Quick & clean: computationally efficient methods for Value of Information measures

Gianluca Baio

University College London
Department of Statistical Science

g.baio@ucl.ac.uk

(Joint work with Anna Heath and Ioanna Manolopoulou)
(Thanks to Mark Strong)

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

Centre for Health Economics Seminars
University of York

Wednesday 19 September 2018

Outline

1. **Value of Information**
   - Basics
2. **EVPPI**
   - EVPPI as a (Gaussian Process) regression problem
   - Faster EVPPI (using INLA/SPDE)
3. **EVSI**
   - Brute force?
   - Moment matching
4. **Conclusions**

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?

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!

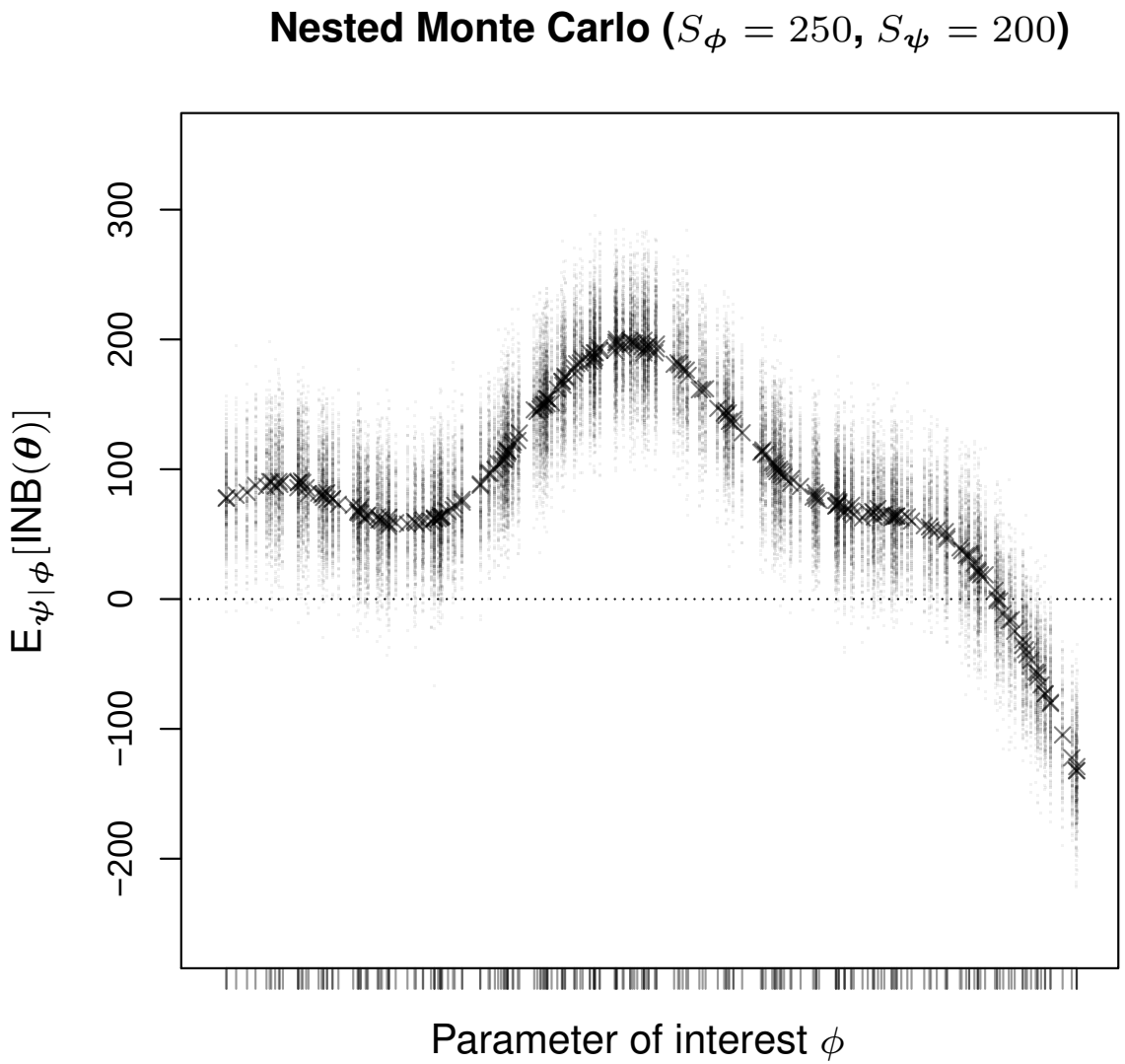
$$\text{EVPPI} = \text{E}_{\phi}\left[\max_t \text{E}_{\psi|\phi}\left[\text{NB}_t(\boldsymbol{\theta})\right]\right] - \max_t \text{E}_{\boldsymbol{\theta}}\left[\text{NB}_t(\boldsymbol{\theta})\right]$$

- **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.

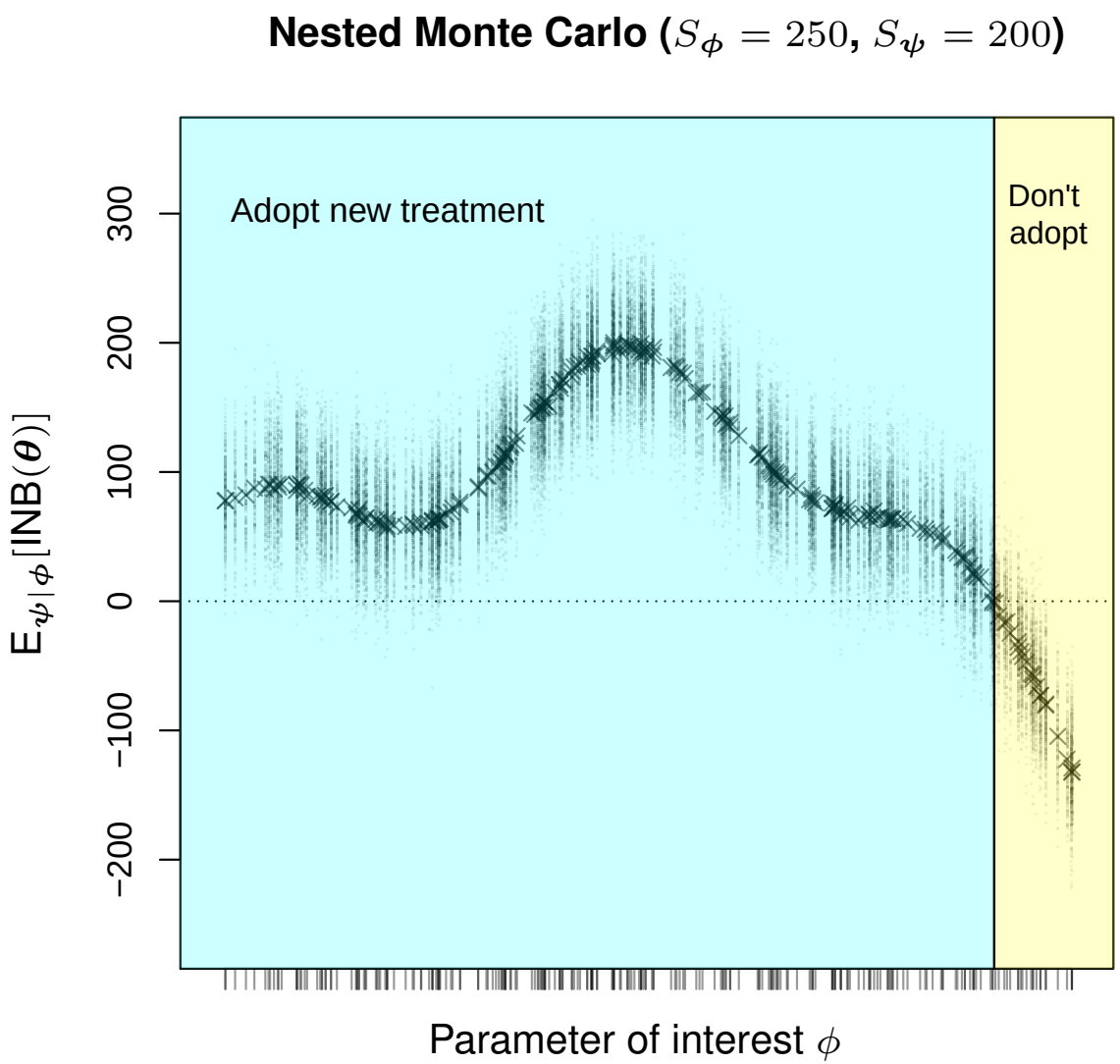
http://savi.shef.ac.uk/SAVI/
https://egon.stats.ucl.ac.uk/projects/EVSI/

Assuming only two interventions, can consider $\text{INB}(\boldsymbol{\theta}) = \text{NB}_1(\boldsymbol{\theta}) - \text{NB}_0(\boldsymbol{\theta})$



Thanks to Mark Strong (slide stolen from "Summer School in Bayesian methods in health economics")
www.statistica.it/gianluca/teaching/summer-school

Assuming only two interventions, can consider $\text{INB}(\boldsymbol{\theta}) = \text{NB}_1(\boldsymbol{\theta}) - \text{NB}_0(\boldsymbol{\theta})$



Thanks to Mark Strong (slide stolen from "Summer School in Bayesian methods in health economics")
www.statistica.it/gianluca/teaching/summer-school

- Can model as a **regression** problem

$$\begin{aligned} \text{NB}_t(\boldsymbol{\theta}) &= \text{E}_{\boldsymbol{\psi}|\boldsymbol{\phi}}\left[\text{NB}_t(\boldsymbol{\theta})\right] + \varepsilon, \qquad \text{with } \varepsilon \sim \text{Normal}(0, \sigma_\varepsilon^2) \\ &= g_t(\boldsymbol{\phi}) + \varepsilon \end{aligned}$$

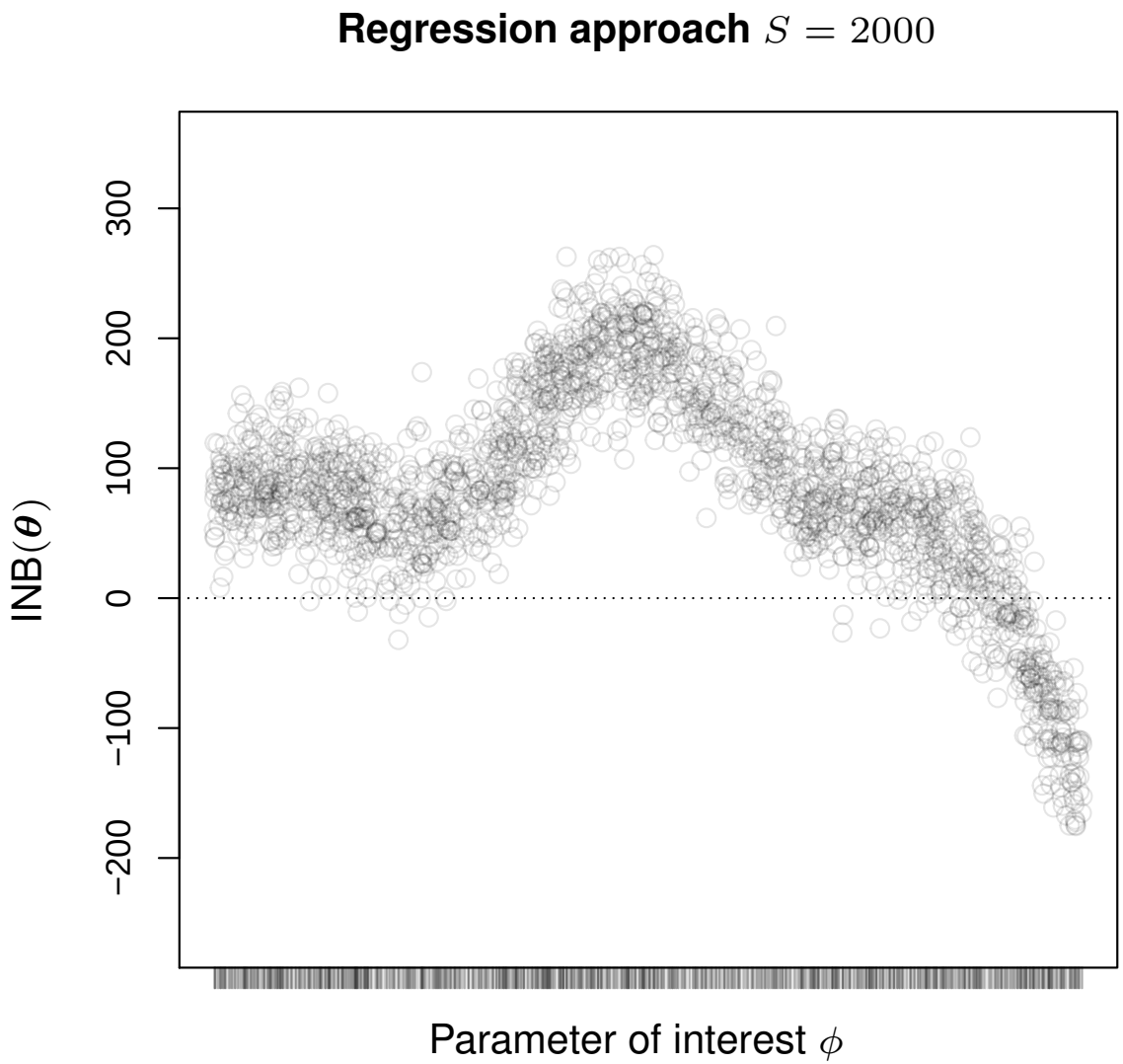
"Data": **simulations** for $\text{NB}_t(\boldsymbol{\theta})$ as "response"
simulations for $\boldsymbol{\phi}$ as "covariates"

