Comparison of mathematical and behavioural elicitation approaches: Application using a decision model for topical negative pressure (TNP) therapy for pressure ulcers.

Laura Bojke, Marta Soares, Jo Dumville

Structure of the presentation

- Background to elicitation & use in decision analysis
- Background to VAC project and need for elicitation
- Pilot elicitation exercise
- Mathematical vs. behavioural elicitation
- Results
- Synthesis issues & discussion

What is expert elicitation?

- An elicitation method is intended to link an experts beliefs to an expression of these in a statistical form.
- Elicitation techniques used in Bayesian statistics because of the need to formulate subjective probabilities.
- Expert elicitation can also be used in decision analysis to quantify unknown parameters in the absence of actual data.
- Decision analysis has typically employed less formal elicitation techniques (consulting experts for ‘best guess’.

Eliciting experts priors

- No standard protocols for the conduct of elicitation assessments.
- Much is context specific, but there are a number of issues to consider, including:
  - General approach to elicitation (behavioural or mathematical)
  - What quantities to elicit
  - Who to elicit from
  - Elicitation method (interval method, histograms)
  - Synthesis approach
  - Assessing adequacy

General approach to elicitation

- Behavioural
  - Focus of achieving consensus
  - Delphi panel and nominal group method are the most popular
  - Can be skewed by dominant individuals
  - Tenancy to produce over-confident estimates
- Mathematical
  - Elicit individually and then synthesise
  - Issue of which synthesis method to use
  - Can weight experts using various methods

What quantities to elicit?

- No consensus on what is most appropriate and performs best.
- Many exercises have sought to elicit probabilities or numbers of events; however costs, utilities and relative risks can also be elicited.
- Generally accepted that experts should not be asked to estimate moments of a distribution, nor should they be asked to estimate unobservable quantities or covariates.

Who to elicit from

- “Substantive expert in the particular area”.
- Not clear if the expert should possess and elicitation skills.
- Some tasks may be easier with statistically trained experts.
- Multiple experts better than a single expert.
- How many?....as many as you can get!
- More complex synthesis methods require a greater number of experts.

Elicitation method

- The simplest elicitation is to ask an expert for a single estimate – direct questioning or gambling methods.
- Usually require an estimate of uncertainty, so we need to elicit a distribution.
  - Direct questioning – experts concept of variance difficult
  - Histograms
  - Interval methods (fixed or variable)

Synthesis approach

- Elicitation using the mathematical approach requires synthesis.
- First fit a distribution then combine distributions.
- Number of approaches available:
  - No synthesis – experts as scenarios
  - Linear/logarithmic pooling
  - Paired comparisons
  - Supra bayes methods
  - And multiple variations
- Histogram distribution allows you to use empiric distribution.

Assessing adequacy

- Many characteristics of elicitation that may or may not be adequate.
- Internal consistency
- Fitness for purpose
- Scoring rules
- Calibration
  - Most commonly used method to assess validity
  - Degree of accordance between assessments and actual observed outcomes

The TNP project

- Topical negative pressure therapy for pressure ulcers: value of information analysis and feasibility.
- Decision model developed and comparators selected with nurses
- Systematic and targeted reviews undertaken — limited data to inform model

TNP pilot elicitation

- Seven tissue viability nurses (both community and hospital nurses all experienced)
- Facilitator-lead session
- A paper based questionnaire with 6 questions was designed to elicit the uncertainty distribution surrounding 6 parameters.
- Only 2 of the questions served to elicit unknown quantities The remaining questions were elicited in order to calibrate experts

- Q1: Elderly (mean age 83 years) patients based in a UK hospital ward with no signs of a pressure ulcer on admittance (including no blanching erythema). All these patients are deemed at risk of developing a pressure ulcer (any grade). From admittance all patients are placed on a visco-polymer foam mattress (i.e. CONFOR-Med).What proportion of patients do you think would develop a pressure ulcer (grade I-IV) by day 14?

- Q2: Hospitalised patients with a range of open surgical wounds that require dressing (mean size of the wounds 5cm^2). These wounds are not burns, or malignant ulcers. Wounds were dressed with gauze alone. What proportion of patients do you think will have healed 6 months after surgery?

