

Bayesian methods in health economics

Gianluca Baio

University College London Department of Statistical Science

g.baio@ucl.ac.uk

Seminar Series of the Master in Advanced Artificial Intelligence Madrid, Universidad Nacional De Educación a Distancia Thursday 11 June 2015



- Health economic evaluation
 - What is health economics?
 - What does health economics do?



- Health economic evaluation
 - What is health economics?
 - What does health economics do?
- Statistical modelling
 - Models for individual-level data
 - Models for aggregated data



- Health economic evaluation
 - What is health economics?
 - What does health economics do?
- Statistical modelling
 - Models for individual-level data
 - Models for aggregated data
- Economic modelling & Decision analysis
 - Cost-effectiveness/cost-utility analysis
 - Criteria for decision-making in health economics



- Health economic evaluation
 - What is health economics?
 - What does health economics do?
- Statistical modelling
 - Models for individual-level data
 - Models for aggregated data
- Economic modelling & Decision analysis
 - Cost-effectiveness/cost-utility analysis
 - Criteria for decision-making in health economics
- Uncertainty analysis
 - Rationale
 - Main ideas



- Health economic evaluation
 - What is health economics?
 - What does health economics do?
- Statistical modelling
 - Models for individual-level data
 - Models for aggregated data
- Economic modelling & Decision analysis
 - Cost-effectiveness/cost-utility analysis
 - Criteria for decision-making in health economics
- Uncertainty analysis
 - Rationale
 - Main ideas
- Conclusions

What does health economics do?

- **Objective**: Combine costs & benefits of a given intervention into a rational scheme for allocating resources
 - Recently, models have been built upon more advanced statistical foundations
 - This problem can be formalised within a statistical decision-theoretic approach. Rational decision-making is effected through the comparison of expected utilities
 - Incremental approach: need to consider at least two interventions

- **Objective**: Combine costs & benefits of a given intervention into a rational scheme for allocating resources
 - Recently, models have been built upon more advanced statistical foundations
 - This problem can be formalised within a statistical decision-theoretic approach. Rational decision-making is effected through the comparison of expected utilities
 - Incremental approach: need to consider at least two interventions
- Increasingly under a **Bayesian framework**, especially in the UK: 5.9.10–12 Dealing with parameter uncertainty in cost-effectiveness analysis (NICE Methods for Technology Assessment)
 - All inputs used in the analysis will be estimated with a degree of imprecision.
 - Probabilistic sensitivity analysis is preferred for translating the imprecision in all input variables into a measure of decision uncertainty in the cost-effectiveness of the options being compared.
 - Appropriate ways of presenting uncertainty include confidence ellipses and scatter plots on the cost-effectiveness plane (when the comparison is restricted to two alternatives) and cost-effectiveness acceptability curves.

Health economic outcome

UCL

- One of the most important characteristic of health economic data is that we have **multivariate outcomes**
 - -e = suitable measure of clinical benefit of an intervention
 - c = suitable costs associated with an intervention

UCL

- One of the most important characteristic of health economic data is that we have **multivariate outcomes**
 - -e = suitable measure of clinical benefit of an intervention
 - c = suitable costs associated with an intervention
- We typically need to assess these quantities jointly
 - Costs and benefit will tend to be correlated
 - Strong positive correlation effective treatments are innovative and result from intensive and lengthy research ⇒ are associated with higher unit costs
 - Negative correlation more effective treatments may reduce total care pathway costs e.g. by reducing hospitalisations, side effects, etc.
 - In any case, the economic evaluation is based on both!

UCL

- One of the most important characteristic of health economic data is that we have **multivariate outcomes**
 - -e = suitable measure of clinical benefit of an intervention
 - c = suitable costs associated with an intervention
- We typically need to assess these quantities jointly
 - Costs and benefit will tend to be correlated
 - Strong positive correlation effective treatments are innovative and result from intensive and lengthy research ⇒ are associated with higher unit costs
 - Negative correlation more effective treatments may reduce total care pathway costs e.g. by reducing hospitalisations, side effects, etc.
 - In any case, the economic evaluation is based on both!
- There are different ways in which we can define (e,c) for a specific problem
 - Direct vs indirect vs intangible costs
 - "Hard-" vs utility-based clinical outcomes
 - Public (e.g. NHS) vs private (e.g. insurance) perspective

Health economic evaluations





- Estimates relevant population parameters
- Varies with the type of available data (& statistical approach!)





