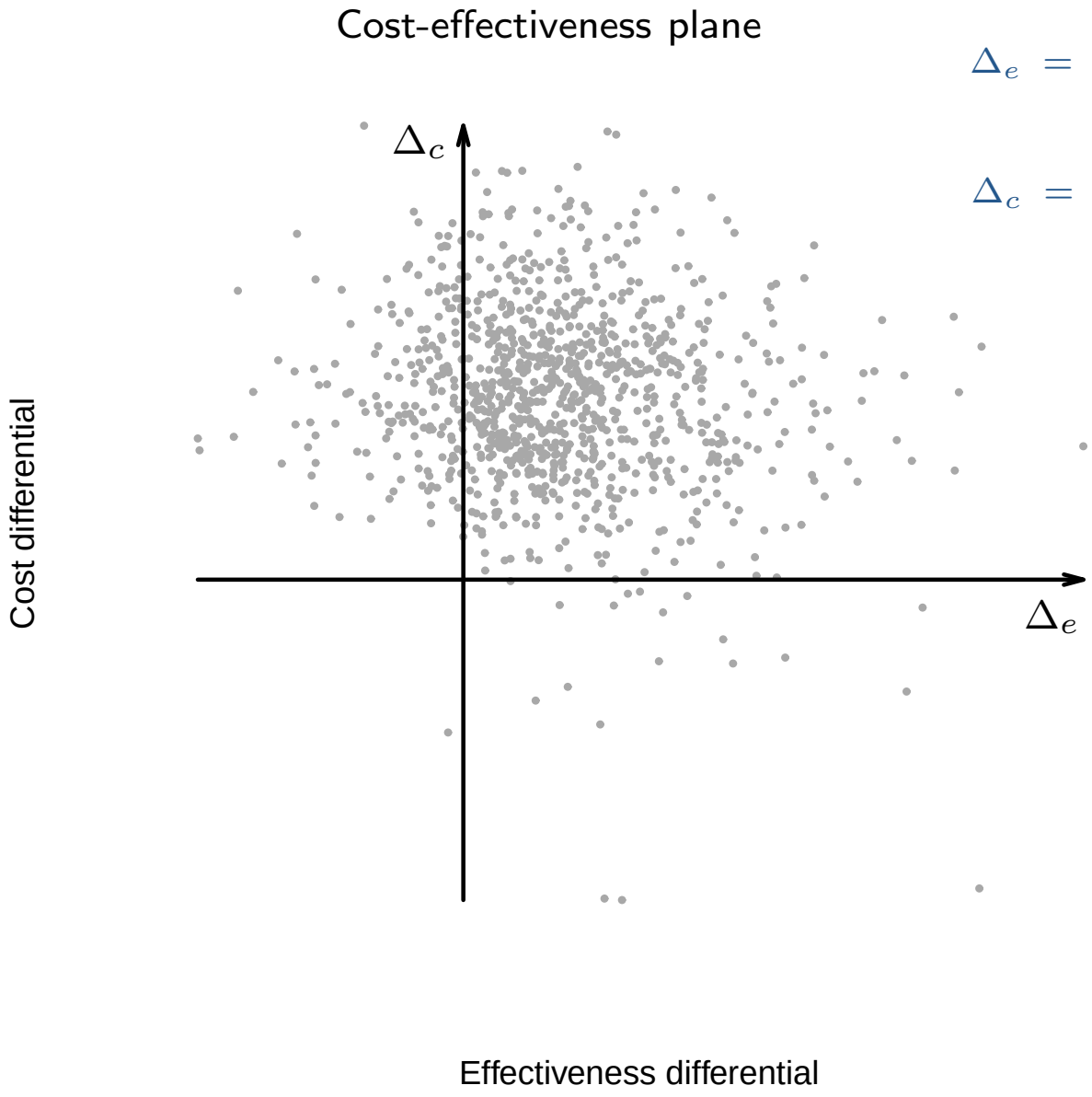
- Assesses the impact of uncertainty (eg in parameters or model structure) on the economic results
- Mandatory in many jurisdictions (including NICE, in the UK)
- **Fundamentally Bayesian!**

- 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

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

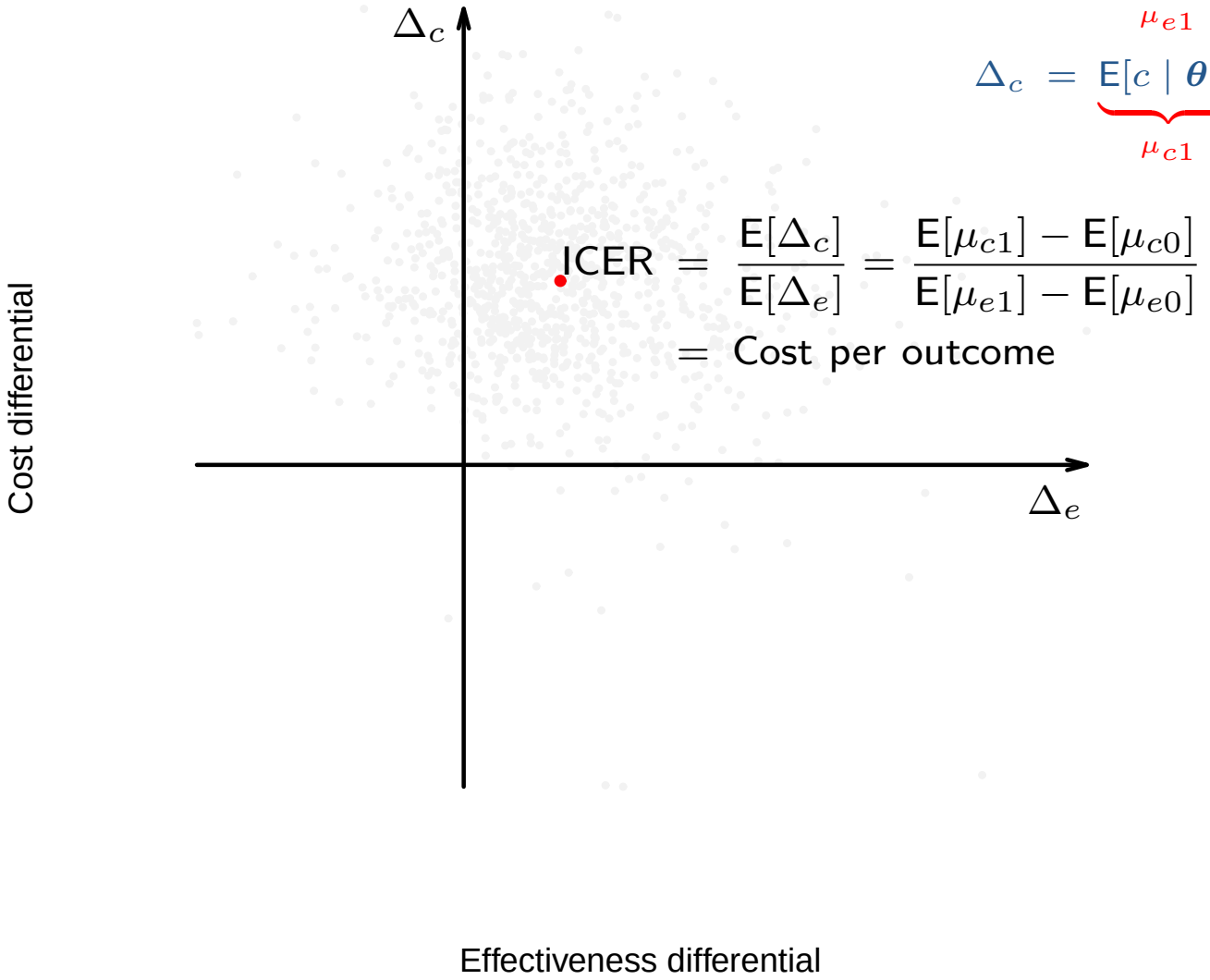
"Integrated approach" Spiegelhalter, Abrams & Myles (2004)
Baio, Berardi & Heath (2017)


Cost-effectiveness plane
$\Delta_e = \underbrace{\mathrm{E}[e \mid \boldsymbol{\theta}_1]}_{\mu_{e1}} - \underbrace{\mathrm{E}[e \mid \boldsymbol{\theta}_0]}_{\mu_{e0}}$
$\Delta_c = \underbrace{\mathrm{E}[c \mid \boldsymbol{\theta}_1]}_{\mu_{c1}} - \underbrace{\mathrm{E}[c \mid \boldsymbol{\theta}_0]}_{\mu_{c0}}$
Δ_c
Δ_e
Cost differential
Effectiveness differential

Cost-effectiveness plane

$$\Delta_e = \underbrace{\mathsf{E}[e \mid \boldsymbol{\theta}_1]}_{\mu_{e1}} - \underbrace{\mathsf{E}[e \mid \boldsymbol{\theta}_0]}_{\mu_{e0}}$$

$$\Delta_c = \underbrace{\mathsf{E}[c \mid \boldsymbol{\theta}_1]}_{\mu_{c1}} - \underbrace{\mathsf{E}[c \mid \boldsymbol{\theta}_0]}_{\mu_{c0}}$$



Bayesians do it better...

- Potential correlation between costs & clinical benefits [Both ILD and ALD]
  - Strong positive correlation — effective treatments are innovative and result from intensive and lengthy research $\Rightarrow$ 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.
  - Because of the way in which standard models are set up, bootstrapping generally only approximates the underlying level of correlation — **MCMC does a better job!**

Bayesians do it better...

- Potential correlation between costs & clinical benefits [Both ILD and ALD]
  - Strong positive correlation — effective treatments are innovative and result from intensive and lengthy research $\Rightarrow$ 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.
  - Because of the way in which standard models are set up, bootstrapping generally only approximates the underlying level of correlation — **MCMC does a better job!**

- Joint/marginal normality not realistic [Mainly ILD]
  - 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
  - Should use more suitable models (e.g. Beta, Gamma or log-Normal) — **generally easier under a Bayesian framework**

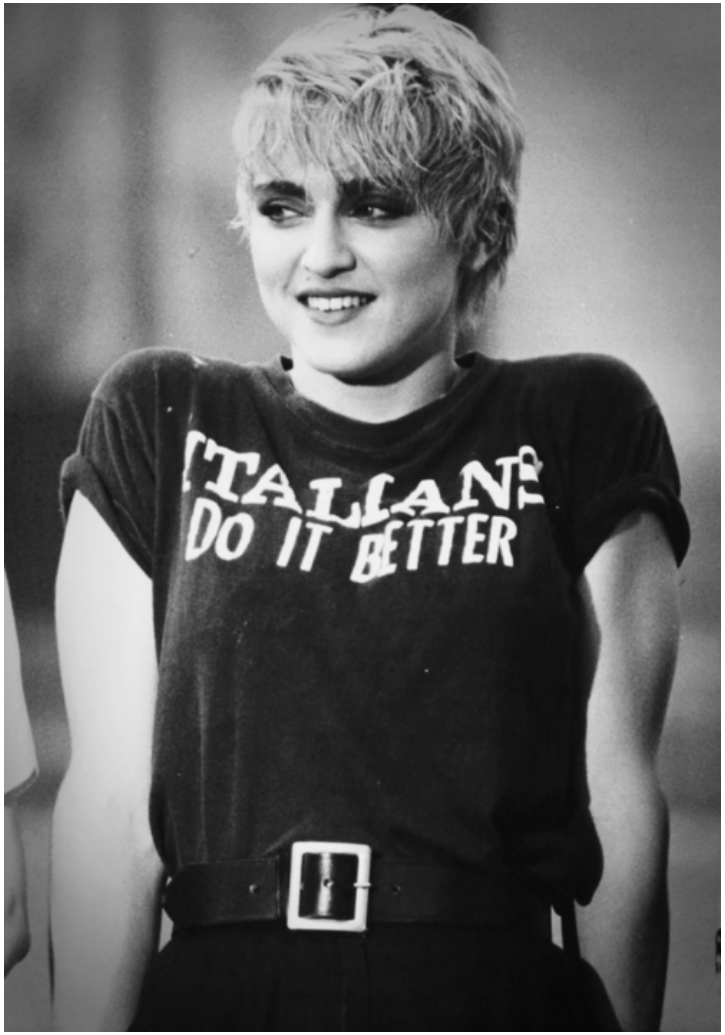
Bayesians do it better...

- Potential correlation between costs & clinical benefits [Both ILD and ALD]
  - Strong positive correlation — effective treatments are innovative and result from intensive and lengthy research $\Rightarrow$ 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.
  - Because of the way in which standard models are set up, bootstrapping generally only approximates the underlying level of correlation — **MCMC does a better job!**

- Joint/marginal normality not realistic [Mainly ILD]
  - 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
  - Should use more suitable models (e.g. Beta, Gamma or log-Normal) — **generally easier under a Bayesian framework**

- Particularly as the focus is on decision-making (rather than just inference), we need to use **all** available evidence to fully characterise current uncertainty on the model parameters and outcomes [Mainly ALD]
  - A Bayesian approach is helpful in combining different sources of information
  - **Propagating uncertainty is a fundamentally Bayesian operation!**

Bayesians do it better...



Outline

1. Health economic evaluation
   - What is it?
   - How does it work?
2. Statistical modelling
   - Individual-level vs aggregated data
   - The importance of being a Bayesian
3. **Some examples**
   - Individual level & partially observed data
   - Survival analysis in HTA
   - Value of information
4. Conclusions

ILD+Missing data

Survival analysis

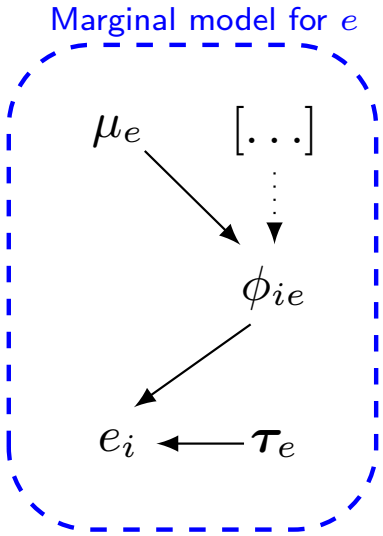
Value of information

Bayesian HTA in action — **Individual level** (& missing) data

- In general, can represent the joint distribution as $p(e, c) = p(e)p(c \mid e) = p(c)p(e \mid c)$

Bayesian HTA in action — **Individual level** (& missing) data

- In general, can represent the joint distribution as $p(e, c) = p(e)p(c \mid e) = p(c)p(e \mid c)$



$$e_i \sim p(e \mid \phi_{ei}, \boldsymbol{\tau}_e)$$
$$g_e(\phi_{ei}) = \alpha_0 \, [+ \ldots]$$
$$\mu_e = g_e^{-1}(\alpha_0)$$

ϕ_{ei} = location
$\boldsymbol{\tau}_e$ = ancillary

Bayesian HTA in action — **Individual level** (& missing) data

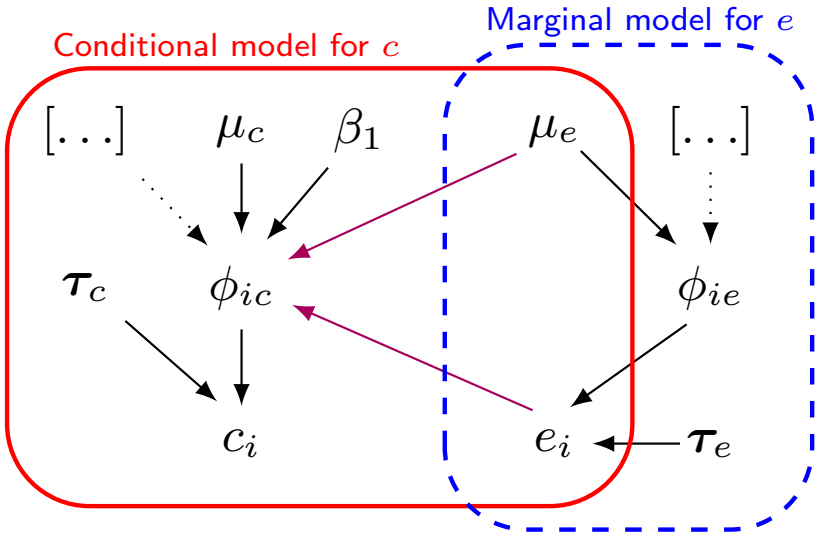
- In general, can represent the joint distribution as $p(e, c) = p(e)p(c \mid e) = p(c)p(e \mid c)$

