Underestimation of Variance of Predicted Mean Health Utilities Derived from Multi-Attribute Utility Instruments: The Use of Multiple Imputation as a Potential Solution.

Kelvin Chan
Feb 10, 2015
Outline

• Multi-attribute utility instruments (MAUI)
  • Scoring algorithms and prediction errors
• Problems with variance underestimation
  • Influence on minimal clinically important differences
• Solution using multiple imputation
  • Proof-of-concept using the US EQ-5D-3L
Multi-attribute utility instruments

- MAUIs include EQ-5D, SF-6D, HUI-3
- Valuation studies conducted on sample of general population
- Developed using regression models
- Indirect measure of health utility, recommended in the UK, Australia and Canada
MAUI – Example EQ-5D

Figure 1. EQ-5D example from http://diabetesclinicevaluation.weebly.com/uploads/9/5/6/7/9567609/6029985.jpg?633
MAUI Valuation Studies

• Scoring algorithms are country-specific and are developed in a single valuation study.
• In a valuation study, respondents are asked to assign utilities to health states using time trade-off (TTO), discrete choice or standard gamble tasks
Better than dead task example

Now you would either live in Life A for 5 years and then die, or you would live in Life B for 10 years and then die. Would you prefer Life A or Life B, or are they the same?

- No problems in walking about
- No problems with self-care
- No problems with performing usual activities
- No pain or discomfort
- Not anxious or depressed

- Some problems in walking about
- No problems with self-care
- No problems with performing usual activities
- Moderate pain or discomfort
- Extremely anxious or depressed
Time Trade-Off (TTO)
EQ-5D scoring algorithm

\[ TTO_i = 1 - disutility_i \]

\[ disutility = X\beta \]

\[ X\beta = \text{int} + \beta_1(MO) + \beta_2(SC) + \beta_3(UA) + \beta_4(PD) + \beta_5(AD) \pm X\beta(\text{miscell.}) \]
Problems with variance underestimation

- Subject to prediction errors
  - $\pm 0.0754$ for EQ-5D
  - $\pm 0.1655$ for SF-6D
- Influence on minimal clinically important differences
  - $0.05 - 0.08$ for EQ-5D
  - $0.01 - 0.09$ for SF-6D

Problems with variance underestimation

• No method to account for variance of the estimated predicted mean health state utilities
  • E.g. subjects with the same health states will always “map” to the same health utility without variation

• Cost-effectiveness analyses based on MAUI do not capture parameter uncertainty in quality-adjusted life years
Solutions

1. Full Bayesian analysis
   • Using posterior predictive distributions of health states using original study data\textsuperscript{4}

   • Challenges:
     • Requires implementation by the original authors
     • Lack of raw data from valuation studies

\textsuperscript{4}Pullenayegum EM, Chan KKW, Feng X. EQ-5D health utilities are estimated subject to considerable uncertainty. Submitted under review.
2. Multiple imputation
   • Approximation to Bayesian treatment of parameter uncertainty, used to handle missing data\textsuperscript{5,6}
   • E.g. true mean utilities of each health state
   • Replaces missing value with a set of plausible values that represent the uncertainty about the right value to impute

\textsuperscript{5}Rubin DB. Multiple imputation for non-response in surveys. John Wiley & Sons; 1987.
Multiple Imputation

MCMC method:

The imputation I-step: draw values of $Y_{i(mis)}$ from a conditional distribution of $Y_{i(mis)}$ given $Y_{i(obs)}$

The posterior P-step: simulates the posterior population mean vector and covariance matrix from the complete sample estimates
Multiple Imputation

• Three phases:

1. Missing data are filled in $m$ times to generate $m$ complete data sets

2. The $m$ complete data sets are analyzed by using standard procedures

3. The results from the $m$ complete data sets are combined for the inference
Multiple Imputation

With $m$ imputations, you can compute $m$ different sets of the point and variance estimates for a parameter $Q$. Let $\hat{Q}_i$ and $\hat{U}_i$ be the point and variance estimates from the $i$th imputed data set, $i=1, 2, \ldots, m$. Then the point estimate for $Q$ from multiple imputations is the average of the $m$ complete-data estimates:

$$\overline{Q} = \frac{1}{m} \sum_{i=1}^{m} \hat{Q}_i$$

Let $\overline{U}$ be the within-imputation variance, which is the average of the $m$ complete-data estimates

$$\overline{U} = \frac{1}{m} \sum_{i=1}^{m} \hat{U}_i$$

And $B$ be the between-imputation variance

$$B = \frac{1}{m-1} \sum_{i=1}^{m} (\hat{Q}_i - \overline{Q})^2$$
Multiple Imputation

The total variance is:

\[ T = \bar{U} + \left(1 + \frac{1}{m}\right)B \]
Multiple imputation method.
Solutions – Multiple Imputation

- Sets are drawn from posterior predictive distributions only once
- If made public, sets can be used in future studies
- Removes requirement for full Bayesian analysis each time an MAUI is used
Aims

• Using the US EQ-5D valuation study,

• Demonstrate that multiple imputation can correct underestimation of variance of mean health utilities
Methods

• US EQ-5D valuation study\(^7\):
  • \(N = 3,773\), 42 health states
  • 2 dummy independent variables
  • Uses D1 model

• Using data from above study, create Bayesian mixed effect model

Bayesian mixed effect model

Let $i$ = respondents (3,773), $j$ = health states (N = 42), $X\beta$ = linear predictor of the US D1 model, which included the 2 dummy variables for each dimension (i.e. a total of 10 dummy variables for 5 dimensions), I3, I2 and I2-squared, $\mu_{ij}$ = time-trade-off utility of the $i^{th}$ respondent and the $j^{th}$ health state.