Health economic evaluations



Health economic evaluations





Standardised process

1. Statistical modelling



- Sampling variability for the health economic outcomes is described by a distribution p(e, c | θ^t), which depends on a set of population parameters θ^t
 - Probability of some clinical outcome
 - Duration in treatment
 - Reduction in the rate of occurrence of some event
 - Unit cost of acquisition of a health technology



- Sampling variability for the health economic outcomes is described by a distribution p(e, c | θ^t), which depends on a set of population parameters θ^t
 - Probability of some clinical outcome
 - Duration in treatment
 - Reduction in the rate of occurrence of some event
 - Unit cost of acquisition of a health technology
- Under the Bayesian approach, parametric uncertainty is modelled using a prior distribution $p(\pmb{\theta}^t)$
 - This describes the level of knowledge in the value of the population parameters
 - Can be based on subjective information, or existing data



- Sampling variability for the health economic outcomes is described by a distribution p(e, c | θ^t), which depends on a set of population parameters θ^t
 - Probability of some clinical outcome
 - Duration in treatment
 - Reduction in the rate of occurrence of some event
 - Unit cost of acquisition of a health technology
- Under the Bayesian approach, parametric uncertainty is modelled using a prior distribution $p(\pmb{\theta}^t)$
 - This describes the level of knowledge in the value of the population parameters
 - Can be based on subjective information, or existing data
- The way in which we construct our statistical model, depends on
 - The characteristic of the available data (individual-level vs aggregated data)
 - The statistical framework (Bayesian vs frequentist)

- Observe vectors $(\boldsymbol{e}, \boldsymbol{c})$ under each intervention being compared
 - May also observe other variables (covariates) *e.g.* individual values for age, sex, co-morbidities, etc
- Use observed data to estimate the relevant population parameters $\pmb{\theta}^t = (\pmb{\theta}_e^t, \pmb{\theta}_c^t)$
 - These are generally vectors, made by several components (*e.g.* means, variances, rates, etc)
- The main interest is in the population average benefits and costs under treatment \boldsymbol{t}

 $\mu_e^t = \mathsf{E}[e \mid \pmb{\theta}^t] \qquad \text{and} \qquad \mu_c^t = \mathsf{E}[c \mid \pmb{\theta}^t]$

- NB: Because of underlying correlation, it is necessary to use some form of joint model
 - But: simple models (such as bivariate Normal) are not suitable, as both e, c tend to be skewed and cost are positive

Can factorise the joint distribution, for example as $p(e, c) = p(c)p(e \mid c)$



Can factorise the joint distribution, for example as $p(e, c) = p(c)p(e \mid c)$



For instance, can model

• $c_{it} \sim \text{Gamma}(\eta_t, \lambda_t)$ [rate & shape] $\Rightarrow \mu_{ct} = \eta_t / \lambda_t$

 $c_{it} \sim \log \operatorname{Normal}(\eta_t, \lambda_t) \ [\log \ mean \ \& \ \log \ sd] \Rightarrow \mu_{ct} = \exp \left(\eta_t + \lambda_t^2/2 \right)$

Can factorise the joint distribution, for example as $p(e, c) = p(c)p(e \mid c)$



For instance, can model

• $c_{it} \sim \text{Gamma}(\eta_t, \lambda_t)$ [rate & shape] $\Rightarrow \mu_{ct} = \eta_t / \lambda_t$ $c_{it} \sim \log \text{Normal}(\eta_t, \lambda_t)$ [log mean & log sd] $\Rightarrow \mu_{ct} = \exp(\eta_t + \lambda_t^2/2)$

UCL

Can factorise the joint distribution, for example as $p(e,c) = p(c)p(e \mid c)$



For instance, can model

• $c_{it} \sim \text{Gamma}(\eta_t, \lambda_t)$ [rate & shape] $\Rightarrow \mu_{ct} = \eta_t / \lambda_t$ $c_{it} \sim \text{logNormal}(\eta_t, \lambda_t)$ [log mean & log sd] $\Rightarrow \mu_{ct} = \exp(\eta_t + \lambda_t^2/2)$ • $\mathsf{E}[e_{it}] = \phi_{it}; \ g(\phi_{it}) = \xi_t + \gamma_t(c_{it} - \mu_{ct}) \Rightarrow \mu_{et} = g^{-1}(\xi_t)$

Gianluca Baio (UCL)

Decision-analytic models



- Often, we do not have access to individual data and all we have is a set of aggregated data on relevant quantities
- These can in turn be used to construct a "population model" to describe the disease history and its implications
 - Decision trees
 - Markov (multistate) models