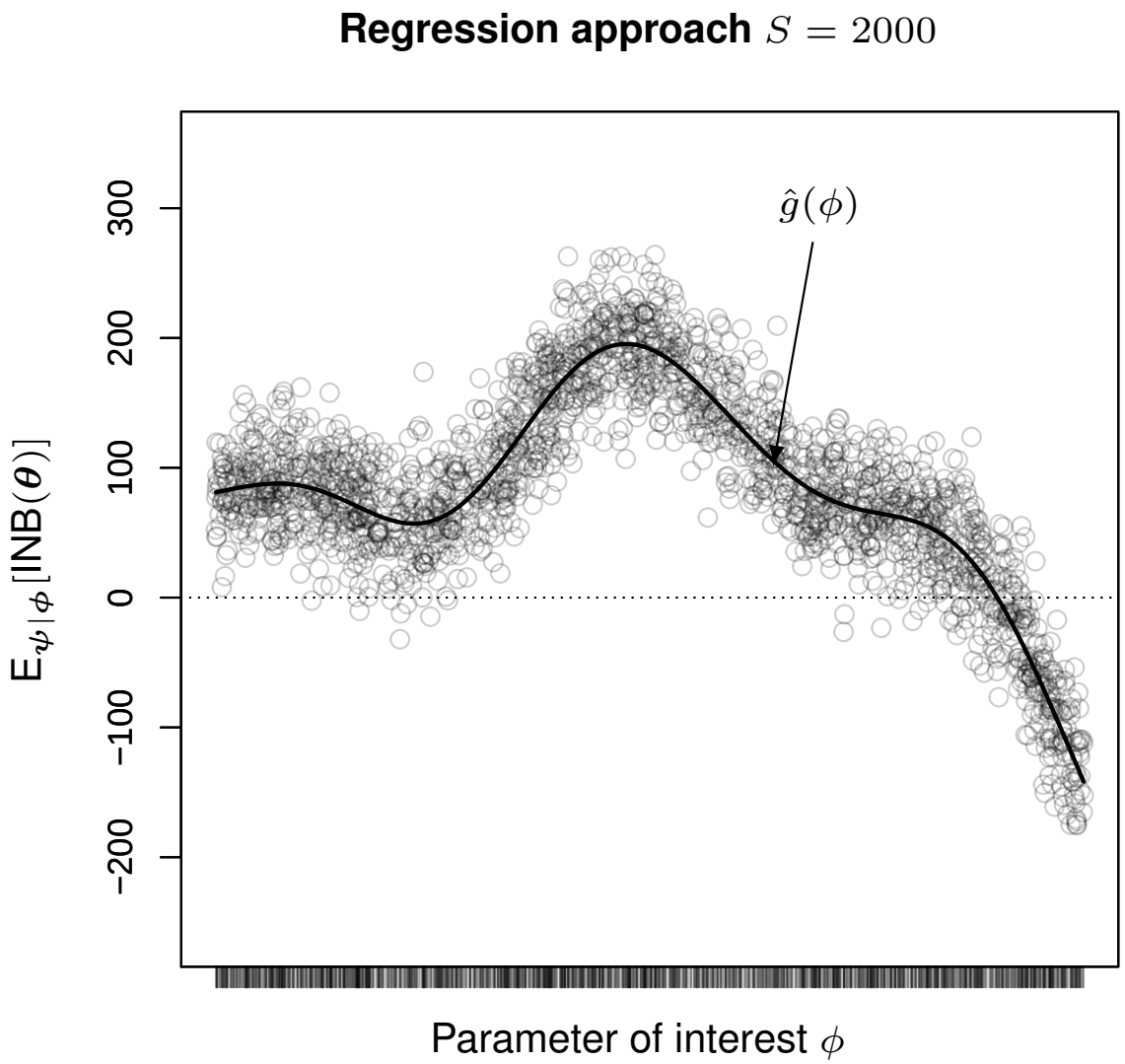
- **NB**: Only need S data points ($=$ PSA simulations), instead of $S_{\boldsymbol{\phi}} \times S_{\boldsymbol{\psi}}$!

π_0	ρ	β_0	...	σ	η	γ	$\text{NB}_0(\boldsymbol{\theta})$	$\text{NB}_1(\boldsymbol{\theta})$
0.365	0.076	0.243	...	0.622	0.001	0.162	19 214 751	19 647 706
0.421	0.024	0.115	...	0.519	0.010	0.134	17 165 526	17 163 407
0.125	0.017	0.420	...	0.482	0.007	0.149	18 710 928	16 458 433
0.117	0.073	0.419	...	0.317	0.003	0.120	16 991 321	18 497 648
0.481	0.008	0.176	...	0.497	0.004	0.191	19 772 898	18 662 329
0.163	0.127	0.227	...	0.613	0.083	0.004	17 106 136	18 983 331
	...		...		...		...	
0.354	0.067	0.318	...	0.519	0.063	0.117	18 043 921	16 470 805

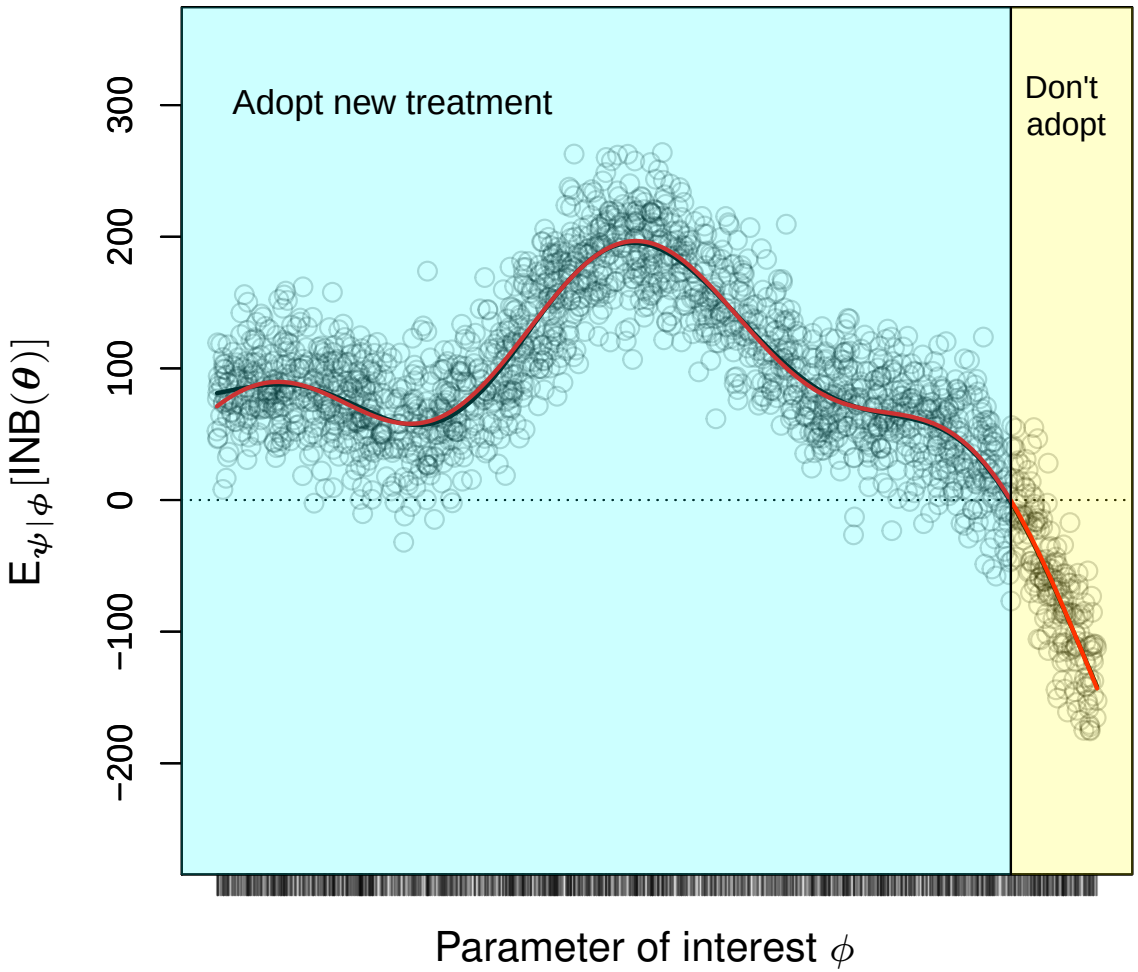
"covariates" "response" "response"

Strong et al *Medical Decision Making*. 2014; **34(3)**: 311-26


Regression approach $S = 2000$
INB(θ)
300
200
100
0
−100
−200
Parameter of interest ϕ


Regression approach $S = 2000$
$\hat{g}(\phi)$
$\mathrm{E}_{\psi|\phi}[\mathrm{INB}(\theta)]$
300
200
100
0
−100
−200
Parameter of interest ϕ

Regression approach $S = 2000$ (True relationship in red)



- Can model as a **regression** problem

$$\begin{aligned} \text{NB}_t(\boldsymbol{\theta}) &= \text{E}_{\boldsymbol{\psi}|\boldsymbol{\phi}}\left[\text{NB}_t(\boldsymbol{\theta})\right] + \varepsilon, \qquad \text{with } \varepsilon \sim \text{Normal}(0, \sigma^2_\varepsilon) \\ &= g_t(\boldsymbol{\phi}) + \varepsilon \end{aligned}$$

"Data": **simulations** for $\text{NB}_t(\boldsymbol{\theta})$ as "response"
simulations for $\boldsymbol{\phi}$ as "covariates"

- Once the functions $g_t(\boldsymbol{\phi})$ are estimated, then can approximate

$$\begin{aligned} \text{EVPPI} &= \text{E}_{\boldsymbol{\phi}}\left[\max_t \text{E}_{\boldsymbol{\psi}|\boldsymbol{\phi}}\left[\text{NB}_t(\boldsymbol{\theta})\right]\right] - \max_t \text{E}_{\boldsymbol{\theta}}\left[\text{NB}_t(\boldsymbol{\theta})\right] \\ &\approx \frac{1}{S}\sum_{s=1}^{S} \max_t \hat{g}_t(\boldsymbol{\phi}_s) - \max_t \frac{1}{S}\sum_{s=1}^{S} \hat{g}_t(\boldsymbol{\phi}_s) \end{aligned}$$

- **NB**: $g_t(\boldsymbol{\phi})$ can be complex, so need to use **flexible** regression methods
  - **GAMs**: $g_t(\boldsymbol{\phi}) = \sum_{q=1}^{Q_{\boldsymbol{\phi}}} h_t(\phi_{sq})$ $\quad h_t(\cdot) =$ smooth functions (cubic polynomials)
    very fast, but only work if number of important parameters $Q_{\boldsymbol{\phi}} \leq 5$ (interactions increase model size exponentially!)
  - If $P > 5$, can use **Gaussian Process** regression

EVPPI via **GP regression**

Model

$$\begin{pmatrix} \text{NB}_t(\boldsymbol{\theta}_1) \\ \text{NB}_t(\boldsymbol{\theta}_2) \\ \vdots \\ \text{NB}_t(\boldsymbol{\theta}_S) \end{pmatrix} := \mathbf{NB}_t \sim \text{Normal}(\boldsymbol{H\beta}, \boldsymbol{\mathcal{C}}_{\text{Exp}} + \sigma_\varepsilon^2 \boldsymbol{I})$$

$$\boldsymbol{H} = \begin{pmatrix} 1 & \phi_{11} & \cdots & \phi_{1P} \\ 1 & \phi_{21} & \cdots & \phi_{2P} \\ \vdots & & \ddots & \\ 1 & \phi_{S1} & \cdots & \phi_{SP} \end{pmatrix} \quad \text{and} \quad \mathcal{C}_{\text{Exp}}(r,s) = \sigma^2 \exp\left[\sum_{p=1}^{P} \left(\frac{\phi_{rp} - \phi_{sp}}{\delta_p}\right)^2\right]$$

- Parameters: $\boldsymbol{\beta}$, $\boldsymbol{\delta}$, σ^2, σ_ε^2
- Very flexible structure — good approximation level
- Can use conjugate priors + numerical optimisation, **but** can still be very slow — computational cost in the order of S^3 (involves inversion of a dense covariance matrix)

EVPPI via GP regression — **but faster**

1. Build from ideas in spatial statistics and use a Matérn covariance function

$$\mathcal{C}_{\mathrm{M}}(r,s) = \frac{\sigma^2}{\Gamma(\nu)2^{\nu-1}}(\kappa\|\boldsymbol{\phi}_r - \boldsymbol{\phi}_s\|)^{\nu}\mathrm{K}_{\nu}(\kappa\|\boldsymbol{\phi}_r - \boldsymbol{\phi}_s\|)$$

   - Fewer parameters, but still implies a dense covariance matrix
   - **But**: can use efficient algorithms to solve **Stochastic Partial Differential Equations** (SPDE) to approximate it — with computational cost $\propto S^{3/2}$!