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

$g_c(\phi_{ci}) = \beta_0 + \beta_1(e_i - \mu_e)\,[+\ldots]$

$\mu_c = g_c^{-1}(\beta_0)$

ϕ_{ci} = location
$\boldsymbol{\tau}_c$ = ancillary



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

$g_e(\phi_{ei}) = \alpha_0\,[+\ldots]$

$\mu_e = g_e^{-1}(\alpha_0)$

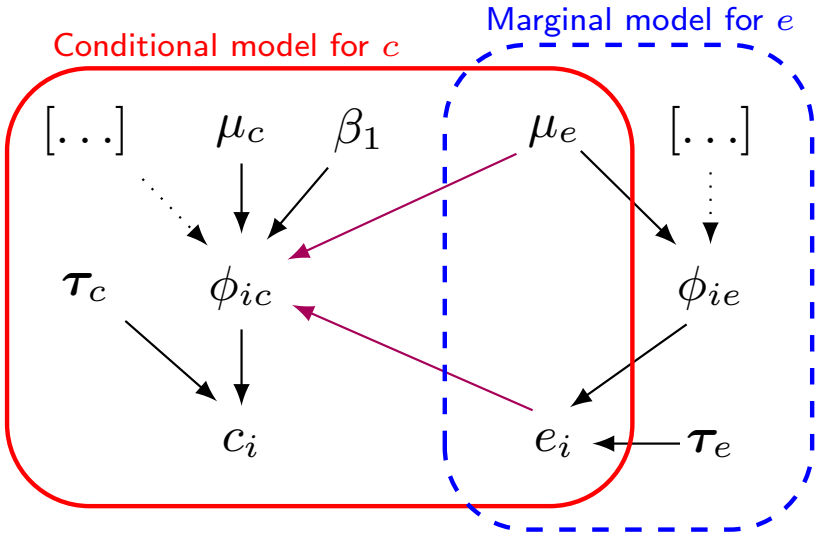
ϕ_{ei} = location
$\boldsymbol{\tau}_e$ = ancillary

Bayesian HTA in action — **Individual level** (& missing) data

- In general, can represent the joint distribution as $p(e, c) = p(e)p(c \mid e) = p(c)p(e \mid c)$

$c_i \sim p(c \mid e, \phi_{ci}, \boldsymbol{\tau}_c)$
$g_c(\phi_{ci}) = \beta_0 + \beta_1(e_i - \mu_e)\,[+\ldots]$
$\mu_c = g_c^{-1}(\beta_0)$

ϕ_{ci} = conditional mean
τ_c = conditional variance



$e_i \sim p(e \mid \phi_{ei}, \boldsymbol{\tau}_e)$
$g_e(\phi_{ei}) = \alpha_0\,[+\ldots]$
$\mu_e = g_e^{-1}(\alpha_0)$

ϕ_{ei} = marginal mean
τ_e = marginal variance

- For example:

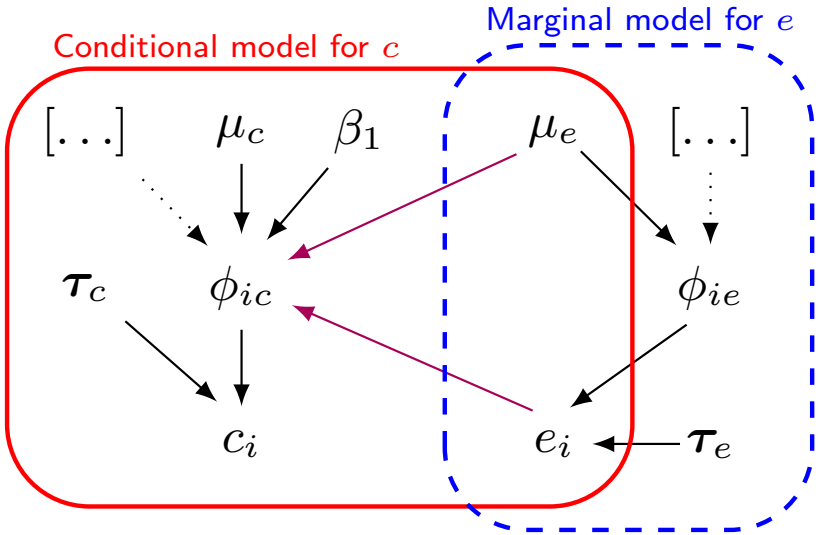
$$e_i \sim \text{Normal}(\phi_{ei}, \tau_e), \qquad \phi_{ei} = \alpha_0\,[+\ldots], \qquad \mu_e = \alpha_0$$
$$c_i \mid e_i \sim \text{Normal}(\phi_{ci}, \tau_c), \qquad \phi_{ci} = \beta_0 + \beta_1(e_i - \mu_e)\,[+\ldots], \qquad \mu_c = \beta_0$$

Bayesian HTA in action — **Individual level** (& missing) data

- In general, can represent the joint distribution as $p(e, c) = p(e)p(c \mid e) = p(c)p(e \mid c)$

$c_i \sim p(c \mid e, \phi_{ci}, \boldsymbol{\tau}_c)$
$g_c(\phi_{ci}) = \beta_0 + \beta_1(e_i - \mu_e)\,[+\ldots]$
$\mu_c = g_c^{-1}(\beta_0)$

ϕ_{ci} = conditional mean
τ_c = shape
τ_c/ϕ_{ci} = rate



$e_i \sim p(e \mid \phi_{ei}, \boldsymbol{\tau}_e)$
$g_e(\phi_{ei}) = \alpha_0\,[+\ldots]$
$\mu_e = g_e^{-1}(\alpha_0)$

ϕ_{ei} = marginal mean
τ_e = marginal scale

- For example:

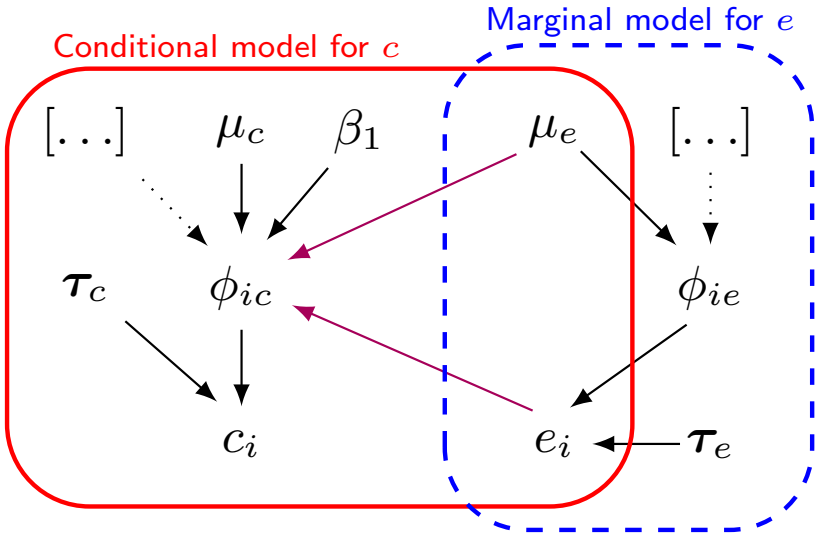
$e_i \sim \text{Beta}(\phi_{ei}\tau_e, (1-\phi_{ei})\tau_e)$, $\quad \text{logit}(\phi_{ei}) = \alpha_0\,[+\ldots]$, $\quad \mu_e = \frac{\exp(\alpha_0)}{1+\exp(\alpha_0)}$
$c_i \mid e_i \sim \text{Gamma}(\tau_c, \tau_c/\phi_{ci})$, $\quad \log(\phi_{ci}) = \beta_0 + \beta_1(e_i - \mu_e)\,[+\ldots]$, $\quad \mu_c = \exp(\beta_0)$

Bayesian HTA in action — **Individual level** (& missing) data

- In general, can represent the joint distribution as $p(e, c) = p(e)p(c \mid e) = p(c)p(e \mid c)$

$c_i \sim p(c \mid e, \phi_{ci}, \tau_c)$
$g_c(\phi_{ci}) = \beta_0 + \beta_1(e_i - \mu_e)\,[+\ldots]$
$\mu_c = g_c^{-1}(\beta_0)$

ϕ_{ci} = conditional mean
τ_c = shape
τ_c/ϕ_{ci} = rate



$e_i \sim p(e \mid \phi_{ei}, \tau_e)$
$g_e(\phi_{ei}) = \alpha_0\,[+\ldots]$
$\mu_e = g_e^{-1}(\alpha_0)$

ϕ_{ei} = marginal mean
τ_e = marginal scale

- For example:

$e_i \sim \text{Beta}(\phi_{ei}\tau_e, (1-\phi_{ei})\tau_e)$, $\quad \text{logit}(\phi_{ei}) = \alpha_0\,[+\ldots]$, $\quad \mu_e = \frac{\exp(\alpha_0)}{1+\exp(\alpha_0)}$
$c_i \mid e_i \sim \text{Gamma}(\tau_c, \tau_c/\phi_{ci})$, $\quad \log(\phi_{ci}) = \beta_0 + \beta_1(e_i - \mu_e)\,[+\ldots]$, $\quad \mu_c = \exp(\beta_0)$

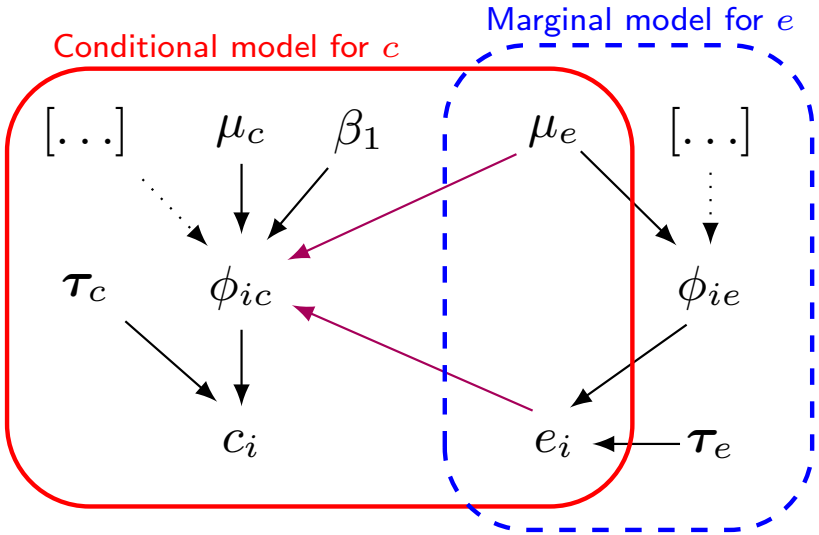
- Combining "modules" and fully characterising uncertainty about deterministic functions of random quantities is relatively straightforward using MCMC

Bayesian HTA in action — **Individual level** (& missing) data

- In general, can represent the joint distribution as $p(e, c) = p(e)p(c \mid e) = p(c)p(e \mid c)$

$c_i \sim p(c \mid e, \phi_{ci}, \tau_c)$
$g_c(\phi_{ci}) = \beta_0 + \beta_1(e_i - \mu_e)\,[+\ldots]$
$\mu_c = g_c^{-1}(\beta_0)$

ϕ_{ci} = conditional mean
τ_c = shape
τ_c/ϕ_{ci} = rate



$e_i \sim p(e \mid \phi_{ei}, \tau_e)$
$g_e(\phi_{ei}) = \alpha_0\,[+\ldots]$
$\mu_e = g_e^{-1}(\alpha_0)$

ϕ_{ei} = marginal mean
τ_e = marginal scale

- For example:

$$e_i \sim \text{Beta}(\phi_{ei}\tau_e, (1-\phi_{ei})\tau_e), \quad \text{logit}(\phi_{ei}) = \alpha_0\,[+\ldots], \quad \mu_e = \frac{\exp(\alpha_0)}{1+\exp(\alpha_0)}$$
$$c_i \mid e_i \sim \text{Gamma}(\tau_c, \tau_c/\phi_{ci}), \quad \log(\phi_{ci}) = \beta_0 + \beta_1(e_i - \mu_e)\,[+\ldots], \quad \mu_c = \exp(\beta_0)$$

- 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

Example: MenSS trial

- 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)

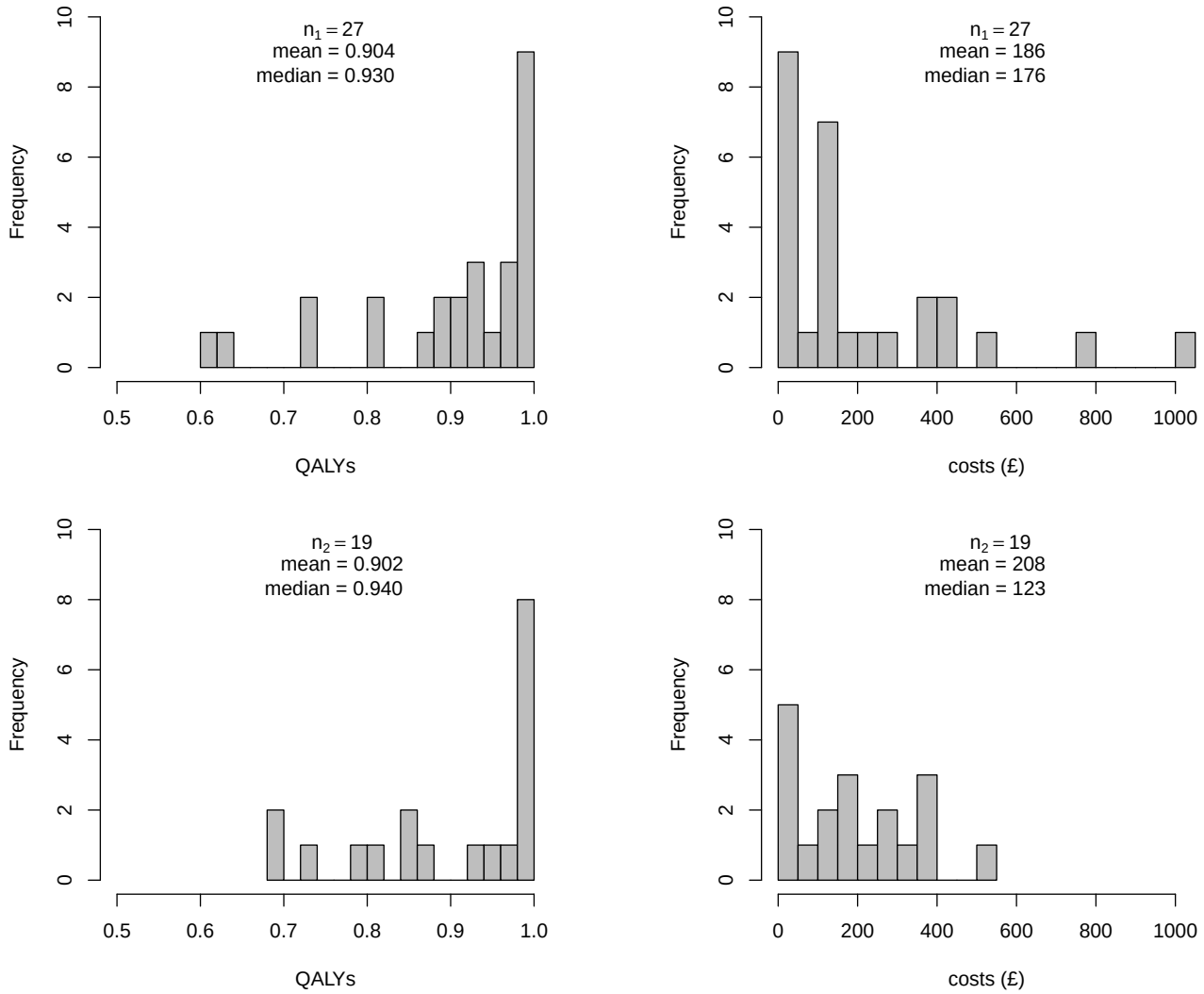
Example: MenSS trial

- 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%)

Example: MenSS trial

- 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)

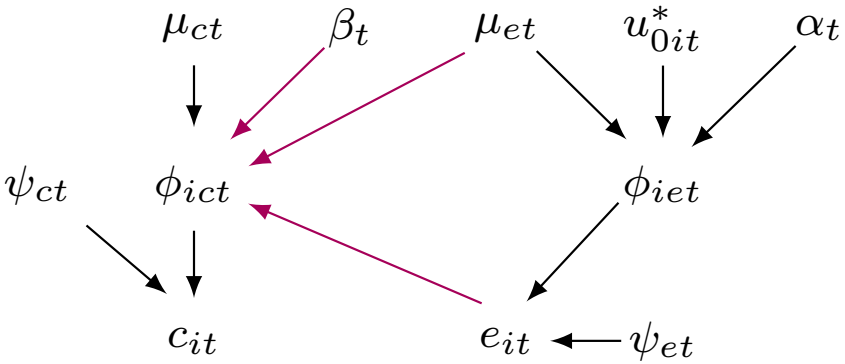


Gabrio et al. (2018). https://arxiv.org/abs/1801.09541

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

Conditional model for $c \mid e$

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



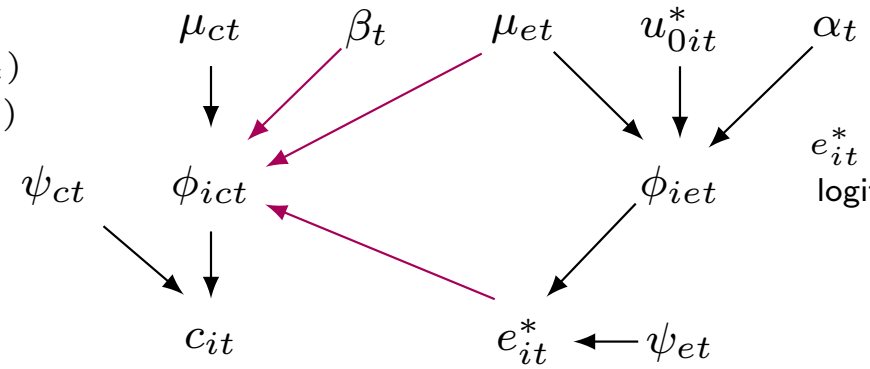
Marginal model for e

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

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^*

$$e^*_{it} \sim \text{Beta}\left(\phi_{eit}\psi_{et}, (1-\phi_{eit})\psi_{et}\right)$$
$$\begin{aligned}\text{logit}(\phi_{eit}) &= \mu_{et} + \alpha_t(u_{0it} - \bar{u}_{0t}) \\ &= \mu_{et} + \alpha_t u^*_{0it}\end{aligned}$$

