THE UNIVERSITY of York

Methods to elicit experts' beliefs over uncertain quantities

in health technology assessment

Marta Soares, Laura Bojke, Jo Dumville, Nicky Cullum, Cynthia Iglesias, Karl Claxton

CHE
Centre For Health Economics

Health technology assessment context

- Decisions regarding access to health care technologies
  based on effectiveness and cost-effectiveness

- Decisions are inevitable:
  - When there is little or no data on some component, expert opinion may be sought for
  - Formal methods: standardised the processes of eliciting experts' opinions, minimise biases and heuristics, contribute to transparency

Decision rules

- NHS is collectively funded and budget constrained
  - Primary purpose is to improve **health** (of all)
- The threshold, λ
  - Given finite budget, the decision to commit to funding a technology implies other patients may loose health
- Metric of value: net health (NH)

$\Delta E - \Delta C * \lambda$ (health units)

or

$\Delta E/\lambda - \Delta C$ (monetary units)

e.g. 3 QALYs gained, £20,000 additional costs = 1 QALY displaced, NH = 2 QALY

decision rule: if NH > 0 the new technology should be adopted

Decision rule: INH >0

Case study

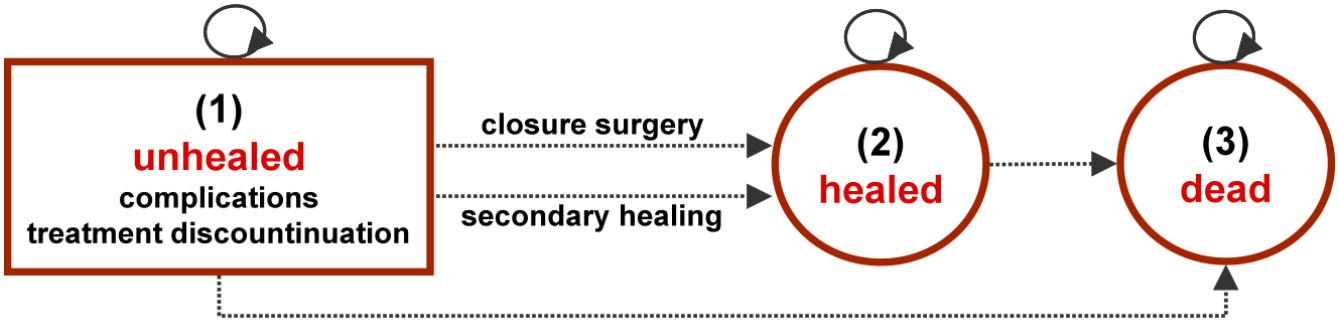
Negative Pressure Wound Therapy(NPWT) for severe pressure ulcers

- limited and sparse evidence base
- NPWT and comparators frequently used in the NHS
  - Substantial practical experience

- In assessing costs and effects, there are aspects for which
  - data existed, but was very uncertain
    - i.e. proportion of patients healing
  - no data existed
    - i.e. use of closure surgery, occurrence of complications

