Statistical issues in small/pilot cost-effectiveness analysis of e-health interventions

Gianluca Baio

(Thanks to Andrea Gabrio and Alexina Mason)

University College London
Department of Statistical Science

g.baio@ucl.ac.uk

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

UCL Health Economics Symposium 2018
University College London

Monday 5 February 2018
Statistics for Health Economic Evaluation

Group profile

The Statistics for Health Economic Evaluation Group, led by Dr Gianluca Baio, is a research group based in the Department of Statistical Science. Our activity revolves around the development and application of Bayesian statistical methodology for health economic evaluation, e.g. cost-effectiveness or cost-utility analysis.

We work in close collaboration with academics both within UCL and at other institutions, including MRC Biostatistics Unit Cambridge, University of Sheffield and University of Rome, as well as with partners in the private sector. We have received a five-year research grant from Mapi, a consultancy company working in the area of economic evaluation. This creates exciting opportunities to develop new methodologies for problems that are relevant to real, practical applications.

Our activities include a series of seminars aimed at statisticians, health economists and clinicians working in economic evaluations. We also aim at seeking potential areas of commonality with other research groups within the department of Statistical Science and statisticians in general.
### “Standard” Statistical modelling — Individual level data in HTA

<table>
<thead>
<tr>
<th>ID</th>
<th>Trt</th>
<th>Demographics</th>
<th>HRQL data</th>
<th>Resource use data</th>
<th>Clinical outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Sex</td>
<td>Age</td>
<td>$u_0$</td>
<td>$u_1$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>M</td>
<td>23</td>
<td>0.32</td>
<td>0.66</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>M</td>
<td>21</td>
<td>0.12</td>
<td>0.16</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>F</td>
<td>19</td>
<td>0.49</td>
<td>0.55</td>
</tr>
</tbody>
</table>

$y_{ij} = \text{Survival time, event indicator (eg CVD), number of events, continuous measurement (eg blood pressure), ...}$

$u_{ij} = \text{Utility-based score to value health (eg EQ-5D, SF-36, Hospital Anxiety \\ & Depression Scale), ...}$

$c_{ij} = \text{Use of resources (drugs, hospital, GP appointments, ...}$}
**“Standard” Statistical modelling — Individual level data in HTA**

<table>
<thead>
<tr>
<th>ID</th>
<th>Trt</th>
<th>Sex</th>
<th>Age</th>
<th>( u_0 )</th>
<th>( u_1 )</th>
<th>( \ldots )</th>
<th>( u_J )</th>
<th>( c_0 )</th>
<th>( c_1 )</th>
<th>( \ldots )</th>
<th>( c_J )</th>
<th>( y_0 )</th>
<th>( y_1 )</th>
<th>( \ldots )</th>
<th>( y_J )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>M</td>
<td>23</td>
<td>0.32</td>
<td>0.66</td>
<td>( \ldots )</td>
<td>0.44</td>
<td>103</td>
<td>241</td>
<td>( \ldots )</td>
<td>80</td>
<td>( y_{10} )</td>
<td>( y_{11} )</td>
<td>( \ldots )</td>
<td>( y_{1J} )</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>M</td>
<td>21</td>
<td>0.12</td>
<td>0.16</td>
<td>( \ldots )</td>
<td>0.38</td>
<td>1204</td>
<td>1808</td>
<td>( \ldots )</td>
<td>877</td>
<td>( y_{20} )</td>
<td>( y_{21} )</td>
<td>( \ldots )</td>
<td>( y_{2J} )</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>F</td>
<td>19</td>
<td>0.49</td>
<td>0.55</td>
<td>( \ldots )</td>
<td>0.88</td>
<td>16</td>
<td>12</td>
<td>( \ldots )</td>
<td>22</td>
<td>( y_{30} )</td>
<td>( y_{31} )</td>
<td>( \ldots )</td>
<td>( y_{3J} )</td>
</tr>
</tbody>
</table>

- \( y_{ij} \) = Survival time, event indicator (eg CVD), number of events, continuous measurement (eg blood pressure), \( \ldots \)
- \( u_{ij} \) = Utility-based score to value health (eg EQ-5D, SF-36, Hospital Anxiety & Depression Scale), \( \ldots \)
- \( c_{ij} \) = Use of resources (drugs, hospital, GP appointments, \( \ldots \))

