Can we believe their beliefs?
Two tales of structured expert elicitation

Lotte Steuten
Associate Member/Professor,
Fred Hutchinson Cancer Research Center/University of Washington

Two tales…

<table>
<thead>
<tr>
<th>Photoacoustic imaging in breast cancer</th>
<th>Negative-pressure wound therapy for pressure ulcers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context: new Dx imaging with unknown accuracy</td>
<td>Context: existing therapy with limited evidence base but wide usage in practice</td>
</tr>
<tr>
<td>Key parameters elicited:</td>
<td>Key parameters elicited: treatment and progression of severe pressure ulcers</td>
</tr>
<tr>
<td>- relative performance Dx in detecting tumor characteristics</td>
<td>Why: inform cost-effectiveness model an value of further research</td>
</tr>
<tr>
<td>- estimates of sensitivity and specificity of new Dx</td>
<td>Why: inform early stage cost-effectiveness model</td>
</tr>
</tbody>
</table>

Why: inform early stage cost-effectiveness model
Rating relative performance of PAM versus MRI

Elicited:
1. relative importance seven tumor characteristics in the examination of images
   - using 0-100 point scale
2. how well MRI and PAM can visualize these characteristics by grading each characteristic with value 0 - 100
   - 0 indicates low performance; 100 indicates high performance.

Expected performance of MRI and PAM was determined by calculating performance score weighted by the relative importance of each attribute, per individual.

Tumor characteristics: mass margins; mass shape; mass size; vascularization; localization; oxygen saturation; and mechanical properties.
Elicitation Results

Elicitation Procedure for sensitivity and specificity
Eliciting the mode, then the upper and lower boundaries and by using the PERT approach a probability distribution was obtained.

Probability distribution of estimations of TPs

Haakma, Steuten, Bojke and IJzerman. Submitted, 2012
Results – Expert elicitation

<table>
<thead>
<tr>
<th></th>
<th>Early adopters</th>
<th>Majority</th>
<th>Overall</th>
<th>Lower</th>
<th>Upper</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sensitivity</strong></td>
<td>67.7%</td>
<td>91.9%</td>
<td>81.7%</td>
<td>67.4%</td>
<td>85.1%</td>
<td>75.6%</td>
</tr>
<tr>
<td><strong>Specificity</strong></td>
<td>70.2%</td>
<td>88.4%</td>
<td>79.1%</td>
<td>52.2%</td>
<td>77.6%</td>
<td>66.5%</td>
</tr>
</tbody>
</table>

Haakma, Steuten, Bojke and IJzerman. Submitted, 2012

Considerations

- Experts considered MRI (sens 90%; spec 70%) the better technology to visualize the most important tumor characteristics (mass margins and mass shape).
- Reflected in elicited TP and TN, with overall calculated sensitivity and specificity of PAM to be lower than MRI
  - Sens between 59% - 85%; mode 76%
  - Spec between 52% - 78%; mode 67%
- Radiologists perceived elicitation exercise as difficult
  - PAM is an early stage technology for which only small scale, experimental experience was available.
- Exercise provided important insights to the developers
  - Revision of the technology and reconsideration of its place in Dx pathway
Negative-Pressure Wound Therapy for Pressure Ulcers

- NPWT is widely used treatment for severe pressure ulcers
  - little robust evidence that it is (cost-) effective
  - uncertainty around cost-effectiveness would potentially be misrepresented using published trial data only
  - broad range of comparators
  - general patterns of care unclear
  - yet, lots of local / practical experience with different therapies

Objectives and design

Questions considered:
- What is the (cost-)effectiveness of NPWT given the range of alternative treatments?
- What further research (design), if any, is worthwhile?

1. Literature search
2. Beliefs elicited from experts
3. Pilot trial
Further research?
Collated inputs for decision analytic model
VOI
Elicitation of:

- all transitions and related events (except death)
  - including beliefs about the impact of the alternative treatments on the occurrence of events (relative effectiveness).
- uncertainty over the quantities of interest.
- no elicitation of resource use or cost parameters
  - to limit burden of exercise

Elicitation procedure: histogram method

Think of UK patients with at least 1 debrided grade 3 or 4 pressure ulcer (>5 cm2 in area):

What proportion of patients do you think would have a grade 3 reference ulcer (rather than a grade 4 reference ulcer)?
Collation of evidence using Bayesian updating:

Results on expected cost-effectiveness, per scenario
CE estimates and decision uncertainty (EX+EL+TR)

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Costs (£)</th>
<th>Effectiveness (QALY)</th>
<th>NHB (QALY)</th>
<th>Next Best ICER (EQALY)</th>
<th>Probability of a Treatment Being Cost-Effective</th>
</tr>
</thead>
<tbody>
<tr>
<td>NPWT</td>
<td>10,399</td>
<td>1.273</td>
<td>0.754</td>
<td>—</td>
<td>0.451</td>
</tr>
<tr>
<td>HC</td>
<td>13,461</td>
<td>1.264</td>
<td>0.501</td>
<td>Dominated</td>
<td>0.304</td>
</tr>
<tr>
<td>ALG</td>
<td>14,898</td>
<td>1.261</td>
<td>0.516</td>
<td>Dominated</td>
<td>0.230</td>
</tr>
<tr>
<td>F</td>
<td>18,969</td>
<td>1.254</td>
<td>0.305</td>
<td>Dominated</td>
<td>0.015</td>
</tr>
</tbody>
</table>

Note: NPWT = negative-pressure wound therapy; HC = spun hydrocolloid; ALG = alginate; F = foam; NHB = net health benefit; ICER = incremental cost-effectiveness ratio; QALY = quality-adjusted life year.

Optimal sample size and value of further research

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Maximum ENBS</td>
<td>Optimal Sample Size, N*</td>
<td>Maximum ENBS</td>
</tr>
<tr>
<td>0.5 years</td>
<td>—</td>
<td>—</td>
<td>£12.3 million</td>
</tr>
<tr>
<td>1 year</td>
<td>£14.0 million</td>
<td>476</td>
<td>£27.2 million</td>
</tr>
<tr>
<td>2 years</td>
<td>£27.1 million</td>
<td>389</td>
<td>£35.2 million</td>
</tr>
</tbody>
</table>
Considerations

• Elicited beliefs can be regarded as a key source of evidence
• Excluding relevant clinical experience would have misrepresented current knowledge about the effectiveness of alternative treatments for severe pressure ulcers.
• In this case study, elicited evidence was used alongside published evidence under the assumption that experts did not consider existing evidence when formulating their judgements.
  – Assumption may not be sustained in other cases, where aggregation of both sources could lead to an incorrect specification of uncertainty (double counting).

Discussion: pros and cons

• Elicitation of beliefs constitutes a reasonably low cost source of evidence;
  – Particularly important in early stage technology assessment when funding is limited, or when a technology is already adopted and there is little incentive to do further research.

• Elicitation is highly subjective and entirely dependent on the sample of experts chosen for the exercise.
  – Particularly problematic when samples are skewed towards including mainly optimists or sceptics;
  – In early stage HTA ‘realistic’ beliefs may not yet exist due to no/limited experience with technology
Discussion: pros and cons

- Can provide preliminary estimates of the importance and extent of uncertainty for particular model parameters or assumptions
  - can help to inform go/no go decisions in early stage HTA and
  - guide decision on whether and what further evidence to acquire
Thank you!

@LotteSteuten
Email: lsteuten@fredhutch.org