Gabrio et al. (2018). https://arxiv.org/abs/1801.09541

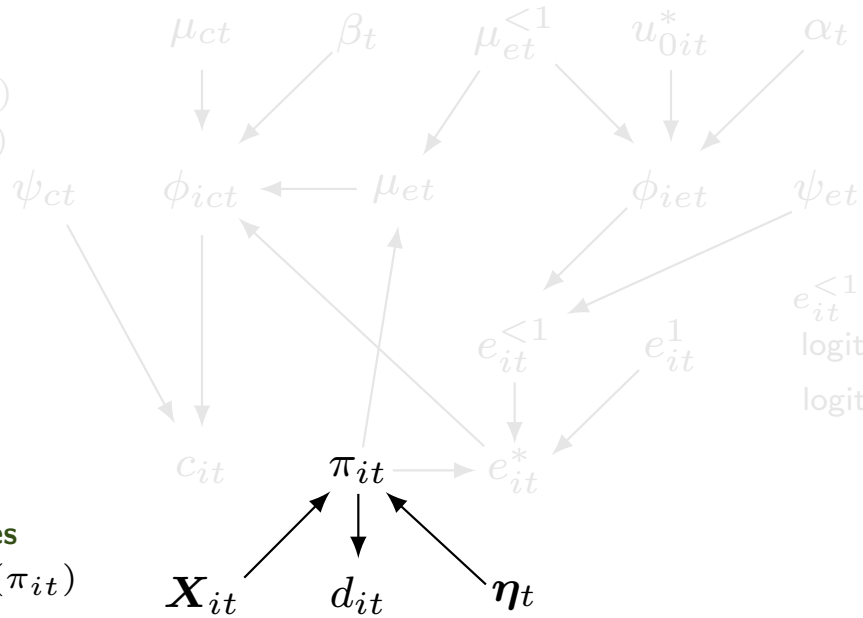
1. **Bivariate Normal**
   - Simpler and closer to "standard" frequentist model
   - Account for correlation between QALYs and costs
2. **Beta-Gamma**
   - Account for correlation between outcomes
   - Model the relevant ranges: QALYs $\in (0,1)$ and costs $\in (0,\infty)$
   - **But**: needs to rescale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

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



Model for the structural ones

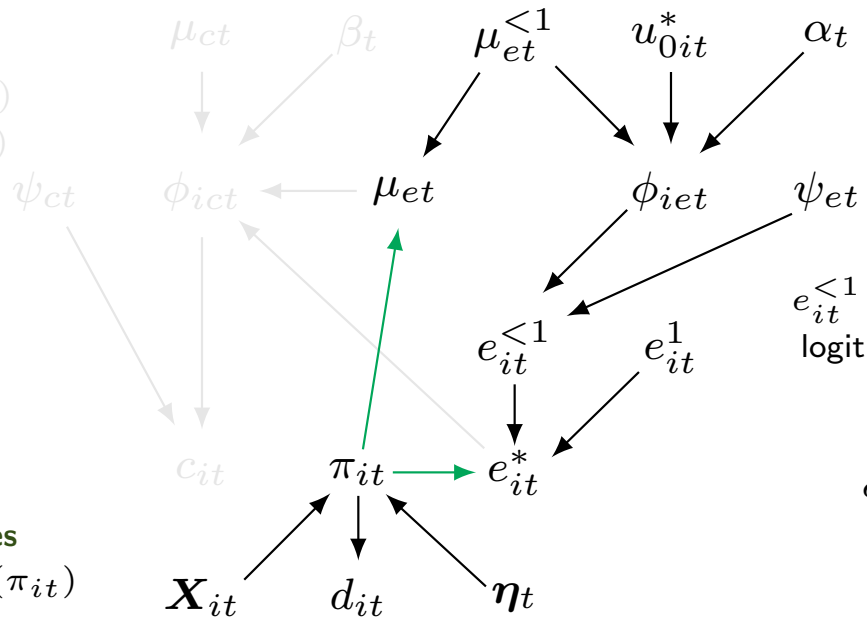
$$d_{it} := \mathbb{I}(e_{it} = 1) \sim \text{Bernoulli}(\pi_{it})$$
$$\text{logit}(\pi_{it}) = \boldsymbol{X}_{it}\boldsymbol{\eta}_t$$

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
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

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

$$e^*_{it} = \pi_{it}e^1_{it} + (1-\pi_{it})e^{<1}_{it}$$

$$\mu_{et} = (1-\bar{\pi}_t)\mu^{<1}_{et} + \bar{\pi}_t$$

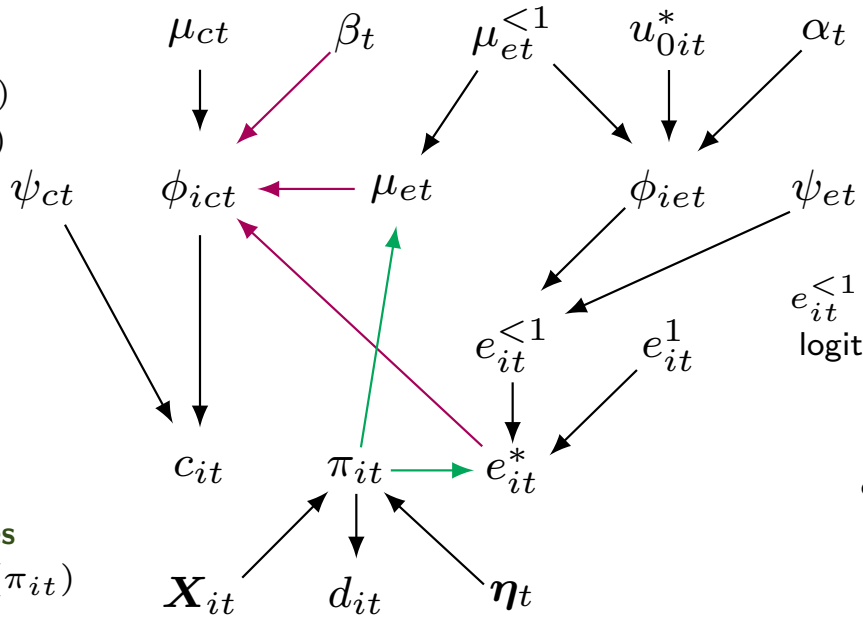
Model for the structural ones

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

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
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

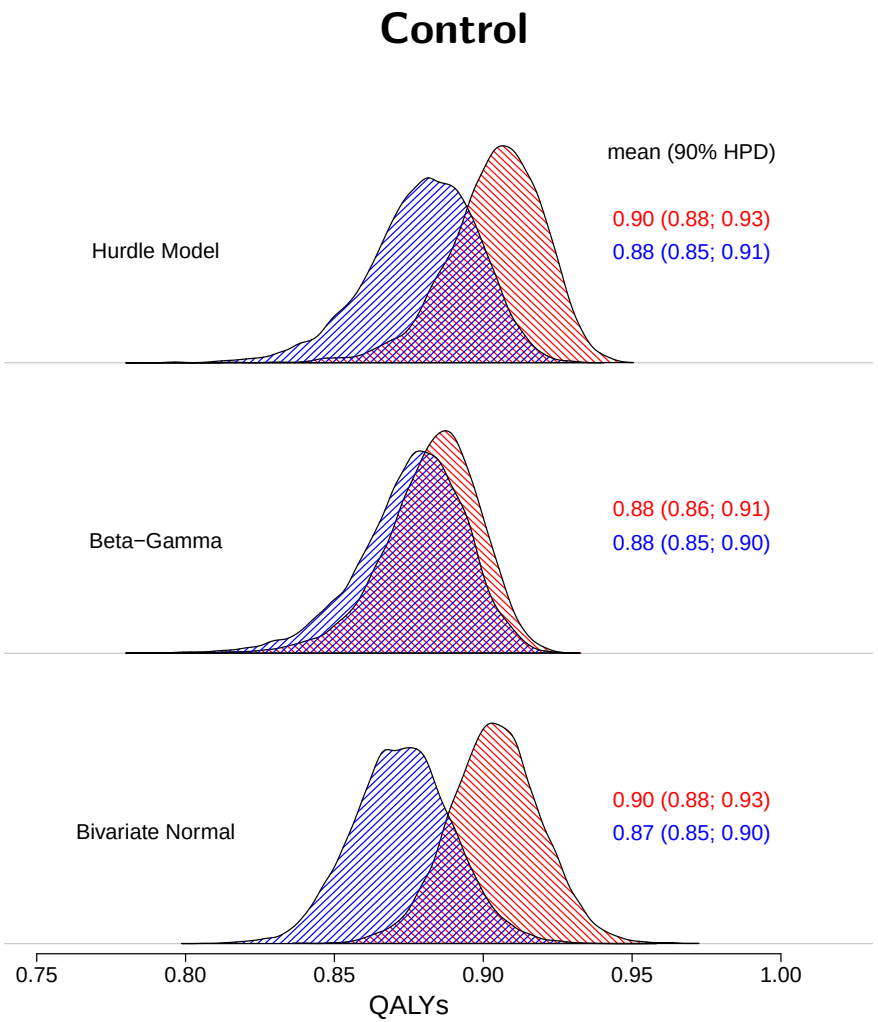
$$e^1_{it} := 1$$
$$e^{<1}_{it} \sim \text{Beta}\left(\phi_{eit}\psi_{et}, (1-\phi_{eit})\psi_{et}\right)$$
$$\text{logit}(\phi_{eit}) = \mu^{<1}_{et} + \alpha_t(u_{0it} - \bar{u}_{0t}) = \mu^{<1}_{et} + \alpha_t u^*_{0it}$$

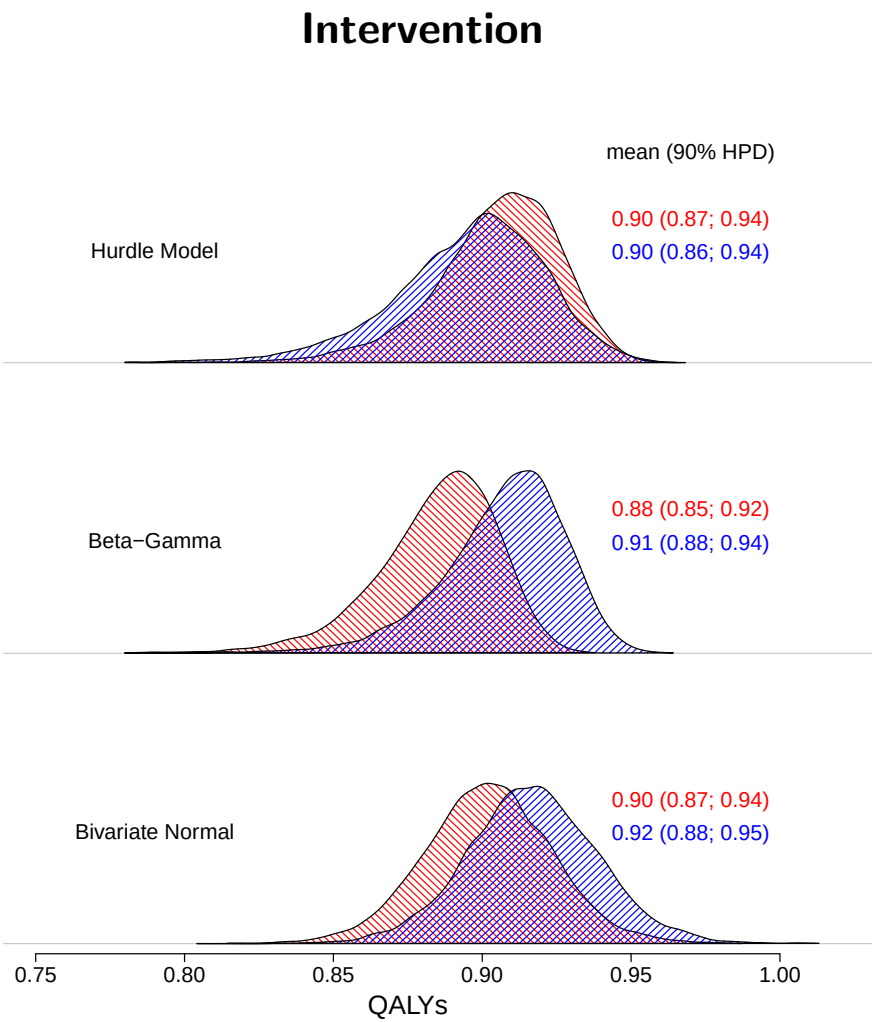
$$e^*_{it} = \pi_{it}e^1_{it} + (1-\pi_{it})e^{<1}_{it}$$

$$\mu_{et} = (1-\bar{\pi}_t)\mu^{<1}_{et} + \bar{\pi}_t$$

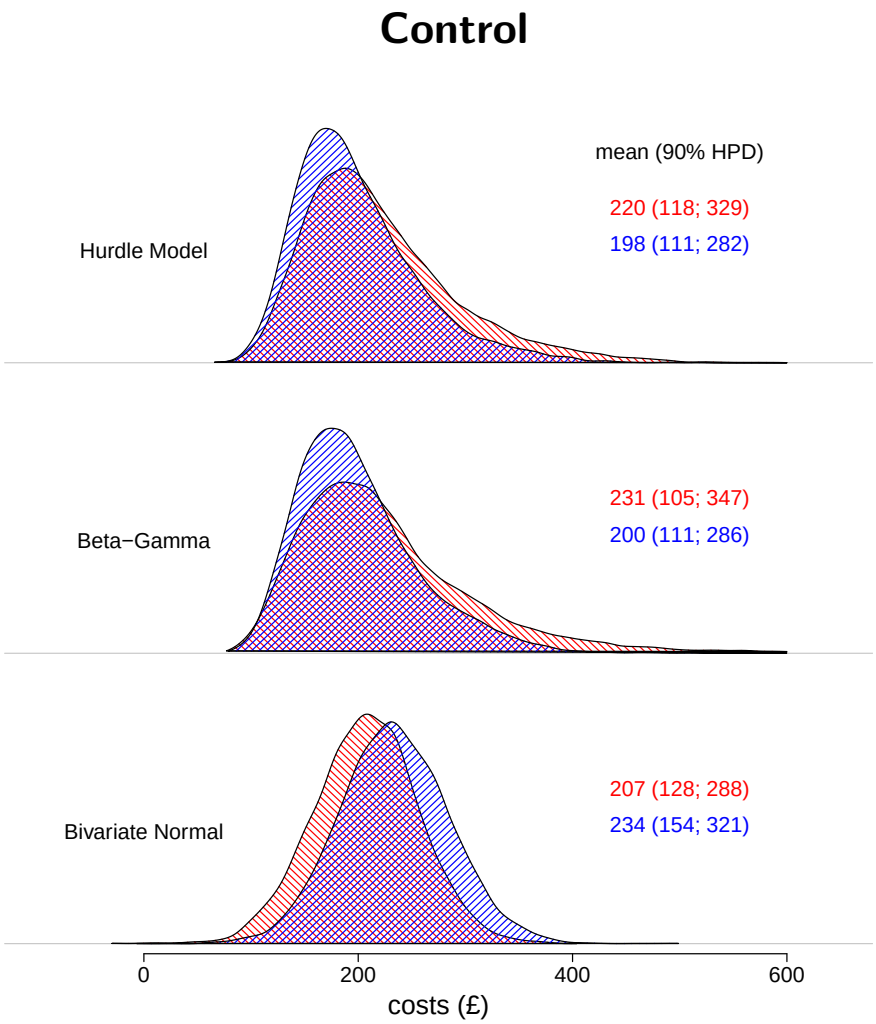
Model for the structural ones

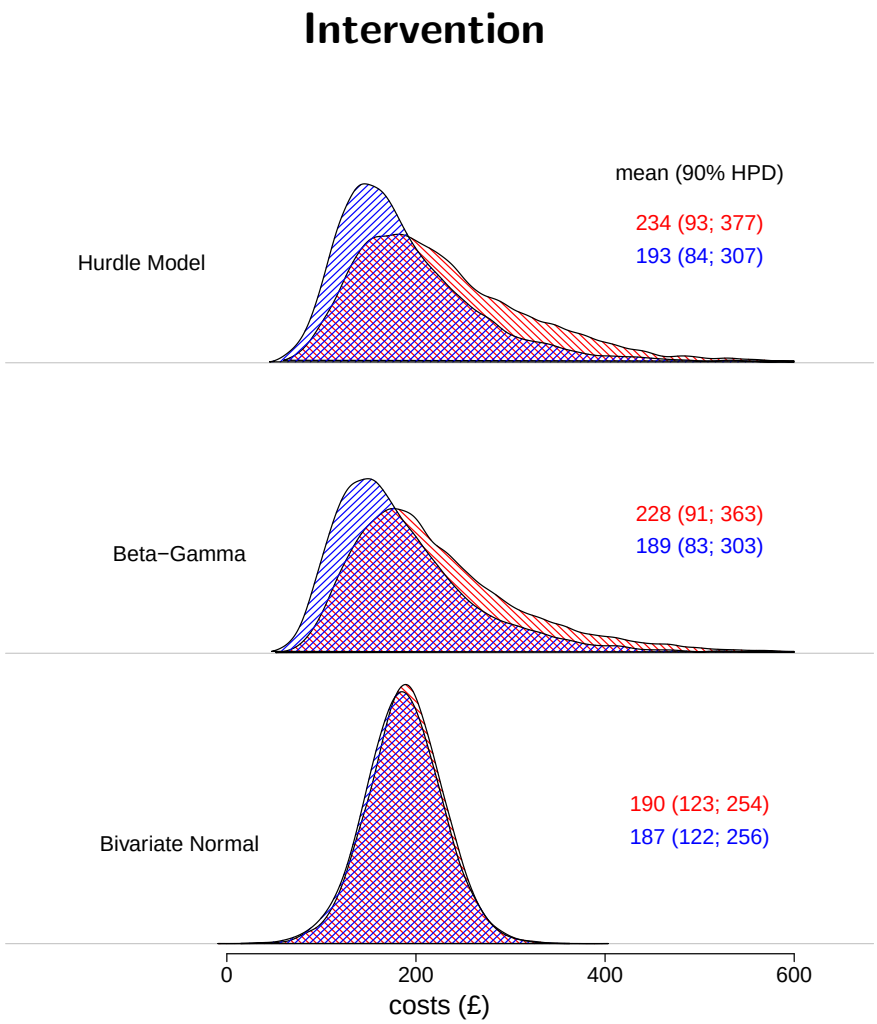
$$d_{it} := \mathbb{I}(e_{it} = 1) \sim \text{Bernoulli}(\pi_{it})$$
$$\text{logit}(\pi_{it}) = \boldsymbol{X}_{it}\boldsymbol{\eta}_t$$