Decision-analytic models

- Often, we do not have access to individual data and all we have is a set of aggregated data on relevant quantities
- These can in turn be used to construct a "population model" to describe the disease history and its implications
 - Decision trees
 - Markov (multistate) models

Example: influenza



Decision-analytic models





- Cost minimisation
 - Assumes that the benefits produced by two interventions are identical \Rightarrow the only dimension of interest is costs
- Cost-benefit analysis
 - Requires that costs and benefits are converted and analysed into monetary terms \Rightarrow difficulties in valuing health outcomes in monetary units

- Cost minimisation
 - Assumes that the benefits produced by two interventions are identical \Rightarrow the only dimension of interest is costs
- Cost-benefit analysis
 - Requires that costs and benefits are converted and analysed into monetary terms \Rightarrow difficulties in valuing health outcomes in monetary units

• Cost-effectiveness analysis (CEA)

- Evaluates cost-per-outcome gained
- Outcomes are usually "hard" measurements (eg death) ⇒ easy to understand for clinicians, but difficult to compare across diseases (may have different main outcome)

• Cost-utility analysis (CUA)

- Considers a common health outcome unit (= QALYs), so easy to compare across diseases
- Often interchangeable with CEA (common methodology!)

2. Economic modelling

Can think of this step as the process of obtaining relevant population summaries for the measures of cost & clinical benefits. For example, when comparing two interventions t = 0, 1, the main focus is on

• The increment in mean benefits

$$\Delta_e = \underbrace{\mathsf{E}[e \mid \theta^1]}_{\mu_1^e} - \underbrace{\mathsf{E}[e \mid \theta^0]}_{\mu_0^e}$$

• The increment in mean costs:

$$\Delta_c = \underbrace{\mathsf{E}[c \mid \theta^1]}_{\mu_1^c} - \underbrace{\mathsf{E}[c \mid \theta^0]}_{\mu_0^c}$$

2. Economic modelling

Can think of this step as the process of obtaining relevant population summaries for the measures of cost & clinical benefits. For example, when comparing two interventions t = 0, 1, the main focus is on

• The increment in mean benefits

$$\Delta_e = \underbrace{\mathsf{E}[e \mid \theta^1]}_{\mu_1^e} - \underbrace{\mathsf{E}[e \mid \theta^0]}_{\mu_0^e}$$

• The increment in mean costs:

$$\Delta_c = \underbrace{\mathsf{E}[c \mid \theta^1]}_{\mu_1^c} - \underbrace{\mathsf{E}[c \mid \theta^0]}_{\mu_0^c}$$

- **NB**: In a Bayesian context, these are functions of θ and thus random variables!
- When using individual-level data, estimation typically directly available from the statistical model; for decision-analytic models, it may be necessary to combine the parameters to obtain these

Gianluca Baio (UCL)

3. Decision analysis

In order to compare the two interventions (t = 0, 1), we define suitable health economic indicators

• The population average increment in benefits

$$\mathsf{E}[\Delta_e] = \bar{e}_1 - \bar{e}_0 = \mathsf{E}[\mu_1^e] - \mathsf{E}[\mu_0^e]$$

• The population average increment in costs:

$$\mathsf{E}[\Delta_c] = \bar{c}_1 - \bar{c}_0 = \mathsf{E}[\mu_1^c] - \mathsf{E}[\mu_0^c]$$

3. Decision analysis

In order to compare the two interventions (t = 0, 1), we define suitable health economic indicators

• The population average increment in benefits

$$\mathsf{E}[\Delta_e] = \bar{e}_1 - \bar{e}_0 = \mathsf{E}[\mu_1^e] - \mathsf{E}[\mu_0^e]$$

• The population average increment in costs:

$$\mathsf{E}[\Delta_c] = \bar{c}_1 - \bar{c}_0 = \mathsf{E}[\mu_1^c] - \mathsf{E}[\mu_0^c]$$

 Generally, economic summaries are computed in the form of "cost per outcome" ratios

$$\mathsf{ICER} = \frac{\mathsf{E}[\Delta_c]}{\mathsf{E}[\Delta_e]} = \mathsf{Additional\ cost\ to\ gain\ 1\ unit\ of\ benefit}$$



Cost-effectiveness plane



Effectiveness differential

Gianluca Baio (UCL)

Cost-effectiveness plane



Effectiveness differential

Gianluca Baio (UCL)

Bayesian methods in health economics

Seminar UNED, 11 Jun 2015 14 / 27

• When considering only two interventions t = 0, 1, can equivalently represent the problem using the Expected Incremental Benefit