- **Q3: Hospitalised patients with one completely debrided grade III/IV pressure ulcer (> 5 cm^2). Treatment with VAC therapy. What proportion of patients do you think will have healed after 6 months?**

- **Q4: Community based patients with one completely debrided grade III/IV pressure ulcer (> 5 cm^2). Treatment with VAC therapy. What proportion of patients do you think will have healed after 6 months?**

- Q5: Community patients with a confirmed venous leg ulcer, eligible for high compression therapy and treated with four-layer compression bandaging. What is the MEDIAN (measure of central tendency) time to healing for those that respond to treatment?

- Q6: Community patients with a confirmed venous leg ulcer, eligible for high compression therapy and treated with four-layer compression bandaging. When can we expect 70% of patients (who have responded to treatment) to be healed?

Q1-4: percentile elicitation

Q5-6: positive values; likely to be skewed

Histogram approach used so same format could be used for all questions.

For each question a discrete scale was predefined and experts asked to place 20 crosses on a frequency chart

Percentage (proportion)	Frequency
0%	
5%	
10%	
15%	
20%	
25%	
30%	
35%	
40%	
45%	
50%	
55%	
60%	
65%	
70%	
75%	
80%	
85%	
90%	
95%	
100%	

Individual method

- Questions completed by experts in isolation

- Calibration/synthesis

Calibration

- Q1-2 and 5-6 asked about values for which there is research data.
- The known parameters were used to calibrate experts in the mathematical approach = weighting index for each expert
- Weights used to adjust the estimates of unknown quantities. Expert with the highest weight contributes the most to the pooled estimate and the expert with the lowest weight contributes the least.

Synthesis

- In the individual approach expert's assessments for the unknown parameters were synthesised using the linear pooling method

Consensus

- Group
- Calibration/synthesis implicit

Comparison of mathematical and behavioural approaches

- Comparing the consensus and mathematical estimates of known parameters graphically
- Feedback from nurses

Results (1) — Elicited and known values

	Q1 (%)	Q2 (%)	Q3 (%)	Q4 (%)	Q5 (days)	Q6 (days)
Consensus	12 (6) 0-25	78 (6) 65-85	77 (8) 65-90	82 (8) 70-95	136 (22) 91-182	91(22) 45-136
Expert 1	20 (8) 5-35	25 (8) 15-40	30 (7) 20-45	31 (12) 15-55	182 (19) 152-213	191 (11) 167-213
Expert 2	10 (4) 5-20	50 (32) 0-100	67 (7) 55-80	46 (6) 35-55	147 (31) 91-213	75 (19) 46-106
Expert 3	65 (8) 50-80	56 (7) 40-65	92 (5) 80-100	78 (7) 65-90	119 (14) 91-137	114 (18) 76-137
Expert 4	10 (4) 5-15	10 (4) 5-20	80 (5) 75-95	76 (5) 70-85	128 (25) 91-182	123 (20) 91-152
Expert 5	7 (3) 5-15	53 (11) 35-70	86 (5) 80-95	87 (5) 80-95	144 (17) 122-167	106 (12) 91-122
Expert 6	4(4) 0-15	88 (13) 50-100	87(9) 65-100	85 (9) 70-100	91 (24) 46-122	158 (19) 122-182
Expert 7	40 (12) 20-65	76 (6) 65-85	68 (7) 55-85	52 (9) 35-70	157 (14) 137-182	228 (12) 213-243
Known quantities	14.98 (7.2)	75.12 (20.7)	-	-	90 (5.7)	160 (13)

