Regional variation in health state preferences
Does it exist, and does it matter?

Eleanor Pullenayegum,
on behalf of the Canadian
EQ-5D-5L valuation team
Acknowledgements

• Canadian EQ-5D-5L Valuation Study Team
  – PIs: Feng Xie & Jeffrey Johnson
  – Co-Is: Stirling Bryan, Nick Bansback, Arto Ohinmaa, Lise Poissant
  – Co-ordinator: Kathy Gaebel
  – RAs: Samantha Pollard, Lidia Engel, Serena Humphries, Fatima Al Sayah, Touria Addou
  – Practicum Student: Kelly Sunderland
• Funding from CIHR and EuroQol Foundation
Why bother asking?

- Health utilities used in decision-making should reflect preferences of the population who fund healthcare
- Between-country variability well established
- Canadian-specific utilities preferred
- BUT
  - Decisions are made provincially
  - Significant differences in demographics, culture and values among provinces
  - If preferences vary among provinces, provincial utilities should be preferred
Does regional variation exist?

- Scoring algorithm
- Model mean TTO as a function of health state attributes
- Yields predicted mean utilities for those 3038 health states not directly valued
- Do provincial-specific algorithms fit better than the Canada-wide algorithm
  - In terms of mean absolute error on cross-validation
### National Scoring Algorithm

Would a separate scoring algorithm for each province fit the data better?

<table>
<thead>
<tr>
<th>Term</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.1351</td>
</tr>
<tr>
<td>Mobility</td>
<td>-0.0389</td>
</tr>
<tr>
<td>Self-care</td>
<td>-0.0458</td>
</tr>
<tr>
<td>Usual Activities</td>
<td>-0.0195</td>
</tr>
<tr>
<td>Pain/Discomfort</td>
<td>-0.0444</td>
</tr>
<tr>
<td>Anxiety/Depression</td>
<td>-0.0376</td>
</tr>
<tr>
<td>Mobility level 4 or worse</td>
<td>-0.0510</td>
</tr>
<tr>
<td>Self-care level 4 or worse</td>
<td>-0.0584</td>
</tr>
<tr>
<td>Usual Activities level 4 or worse</td>
<td>-0.1103</td>
</tr>
<tr>
<td>Pain/Discomfort level 4 or worse</td>
<td>-0.1409</td>
</tr>
<tr>
<td>Anxiety/Depression level 4 or worse</td>
<td>-0.1277</td>
</tr>
<tr>
<td>Num45^2</td>
<td>0.0085</td>
</tr>
</tbody>
</table>

Example: Utility for this state would be

\[
1.1351 - 2 \times 0.0389 - 0.0458 - 0.0195 - 4 \times 0.0444 - 2 \times 0.0376 - 0.1409 = 0.60
\]
Does regional variation exist?

Model stratified by province vs. Model ignoring province

Incorporating provinces:

- Large improvement in AIC
- $p<0.0001$ for likelihood ratio test
- Improved goodness-of-fit

<table>
<thead>
<tr>
<th>Model</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>No region interactions</td>
<td>10880</td>
</tr>
<tr>
<td>Fully stratified</td>
<td>10787</td>
</tr>
<tr>
<td>Selected interactions</td>
<td>10785</td>
</tr>
</tbody>
</table>
Do provincial algorithms fit better?

<table>
<thead>
<tr>
<th>Observed mean TTO utilities</th>
<th>Predicted utilities:</th>
<th>Mean absolute Error (MAE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provincial</td>
<td>Provincial National</td>
<td>MAE provincial</td>
</tr>
<tr>
<td>National</td>
<td></td>
<td>MAE national</td>
</tr>
</tbody>
</table>

Province-specific scoring algorithms
- Use all data, not just province-specific
- Computed using shrinkage
- Borrows strength across provinces
- Better mean-square error than provincial data alone
Regional Scoring Algorithms
Bland-Altman Plot

Can one method of measurement replace another?
Poor health states have lower utilities than national average in ON.

Better health states have higher utilities than national average in ON.

BC utilities lower than national average.

Poor health states have higher utilities than national average in QC.
Does it matter?

- Large sample $\rightarrow$ excessive power
  - Detect differences that are not important
- In the HTA context, EQ-5D is used to calculate
  - Mean utilities in selected populations
  - Incremental mean utilities
- MID for EQ-5D-3L is in the range 0.03-0.08
Does it matter? (Simulation)

- Simulated health state distributions
  - Probability of occupying each of the 3125 EQ-5D-5L health states
  - Health states that are similar had similar probabilities
  - Multivariate Normal with a probit transformation
- Simulated health states from a population of size 1000
- Applied provincial and national scoring algorithms
- Compared mean & incremental mean utilities
  - National vs. Provincial
Bland-Altman plots for mean utilities

Very sick populations – overestimate

Healthy populations - underestimate

Always overestimate

Mostly underestimate
Bland-Altman plots for incremental mean utilities
Discussion

- Regional variation exists & is important
- Canadian 5L study was not intended to provide provincial algorithms
  - Inadequate precision
  - Single city per province
  - Not all provinces included
- Pan-Canadian study would be expected to do better than province-specific studies
  - Shrinkage analysis borrows strength across provinces
- Implications for HTA in Canada?