2. Re-formulate the model as

$$\begin{aligned} \mathbf{NB}_t &\sim \text{Normal}(\boldsymbol{H\beta}, \boldsymbol{\mathcal{C}}_{\mathrm{M}} + \sigma^2_{\varepsilon}\boldsymbol{I}) \\ &\sim \text{Normal}(\boldsymbol{H\beta} + f(\boldsymbol{\omega}), \sigma^2_{\varepsilon}\boldsymbol{I}) \end{aligned}$$

   - $f(\boldsymbol{\omega})$ are a set of "spatially structured" effects, with $\boldsymbol{\omega} \sim \text{Normal}\left(0, \boldsymbol{Q}^{-1}(\xi)\right)$
   - $\boldsymbol{Q}(\xi)$ is a **sparse** precision matrix determined by the SPDE solution

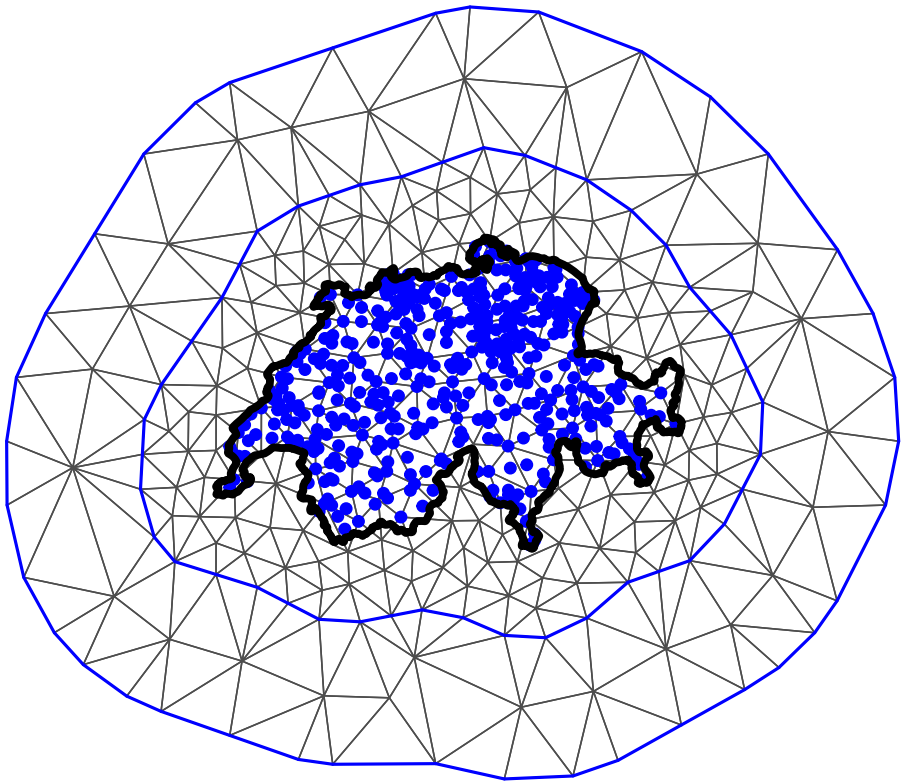
3. Crucially, if we set a sparse Gaussian prior on $\boldsymbol{\beta}$, this is a Latent Gaussian model $\Rightarrow$ can be estimated using super-fast **Integrated Nested Laplace Approximation** (INLA)

NB: Both methods implemented in the `R` package **`BCEA`** (Bayesian Cost-Effectiveness Analysis)
`http://www.statistica.it/gianluca/BCEA` `https://github.com/giabaio/BCEA`

Heath et al *Stats in Med.* 2016; **35(23)**: 4264-4280; Lindgren et al *JRSS/B.* 2011; **73(4)**: 423-498; Rue et al *JRSS/B.* 2009; **71**: 319-392

Lost in space

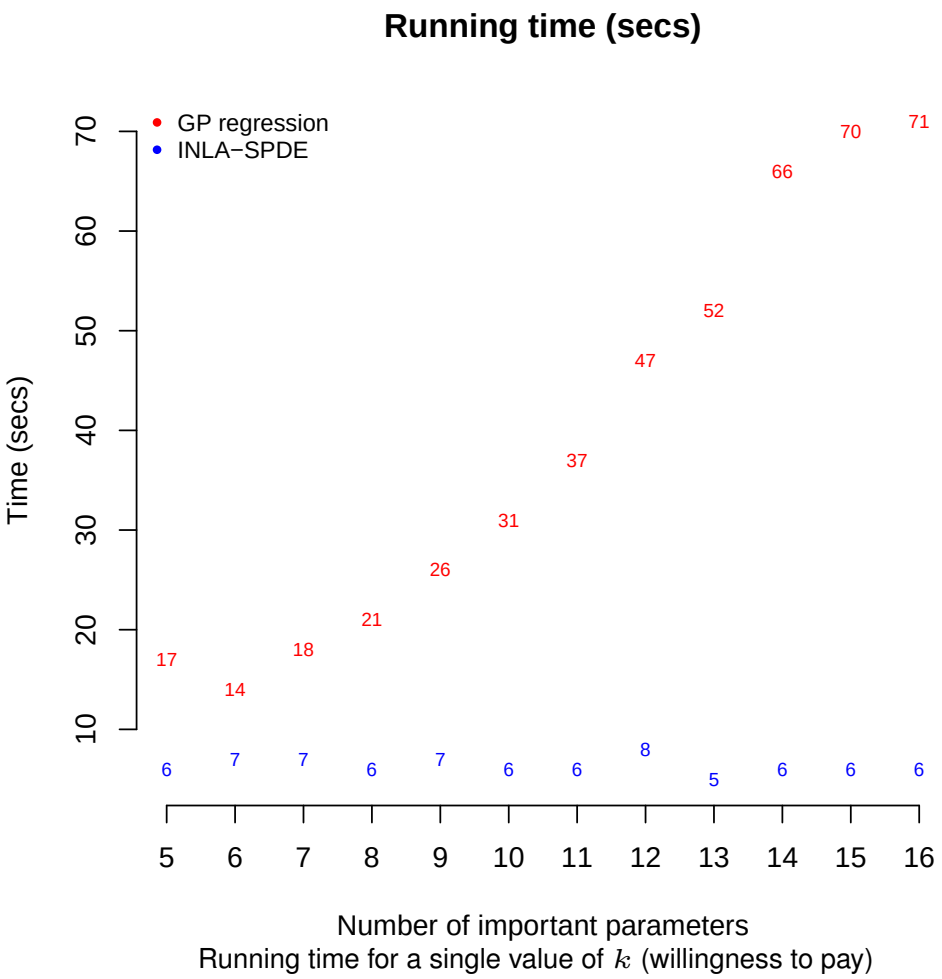
- In a "proper" spatial problem, data are observed at a bivariate grid of points
  - Points that are closer tend to be more correlated than points further apart
  - The INLA-SPDE procedure builds a grid approximation of the underlying bidimensional space
  - Points not on the grid are estimated by interpolation — deriving a full surface

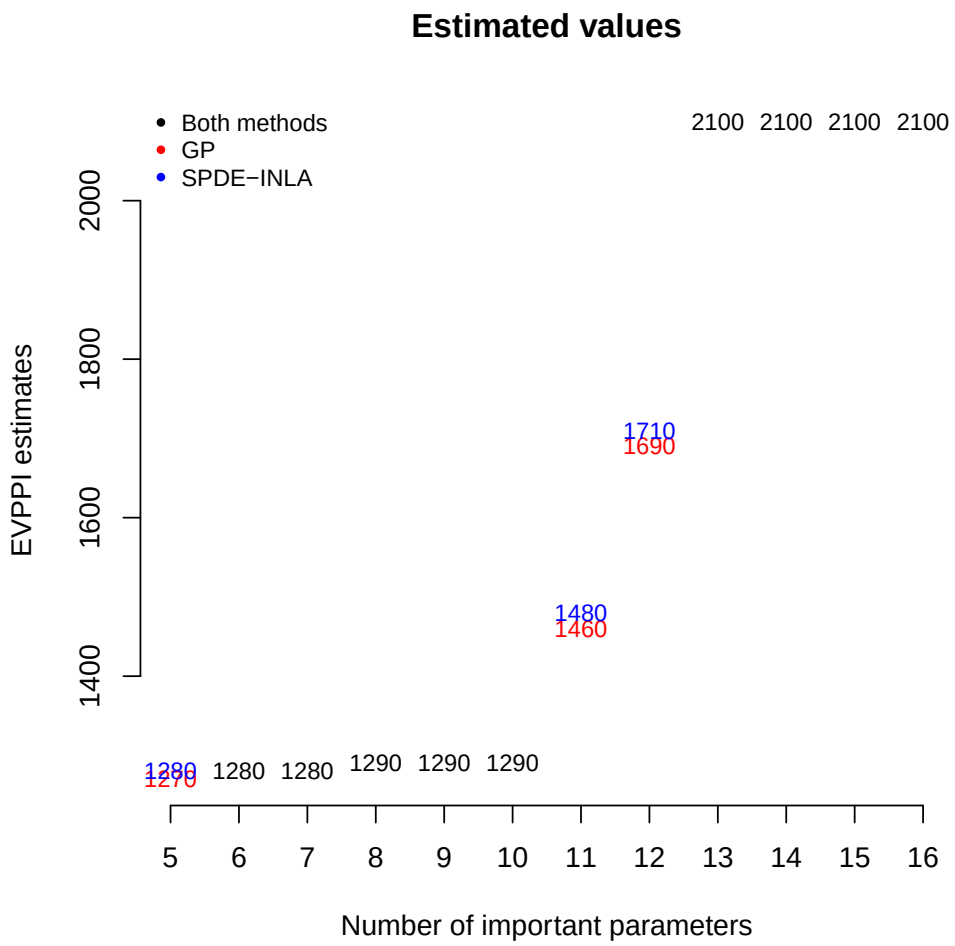


Lost in space

- In a "proper" spatial problem, data are observed at a bivariate grid of points
  - Points that are closer tend to be more correlated than points further apart
  - The INLA-SPDE procedure builds a grid approximation of the underlying bidimensional space
  - Points not on the grid are estimated by interpolation — deriving a full surface
- In our case, data are observed on a high-dimensional space, with no proper "spatial" interpretation!
- Need to use some form of **dimensionality reduction** to project the P-dimensional space of $\boldsymbol{\phi}$ to a 2-dimensional space
  - Simple solution: use PCA to preserve Euclidean distances and thus capture the "spatial" correlation across the elements of $\boldsymbol{\phi}$
  - Even better, regression-based dimension reduction method: **Principal Fitted Components**
    1. Estimate the function $R(\boldsymbol{\phi}) : P \rightarrow d$ so that $\mathbf{NB}_t \perp\!\!\!\perp \boldsymbol{\phi} \mid R(\boldsymbol{\phi})$
    2. "**Project**" the $P-$dimensional information contained in $\boldsymbol{\phi}$ to the $d-$dimensional function $R(\cdot)$
    3. Ideally, $d << P$ — in fact, would like $d \leq 2$
  - Computational cost is negligible
  - Can use model-fitting statistics (eg AIC) to determine the "best" model for given choices of d $(= 2, 3, \ldots)$
  - **NB**: if the AIC suggests $d > 2$ then EVPPI estimates likely to be biased!

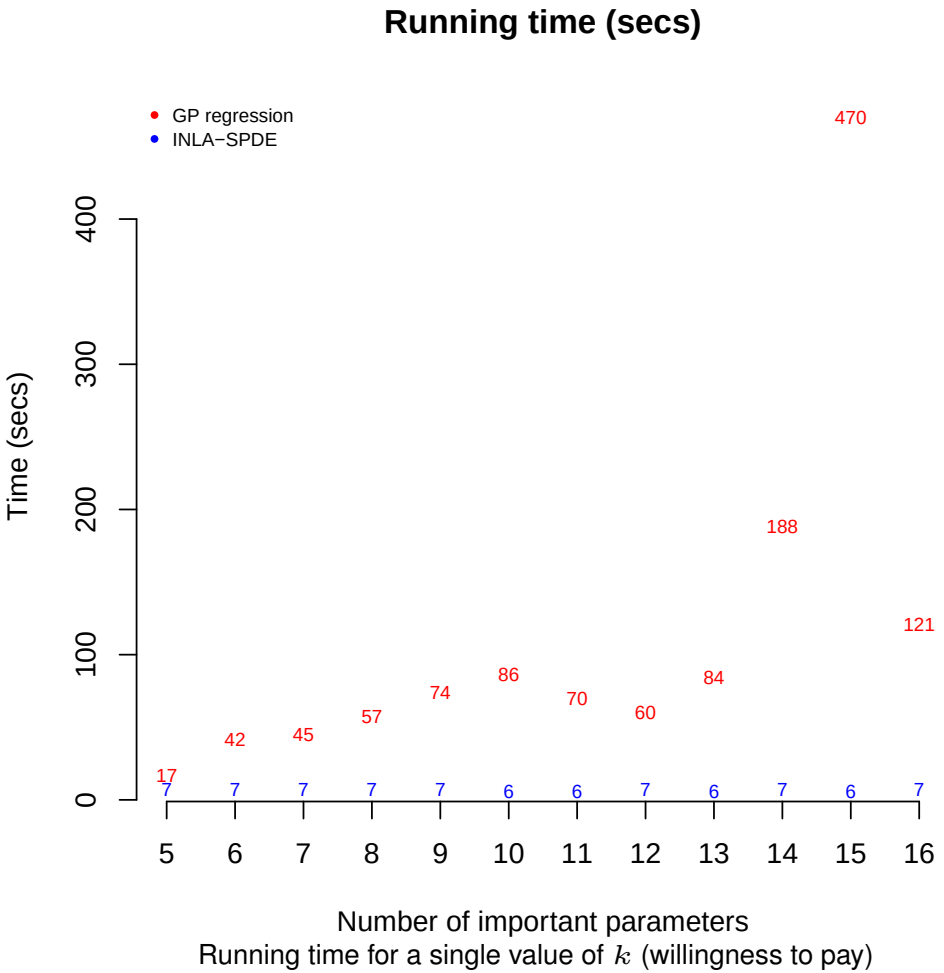
Examples

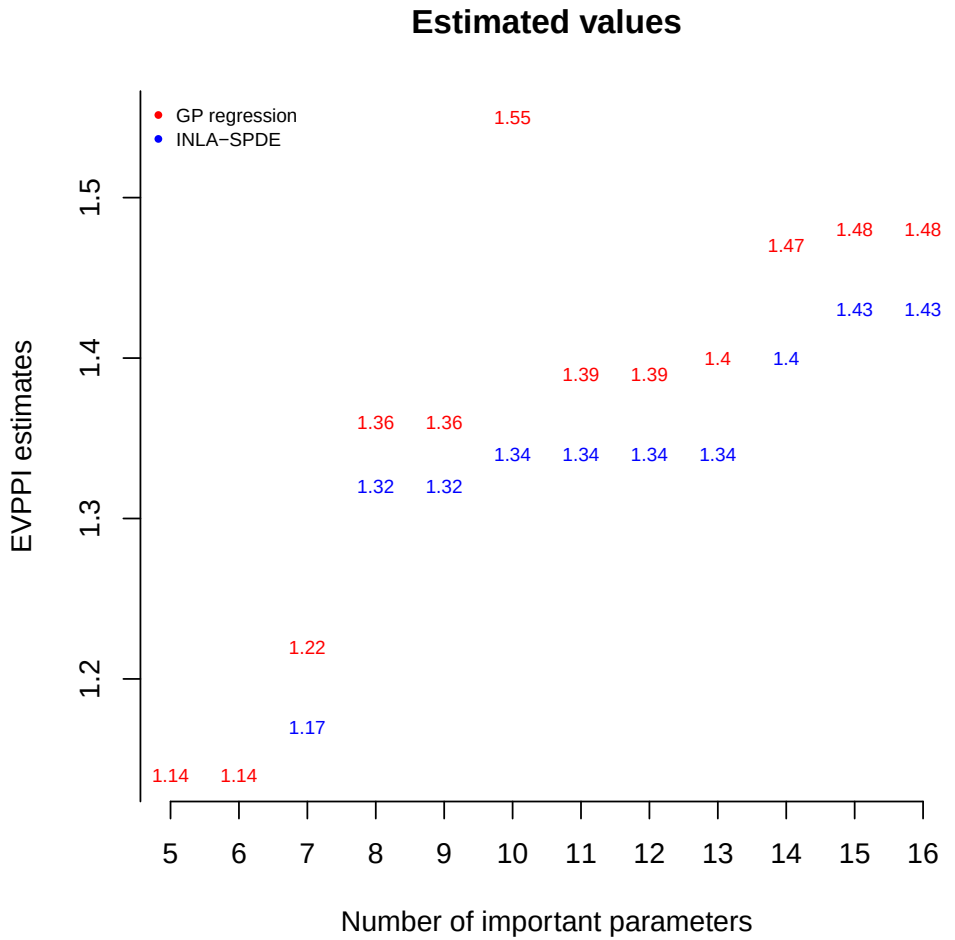




- Fictional decision tree model with correlated parameters
- 2 treatment options and overall 19 parameters
- Parameters simulated from multivariate Normal distribution, so can compute exact EVPPI