Complete cases only
All cases (missing at random, MAR)

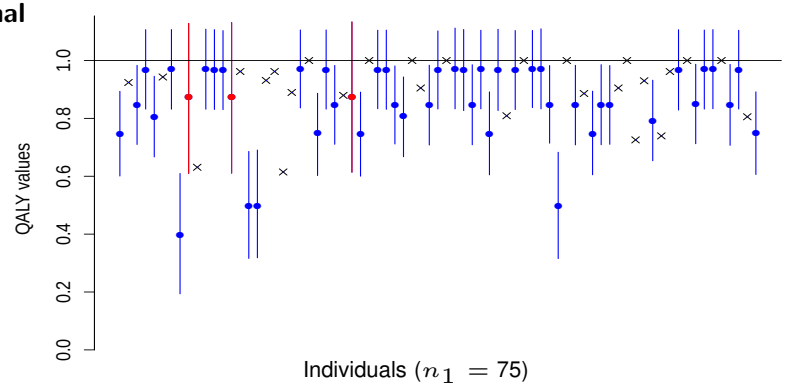


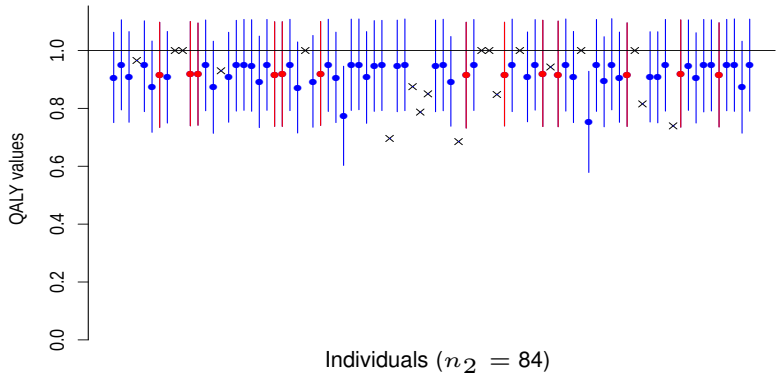


Complete cases only
All cases (missing at random, MAR)

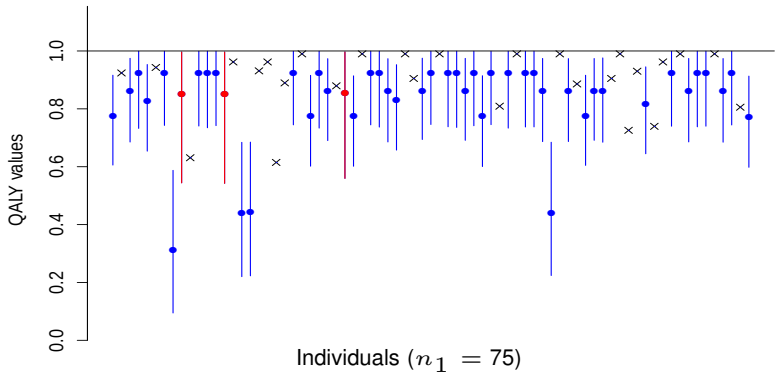
Bayesian multiple imputation (under MAR)

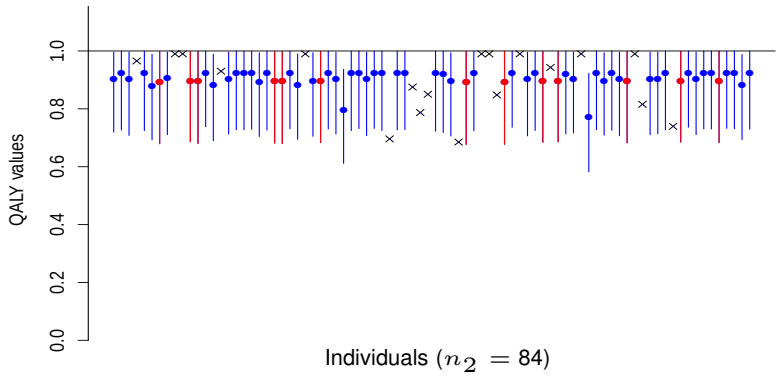
Bivariate Normal



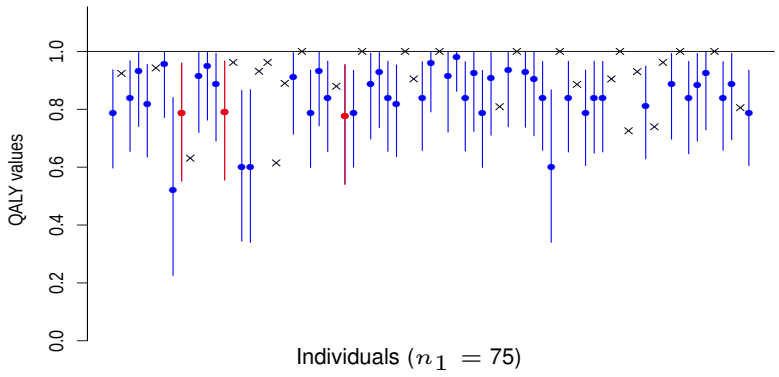


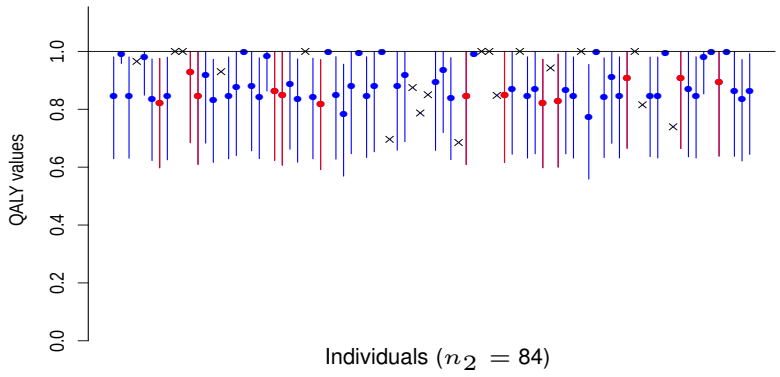
Beta-Gamma





Hurdle model





Imputed, observed baseline
Imputed, missing baseline
× Observed

End

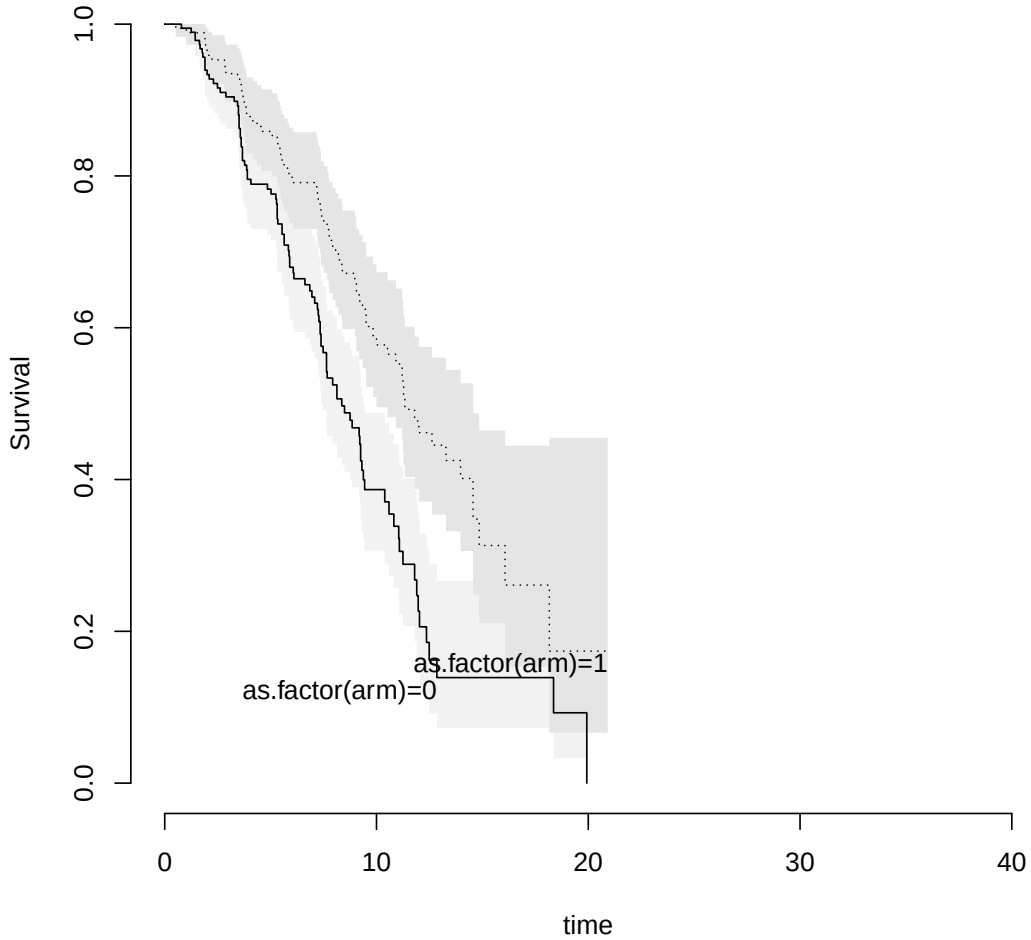
Outline

1. Health economic evaluation
   - What is it?
   - How does it work?
2. Statistical modelling
   - Individual-level vs aggregated data
   - The importance of being a Bayesian
3. **Some examples**
   - Individual level & partially observed data
   - Survival analysis in HTA
   - Value of information
4. Conclusions

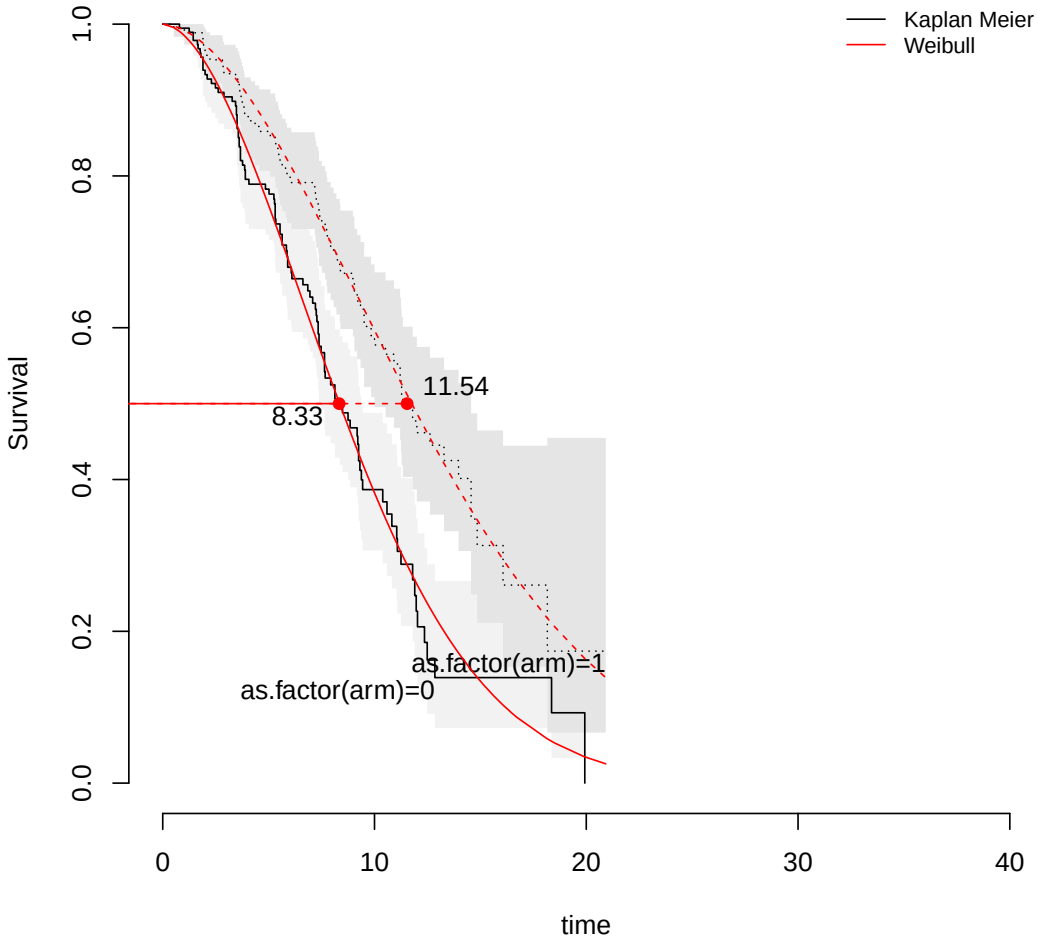
Survival analysis

Value of information

- Survival data are often the main outcome in clinical studies relevant for HTA
  - Cancer drugs (progression-free/overall survival time): $\approx$ 40% of NICE appraisals!
  - Need to **extrapolate**, for economic modelling purposes. **BUT**: Limited follow up from trials, not consistent with time horizon of economic model

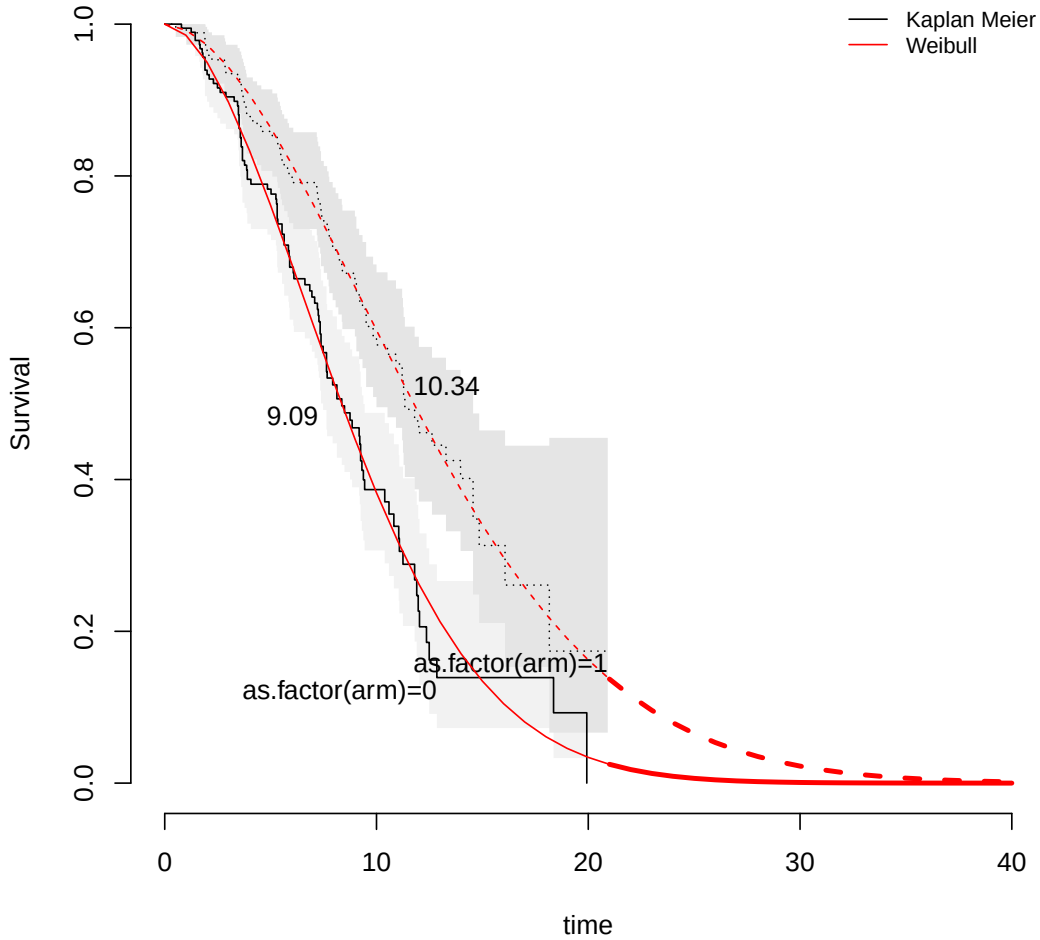


- Survival data are often the main outcome in clinical studies relevant for HTA
  - Cancer drugs (progression-free/overall survival time): $\approx$ 40% of NICE appraisals!
  - Need to **extrapolate**, for economic modelling purposes. **BUT**: Limited follow up from trials, not consistent with time horizon of economic model