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

where k is the willingness to pay

- Puts costs and benefits on the same scale
- Represents the amount of the decision-maker is willing to invest to increment the benefits by 1 unit



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

where k is the willingness to pay

- Puts costs and benefits on the same scale
- Represents the amount of the decision-maker is willing to invest to increment the benefits by 1 unit
- One-to-one relationship between ICER and EIB

$$\mathsf{EIB} > 0 \Rightarrow k > \frac{\mathsf{E}[\Delta_c]}{\mathsf{E}[\Delta_e]} = \mathsf{ICER}$$



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

where k is the willingness to pay

- Puts costs and benefits on the same scale
- Represents the amount of the decision-maker is willing to invest to increment the benefits by 1 unit
- One-to-one relationship between ICER and EIB

$$\mathsf{EIB} > 0 \Rightarrow k > \frac{\mathsf{E}[\Delta_c]}{\mathsf{E}[\Delta_e]} = \mathsf{ICER}$$

- The EIB is also more directly linked to a (Bayesian) decision-theoretic approach
 - Define a utility function to quantify the "value" of an intervention
 - Compute the expected utility (wrt to both individual & population variations)
 - Choose the intervention with the highest expected utility

Cost-effectiveness plane vs EIB vs ICER



Cost effectiveness plane New Chemotherapy vs Old Chemotherapy



Cost-effectiveness plane vs EIB vs ICER







Cost-effectiveness plane vs EIB vs ICER



Expected Incremental Benefit



So: problem solved?





So: problem solved?... Well, not really!



So: problem solved?... Well, not really!

- The quality of the current evidence is often limited
 - During the pre-market authorisation phase, the regulator should decide whether to grant reimbursement to a new product — and in some countries also set the price — on the basis of uncertain evidence, regarding both clinical and economic outcomes
 - Although it is possible to answer some unresolved questions after market authorisation, relevant decisions such as that on reimbursement (which determines the overall access to the new treatment) have already been taken



So: problem solved?... Well, not really!

- The quality of the current evidence is often limited
 - During the pre-market authorisation phase, the regulator should decide whether to grant reimbursement to a new product — and in some countries also set the price — on the basis of uncertain evidence, regarding both clinical and economic outcomes
 - Although it is possible to answer some unresolved questions after market authorisation, relevant decisions such as that on reimbursement (which determines the overall access to the new treatment) have already been taken
- This leads to the necessity of performing (probabilistic) sensitivity analysis (PSA)
 - Formal quantification of the impact of uncertainty in the parameters on the results of the economic model
 - Standard requirement in many health systems (e.g. for NICE in the UK), but still not universally applied
 - Often limited to *parametric* uncertainty, but should be extended to *structural* uncertainty too

Uncertainty analysis — Frequentist vs Bayesian approach



1. Estimation (base-case)





 \Rightarrow

Uncertainty analysis — Frequentist vs Bayesian approach

ach **AUCI** 2. PSA

1. Estimation (base-case)













Summarising PSA — CEAC



Cost Effectiveness Acceptability Curve



Is this all we need?

- The CEAC only deals with the **probability** of making the "right decision"
- But it does not account for the **payoff/penalty** associated with making the "wrong" one!

Is this all we need?

- The CEAC only deals with the **probability** of making the "right decision"
- But it does not account for the **payoff/penalty** associated with making the "wrong" one!
- **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 = $\pounds 3$

Decrease in effectiveness = 0.000001 QALYs

Is this all we need?

- The CEAC only deals with the **probability** of making the "right decision"
- But it does not account for the **payoff/penalty** associated with making the "wrong" one!
- **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 = $\pounds 3$

Decrease in effectiveness = 0.000001 QALYs

- **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 = $\pounds 1\,000\,000\,000$

Decrease in effectiveness = 999999 QALYs

Expected value of information

- Basically quantifies the value (in terms of **utility**) of reducing uncertainty in the parameters
 - NB: Information has value only if can modify your current decision



Expected value of information

- Basically quantifies the value (in terms of **utility**) of reducing uncertainty in the parameters
 - NB: Information has value only if can modify your current decision
- Obtained by comparing
 - 1 The decision made using currently available evidence
 - 2 The decision that would be made if we could gather "perfect" information on the uncertain parameters