Heath et al *Statistics in Medicine*. 2016; **35(23)**: 4264-4280





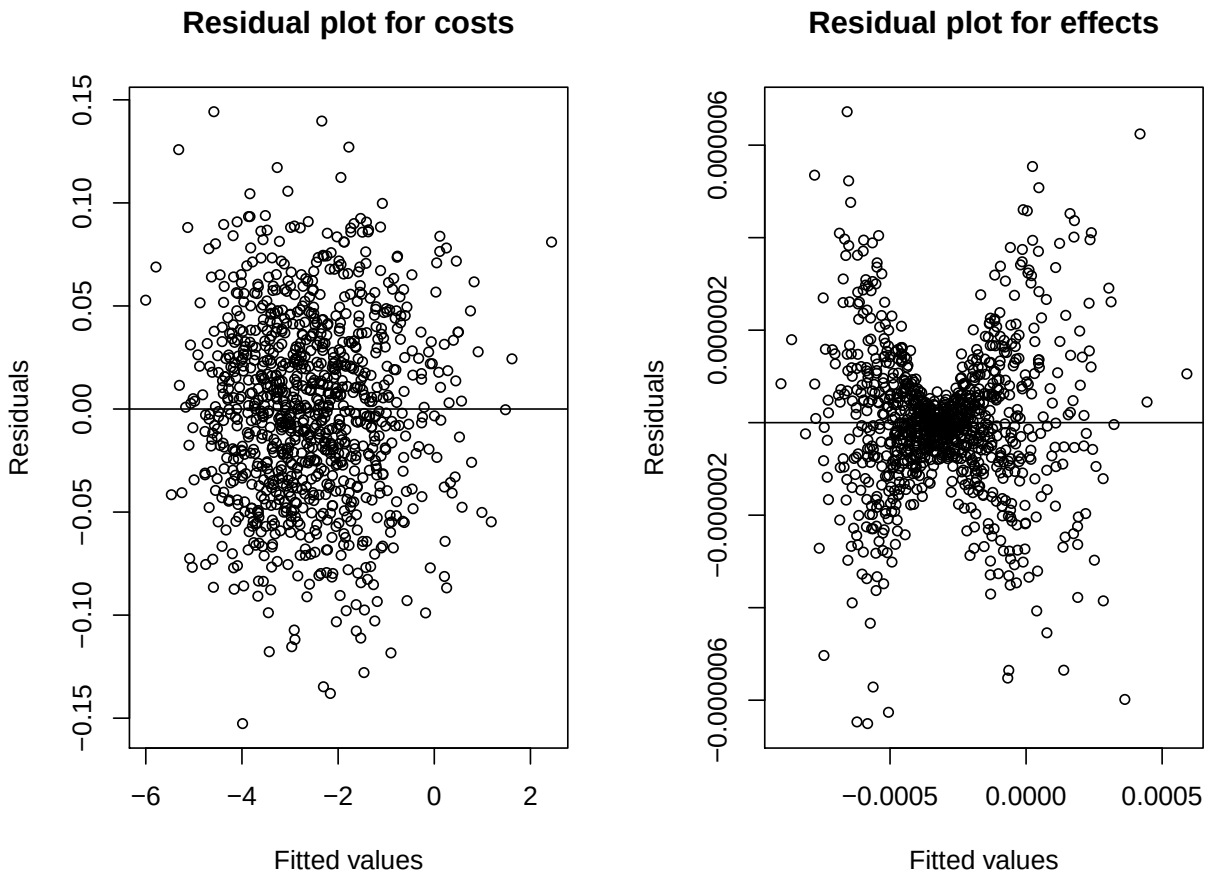
- Cost-effectiveness model for influenza vaccine based on evidence synthesis
- 2 treatment options and overall 63 parameters
- Model not available in closed form (needs MCMC simulations)

Heath et al *Statistics in Medicine*. 2016; **35(23)**: 4264-4280

Breaking bad...

Breast cancer screening (Welton et al. 2008. *JRSS/A*)

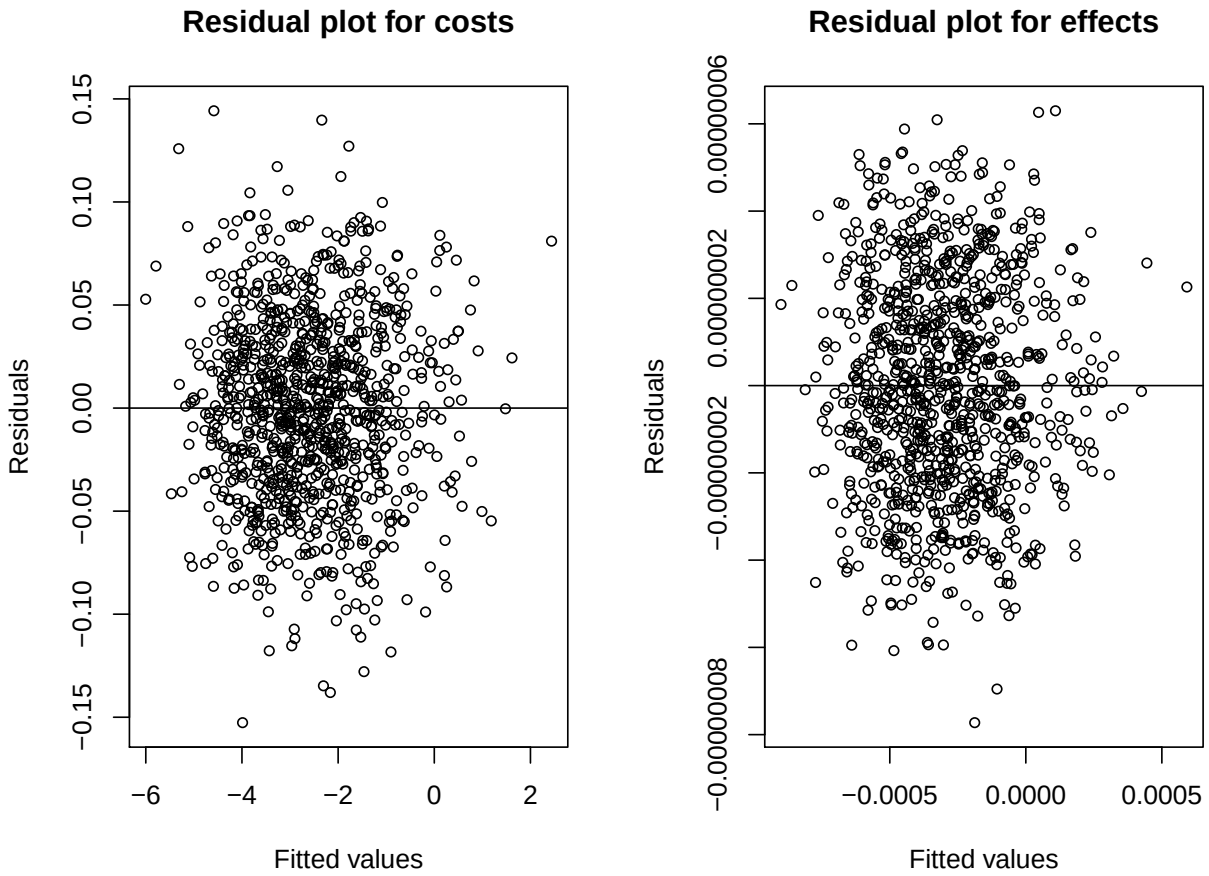
- Multi-decision model developed for the UK setting, with 4 interventions
- Complex evidence synthesis for 6 parameters — highly structured!



The fix!

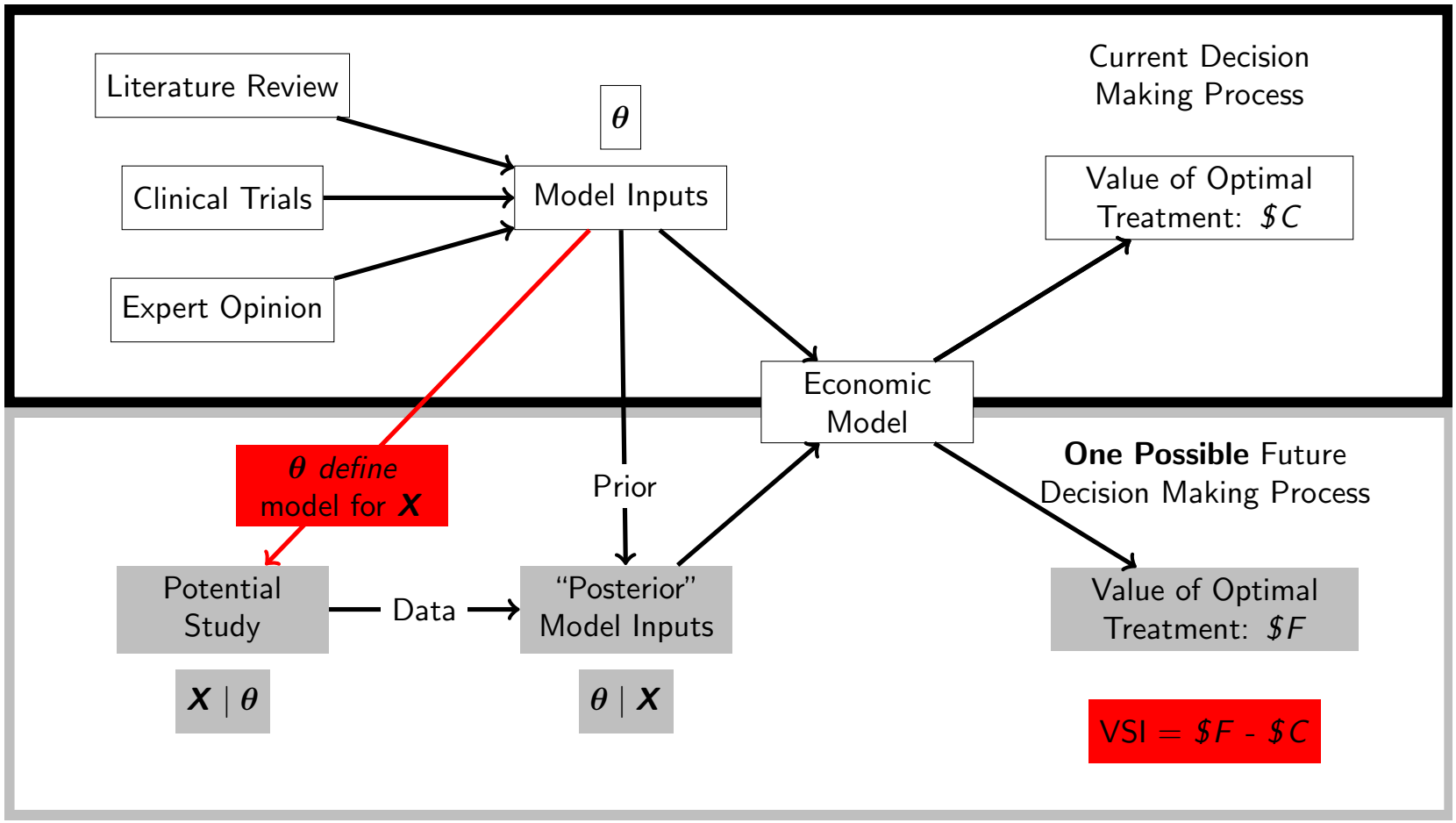
- Can relatively easily modify the basic structure of the model, e.g. include interaction terms to make $\boldsymbol{H\beta}$ non-linear

$$\beta_0 + \beta_1\phi_{1s} + \beta_2\phi_{2s} + \beta_3\phi_{3s} + \beta_4\phi_{1s}\phi_{2s} + \beta_5\phi_{1s}\phi_{3s} + \beta_6\phi_{2s}\phi_{3s}$$