Mean time: $\int_0^\infty S(t)dt$

- Survival data are often the main outcome in clinical studies relevant for HTA
  - Cancer drugs (progression-free/overall survival time): $\approx$ 40% of NICE appraisals!
  - Need to **extrapolate**, for economic modelling purposes. **BUT**: Limited follow up from trials, not consistent with time horizon of economic model



- Survival data are often the main outcome in clinical studies relevant for HTA
  - Cancer drugs (progression-free/overall survival time): $\approx$ 40% of NICE appraisals!
  - Need to **extrapolate**, for economic modelling purposes. **BUT**: Limited follow up from trials, not consistent with time horizon of economic model

- When there is strong correlation among the survival parameters, the results of uncertainty analysis may be (strongly) biased under a more simplistic frequentist model
  - This matters most in health economics, because this bias carries over the economic modelling, optimal decision making and assessment of the impact of parametric uncertainty!
  - **A full Bayesian approach propagates directly correlation and uncertainty in the model parameters through to the survival curves and the economic model**

- For more complex models, MLE-based estimates may fail to converge
  - This may be an issue for multi-parameter models, where limited data (not compounded by relevant prior information) are not enough to fit all the model parameters
  - **NB**: you would normally need to fit more complex models for cases where the survival curves are "strange" and so the usual parametric models fail to provide sufficient fit

- Survival data are often the main outcome in clinical studies relevant for HTA
  - Cancer drugs (progression-free/overall survival time): $\approx$ 40% of NICE appraisals!
  - Need to **extrapolate**, for economic modelling purposes. **BUT**: Limited follow up from trials, not consistent with time horizon of economic model

```
Model fit for the Generalised F model, obtained using Flexsurvreg
(Maximum Likelihood Estimate). Running time: 1.157 seconds

                       mean         se         L95%         U95%
mu                2.29139696 0.0798508 2.13489e+00 2.44790e+00
sigma             0.58729598 0.0725044 4.61076e-01 7.48069e-01
Q                 0.84874994 0.2506424 3.57500e-01 1.34000e+00
P                 0.00268265 0.0902210 6.33197e-32 1.13655e+26
as.factor(arm)1 0.34645851 0.0877892 1.74395e-01 5.18522e-01
```

```
Model fit for the Generalised F model, obtained using Stan
(Bayesian inference via Hamiltonian Monte Carlo). Running time: 26.692 seconds

                     mean         se       L95%       U95%
mu                2.256760 0.3455163 0.0897086 0.0865904
sigma             0.507861 0.0762112 0.3608566 0.6582047
Q                 0.700062 0.3358360 0.0786118 1.3880582
P                 1.131968 0.5837460 0.3908284 2.6342762
as.factor(arm)1 0.345516 0.0865904 0.1745665 0.5176818
```

Set up/interventions

- ICD (Implantable Cardioverter Defibrillators) compared to anti-arrhythmic drugs (AAD) for prevention of sudden cardiac death in people with cardiac arrhythmia

Set up/interventions

- ICD (Implantable Cardioverter Defibrillators) compared to anti-arrhythmic drugs (AAD) for prevention of sudden cardiac death in people with cardiac arrhythmia

Data

- Individual data from cohort of 535 UK cardiac arrhythmia patients implanted with ICDs between 1991 and 2002
- Meta-analysis of three (non-UK) RCTs providing published HRs
  - Relatively short-term follow-up: approximately 75% people, followed for less than 5 years, maximum 10 years
- UK population mortality statistics by age, sex, cause of death

Set up/interventions

- ICD (Implantable Cardioverter Defibrillators) compared to anti-arrhythmic drugs (AAD) for prevention of sudden cardiac death in people with cardiac arrhythmia

Data

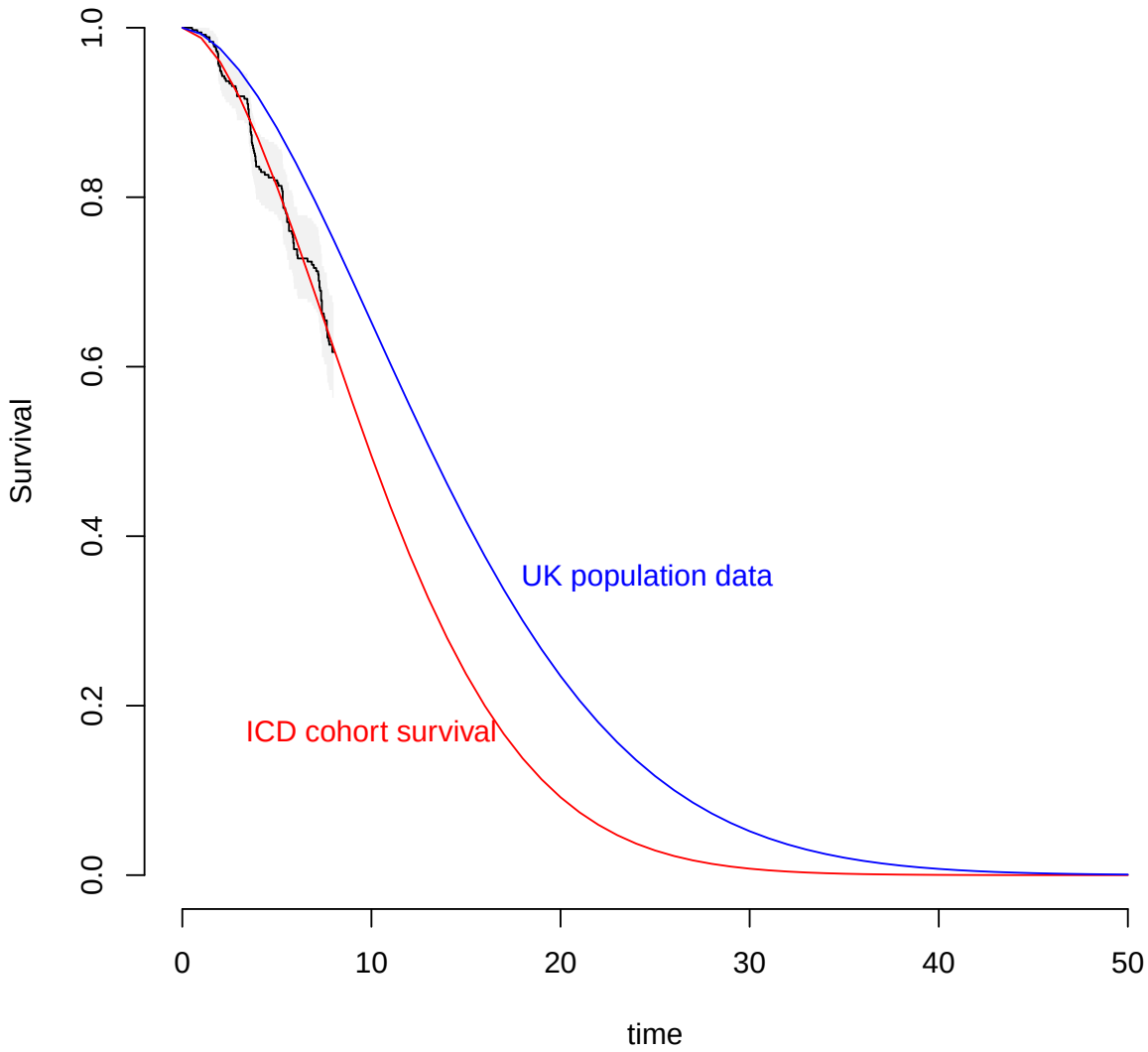
- Individual data from cohort of 535 UK cardiac arrhythmia patients implanted with ICDs between 1991 and 2002
- Meta-analysis of three (non-UK) RCTs providing published HRs
  - Relatively short-term follow-up: approximately 75% people, followed for less than 5 years, maximum 10 years
- UK population mortality statistics by age, sex, cause of death

Objective

- Estimate the survival curve **over the lifetime** of ICD and AAD patients in UK
- Extrapolate the output to inform the wider economic model

Basic idea

Use UK population data (matched by age/sex) to "**anchor**" the ICD population risk



Basic idea

Use UK population data (matched by age/sex) to "**anchor**" the ICD population risk

- Perhaps the easiest way of doing this anchoring is to relate the hazards between the two populations — eg **proportional hazard ratio**:

$$h_{\text{ICD}}(t) = e^{\beta} h_{\text{UK}}(t) \qquad \Leftrightarrow \qquad \text{HR} = \frac{h_{\text{ICD}}(t)}{h_{\text{UK}}(t)} = e^{\beta} = \text{Constant}$$

- Relatively easy to model — but probably very unrealistic!
  - ICD patients are at (much?) greater risk of arrhythmia death
  - If the proportion of deaths caused by arrythmia changes over time, we would induce bias, because we would be extrapolate a constant HR for all causes mortality

Use UK population data (matched by age/sex) to "**anchor**" the ICD population risk

- Perhaps the easiest way of doing this anchoring is to relate the hazards between the two populations — eg **proportional hazard ratio**:

$$h_{\text{ICD}}(t) = e^{\beta} h_{\text{UK}}(t) \qquad \Leftrightarrow \qquad \text{HR} = \frac{h_{\text{ICD}}(t)}{h_{\text{UK}}(t)} = e^{\beta} = \text{Constant}$$

- Relatively easy to model — but probably very unrealistic!
  - ICD patients are at (much?) greater risk of arrhythmia death
  - If the proportion of deaths caused by arrythmia changes over time, we would induce bias, because we would be extrapolate a constant HR for all causes mortality

- Formally account for multiple mortality causes (**Poly-Weibull** model):

$$\begin{aligned} h_{\text{ICD}}(t) &= h^{\text{arr}}_{\text{ICD}}(t) + h^{\text{oth}}_{\text{ICD}}(t) \\ &= e^{\beta} h^{\text{arr}}_{\text{UK}}(t) + h^{\text{oth}}_{\text{UK}}(t) \\ &= e^{\beta} \alpha_1 \mu_1 t^{\alpha_1 - 1} + \alpha_2 \mu_2 t^{\alpha_2 - 1} \end{aligned}$$

- This assumes that:
  - Arrhythmia hazard is **proportional** to matched UK population
  - Other causes hazard is **identical** to matched UK population

♪ You don't know what you're doing!... ♪

https://youtu.be/fOSU59DP760

- To set up a full Bayesian model including a reasonable specification of the priors can be a hard task
- Often people claim that they have "no prior information".

♪ You don't know what you're doing!... ♪

https://youtu.be/fOSU59DP760

- To set up a full Bayesian model including a reasonable specification of the priors can be a hard task
- Often people claim that they have "no prior information". **But don't they?...**

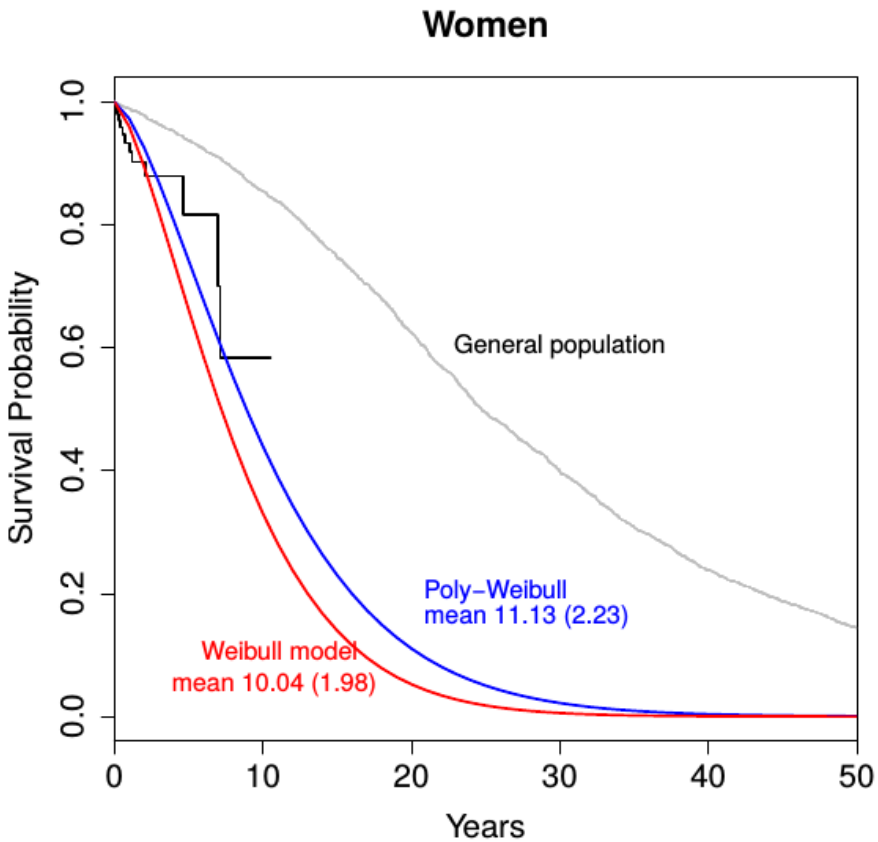
- To set up a full Bayesian model including a reasonable specification of the priors can be a hard task
- Often people claim that they have "no prior information". **But don't they?...**

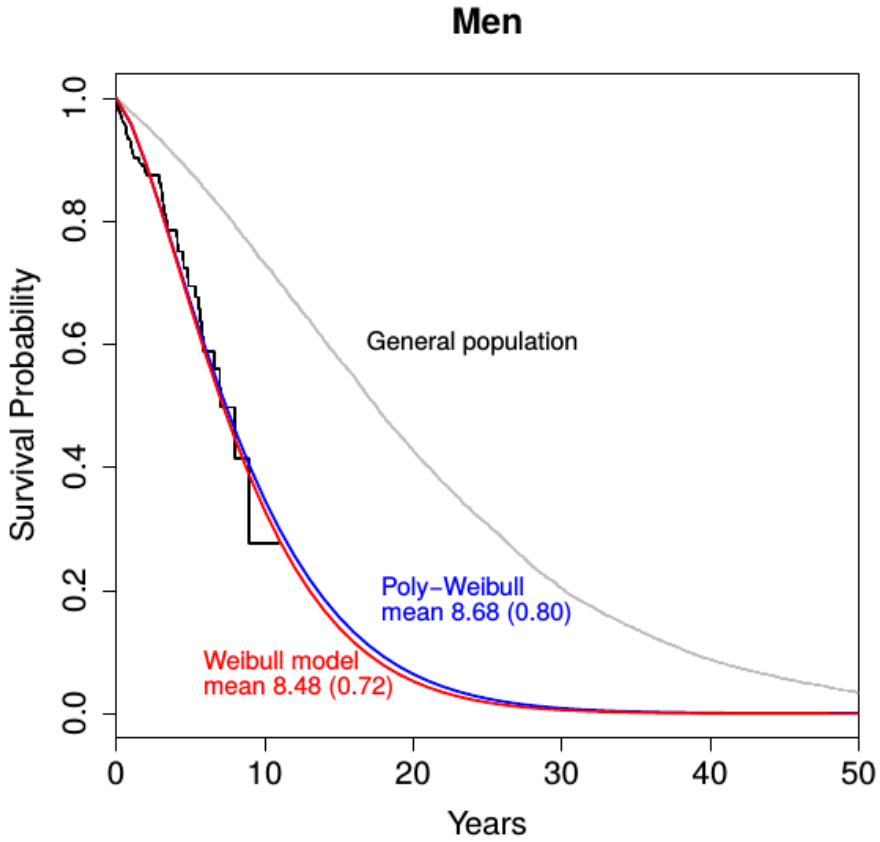
- In the ICD case, age at entry is around 60 — we **know** that people won't survive for more than other 60 years
  - Setting a prior for the scale $\mu_i \sim \text{Uniform}(0, 100)$ implies that the prior mean survival of the resulting Weibull distribution is

$$\mu_i \Gamma\left(1 + \frac{1}{\alpha}\right) < 60$$

- Can also include some knowledge on the shape α and the coefficient β to limit their variations in reasonable ranges