Compute individual QALYs and total costs as

\[
e_i = \sum_{j=1}^{J} (u_{ij} + u_{ij-1}) \frac{\delta_j}{2} \quad \text{and} \quad c_i = \sum_{j=0}^{J} c_{ij},
\]

with: \( \delta_j = \frac{\text{Time}_j - \text{Time}_{j-1}}{\text{Unit of time}} \)
“Standard” Statistical modelling — **Individual level** data in HTA

### Example Table

<table>
<thead>
<tr>
<th>ID</th>
<th>Trt</th>
<th>Sex</th>
<th>Age</th>
<th>ID</th>
<th>Trt</th>
<th>Sex</th>
<th>Age</th>
<th>ID</th>
<th>Trt</th>
<th>Sex</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>M</td>
<td>23</td>
<td>1</td>
<td>0</td>
<td>M</td>
<td>23</td>
<td>2</td>
<td>0</td>
<td>F</td>
<td>19</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>M</td>
<td>21</td>
<td>1</td>
<td>0</td>
<td>M</td>
<td>21</td>
<td>2</td>
<td>1</td>
<td>M</td>
<td>21</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>F</td>
<td>19</td>
<td>2</td>
<td>0</td>
<td>F</td>
<td>19</td>
<td>3</td>
<td>2</td>
<td>F</td>
<td>19</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ID</th>
<th>Trt</th>
<th>Sex</th>
<th>Age</th>
<th>ID</th>
<th>Trt</th>
<th>Sex</th>
<th>Age</th>
<th>ID</th>
<th>Trt</th>
<th>Sex</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>3</td>
<td>M</td>
<td>22</td>
<td>3</td>
<td>1</td>
<td>M</td>
<td>22</td>
<td>4</td>
<td>2</td>
<td>M</td>
<td>22</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>F</td>
<td>20</td>
<td>4</td>
<td>0</td>
<td>F</td>
<td>20</td>
<td>5</td>
<td>1</td>
<td>F</td>
<td>20</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>F</td>
<td>21</td>
<td>5</td>
<td>0</td>
<td>F</td>
<td>21</td>
<td>6</td>
<td>1</td>
<td>F</td>
<td>21</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>M</td>
<td>20</td>
<td>6</td>
<td>0</td>
<td>M</td>
<td>20</td>
<td>7</td>
<td>1</td>
<td>M</td>
<td>20</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>F</td>
<td>20</td>
<td>7</td>
<td>0</td>
<td>F</td>
<td>20</td>
<td>8</td>
<td>1</td>
<td>F</td>
<td>20</td>
</tr>
</tbody>
</table>

| 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, ...) |

--

#### Compute individual QALYs and total costs as

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

#### (Often implicitly) assume normality and linearity and model independently individual QALYs and total costs by controlling for baseline values

\[
e_i = \alpha e_0 + \alpha e_1 u_{0i} + \alpha e_2 Trt_i + \varepsilon_{ei} [+ ...], \quad \varepsilon_{ei} \sim \text{Normal}(0, \sigma_e) \]
\[
c_i = \alpha c_0 + \alpha c_1 c_{0i} + \alpha c_2 Trt_i + \varepsilon_{ci} [+ ...], \quad \varepsilon_{ci} \sim \text{Normal}(0, \sigma_c) \]


"Standard" Statistical modelling — **Individual level** data in HTA

| 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   | ... | 8 77 | $y_{20}$ | $y_{21}$ | ... | $y_{2J}$ |
| 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, ...)

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}, \quad \text{[with: } \delta_j = \frac{\text{Time}_j - \text{Time}_{j-1}}{\text{Unit of time}}]\]

(Often implicitly) assume **normality** and linearity and model **independently** individual QALYs and total costs by **controlling for baseline values**

$$e_i = \alpha e_0 + \alpha e_1 u_0 i + \alpha e_2 Trt_i + \varepsilon_{ei} [\ldots], \quad \varepsilon_{ei} \sim \text{Normal}(0, \sigma_e)$$

$$c_i = \alpha c_0 + \alpha c_1 c_0 i + \alpha c_2 Trt_i + \varepsilon_{ci} [\ldots], \quad \varepsilon_{ci} \sim \text{Normal}(0, \sigma_c)$$

Estimate population average cost and effectiveness differentials and use bootstrap to quantify uncertainty
What's wrong with this?...