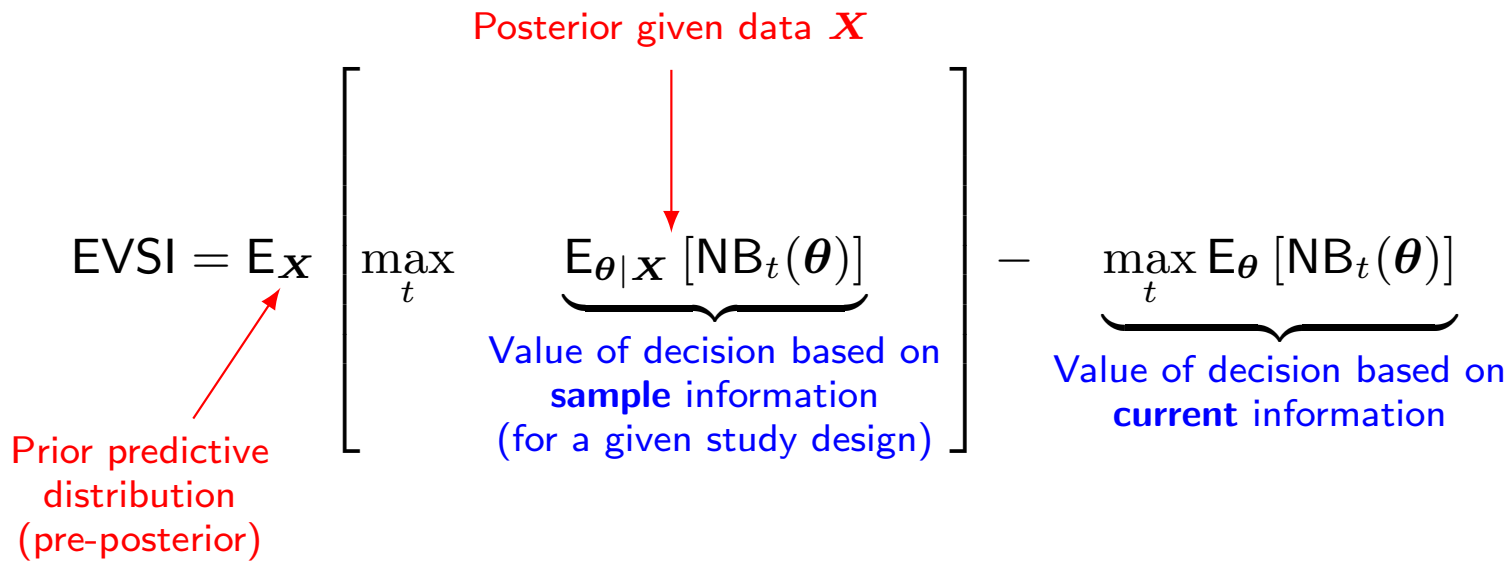


Baio et al (2017). *Bayesian Cost-Effectiveness Analysis with the R package BCEA*

- EVSI measures the value of reducing uncertainty by running a study of a given design
- Can compare the benefits and costs of a study with given design
  - To see if a proposed study likely to be a good use of resources
  - To find the optimal study design



- EVSI measures the value of reducing uncertainty by running a study of a given design
- Can compare the benefits and costs of a study with given design
  - To see if a proposed study likely to be a good use of resources
  - To find the optimal study design



- 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)

- EVSI measures the value of reducing uncertainty by running a study of a given design
- Can compare the benefits and costs of a study with given design
  - To see if a proposed study likely to be a good use of resources
  - To find the optimal study design

Posterior given data $\boldsymbol{X}$

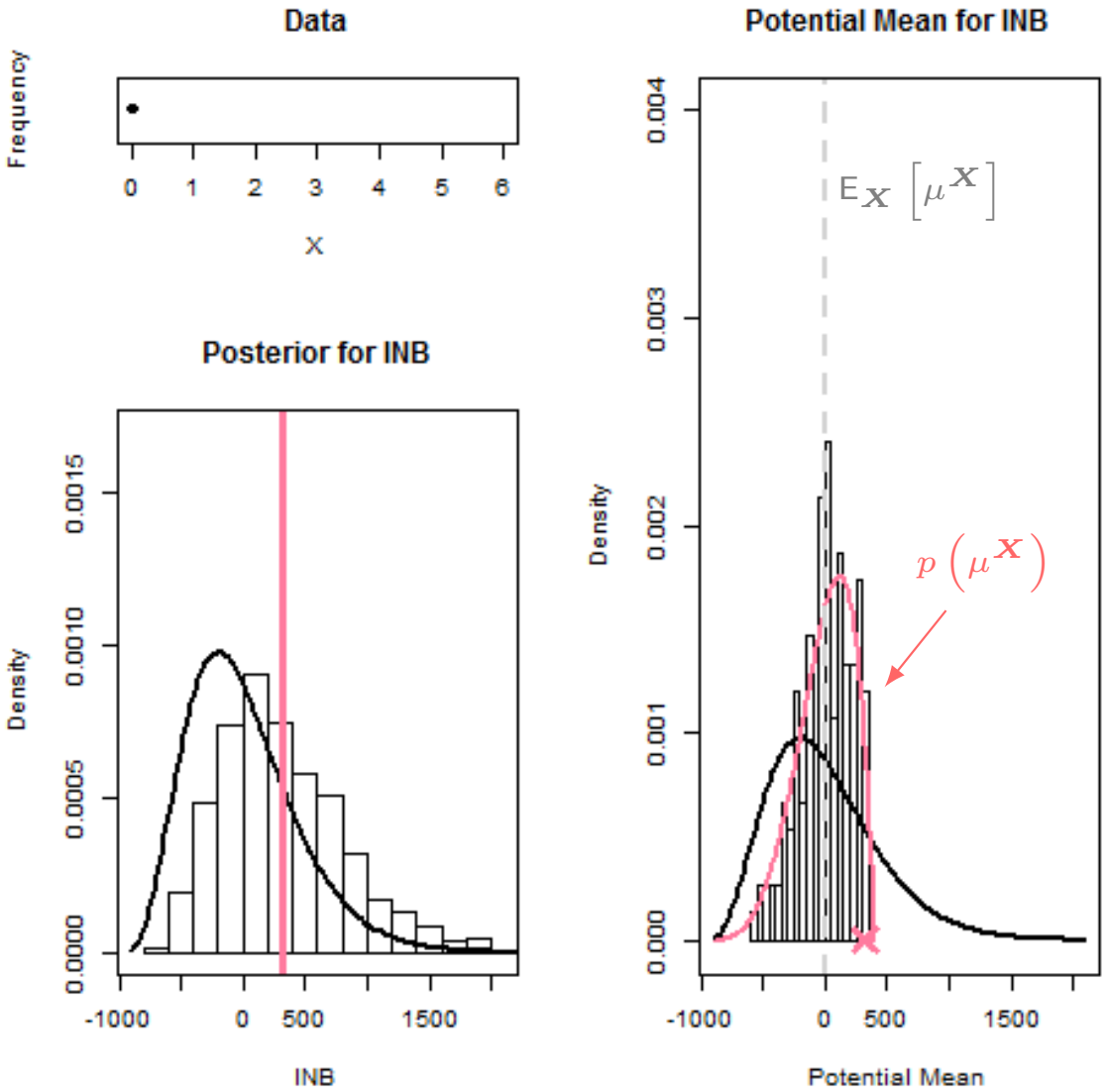
$$\text{EVSI} = \mathrm{E}_{\boldsymbol{X}}\left[\max_t \underbrace{\mathrm{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\left[\text{NB}_t(\boldsymbol{\theta})\right]}_{\text{Value of decision based on } \textbf{sample} \text{ information (for a given study design)}}\right] - \underbrace{\max_t \mathrm{E}_{\boldsymbol{\theta}}\left[\text{NB}_t(\boldsymbol{\theta})\right]}_{\text{Value of decision based on } \textbf{current} \text{ information}}$$

Prior predictive distribution (pre-posterior)

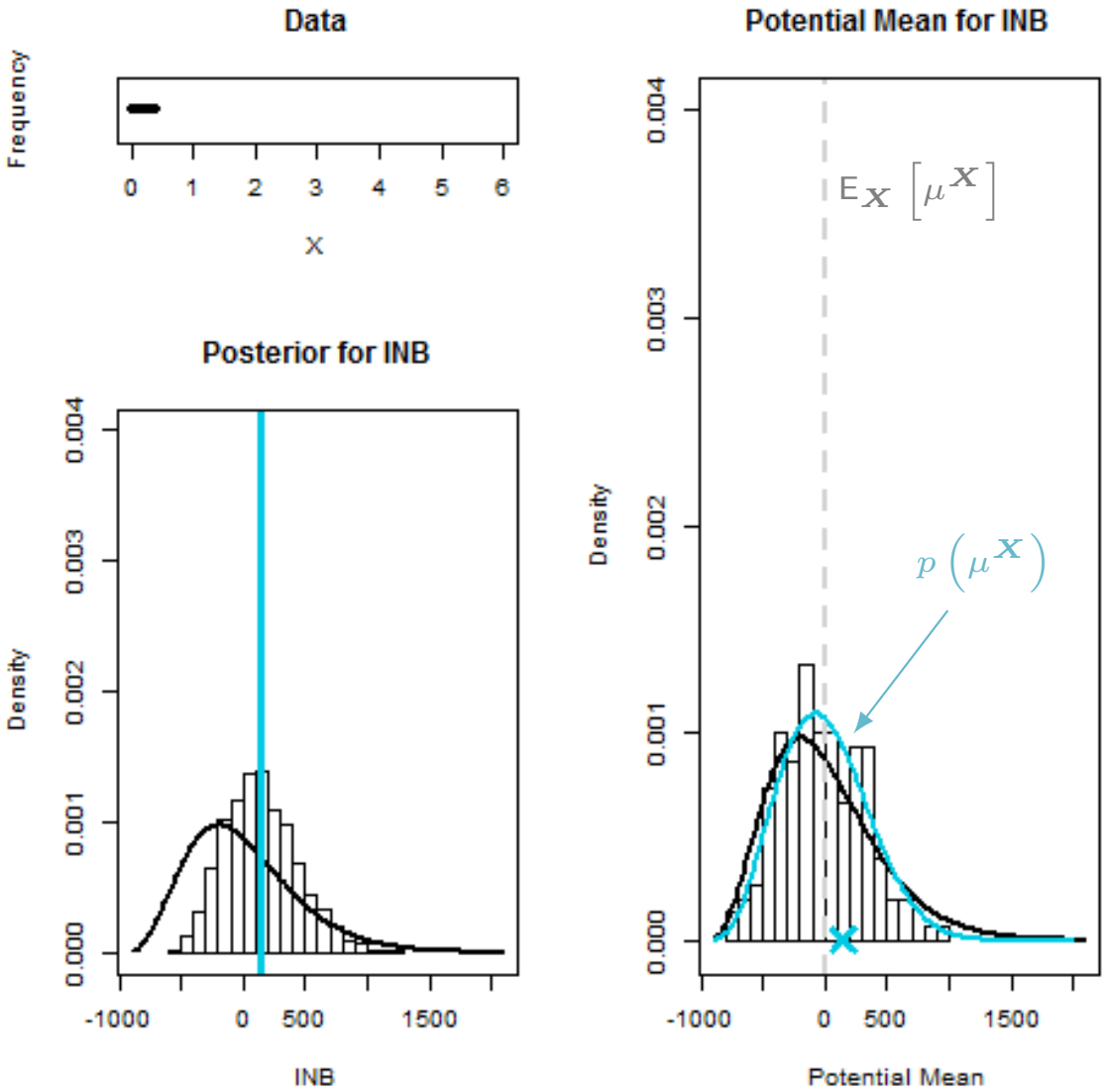
- Assuming only two interventions, can re-express as

$$\text{EVSI} = \mathrm{E}_{\boldsymbol{X}}\left[\max\left\{0, \underbrace{\mathrm{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\left[\text{INB}(\boldsymbol{\theta})\right]}_{\mu^{\boldsymbol{X}}}\right\}\right] - \max\left\{0, \mathrm{E}_{\boldsymbol{\theta}}\left[\text{INB}(\boldsymbol{\theta})\right]\right\}$$

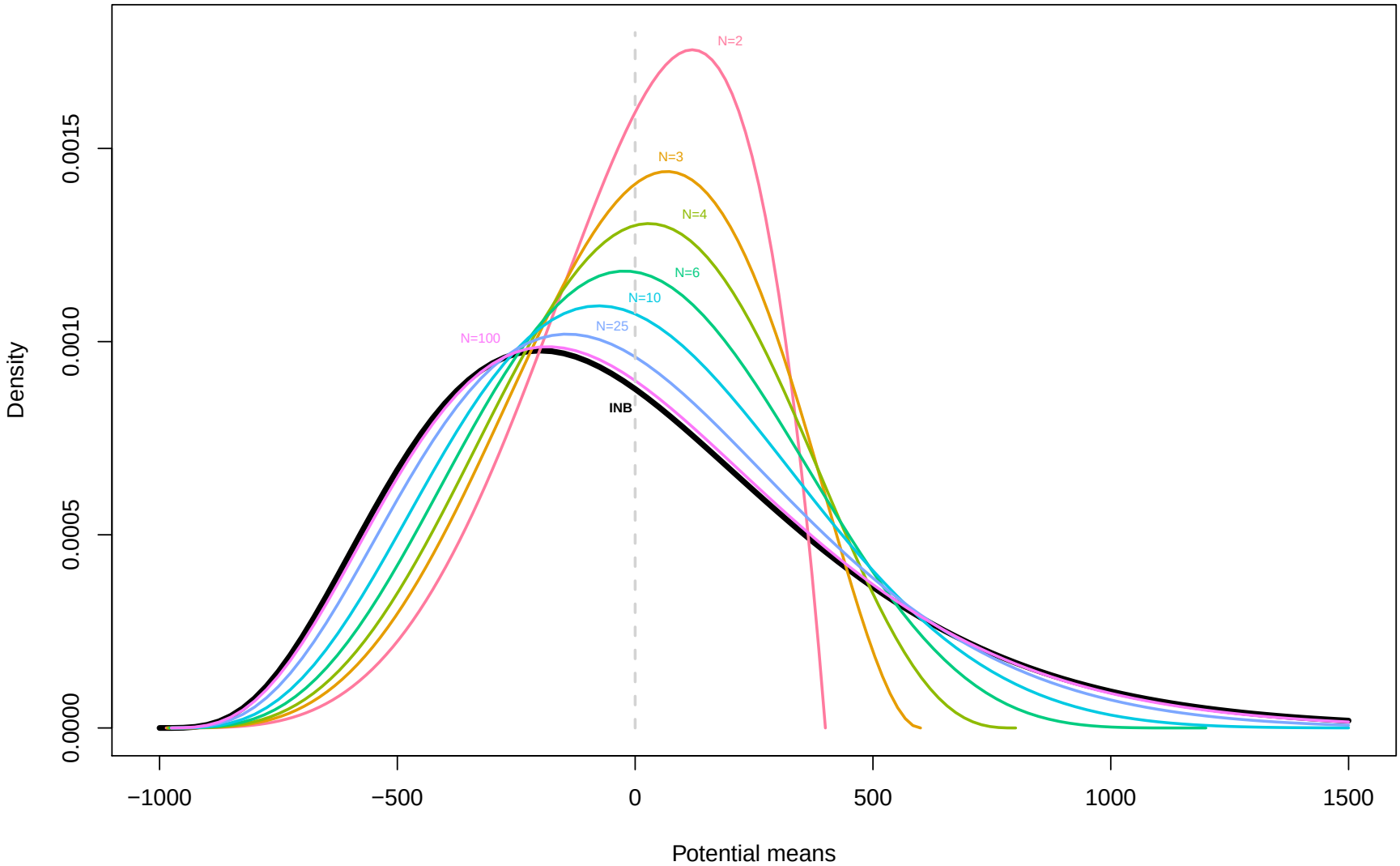
New study sample size: $N = 2$



New study sample size: $N = 10$



A counter intuitive relationship...