Expected value of information

- Basically quantifies the value (in terms of **utility**) of reducing uncertainty in the parameters
 - NB: Information has value only if can modify your current decision
- Obtained by comparing
 - 1 The decision made using currently available evidence
 - 2 The decision that would be made if we could gather "perfect" information on the uncertain parameters
- Comments:
 - "Perfect" information is just a hypothetical concept that is why we consider the "expected value"
 - If the optimal treatment is not dominated at any point in the parameter space, the EVPI is equal to 0 and the uncertainty in the parameters has no impact on the decision process



- Non-intuitive interpretation: when is the EVPI "low enough"?
 - Links with research prioritisation compare the EVPI with the cost of buying information (e.g. a trial) and decide whether it is worth deferring the decision
 - Depends also on the size of the target population



- Non-intuitive interpretation: when is the EVPI "low enough"?
 - Links with research prioritisation compare the EVPI with the cost of buying information (e.g. a trial) and decide whether it is worth deferring the decision
 - Depends also on the size of the target population
- Usually, it is impossible to buy information on **all** the model parameters
 - Some parameters are not even *that* interesting e.g. fixed costs, "things" we cannot change, ...
 - Some other though, are interesting, because we can conduct a study to learn more and thus potentially change the optimal decision
 - Can consider the Expected Value of Partial Perfect Information

Expected Value of Partial Information

- Suppose the parameters of your model are collected in a vector ${m heta}$
- And that you can split them into two subsets
 - The "important" parameters ϕ and the "unimportant" parameters ψ
- We are interested in quantifying the value of gaining more information on ϕ , while leaving the current level of uncertainty on ψ unchanged

Expected Value of Partial Information

- Suppose the parameters of your model are collected in a vector ${m heta}$
- And that you can split them into two subsets
 - The "important" parameters ϕ and the "unimportant" parameters ψ
- We are interested in quantifying the value of gaining more information on ϕ , while leaving the current level of uncertainty on ψ unchanged
- **Technical issue**: because ϕ and ψ are typically correlated, we cannot make easy computations for the EVPPI (certainly not in Excel!)
 - Nested Monte Carlo simulations
 - (1) Simulate a large number of values for ϕ
 - $oldsymbol{2}$ For each of the simulated values of ϕ , simulate a large values of ψ
 - (3) This means we may need to run a PSA with 10 000s \times 10 000s iterations too big!

Expected Value of Partial Information

- Suppose the parameters of your model are collected in a vector ${m heta}$
- And that you can split them into two subsets
 - The "important" parameters ϕ and the "unimportant" parameters ψ
- We are interested in quantifying the value of gaining more information on ϕ , while leaving the current level of uncertainty on ψ unchanged
- Technical issue: because ϕ and ψ are typically correlated, we cannot make easy computations for the EVPPI (certainly not in Excel!)
 - Nested Monte Carlo simulations
 - (1) Simulate a large number of values for ϕ
 - 2 For each of the simulated values of ϕ , simulate a large values of ψ
 - (3) This means we may need to run a PSA with 10 000s \times 10 000s iterations too big!
 - New methods based on "non-parametric regression" (fancy stats)
 - Can use a standard run of 1 000 PSA simulations and can approximate the true value of the EVPPI very accurately!

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"



Thank you!

Some references



Baio, G. (2012),

Bayesian Methods in Health Economics. Chapman Hall, CRC Press, Boca Raton, FL.

Briggs, A., M. Sculpher, and K. Claxton (2006). Decision modelling for health economic evaluation.

Oxford University Press, Oxford, UK,



Heath, A., I. Manolopoulou, and G. Baio (2015).

Efficient High-Dimensional Gaussian Process Regression to calculate the Expected Value of Partial Perfect Information in Health Economic Evaluations.

http://arxiv.org/abs/1504.05436v1.

Jackson, C., L. Sharples, and S. Thompson (2010).

Structural and parameter uncertainty in Bayesian cost-effectiveness analysis. Journal of the Royal Statistical Society, C 59, 233-253.

Spiegelhalter, D., K. Abrams, and J. Myles (2004). Bayesian Approaches to Clinical Trials and Health-Care Evaluation. John Wiley and Sons. Chichester, UK.



Strong, M., J. Oakley, and A. Brennan (2014).

Estimating Multiparameter Partial Expected Value of Perfect Information from a Probabilistic Sensitivity Analysis Sample A Nonparametric Regression Approach. Medical Decision Making 34(3), 311-326.

Welton, N., A. Sutton, N. Cooper, K. Abrams, and A. Ades (2012). Evidence Synthesis for Decision Making in Healthcare John Wiley and Sons, Chichester, UK.

Gianluca Baio (UCL)