- Potential correlation between costs & clinical benefits
  - Strong positive correlation — effective treatments are innovative and result from intensive and lengthy research ⇒ are associated with higher unit costs
  - Negative correlation — more effective treatments may reduce total care pathway costs e.g. by reducing hospitalisations, side effects, etc.
  - Because of the way in which standard models are set up, bootstrapping generally only approximates the underlying level of correlation — **Bayesian methods usually do a better job!**
What’s wrong with this?...

• Potential correlation between costs & clinical benefits
  – Strong positive correlation — effective treatments are innovative and result from intensive and lengthy research ⇒ are associated with higher unit costs
  – Negative correlation — more effective treatments may reduce total care pathway costs e.g. by reducing hospitalisations, side effects, etc.
  – Because of the way in which standard models are set up, bootstrapping generally only approximates the underlying level of correlation — **Bayesian methods usually do a better job!**

• Joint/marginal normality not realistic
  – Costs usually skewed and benefits may be bounded in [0; 1]
  – Can use transformation (e.g. logs) — but care is needed when back transforming to the natural scale
  – Should use more suitable models (e.g. Beta, Gamma or log-Normal) — **generally easier under a Bayesian framework**
What’s wrong with this?...

- Potential correlation between costs & clinical benefits
  - Strong positive correlation — effective treatments are innovative and result from intensive and lengthy research ⇒ are associated with higher unit costs
  - Negative correlation — more effective treatments may reduce total care pathway costs e.g. by reducing hospitalisations, side effects, etc.
  - Because of the way in which standard models are set up, bootstrapping generally only approximates the underlying level of correlation — Bayesian methods usually do a better job!

- Joint/marginal normality not realistic
  - Costs usually skewed and benefits may be bounded in [0; 1]
  - Can use transformation (e.g. logs) — but care is needed when back transforming to the natural scale
  - Should use more suitable models (e.g. Beta, Gamma or log-Normal) — generally easier under a Bayesian framework

- ... and of course Partially Observed data
  - Can have item and/or unit non-response
  - Missingness may occur in either or both benefits/costs
  - The missingness mechanisms may also be correlated
  - What exactly to adjust for, at baseline — available vs complete cases!
Known unknowns?...


**Missing cost (2003-2009)**
- Unclear: 20
- Others: 12
- MI: 8
- Cond: 8
- Mean: 9
- Lin Ext: 2
- LVCF: 2
- CCA: 27

**Missing cost (2009-2015)**
- Unclear: 22
- Others: 9
- MI: 27
- Cond: 2
- Mean: 4
- Lin Ext: 1
- LVCF: 4
- CCA: 12

**Missing effectiveness (2003-2009)**
- Unclear: 21
- Others: 8
- MI: 13
- Cond: 2
- Mean: 7
- Lin Ext: 4
- LVCF: 9
- CCA: 24

**Missing effectiveness (2009-2015)**
- Unclear: 8
- Others: 11
- MI: 28
- Cond: 3
- Mean: 7
- Lin Ext: 1
- LVCF: 6
- CCA: 17
Motivating 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)
Motivating 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)

<table>
<thead>
<tr>
<th>Time</th>
<th>Type of outcome</th>
<th>observed (%)</th>
<th>observed (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control (n_1=75)</td>
<td>Intervention (n_2=84)</td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>utilities</td>
<td>72 (96%)</td>
<td>72 (86%)</td>
</tr>
<tr>
<td>3 months</td>
<td>utilities and costs</td>
<td>34 (45%)</td>
<td>23 (27%)</td>
</tr>
<tr>
<td>6 months</td>
<td>utilities and costs</td>
<td>35 (47%)</td>
<td>23 (27%)</td>
</tr>
<tr>
<td>12 months</td>
<td>utilities and costs</td>
<td>43 (57%)</td>
<td>36 (43%)</td>
</tr>
<tr>
<td><strong>Complete cases</strong></td>
<td>utilities and costs</td>
<td>27 (44%)</td>
<td>19 (23%)</td>
</tr>
</tbody>
</table>
Motivating 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)
Motivating 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)