Results





- Ignoring cause-specific mortality (Weibull) results in larger bias, especially for females (because the arrhythmia proportion of deaths does vary over time in that subgroup)

End

Outline

1. Health economic evaluation
   - What is it?
   - How does it work?
2. Statistical modelling
   - Individual-level vs aggregated data
   - The importance of being a Bayesian
3. **Some examples**
   - Individual level & partially observed data
   - Survival analysis in HTA
   - Value of information
4. Conclusions

ILD+Missing data

Value of information

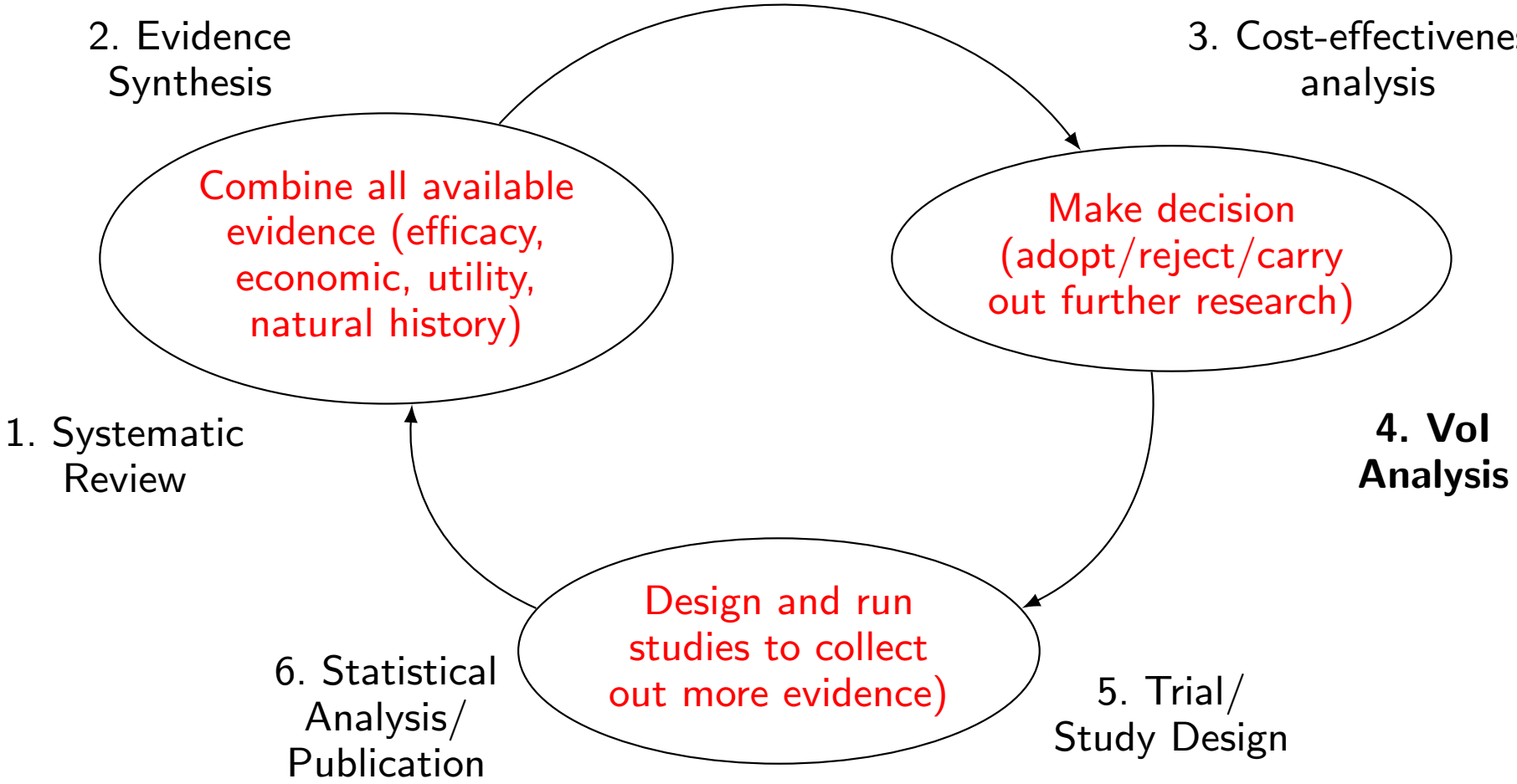
Knowledge *is* power?...

(A tale of two stupid examples)

- **Example 1**: Intervention $t = 1$ is the most cost-effective, given current evidence
  - $\Pr(t = 1 \text{ is cost-effective}) = 0.51$
  - If we get it wrong: Increase in costs = £3
    Decrease in effectiveness = 0.000001 QALYs
  - Large uncertainty/negligible consequences ⇒ **can afford uncertainty**

- **Example 1**: Intervention $t = 1$ is the most cost-effective, given current evidence
  - $\Pr(t = 1 \text{ is cost-effective}) = 0.51$
  - If we get it wrong: Increase in costs $=$ £3
    Decrease in effectiveness $= 0.000001$ QALYs
  - Large uncertainty/negligible consequences $\Rightarrow$ **can afford uncertainty**
- **Example 2**: Intervention $t = 1$ is the most cost-effective, given current evidence
  - $\Pr(t = 1 \text{ is cost-effective}) = 0.999$
  - If we get it wrong: Increase in costs $=$ £1 000 000 000
    Decrease in effectiveness $= 999999$ QALYs
  - Tiny uncertainty/dire consequences $\Rightarrow$ **probably should think about it...**

Evidence Based Decision-Making and Value of Information (VoI)



Process inherently Bayesian!

VoI: Basic idea

- A new study will provide new data
  - Reducing (or even eliminating) uncertainty in a subset of model parameters
- Update the cost-effectiveness model
  - If the optimal decision changes, gain in monetary net benefit (NB = utility) from using new optimal treatment
  - If optimal decision unchanged, no gain in NB
- **Expected** VOI is the average gain in NB

VoI: Basic idea and relevant measures

- A new study will provide new data
  - Reducing (or even eliminating) uncertainty in a subset of model parameters
- Update the cost-effectiveness model
  - If the optimal decision changes, gain in monetary net benefit (NB = utility) from using new optimal treatment
  - If optimal decision unchanged, no gain in NB
- **Expected** VOI is the average gain in NB

1. **Expected Value of Perfect Information** (EVPI)
   - Value of completely resolving uncertainty in all input parameters to decision model
   - Infinite-sized long-term follow-up trial measuring everything!
   - Gives an upper-bound on the value of new study — if EVPI is low, suggests we can make our decision based on existing information

VoI: Basic idea and relevant measures

- A new study will provide new data
  - Reducing (or even eliminating) uncertainty in a subset of model parameters
- Update the cost-effectiveness model
  - If the optimal decision changes, gain in monetary net benefit (NB = utility) from using new optimal treatment
  - If optimal decision unchanged, no gain in NB
- **Expected** VOI is the average gain in NB

1. **Expected Value of Perfect Information** (EVPI)
   - Value of completely resolving uncertainty in all input parameters to decision model
   - Infinite-sized long-term follow-up trial measuring everything!
   - Gives an upper-bound on the value of new study — if EVPI is low, suggests we can make our decision based on existing information
2. **Expected Value of Partial Perfect Information** (EVPPI)
   - Value of eliminating uncertainty in subset of input parameters to decision model
   - Infinite-sized trial measuring relative effects on 1-year survival
   - Useful to identify which parameters responsible for decision uncertainty

VoI: Basic idea and relevant measures

- A new study will provide new data
  - Reducing (or even eliminating) uncertainty in a subset of model parameters
- Update the cost-effectiveness model
  - If the optimal decision changes, gain in monetary net benefit (NB = utility) from using new optimal treatment
  - If optimal decision unchanged, no gain in NB
- **Expected** VOI is the average gain in NB

1. **Expected Value of Perfect Information** (EVPI)
   - Value of completely resolving uncertainty in all input parameters to decision model
   - Infinite-sized long-term follow-up trial measuring everything!
   - Gives an upper-bound on the value of new study — if EVPI is low, suggests we can make our decision based on existing information
2. **Expected Value of Partial Perfect Information** (EVPPI)
   - Value of eliminating uncertainty in subset of input parameters to decision model
   - Infinite-sized trial measuring relative effects on 1-year survival
   - Useful to identify which parameters responsible for decision uncertainty
3. **Expected Value of Sample Information** (EVSI)
   - Value of reducing uncertainty by conducting a study of given design
   - Can compare the benefits and costs of a study with given design
   - Is the proposed study likely to be a good use of resources? What is the optimal design?

Parameters simulations

Iter/n	π_0	ρ	...	γ
1	0.365	0.076	...	0.162
2	0.421	0.024	...	0.134
3	0.125	0.017	...	0.149
4	0.117	0.073	...	0.120
5	0.481	0.008	...	0.191
6	0.163	0.127	...	0.004
...		...		
1000	0.354	0.067	...	0.117

- Characterise uncertainty in the model parameters
  - In a full Bayesian setting, these are draws from the joint posterior distribution of $\boldsymbol{\theta}$
  - In a frequentist setting, these are typically Monte Carlo draws from a set of univariate distributions that describe some level of uncertainty around MLEs (two-step/hybrid)

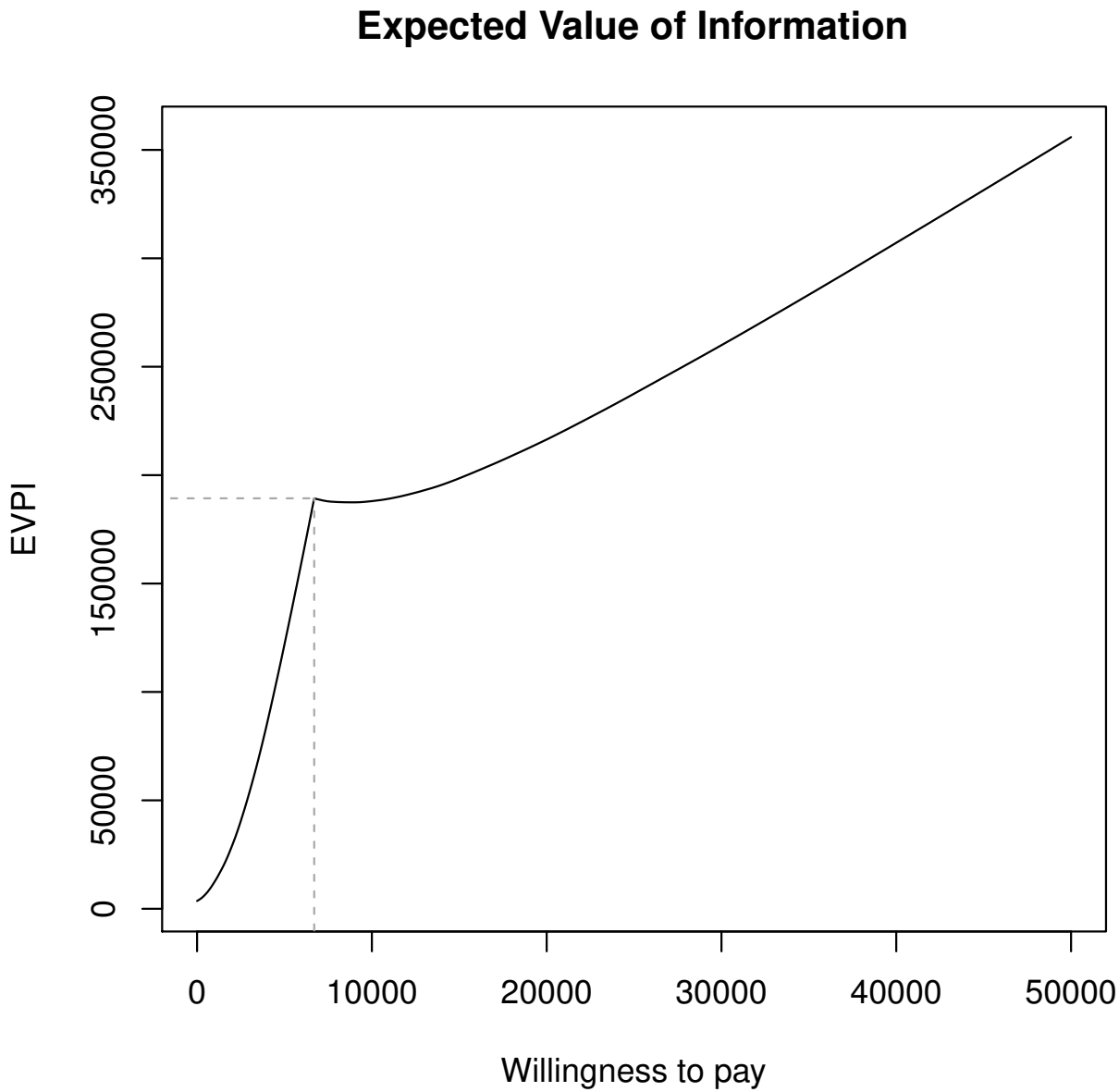
	Parameters simulations				Expected utility	
Iter/n	π_0	ρ	...	γ	$\text{NB}_0(\boldsymbol{\theta})$	$\text{NB}_1(\boldsymbol{\theta})$
1	0.365	0.076	...	0.162	19 214 751	*19 647 706*
2	0.421	0.024	...	0.134	*17 165 526*	17 163 407
3	0.125	0.017	...	0.149	18 710 928	16 458 433
4	0.117	0.073	...	0.120	16 991 321	*18 497 648*
5	0.481	0.008	...	0.191	*19 772 898*	18 662 329
6	0.163	0.127	...	0.004	17 106 136	*18 983 331*
...		...				...
1000	0.354	0.067	...	0.117	*18 043 921*	16 470 805
				Average	18 659 238	*19 515 004*

- Uncertainty in the parameters induces a distribution of decisions
  - Typically based on the **net benefits**: $\text{NB}_t(\boldsymbol{\theta}) = k\mu_{et} - \mu_{ct}$
  - In each parameters configuration can identify the *optimal strategy*
- Averaging over the uncertainty in $\boldsymbol{\theta}$ provides t^*, the overall optimal decision *given current uncertainty* (= choose the intervention associated with ***highest* expected utility**)

	Parameters simulations				Expected utility		Maximum	Opportunity
Iter/n	π_0	ρ	...	γ	$\text{NB}_0(\boldsymbol{\theta})$	$\text{NB}_1(\boldsymbol{\theta})$	net benefit	loss
1	0.365	0.076	...	0.162	19 214 751	*19 647 706*	19 647 706	0
2	0.421	0.024	...	0.134	*17 165 526*	17 163 407	17 165 526	2 119.3
3	0.125	0.017	...	0.149	18 710 928	16 458 433	18 710 928	2 252 495.5
4	0.117	0.073	...	0.120	16 991 321	*18 497 648*	18 497 648	0
5	0.481	0.008	...	0.191	*19 772 898*	18 662 329	19 772 898	1 110 569.3
6	0.163	0.127	...	0.004	17 106 136	*18 983 331*	18 983 331	0
...		...				...	...	
1000	0.354	0.067	...	0.117	*18 043 921*	16 470 805	18 043 921	1 573 116.0
				Average	18 659 238	*19 515 004*	**19 741 589**	**226 585**

- **Expected Value of "Perfect" Information** (EVPI) summarises uncertainty in the decision
  - Defined as the **average Opportunity Loss**
  - Can also be computed as the difference between the **average maximum expected utility under "perfect" information** and the **maximum expected utility overall**:

$$\text{EVPI} = \text{E}_{\boldsymbol{\theta}}\underbrace{\left[\max_t \text{NB}_t(\boldsymbol{\theta})\right]}_{\substack{\text{Value of decision}\\ \text{if we knew } \boldsymbol{\theta}}} - \underbrace{\max_t \text{E}_{\boldsymbol{\theta}}\left[\text{NB}_t(\boldsymbol{\theta})\right]}_{\substack{\text{Value of decision based}\\ \text{on current information}}} = \text{E}_{\boldsymbol{\theta}}\underbrace{\left[\max_t \text{NB}_t(\boldsymbol{\theta}) - \text{NB}_{t^*}(\boldsymbol{\theta})\right]}_{\substack{\text{Opportunity lost from using } t^*\\ \text{instead of the optimal } t \text{ for } \boldsymbol{\theta}}}$$


Expected Value of Information
EVPI
0
50000
150000
250000
350000
0
10000
20000
30000
40000
50000
Willingness to pay

Summarising PSA + Research priority: Expected Value of **Partial** Perfect Information

- $\boldsymbol{\theta}$ = all the model parameters; can be split into two subsets
  - The "**parameters of interest**" ϕ, e.g. prevalence of a disease, HRQL measures, length of stay in hospital, ...
  - The "**remaining parameters**" ψ, e.g. cost of treatment with other established medications,
- We are interested in quantifying the value of gaining more information on ϕ, while leaving the current level of uncertainty on ψ unchanged

Summarising PSA + Research priority: Expected Value of **Partial** Perfect Information

