A comparison of direct and indirect methods for the estimation of health utilities from clinical outcomes

Mónica Hernández Alava, Allan Wailoo, Fred Wolfe and Kaleb Michaud
Introduction

• Mapping (or ‘cross-walking’) is used to estimate a utility score/index from a different outcome measure
  - clinical trials without a preference based measure
  - Within PROMS agenda as performance indicators
  - Essential element of VBP

• Mapping involves:
  • Estimating a relationship using a statistical model
  • Predicting using the estimated model

• THIS IS ESSENTIALLY A STATISTICAL ISSUE!
EQ-5D-3L UK-tariff

• Descriptive system 11111, 21123
  • 5 dimensions - mobility, self care, usual activities, pain, anxiety and depression
  • 3 levels in each dimension- no problems, some problems, extreme problems
  • 243 combinations

• Valuation (Dolan et al 1995) – utility scores
  • Analysis of preference data: 3000 individuals
Two general methods

• **Direct**: dependent variable – utility/index scores
  • 11213 -> 0.378

• **Indirect**: dependent variables – levels of descriptive system.
  • Expected index score is calculated as a second step
  • “Response mapping”
AIM: Estimate EQ-5D as a function of HAQ, Pain and other covariates – direct & indirect methods

- US not-for-profit database
- N=100,398 (16k patients)
- Adults with RA diagnosis
- Classic EQ-5D (UK tariff) distribution
- Multimodal
- Peak at 1
- Bounded top and bottom
- Gap between 1 and 0.883
Existing evidence

- Direct methods:
  - Linear regression
  - Tobit (often incorrectly applied!)
  - CLAD
  - Two-part models

Biased estimates of treatment effect
Methods and models

- **Direct methods:**
  - Adjusted Limited Dependent Variable Mixture Model
    (development of Hernández Alava et al 2012)
  - RE linear regression

- **Indirect method:**
  - Set of Generalised Ordered Probits
    (development of “Response Mapping” Gray et al 2006)
Direct method: Finite Mixture Modelling

- Useful where simple models don’t fit complex data
- Model data as a finite mixture of component models (usually of the same type)
- Often used where interest is in identifying clusters of groups
- But here we are interested in approach because of flexibility
- **Any** continuous distribution can be approximated by a mixture of normals
Mixture model example - more components
• Don’t need to use normal distributions
• More appropriate bespoke distribution
• Each component reflects EQ-5D properties
• Overcomes need for a class of “1”s
• Combination of:
  a) Adjusted dist
  AND
  b) Mixture framework
Indirect method: Random Effects Generalised Ordered Probit

- 3 point ordered discrete dependent variable for each of the five dimensions of EQ-5D
- (RE) Ordered Probit – implicit parallel regression assumption too restrictive
- Multinomial logit model BUT ignores ordinality of the dependent variable
Indirect method: Random Effects Generalised Ordered Probits

- \( q_{it}^s \) discrete dependent variables for \( s = \{ \text{mobility, self care, usual activities, pain, anxiety and depression} \} \)

\[
\begin{align*}
P(q_{it}^s = 1|x_{it}, u_i^s) &= 1 - \Phi(x_{it}\beta_1^s + u_i^s) \\
P(q_{it}^s = 2|x_{it}, u_i^s) &= \Phi(x_{it}\beta_1^s + u_i^s) - \Phi(x_{it}\beta_2^s + u_i^s) \\
P(q_{it}^s = 3|x_{it}, u_i^s) &= \Phi(x_{it}\beta_2^s + u_i^s)
\end{align*}
\]

- Expected value calculated mathematically – average of all 243 utility values weighted by their estimated probabilities
Model selection and comparisons

• Explanatory variables:
  HAQ, HAQ², pain, gender, age and age²
• BIC to choose number of mixture components – 4
• MAE & RMSE (insensitive but widely used)
• Monte Carlo simulation to generate data from models and compare to observed data
• Only 1% with HAQ > 2.5
Values exceed 1!
Conclusion/Discussion

• Linear models are **not appropriate** for mapping
  • Response mapping and mixture model approaches substantially better in all regards
  • …and it matters!

• Generalized ordered probit can be used for response mapping
  • Respects ordered nature of data
Conclusion/Discussion

• Bespoke mixture model performs best overall in this example

• Further work
  • Develop response mapping (correlations, more flexible functional forms)
  • Compare methods in other datasets/simulation/outcomes
  • How will it work with EQ-5D-5L?
    - Depends how valuations are modelled
To Discover And Understand.