Moment matching

Objective: Estimate the distribution $p(\mu^{\boldsymbol{X}})$ with $\mu^{\boldsymbol{X}} = \mathrm{E}_{\boldsymbol{\theta}|\boldsymbol{X}}[\mathrm{INB}(\boldsymbol{\theta})]$

- That's the hard part to estimate the EVSI

We know that

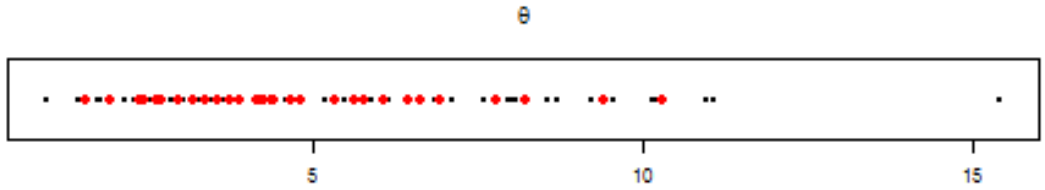
1. As $n \to \infty$, $p(\mu^{\boldsymbol{X}})$ is "similar" to the PSA distribution of $\mathrm{INB}(\boldsymbol{\theta})$
2. $\mathrm{E}_{\boldsymbol{X}}\left[\mu^{\boldsymbol{X}}\right] = \mathrm{E}_{\boldsymbol{X}}\left[\mathrm{E}_{\boldsymbol{\theta}|\boldsymbol{X}}\left[\mathrm{INB}(\boldsymbol{\theta})\right]\right] = \mathrm{E}_{\boldsymbol{\theta}}\left[\mathrm{INB}(\boldsymbol{\theta})\right]$
3. $\mathrm{Var}_{\boldsymbol{X}}\left[\mu^{\boldsymbol{X}}\right] = \underbrace{\mathrm{Var}_{\boldsymbol{\theta}}\left[\mathrm{INB}(\boldsymbol{\theta})\right]}_{\text{PSA variance for INB}(\boldsymbol{\theta})} - \underbrace{\mathrm{E}_{\boldsymbol{X}}\left[\mathrm{Var}_{\boldsymbol{\theta}|\boldsymbol{X}}\left[\mathrm{INB}(\boldsymbol{\theta})\right]\right]}_{\text{Posterior variance for INB}(\boldsymbol{\theta})}$

Idea: can approximate the unknown distribution $p(\mu^{\boldsymbol{X}})$ by rescaling the PSA distribution for $\mathrm{INB}(\boldsymbol{\theta})$, moment-matching it to the mean and variance defined above

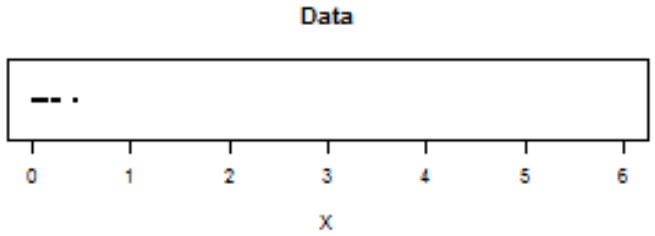
- All we need is to estimate the expected posterior variance...
- Can do this efficiently by only using $Q \approx 30$ to $50 << S$ PSA simulations!

Heath et al. 2017. *Medical Decision Making*. **38(2)**: 163-173

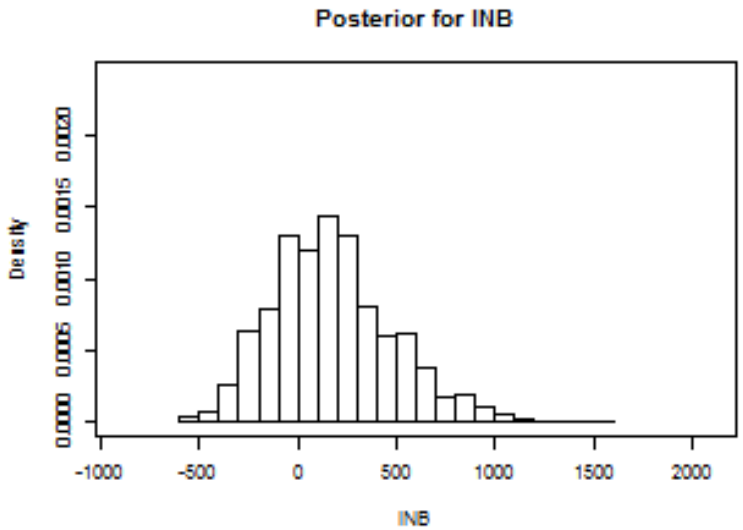
1. Select $q = 1, \ldots, Q$ values out of the S PSA simulations for $\boldsymbol{\theta}$



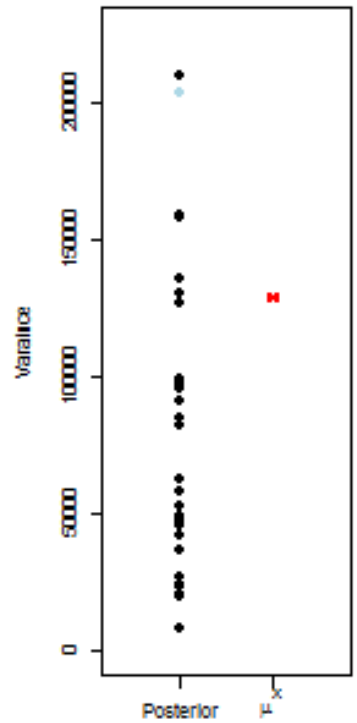
2. Simulate data $\boldsymbol{X}_q$ from $p(\boldsymbol{X} \mid \boldsymbol{\theta})$



3. Run the model to estimate $p(\boldsymbol{\theta} \mid \boldsymbol{X})$ and simulate values for $\text{INB}(\boldsymbol{\theta} \mid \boldsymbol{X}_q)$



4. Estimate the sample variance $\sigma^2_q = \text{Var}_{\boldsymbol{\theta} \mid \boldsymbol{X}_q}[\text{INB}(\boldsymbol{\theta})]$



$$\underbrace{\text{Var}_{\boldsymbol{X}}\left[\text{E}_{\boldsymbol{\theta} \mid \boldsymbol{X}}[\text{INB}(\boldsymbol{\theta})]\right]}_{\sigma^2_{\boldsymbol{X}}} = \underbrace{\text{Var}_{\boldsymbol{\theta}}[\text{INB}(\boldsymbol{\theta})]}_{\sigma^2} - \underbrace{\text{E}_{\boldsymbol{X}}\left[\text{Var}_{\boldsymbol{\theta} \mid \boldsymbol{X}}[\text{INB}(\boldsymbol{\theta})]\right]}_{\frac{1}{Q}\sum_{q=1}^{Q}\sigma^2_q}$$

5. Use the Q estimates for σ^2_q to estimate the expected posterior variance

- Can now rescale the original PSA samples for INB($\boldsymbol{\theta}$) to ensure that mean and variance now match the computed values

$$\eta^{\mathbf{x}} = f(\mu^{\mathbf{x}}) = \text{INB}\left(\boldsymbol{\theta}^{(s)}\right)\sqrt{\frac{\sigma^2_{\mathbf{x}}}{\sigma^2}} + \mu\left(1 - \sqrt{\frac{\sigma^2_{\mathbf{x}}}{\sigma^2}}\right)$$

  - $\text{INB}\left(\boldsymbol{\theta}^{(s)}\right) = s-$th PSA simulation for the INB
  - $\mu = \text{E}_{\boldsymbol{\theta}}\left[\text{INB}(\boldsymbol{\theta})\right]$ = PSA average INB
  - σ^2 = PSA variance of the INB

  and finally estimate the EVSI as

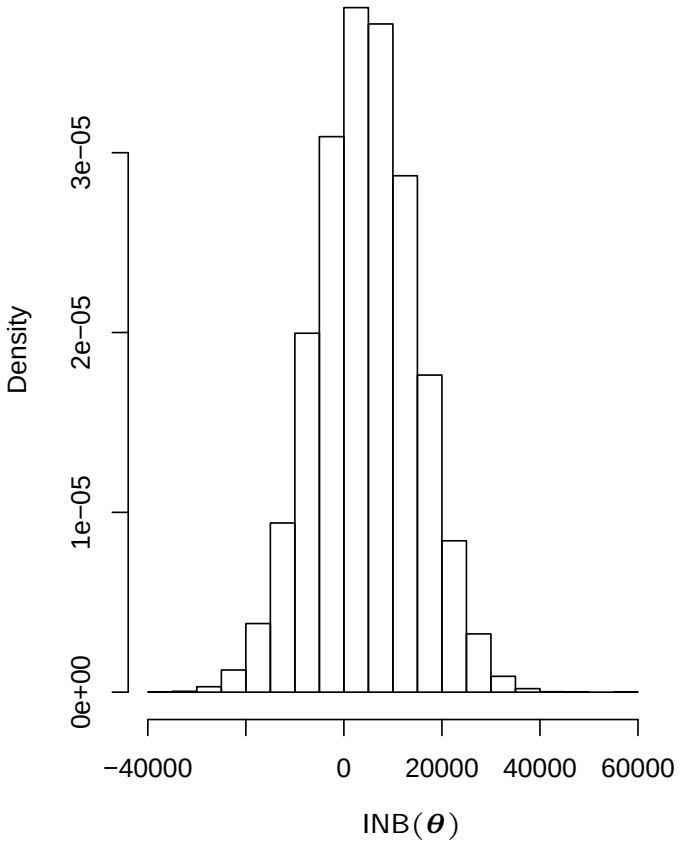
$$\text{EVSI} = \frac{1}{S}\sum_{s=1}^{S}\max\left\{0, \eta^{\mathbf{x}}\right\} - \max\left\{0, \mu\right\}$$

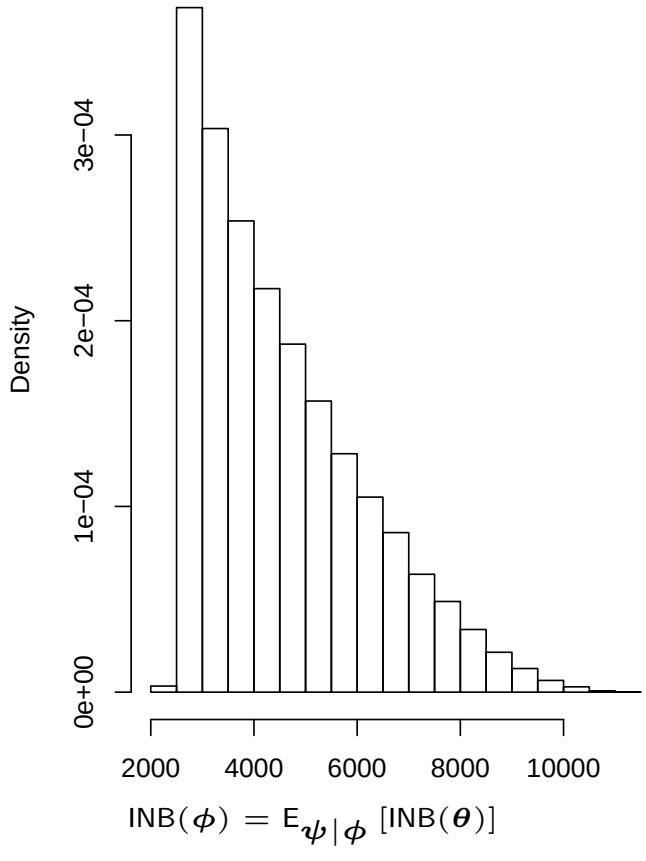
- Can also compute conditional version for $\boldsymbol{\phi} \in \boldsymbol{\theta}$. "Simply" substitute
  - σ^2 with $\sigma^2_{\boldsymbol{\phi}}$ = PSA variance for conditional INB (obtained using analysis of EVPPI)
  - $\text{INB}\left(\boldsymbol{\theta}^{(s)}\right)$ with $\text{INB}\left(\boldsymbol{\phi}^{(s)}\right) = \text{E}_{\boldsymbol{\psi}|\boldsymbol{\phi}}\left[\text{INB}\left(\boldsymbol{\theta}^{(s)}\right)\right]$

Moment matching

A Small Technicality...

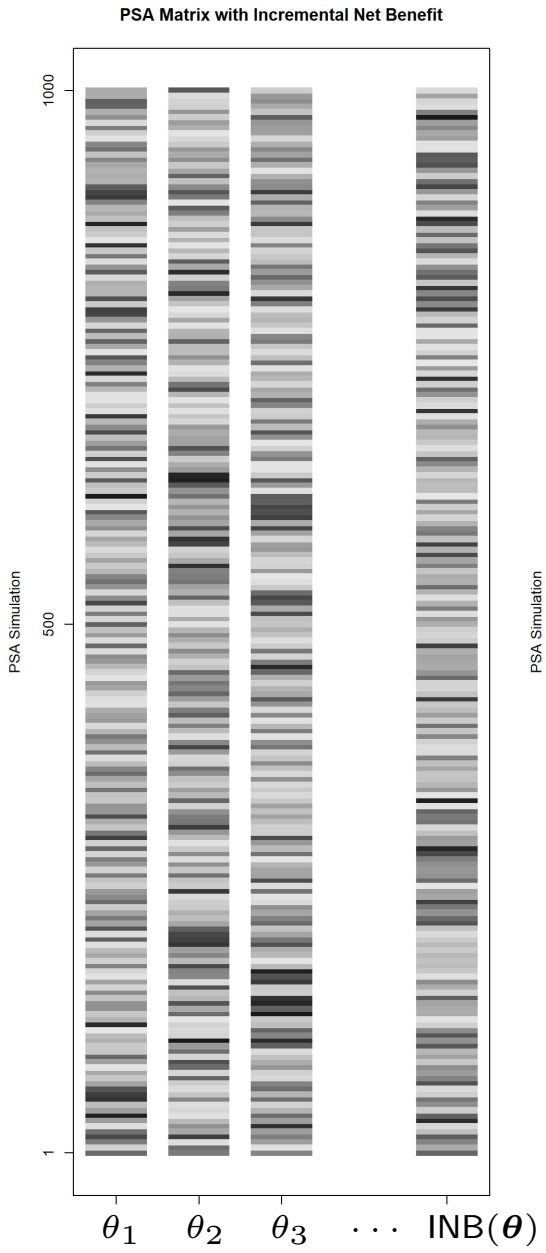
- Only the focal parameters ϕ will be informed by the future study
- The distribution of $\mu^{\boldsymbol{X}}$ is similar to that induced by the EVPPI analysis!