- $\boldsymbol{\theta}$ = all the model parameters; can be split into two subsets
  - The "**parameters of interest**" ϕ, e.g. prevalence of a disease, HRQL measures, length of stay in hospital, ...
  - The "**remaining parameters**" ψ, e.g. cost of treatment with other established medications,
- We are interested in quantifying the value of gaining more information on ϕ, while leaving the current level of uncertainty on ψ unchanged
- In formulæ:
  - First, consider the expected utility (EU) **if** we were able to learn ϕ but not ψ

$$\mathrm{E}_{\psi|\phi}\left[\mathrm{NB}_t(\boldsymbol{\theta})\right]$$

Summarising PSA + Research priority: Expected Value of **Partial** Perfect Information

- $\boldsymbol{\theta}$ = all the model parameters; can be split into two subsets
  - The "**parameters of interest**" ϕ, e.g. prevalence of a disease, HRQL measures, length of stay in hospital, ...
  - The "**remaining parameters**" ψ, e.g. cost of treatment with other established medications,
- We are interested in quantifying the value of gaining more information on ϕ, while leaving the current level of uncertainty on ψ unchanged
- In formulæ:
  - First, consider the expected utility (EU) **if** we were able to learn ϕ but not ψ
  - **If** we knew ϕ perfectly, best decision = the maximum of this EU

$$\max_t \mathrm{E}_{\psi|\phi}\left[\mathrm{NB}_t(\boldsymbol{\theta})\right]$$

Summarising PSA + Research priority: Expected Value of **Partial** Perfect Information

- $\boldsymbol{\theta}$ = all the model parameters; can be split into two subsets
  - The "**parameters of interest**" ϕ, e.g. prevalence of a disease, HRQL measures, length of stay in hospital, ...
  - The "**remaining parameters**" ψ, e.g. cost of treatment with other established medications,
- We are interested in quantifying the value of gaining more information on ϕ, while leaving the current level of uncertainty on ψ unchanged
- In formulæ:
  - First, consider the expected utility (EU) **if** we were able to learn ϕ but not ψ
  - **If** we knew ϕ perfectly, best decision = the maximum of this EU
  - Of course we cannot learn ϕ **perfectly**, so take the expected value

$$\mathrm{E}_{\phi}\left[\max_{t} \mathrm{E}_{\psi|\phi}\left[\mathrm{NB}_{t}(\boldsymbol{\theta})\right]\right]$$

Summarising PSA + Research priority: Expected Value of **Partial** Perfect Information

- $\boldsymbol{\theta}$ = all the model parameters; can be split into two subsets
  - The "**parameters of interest**" ϕ, e.g. prevalence of a disease, HRQL measures, length of stay in hospital, ...
  - The "**remaining parameters**" ψ, e.g. cost of treatment with other established medications,
- We are interested in quantifying the value of gaining more information on ϕ, while leaving the current level of uncertainty on ψ unchanged
- In formulæ:
  - First, consider the expected utility (EU) **if** we were able to learn ϕ but not ψ
  - **If** we knew ϕ perfectly, best decision = the maximum of this EU
  - Of course we cannot learn ϕ **perfectly**, so take the expected value
  - And compare this with the **maximum expected utility overall**

$$\mathrm{E}_{\phi}\left[\max_{t}\mathrm{E}_{\psi|\phi}\left[\mathrm{NB}_{t}(\boldsymbol{\theta})\right]\right]-\max_{t}\mathrm{E}_{\boldsymbol{\theta}}\left[\mathrm{NB}_{t}(\boldsymbol{\theta})\right]$$

Summarising PSA + Research priority: Expected Value of **Partial** Perfect Information

- $\boldsymbol{\theta}$ = all the model parameters; can be split into two subsets
  - The "**parameters of interest**" ϕ, e.g. prevalence of a disease, HRQL measures, length of stay in hospital, ...
  - The "**remaining parameters**" ψ, e.g. cost of treatment with other established medications,
- We are interested in quantifying the value of gaining more information on ϕ, while leaving the current level of uncertainty on ψ unchanged
- In formulæ:
  - First, consider the expected utility (EU) **if** we were able to learn ϕ but not ψ
  - **If** we knew ϕ perfectly, best decision = the maximum of this EU
  - Of course we cannot learn ϕ **perfectly**, so take the expected value
  - And compare this with the **maximum expected utility overall**
  - This is the EVPPI!

$$\text{EVPPI} = \text{E}_{\boldsymbol{\phi}}\left[\max_t \text{E}_{\boldsymbol{\psi}|\boldsymbol{\phi}}\left[\text{NB}_t(\boldsymbol{\theta})\right]\right] - \max_t \text{E}_{\boldsymbol{\theta}}\left[\text{NB}_t(\boldsymbol{\theta})\right]$$

Summarising PSA + Research priority: Expected Value of **Partial** Perfect Information

- $\boldsymbol{\theta}$ = all the model parameters; can be split into two subsets
  - The "**parameters of interest**" ϕ, e.g. prevalence of a disease, HRQL measures, length of stay in hospital, ...
  - The "**remaining parameters**" ψ, e.g. cost of treatment with other established medications,
- We are interested in quantifying the value of gaining more information on ϕ, while leaving the current level of uncertainty on ψ unchanged
- In formulæ:
  - First, consider the expected utility (EU) **if** we were able to learn ϕ but not ψ
  - **If** we knew ϕ perfectly, best decision = the maximum of this EU
  - Of course we cannot learn ϕ **perfectly**, so take the expected value
  - And compare this with the **maximum expected utility overall**
  - This is the EVPPI!

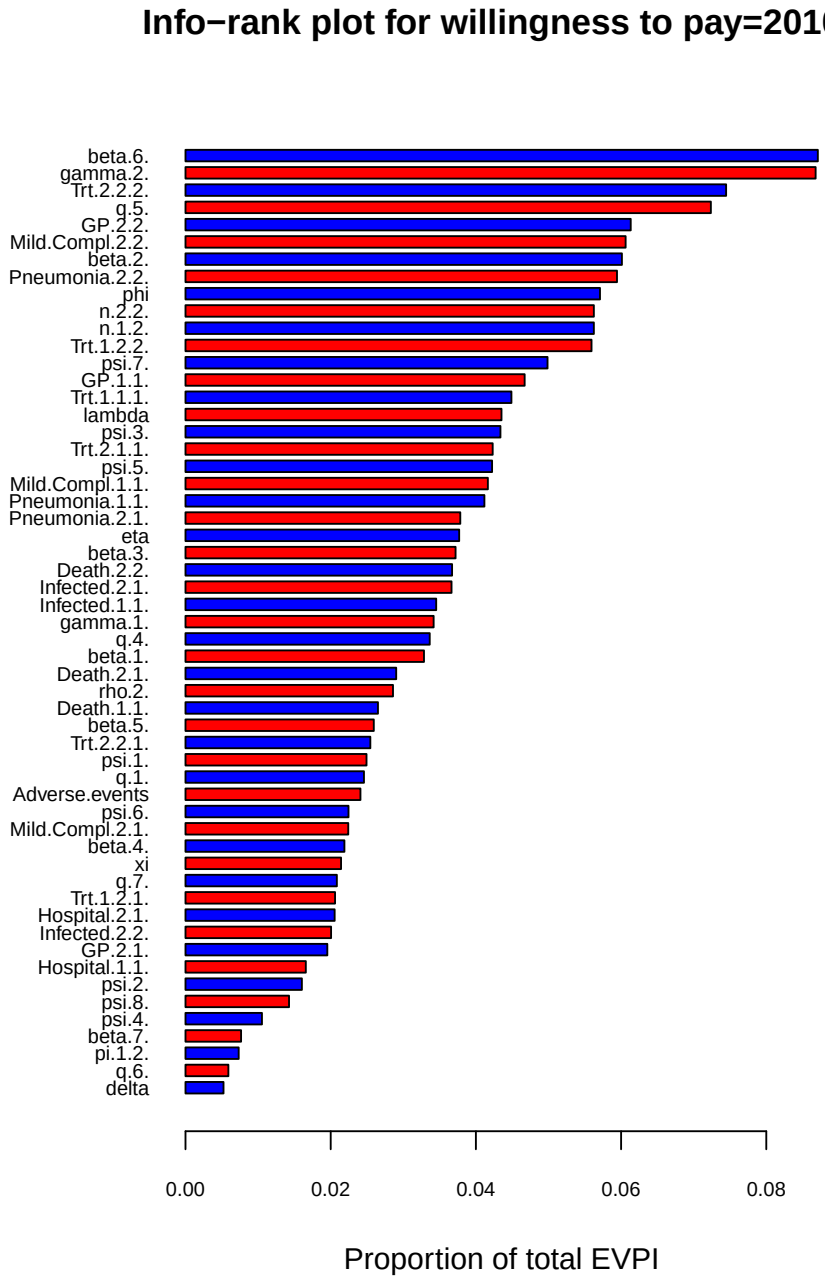
$$\text{EVPPI} = \text{E}_{\phi}\left[\max_t \text{E}_{\psi|\phi}\left[\text{NB}_t(\boldsymbol{\theta})\right]\right] - \max_t \text{E}_{\boldsymbol{\theta}}\left[\text{NB}_t(\boldsymbol{\theta})\right]$$

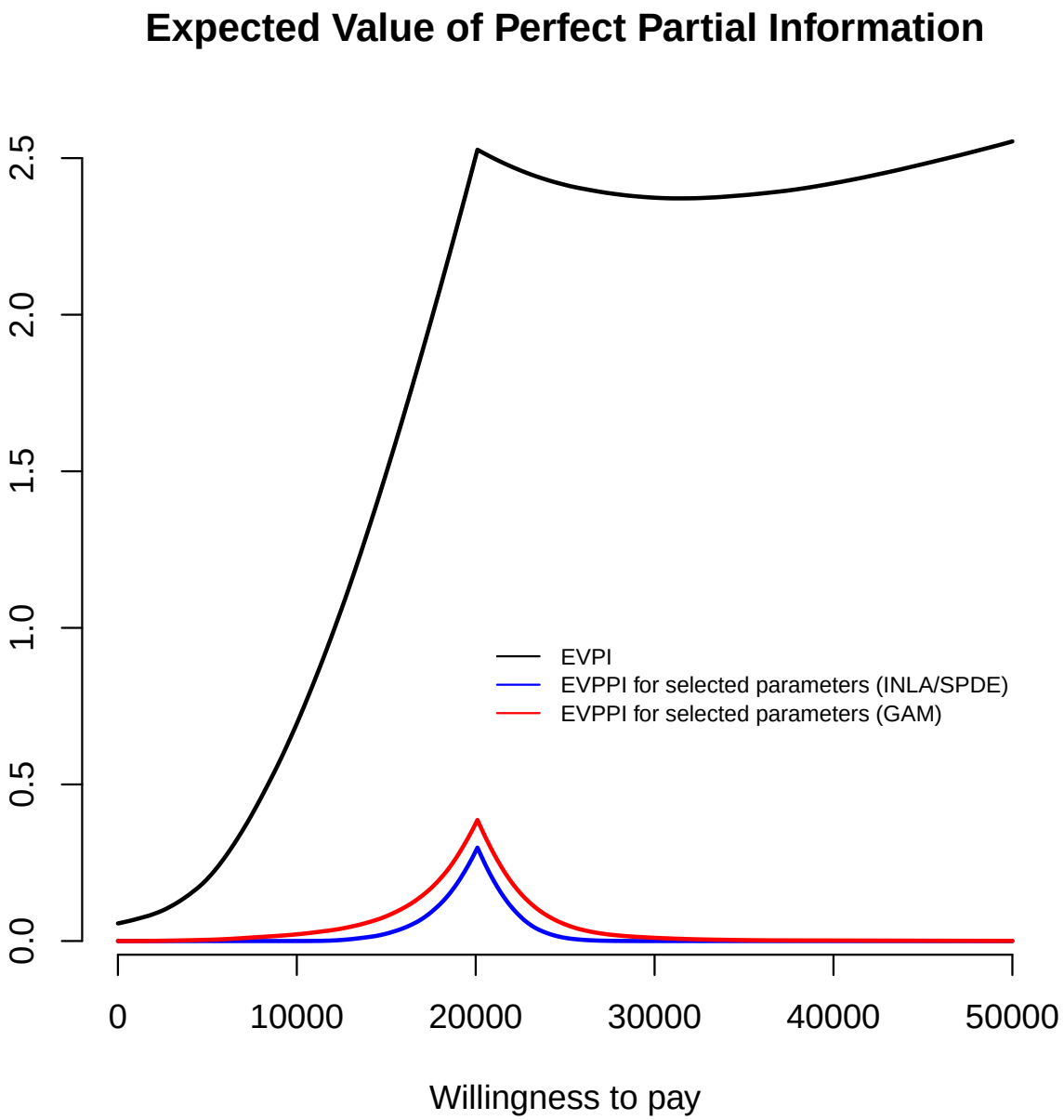
- **That's** the difficult part!
  - Can do nested Monte Carlo, but takes forever to get accurate results
  - **Recent methods** based on **Gaussian Process regression** very efficient & quick!

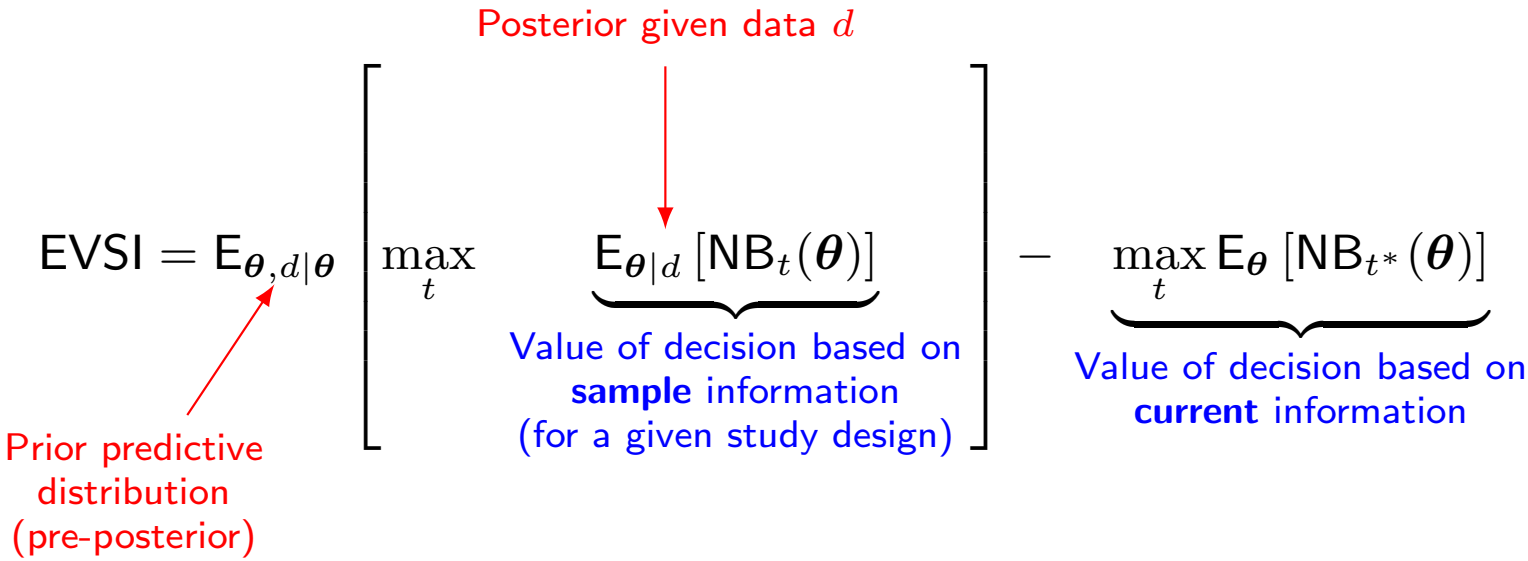
Strong et al *Medical Decision Making*. 2014; **34(3)**: 311-26.
Heath et al *Statistics in Medicine*. 2016; **35(23)**: 4264-4280.