Calibration values

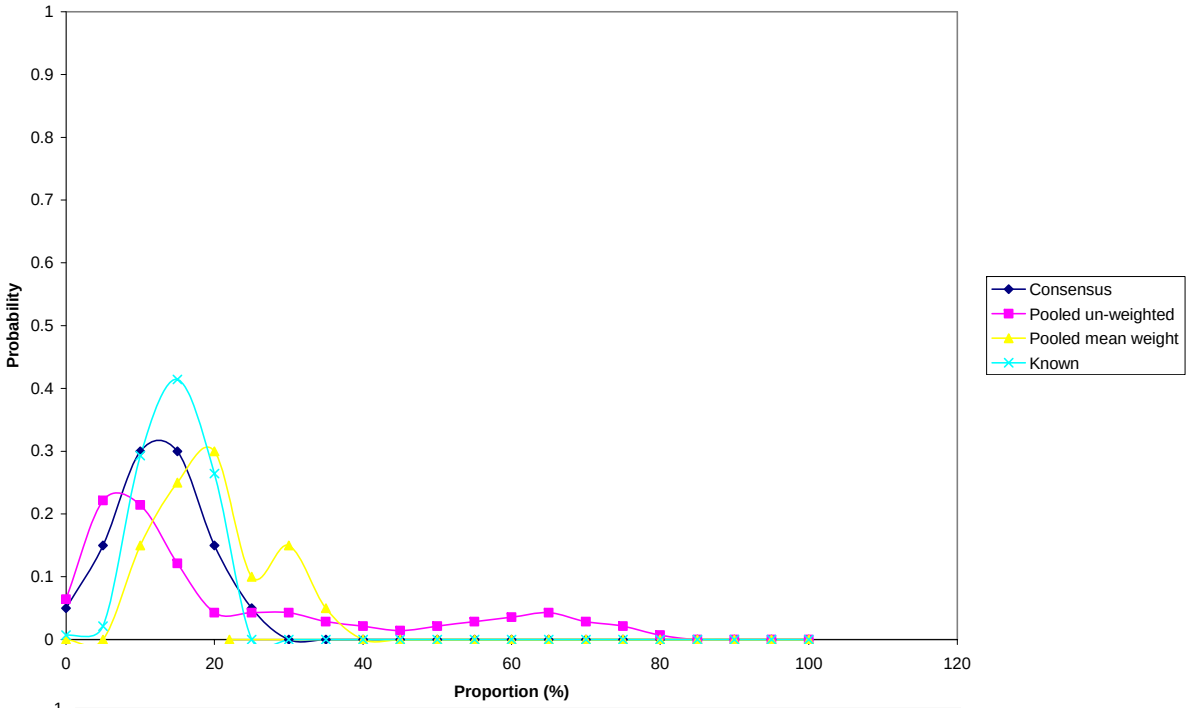
	Q1	Q2	Q5	Q6	Mean weight
Expert 1	0.20	0.00	0.00	0.21	0.10
Expert 2	0.28	0.11	0.10	0.00	0.12
Expert 3	0.00	0.01	0.23	0.09	0.08
Expert 4	0.32	0.00	0.22	0.19	0.18
Expert 5	0.12	0.06	0.02	0.02	0.05
Expert 6	0.08	0.05	0.43	0.48	0.26
Expert 7	0.01	0.77	0.00	0.00	0.19

Is calibration being successful?

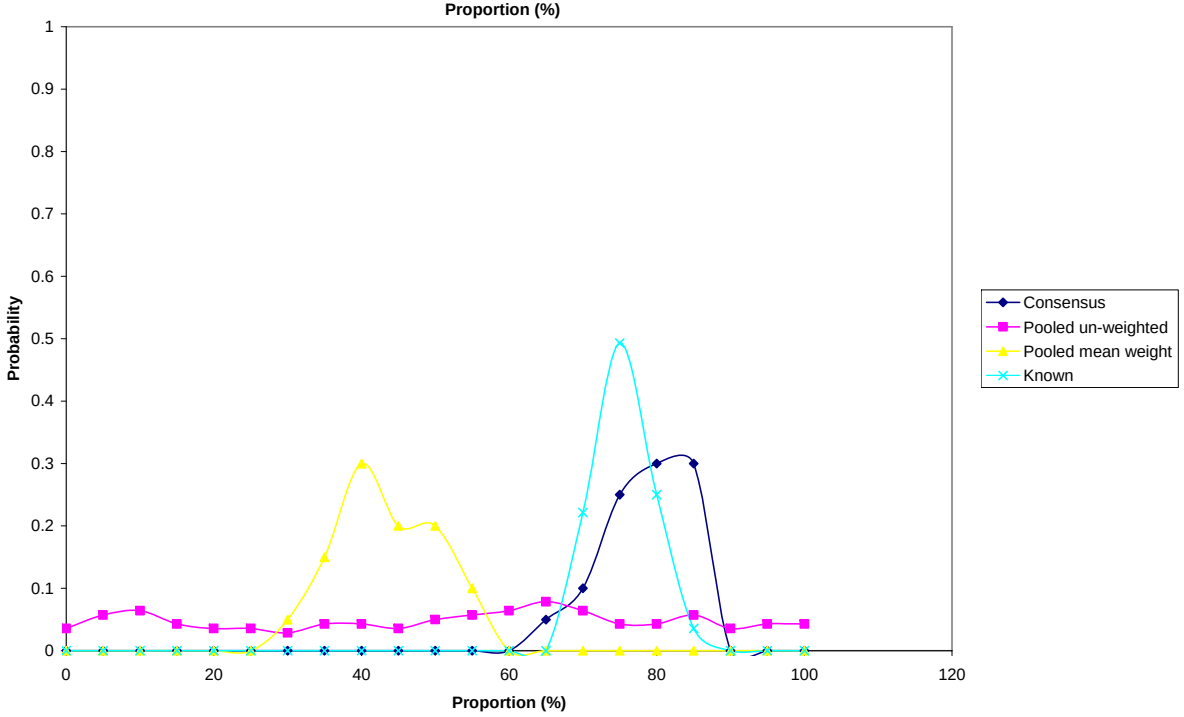
Evaluation of accuracy of performance?