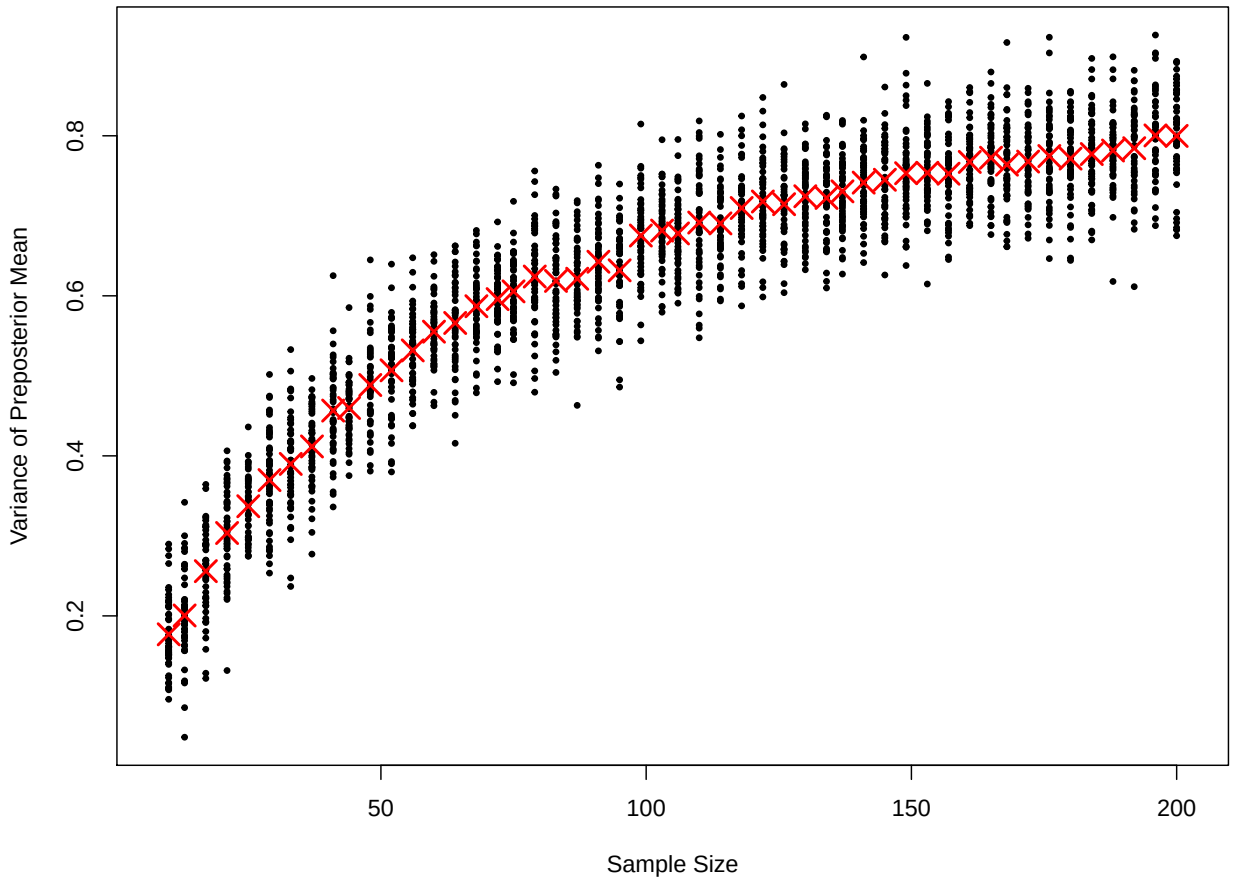


Multiple Focal Parameters

A slight complication...



To estimate EVSI across different sample sizes we could simulate $Q \times N$ samples from hypothetical posteriors



... But we'd lose all the computational efficiency of the moment matching approach...

Moment matching across different sample sizes

- Consider a set of sample sizes $\boldsymbol{N} = (N_1, \ldots, N_Q)$
- For each $q = 1, \ldots, Q$
  1. Randomly select $\boldsymbol{\theta}_q$ out of the S PSA samples
  2. Set $N = N_q$
  3. Simulate **one** sample $\boldsymbol{X}_q$ from $p(\boldsymbol{X} \mid \boldsymbol{\theta}_q, N_q)$
  4. Estimate the posterior distribution $p(\boldsymbol{\theta} \mid \boldsymbol{X}_q, N_q)$ and $\text{INB}(\boldsymbol{\theta} \mid \boldsymbol{X}_q, N_q)$
  5. Estimate the variance σ^2_q associated with a given design (size N_q) and data ($\boldsymbol{X}_q$)
- **NB**: Now we need to estimate $\sigma^2_{\boldsymbol{X}}$ as a function of the sample size: $\sigma^2_{\boldsymbol{X}}(N) = f(N)$

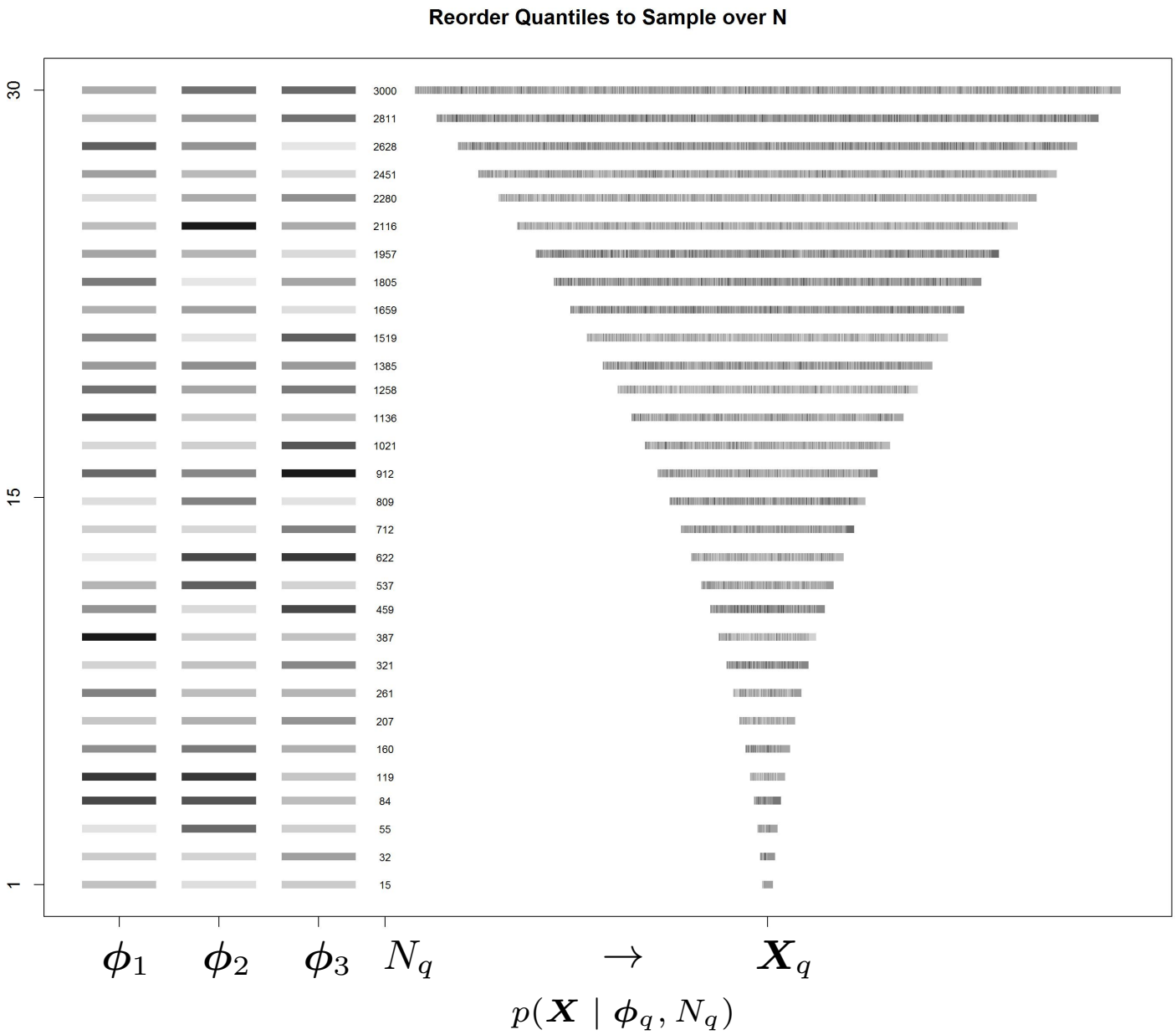
$$\sigma^2_{\boldsymbol{X}}(N_q) = \sigma^2 - \sigma^2_q = f(N_q) + \varepsilon_q$$

- Use Bayesian non-linear regression and model

$$f(N_q) = \sigma^2_{\boldsymbol{\phi}} \frac{N_q}{N_q + h} \qquad \varepsilon_q \sim \text{Normal}(0, \sigma^2_\varepsilon)$$

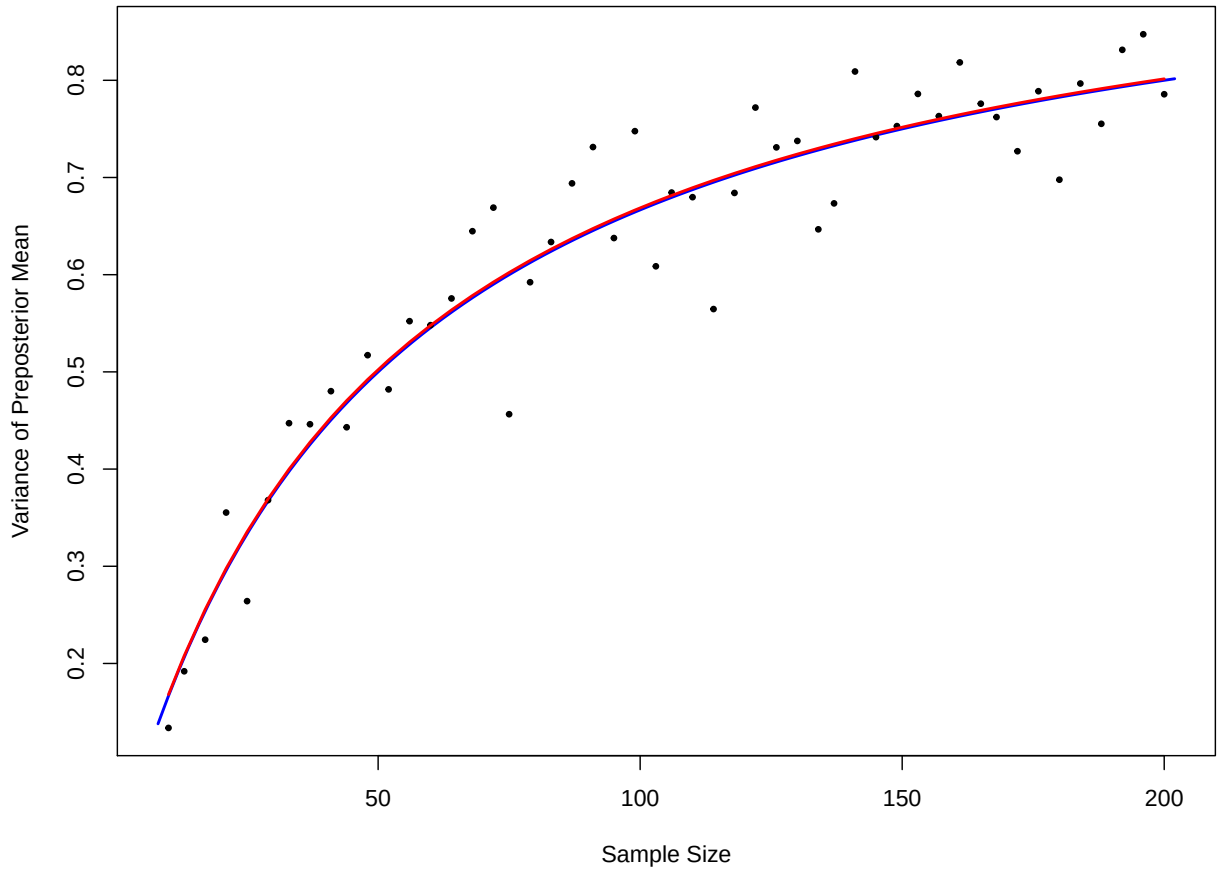
  - $\sigma^2_{\boldsymbol{\phi}} = \text{Var}_{\boldsymbol{\phi}}\left[\text{INB}(\boldsymbol{\phi})\right]$
  - $h =$ Regression parameter
  - $\varepsilon_q =$ error term