**Control**

- CC mean = 0.919
- AC mean = 0.881

**Intervention**

- CC mean = 0.831
- AC mean = 0.868
Cost-effectiveness analysis (1)

Normal & Independent outcomes

Cost–Effectiveness Plane

Cost–Effectiveness Acceptability Curve

Δ\(e\) = −0.002 (−0.045, 0.042)
Δ\(c\) = −18 (−125, 89)

Δ\(e\) = 0.041 (−0.004, 0.086)
Δ\(c\) = −18 (−125, 89)

Gianluca Baio (UCL)
**Bivariate Normal**

- Simpler and closer to “standard” frequentist model
- Account for correlation between QALYs and costs

Conditional model for $c | e$

$$c_{it} | e_{it} \sim \text{Normal}(\phi_{cit}, \psi_{ct})$$

$$\phi_{cit} = \mu_{ct} + \beta_t(e_{it} - \mu_{et})$$

Marginal model for $e$

$$e_{it} \sim \text{Normal}(\phi_{eit}, \psi_{et})$$

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

$$= \mu_{et} + \alpha_t u^*_{0it}$$
Bivariate Normal
- Simpler and closer to “standard” frequentist model
- Account for correlation between QALYs and costs

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 $e^*$
$e_{it}^* \sim \text{Beta}(\phi_{eit}\psi_{et}, (1 - \phi_{eit})\psi_{et})$
$logit(\phi_{eit}) = \mu_{et} + \alpha_t(u_{0it} - \bar{u}_{0t})$
$= \mu_{et} + \alpha_t u_{0it}^*$
Bivariate Normal
- Simpler and closer to "standard" frequentist model
- Account for correlation between QALYs and costs

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

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

Model for the structural ones
$d_{it} := 1(e_{it} = 1) \sim \text{Bernoulli}(\pi_{it})$
$logit(\pi_{it}) = X_{it} \eta_t$
Complete cases only
All cases (missing at random, MAR)
Results

Complete only vs all cases

Control

Hurdle Model

mean (90% HPD)

220 (118; 329)
198 (111; 282)

Beta–Gamma

231 (105; 347)
200 (111; 286)

Bivariate Normal

207 (128; 288)
234 (154; 321)

Intervention

Hurdle Model

mean (90% HPD)

234 (93; 377)
193 (84; 307)

Beta–Gamma

228 (91; 363)
189 (83; 303)

Bivariate Normal

190 (123; 254)
187 (122; 256)

costs (£)

Complete cases only
All cases (missing at random, MAR)
Bayesian multiple imputation (under MAR)

Bivariate Normal

Individuals ($n_1 = 75$)

Beta-Gamma

Individuals ($n_1 = 75$)

Hurdle model

Individuals ($n_1 = 75$)

Individuals ($n_2 = 84$)

Individuals ($n_2 = 84$)

- Imputed, observed baseline
- Imputed, missing baseline
- Observed
Cost-effectiveness analysis (2)

More complex modelling

Gianluca Baio (UCL)

Statistical issues in small/pilot CEAs

UCL HE Symposium, 5 Feb 2018
Conclusions

• Missing data are common in all experimental settings
  – Even more so in cases of small/pilot studies!
  – Sensitivity to modelling assumptions that cannot be empirically tested should be considered more thoroughly
  – Results can be wildly different — even adjustment for (different) baseline values can have a massive impact

• e-health or complex interventions there may be extra complexities
  – Target population generally “healthy” ⇒ “structural” values for cost or utilities

• Generally speaking, “standard” methods/approaches may fail to handle of these complexities!

• A full Bayesian approach to handling missing data extends standard “imputation methods”
  – Can consider MAR and MNAR with relatively little expansion to the basic model

• Particularly helpful in cost-effectiveness analysis, to account for
  – Asymmetrical distributions for the main outcomes
  – Correlation between costs & benefits
  – Structural values (eg spikes at 1 for utilities or spikes at 0 for costs)

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

Gianluca Baio (UCL)
Thank you!