Synthesis

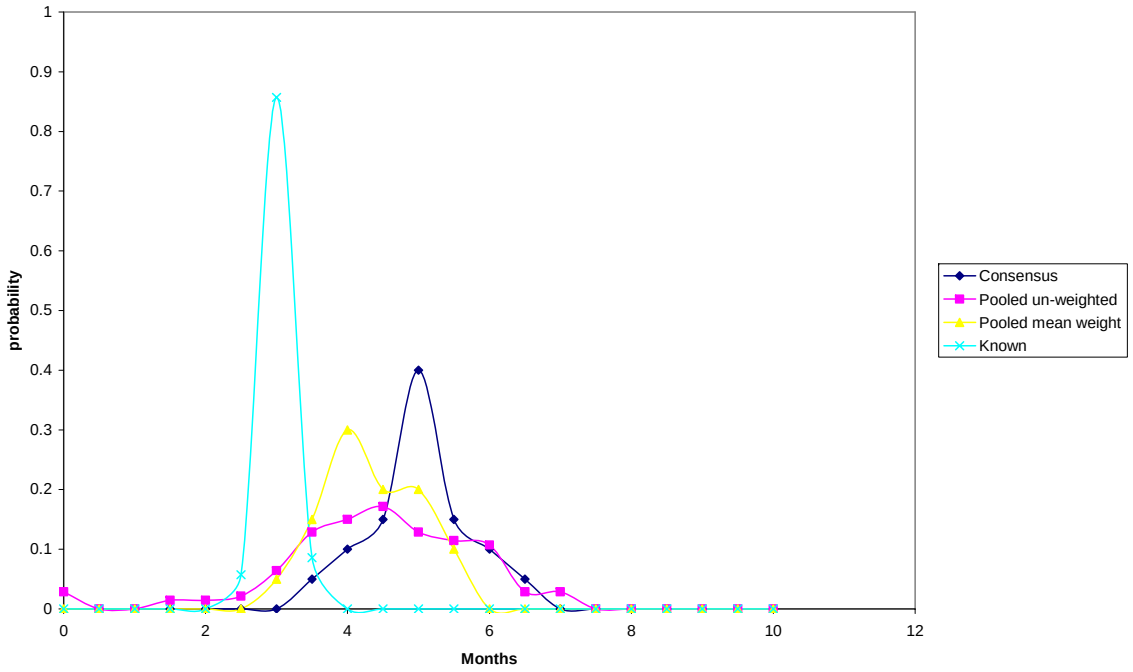
	Q1 (%)	Q2 (%)	Q3 (%)	Q4 (%)	Q5 (days)	Q6 (days)
Linear pooling – without weighting	22.5 (22)	51.4 (28.9)	73.6 (20.8)	65.4 (21.5)	138.6 (34.8)	142.3 (52.1)
Linear pooling – with weighting (Q1)	-	-	68.9 (6.4)	60.5 (7.1)	-	-
Linear pooling – with weighting (Q2)	-	-	70.2 (7.3)	55.7 (8.5)	-	-
Linear pooling – with weighting (Q5)	-	-	85.6 (7.1)	78 (7.3)	-	-
Linear pooling – with weighting (Q6)	-	-	75.1 (7.6)	71.7 8.8)	-	-
Linear pooling – with weighting (mean)	20 (6.4)	56.1 (10.9)	74.9 (7.1)	66.5 (7.9)	132.2 (21.3)	151.9 (16.6)
Consensus	12.5 (6.1)	78.5 (5.9)	77.5 (8.3)	82.5 (8.3)	4.5 (0.72)	3.0 (0.72)
Known quantities	14.98 (7.2)	75.12 (20.7)	-	-	90 (5.7)	160 (13)