$$\mu_{ij} = 1 - (fixed_{ij} + \text{random}_{ij})$$

$$fixed_{ij} = X_{ij}\beta$$

$$random_{ij} = u_i + \delta_j$$

where $u_i$ represents the random effects for respondents ($u_i \sim N(0, \tau_{\text{respondents}})$), and $\delta_j$ represents the random effects for health states ($\delta_j \sim N(0, \tau_{\text{states}})$).
Bayesian mixed effect model

- Take posterior predictive joint distributions of the mean utility to
  - Capture parameter uncertainty
  - Perform multiple imputations (imputed sets drawn randomly from the Gibbs sampler)
Full Bayesian Analysis

- Analysis conducted using US EQ-5D study, divided into 2 sets
  - Derivation set: N = 3,273
  - Application set: N = 500

- Derivation set used D1 model, which was fitted to Bayesian mixed effect model
  - Obtained posterior predictive distribution of the mean utility attached to each health state
  - Applied to application set to compute mean and variance
Multiple imputation

- Multiple imputation with Monte Carlo Markov chain models
  - Performed using derivation set
  - Randomly drawn multiple imputed sets applied to application dataset

- Mean, variance and standard error across imputed sets calculated using Rubin’s rule\textsuperscript{10}

\textsuperscript{10}Rubin DB. Multiple imputation for non-response in surveys. John Wiley & Sons; 1987.
Methods to illustrate that multiple imputation can be used to correct for the underestimation of variance of mean health utilities of the sample.
Results

Comparisons of the sample mean and sample standard error of the mean health utility of the application set (N = 500) based on (i) the regression coefficients (i.e. scoring algorithm), (ii) the full Bayesian model's posterior predictive distribution and multiple imputation

<table>
<thead>
<tr>
<th></th>
<th>Traditional Method (based on coefficients)</th>
<th>Multiple Imputation</th>
<th>Full Bayesian Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.873</td>
<td>0.876</td>
<td>0.873</td>
</tr>
<tr>
<td>Variance</td>
<td>5.30 × 10⁻⁵</td>
<td>1.23 × 10⁻⁴</td>
<td>1.19 × 10⁻⁴</td>
</tr>
<tr>
<td>SE</td>
<td>7.28 × 10⁻³</td>
<td>1.11 × 10⁻²</td>
<td>1.09 × 10⁻²</td>
</tr>
</tbody>
</table>
95% of confidence intervals (CI)/credible regions (CR) of sample mean utility of the application set
Discussion

- Multiple imputation provides “middle ground”
- Researchers do not have to learn Bayesian methods
- Variance and standard error reflect appropriate degree of parameter uncertainty
Limitations

- Need original publishers of MAUI studies to apply this imputation method
- Some imputed value sets appear to be logically inconsistent
- Did not take population weights of US D1 model into account
Conclusions

MI is a potential method to account for the underestimation of variance of predicted health utilities.
Acknowledgement

- Eleanor Pullenayegum
- Andy Willan
- Wendy Lou
Simulation Study

- Used 100 simulated sets sampled from the 42 health states in the US EQ-5D valuation study

- Full Bayesian analysis convergence confirmed using methods by Raftery and Lewis\(^8\) and Heidelberg and Welch\(^9\)

---


Simulation Study

- Using the posterior predictive distributions of each of the 42 health states,
  - 10 samples drawn randomly to construct 95% CI for mean utility
  - 9 d.f.

- Compare the 100 CIs generated from simulation with those of the original study
Results

- Simulation study: >95% coverage of 95% CI
- 38/42 health states covered

Table 1. Parameter estimates of the coefficients and 95% credible regions (CR) of the Bayesian mixed effects model using the dataset of US EQ-5D-3L valuation study.

<table>
<thead>
<tr>
<th>D1 model parameter</th>
<th>Coefficients</th>
<th>95% CR of coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>MO2</td>
<td>0.157</td>
<td>0.120 0.197</td>
</tr>
<tr>
<td>MO3</td>
<td>0.526</td>
<td>0.468 0.586</td>
</tr>
<tr>
<td>SC2</td>
<td>0.178</td>
<td>0.146 0.211</td>
</tr>
<tr>
<td>SC3</td>
<td>0.443</td>
<td>0.391 0.496</td>
</tr>
<tr>
<td>UA2</td>
<td>0.136</td>
<td>0.093 0.183</td>
</tr>
<tr>
<td>UA3</td>
<td>0.351</td>
<td>0.309 0.394</td>
</tr>
<tr>
<td>PD2</td>
<td>0.174</td>
<td>0.141 0.208</td>
</tr>
<tr>
<td>PD3</td>
<td>0.503</td>
<td>0.460 0.547</td>
</tr>
<tr>
<td>AD2</td>
<td>0.159</td>
<td>0.124 0.193</td>
</tr>
<tr>
<td>AD3</td>
<td>0.413</td>
<td>0.372 0.454</td>
</tr>
<tr>
<td>D1</td>
<td>-0.139</td>
<td>-0.178 -0.105</td>
</tr>
<tr>
<td>I2-squared</td>
<td>0.008</td>
<td>0.000 0.015</td>
</tr>
<tr>
<td>I3</td>
<td>-0.107</td>
<td>-0.168 -0.047</td>
</tr>
<tr>
<td>I3-squared</td>
<td>-0.009</td>
<td>-0.019 0.000</td>
</tr>
</tbody>
</table>