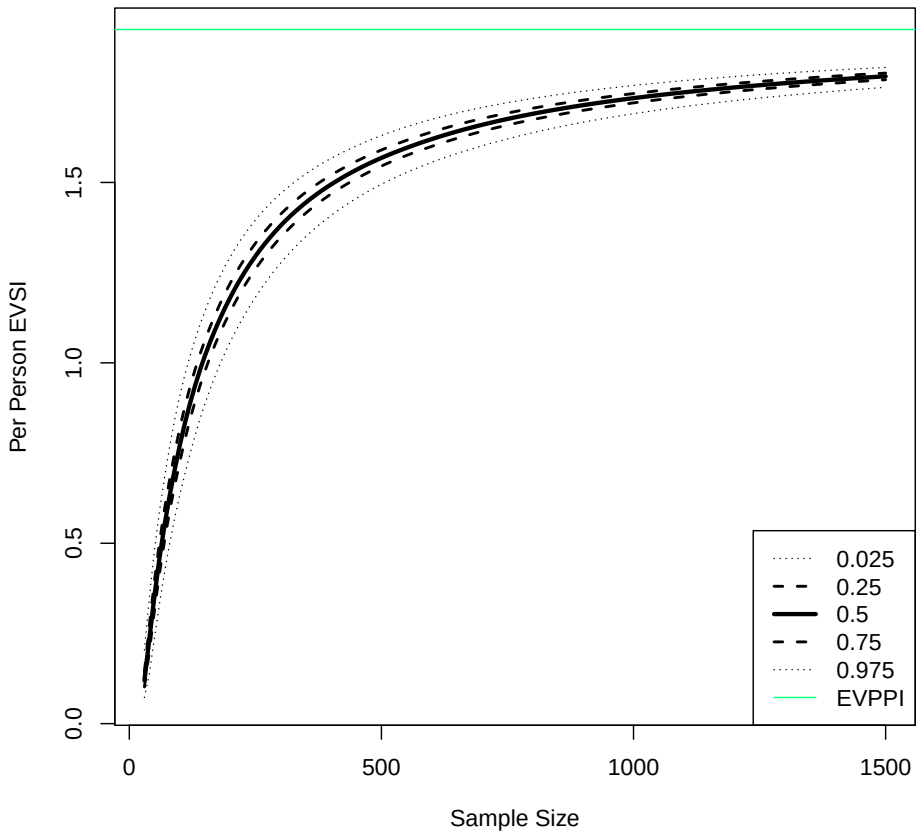
Moment matching across different sample sizes

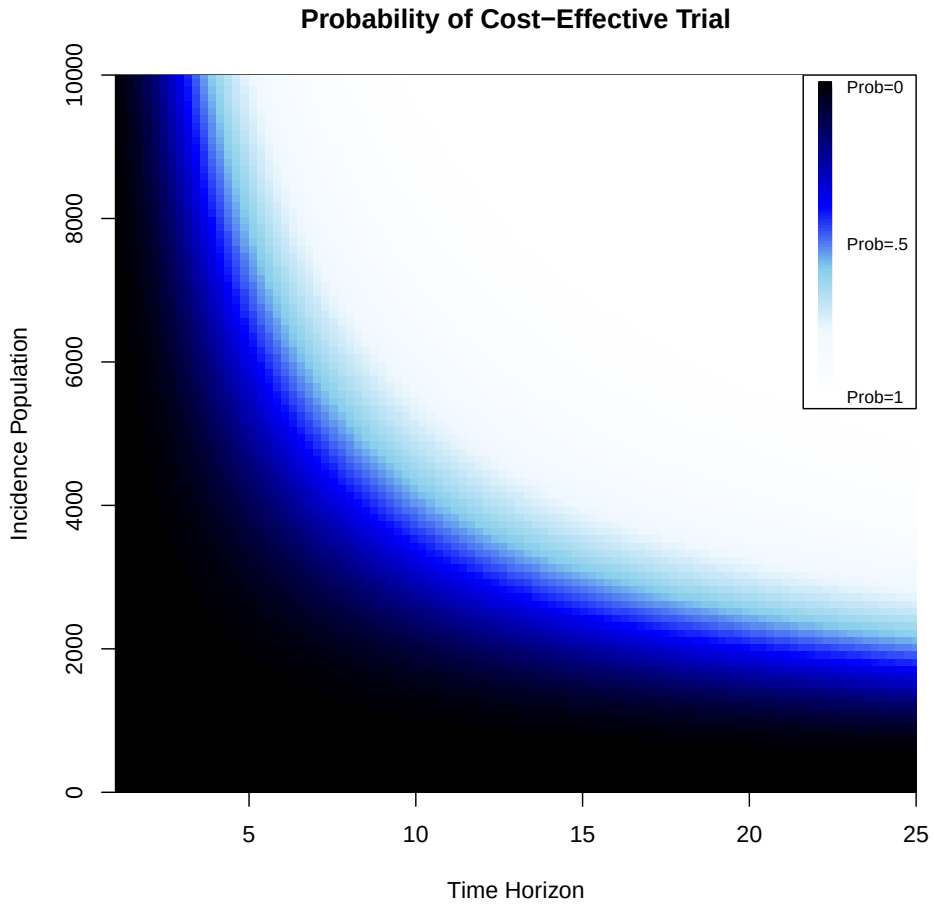


Moment matching across different sample sizes

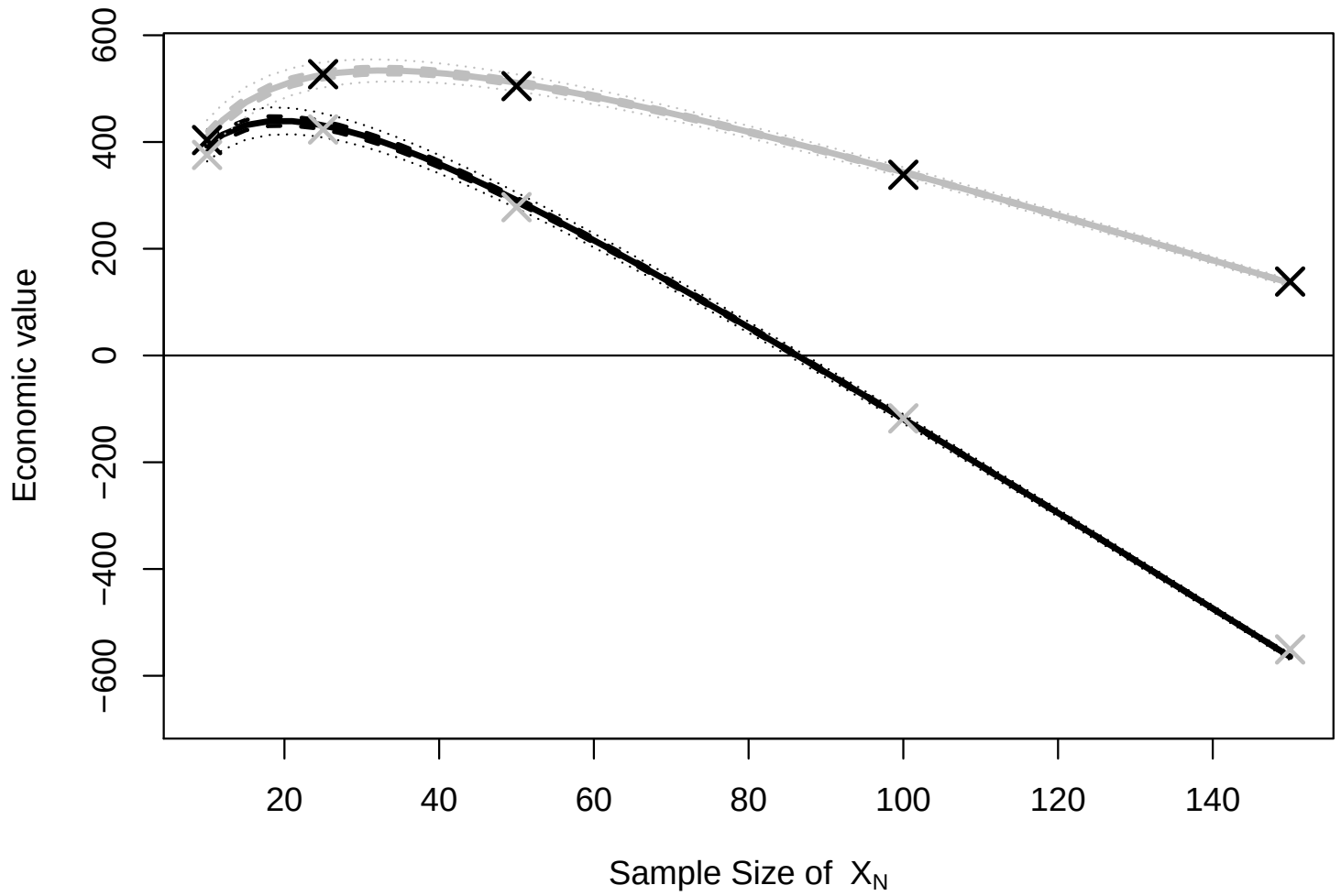
- $\sigma^2_{\boldsymbol{X}}(N)$ increases as N does + $f(N)$ is a monotonically increasing function
- If $N \to \infty$, then EVSI $\to$ EVPPI and so $\sigma^2_{\boldsymbol{X}}(\infty) \to \sigma^2_{\boldsymbol{\phi}}$, because $\mu^{\boldsymbol{X}} \to \text{INB}(\boldsymbol{\phi})$
- Can use weakly informative priors for the parameters
  - $h \sim \text{Normal}\left(N_Q/2, 200N_Q\right)\mathbb{I}(0,)$
  - $\sigma^2_\varepsilon \sim \text{t}(m, s, 3)\mathbb{I}(0,)$, with m, s defined as function of σ^2_q for generality







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

Conclusions

- VoI can be very valuable in driving the whole economic evaluation process
  - Summarising PSA (in **addition** to standard tools, eg CEAC)
  - Research priority (in **place** of standard tools, eg sample size calculations?)
- Historically limited use — also for computational complexity
  - Computation still a crucial component — but this is the price to pay for increasingly realistic and complex models?
  - Things can only get better(?) — recent research has improved this **massively**!
- Need standardised softward to enable practitioners to use the new tools
  - **And to move from Excel-based modelling to using fully proper statistical software (eg R)**
  - Packages and web-applications exist to do this: `SAVI`, `BCEA`, `BCEAweb`, ...

Thank you!

Principal Fitted Components

- Objective: find a **sufficient** dimensionality reduction
  - Estimate the function $R(\boldsymbol{\phi}) : P \to d$ so that $\mathbf{NB}_t \perp\!\!\!\perp \boldsymbol{\phi} \mid R(\boldsymbol{\phi})$
  - "**Project**" the $P-$dimensional information contained in $\boldsymbol{\phi}$ to the $d-$dimensional function $R(\cdot)$
  - Ideally, $d << P$ — in fact, would like $d \leq 2$

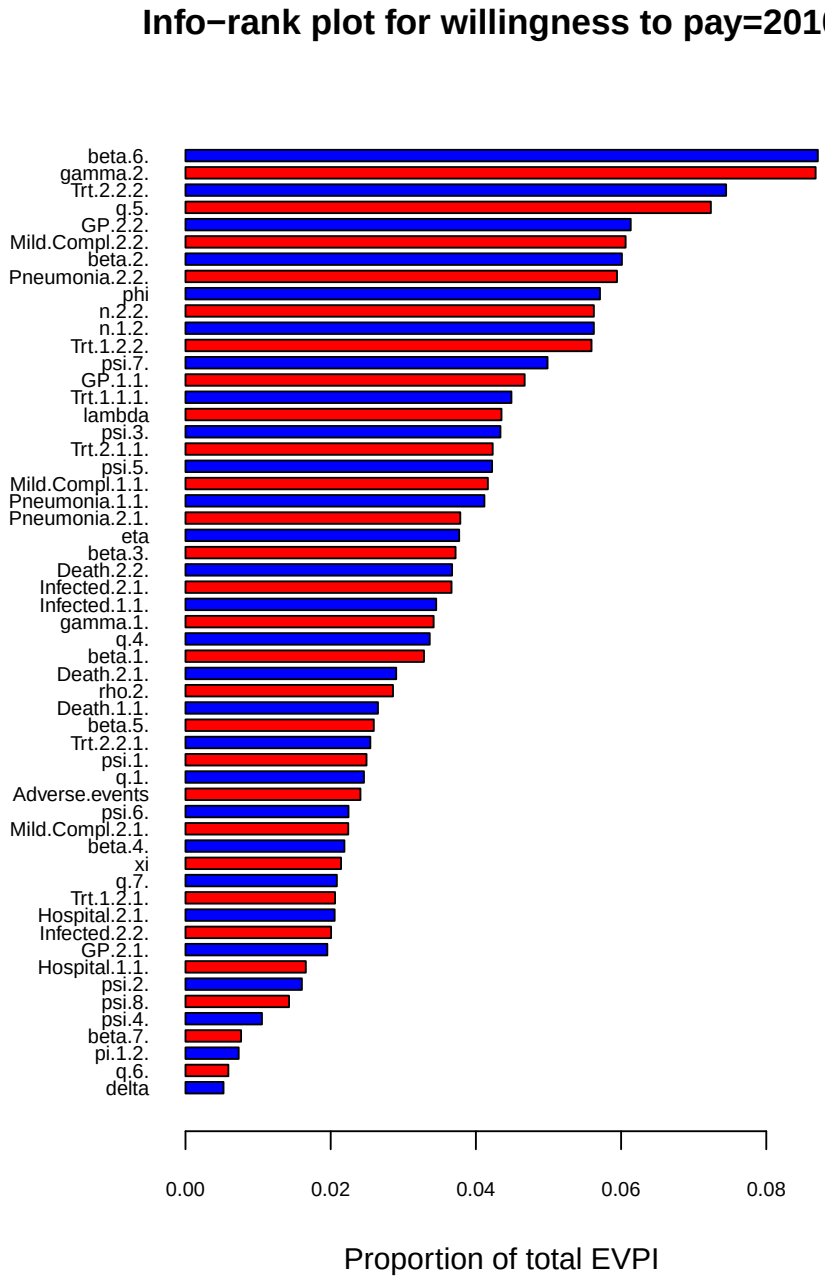
- "Inverse regression" model

$$\boldsymbol{\phi} = \boldsymbol{\mu} + \boldsymbol{\Upsilon}\boldsymbol{f}(\mathbf{NB}_t) + \boldsymbol{\epsilon}$$

  with
  - $\boldsymbol{\mu}$ = intercept
  - $\boldsymbol{\Upsilon} = P \times d$ dimensionality reduction matrix
  - $\boldsymbol{f}(\mathbf{NB}_t)$ = vector-valued function of the "response"
  - $\boldsymbol{\epsilon}$ = error term

- Main advantages
  - Computational cost is negligible
  - Can use model-fitting statistics (eg AIC) to determine the "best" model for given choices of d $(= 2, 3, \ldots)$
  - **NB**: if the AIC suggests $d > 2$ then EVPPI estimates likely to be biased!

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



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