Probability
Proportion (%)
Consensus
Pooled un-weighted
Pooled mean weight
Known

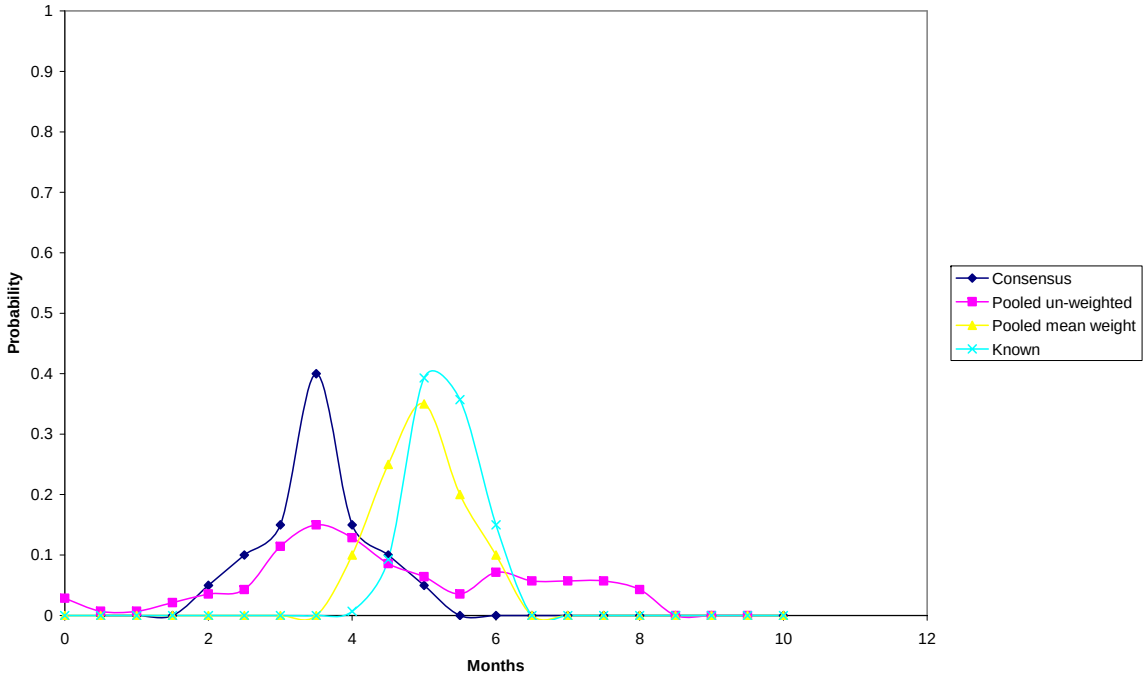
Q1


Probability
Proportion (%)
Consensus
Pooled un-weighted
Pooled mean weight
Known

Q2


probability
Months
Consensus
Pooled un-weighted
Pooled mean weight
Known

Q5


Probability
Months
Consensus
Pooled un-weighted
Pooled mean weight
Known

Q6

Feedback from experts

- Consensus method allowed more full understand of individual questions.
- Also felt that they were somewhat pushed into agreeing distributions for each of the questions
- some of the experts felt that the consensus method gave them greater confidence in expressing their opinions where as others felt that they were not qualified to contradict the opinions of experts that were perhaps considered more qualified or experienced

Conclusions

- Nurses were able to correctly complete the histograms (there was no data that could not be included)

- No clear ‘winner’

Issues of synthesis

- Which synthesis approach to use
  - Linear pooling
  - Supra bayes
  - Others
- Are methods used to synthesise trials e.g. meta analysis appropriate?
- To calibrate or not to calibrate
  - Classical method
  - Others
- Are methods used to quality adjust trials appropriate?

	Individual	Consensus
Advantages		
Disadvantages		

What next

- Larger elicitation exercise for ** parameters