http://savi.shef.ac.uk/SAVI/
https://egon.stats.ucl.ac.uk/projects/EVSI/

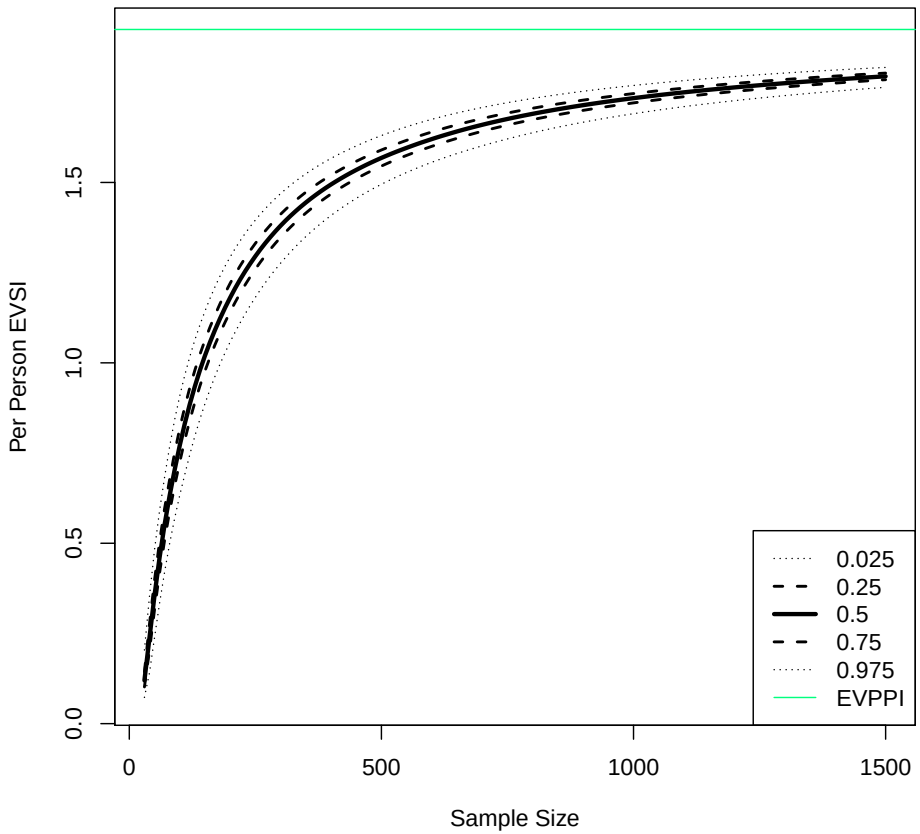
Summarising PSA + Research priority: Expected Value of **Partial** Perfect Information

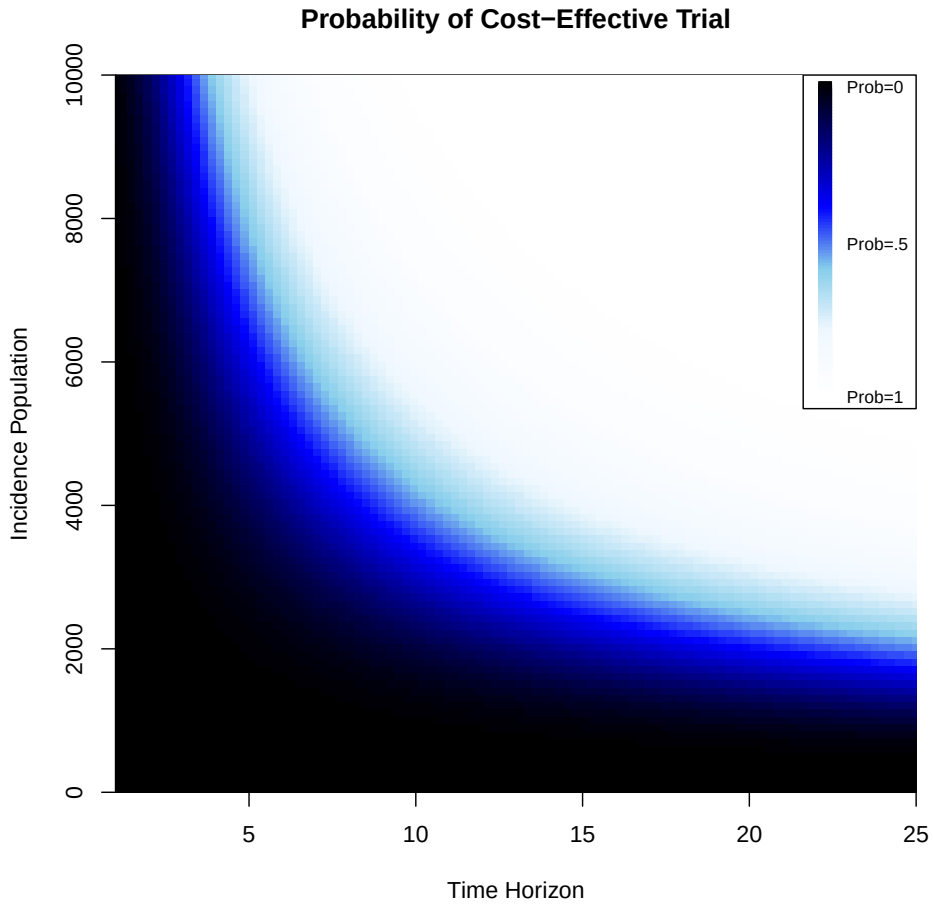



Expected Value of Perfect Partial Information
0.0
0.5
1.0
1.5
2.0
2.5
0
10000
20000
30000
40000
50000
Willingness to pay
EVPI
EVPPI for selected parameters (INLA/SPDE)
EVPPI for selected parameters (GAM)

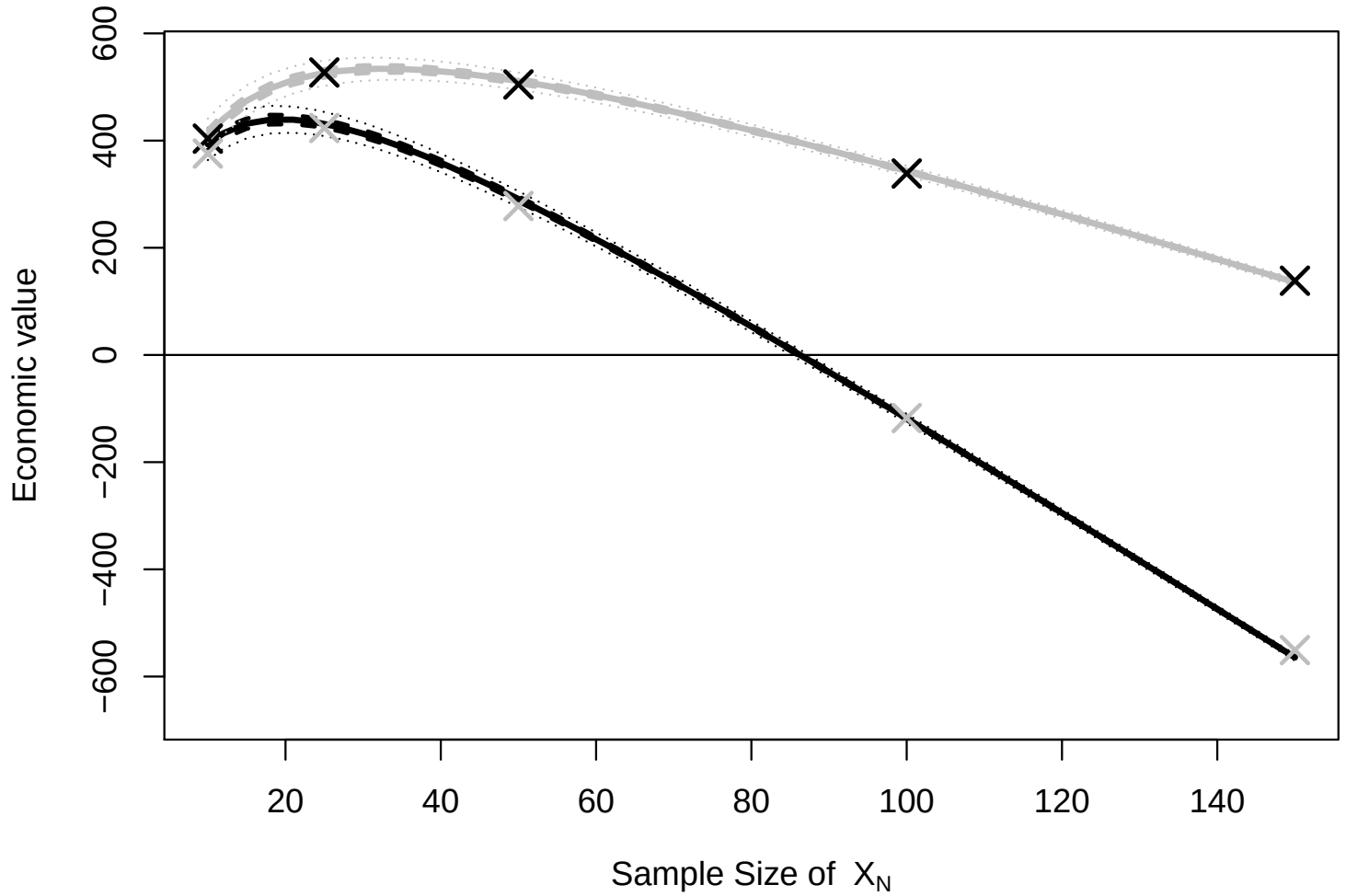


- Computationally complex
  - Requires specific knowledge of the model for (future/hypothetical) data collection
  - Again, recent methods have improved efficiency
- Can be used to drive design of new study (eg sample size calculations)





https://github.com/giabaio/EVSI
https://egon.stats.ucl.ac.uk/projects/EVSI
Heath et al (2018). https://arxiv.org/abs/1804.09590
Heath et al *Medical Decision Making*. 2017. **38(2)**: 163-173


Economic value
600
400
200
0
−200
−400
−600
20
40
60
80
100
120
140
Sample Size of X_N

End

Outline

1. Health economic evaluation
   - What is it?
   - How does it work?
2. Statistical modelling
   - Individual-level vs aggregated data
   - The importance of being a Bayesian
3. **Some examples**
   - Individual level & partially observed data
   - Survival analysis in HTA
   - Value of information
4. Conclusions

ILD+Missing data

Survival analysis

Conclusions

- Bayesian modelling particularly effective in health economic evaluations
- Allows the incorporation of external, additional information to the current analysis
  - Previous studies
  - Elicitation of expert opinions

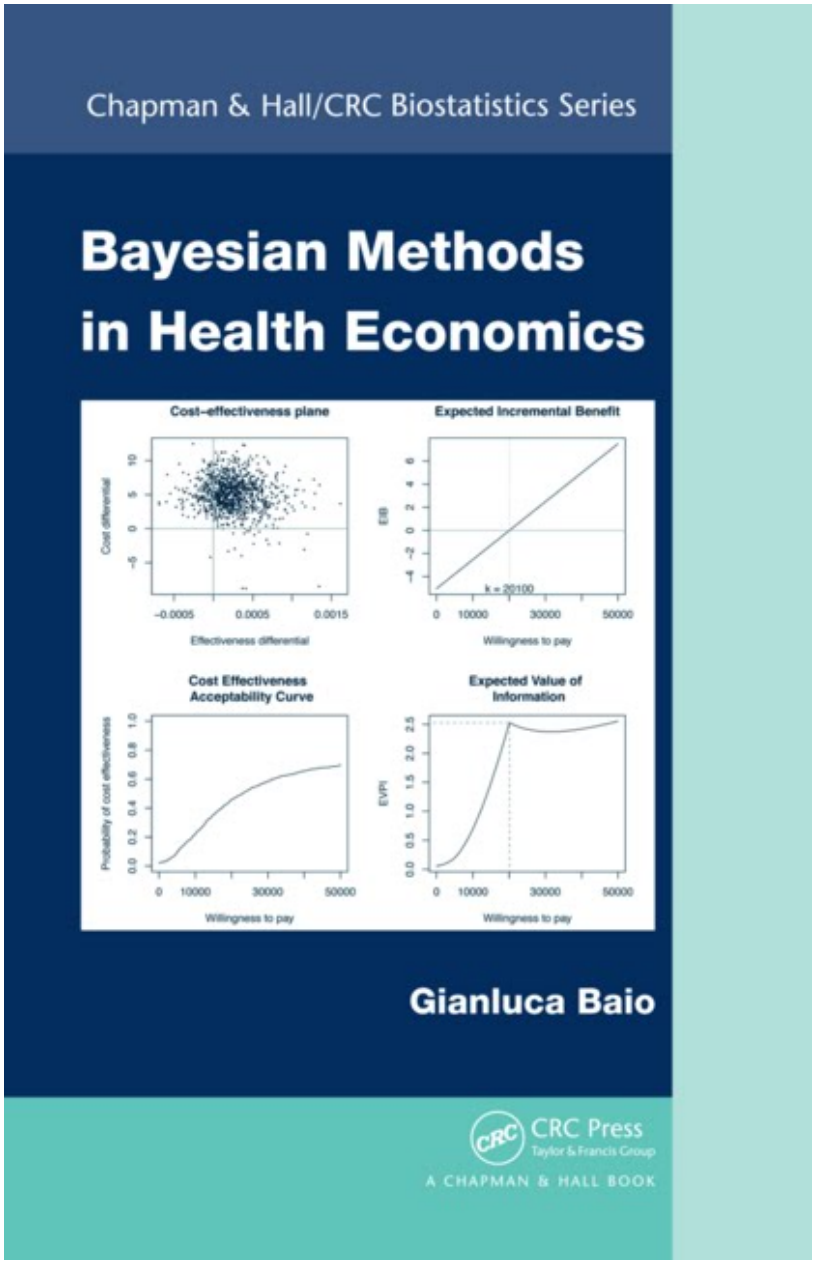
Conclusions

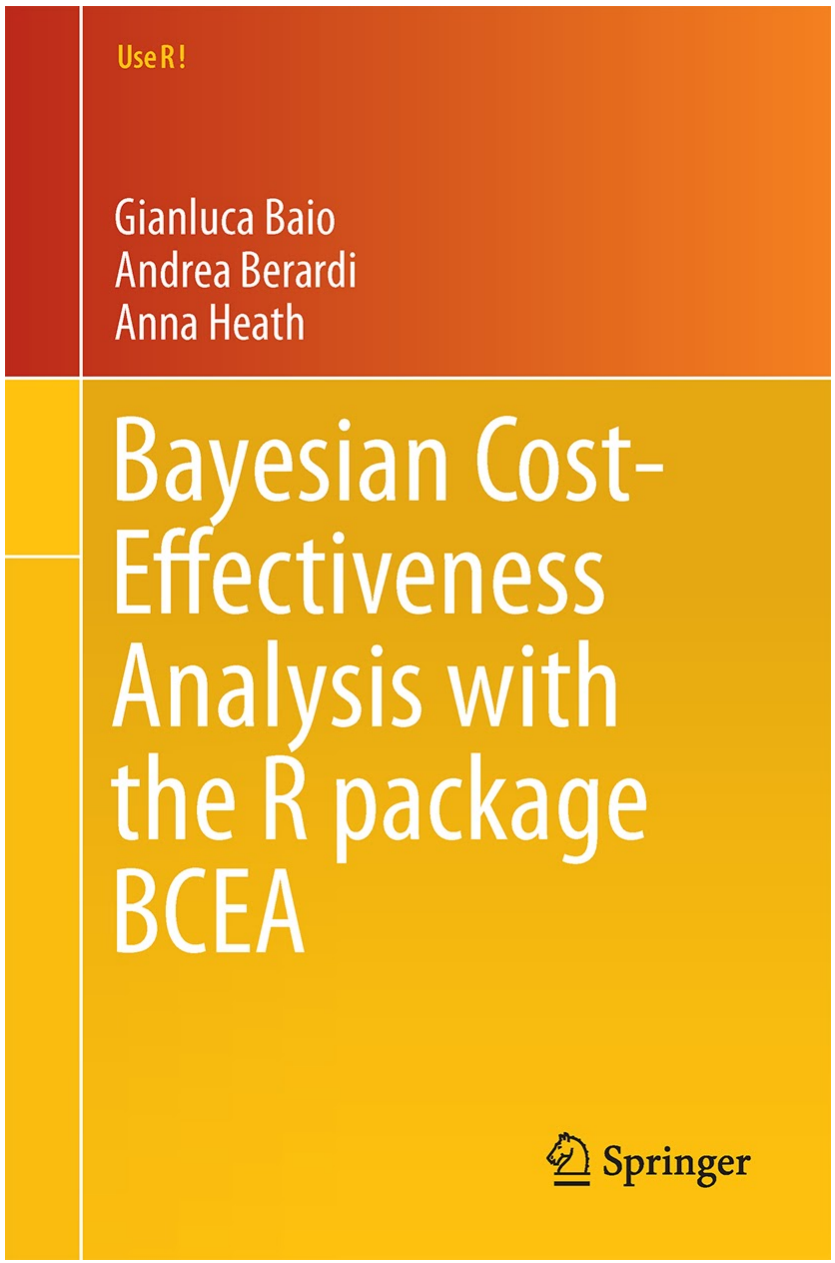
- Bayesian modelling particularly effective in health economic evaluations
- Allows the incorporation of external, additional information to the current analysis
  - Previous studies
  - Elicitation of expert opinions
- In general, Bayesian models are more flexible and allow the inclusion of complex relationships between variables and parameters
  - This is particularly effective in decision-models, where information from different sources may be combined into a single framework
  - Useful in the case of individual-level data (eg from Phase III RCT)

Conclusions

- Bayesian modelling particularly effective in health economic evaluations
- Allows the incorporation of external, additional information to the current analysis
  - Previous studies
  - Elicitation of expert opinions
- In general, Bayesian models are more flexible and allow the inclusion of complex relationships between variables and parameters
  - This is particularly effective in decision-models, where information from different sources may be combined into a single framework
  - Useful in the case of individual-level data (eg from Phase III RCT)
- Using MCMC methods, it is possible to produce the results in terms of simulations from the posterior distributions
  - These can be used to build suitable variables of cost and benefit
  - Particularly effective for running "probabilistic sensitivity analysis"

Shameless self marketing





Ευχαριστώ!