Case study

- Decision model



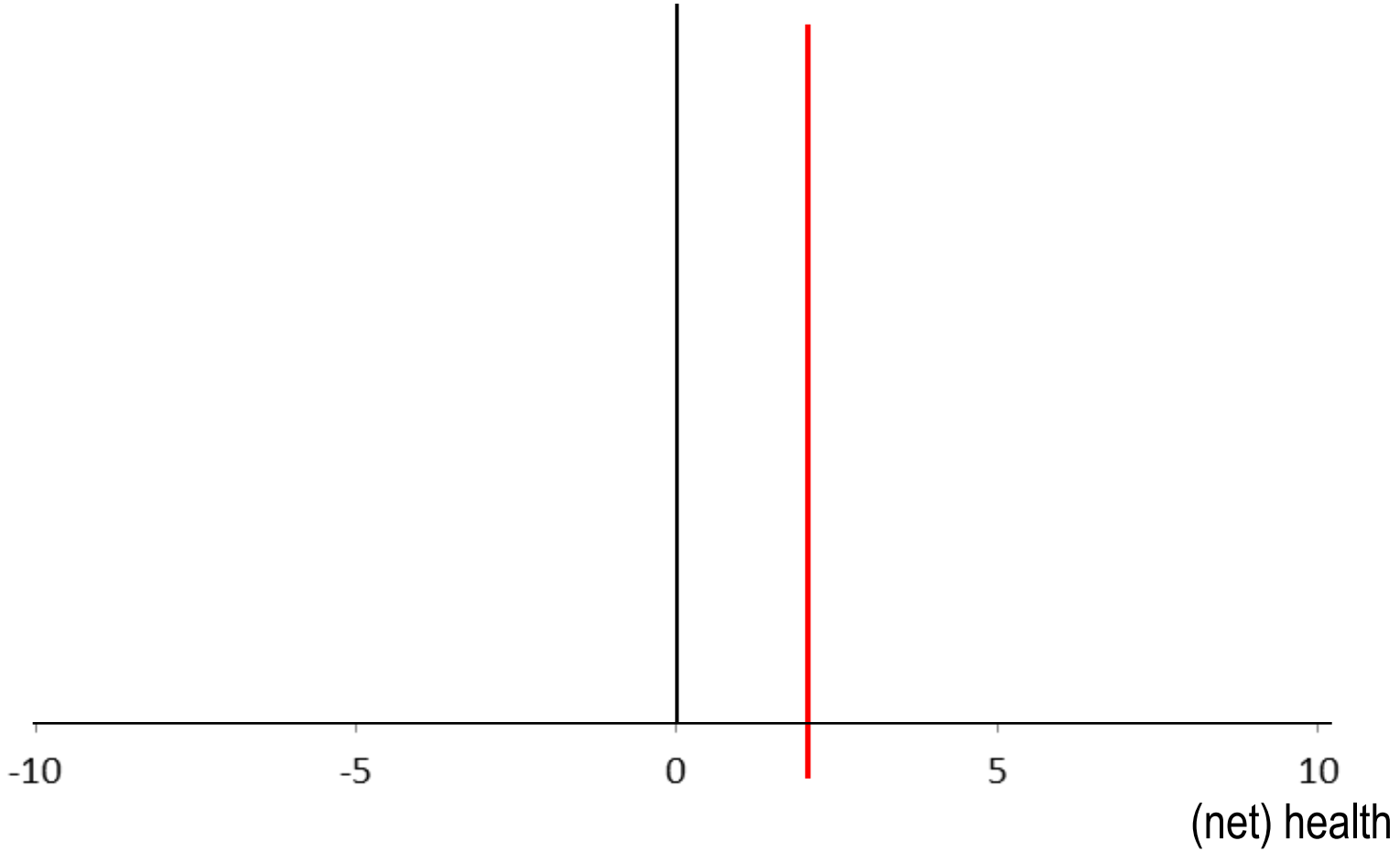
Value

Incremental (net) health, 2 alternatives

Assessment of value

Expected INH = 2 health units , >0

ADOPT

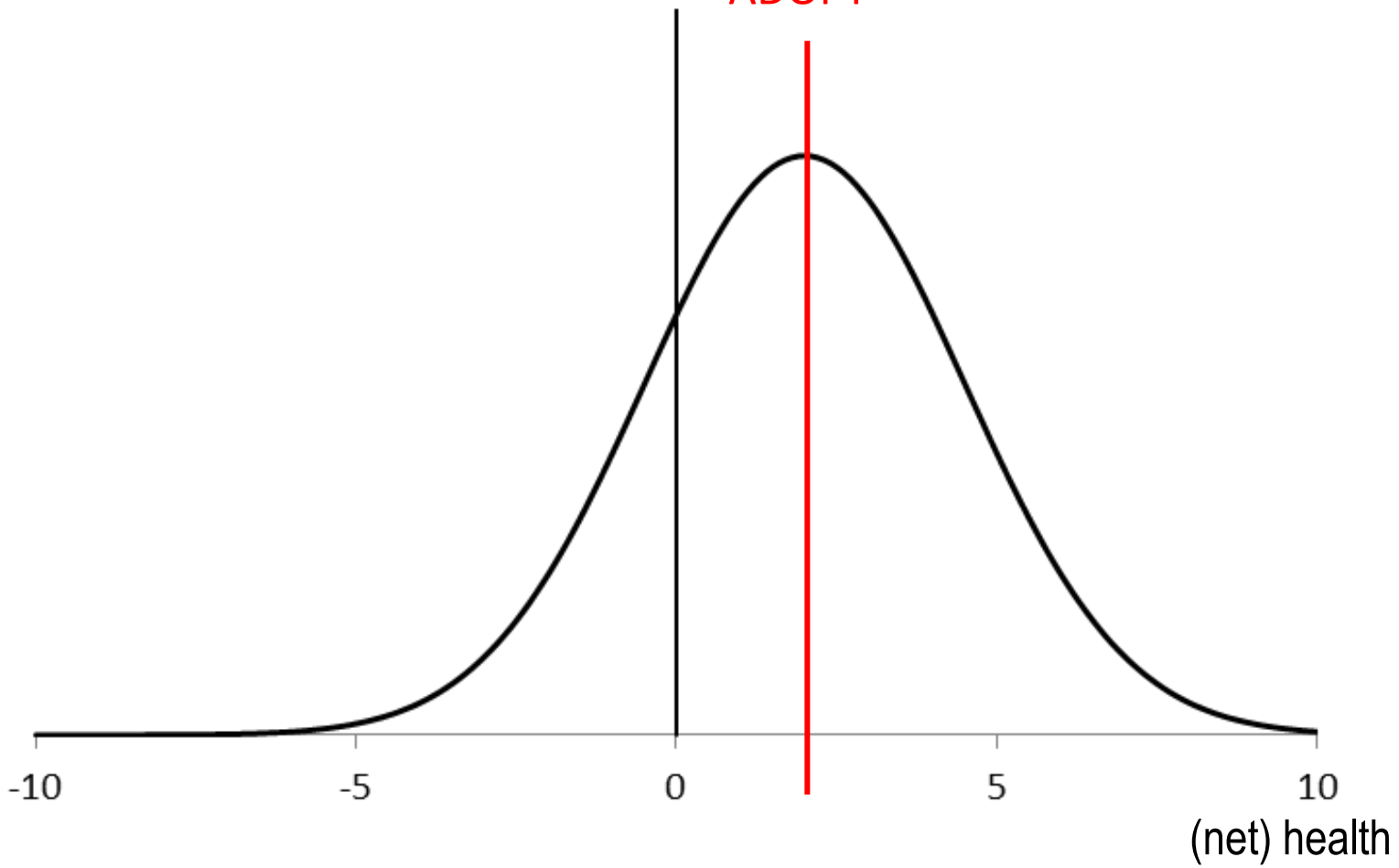


Uncertainty over value

Assessment of value

Expected value = 2 health units

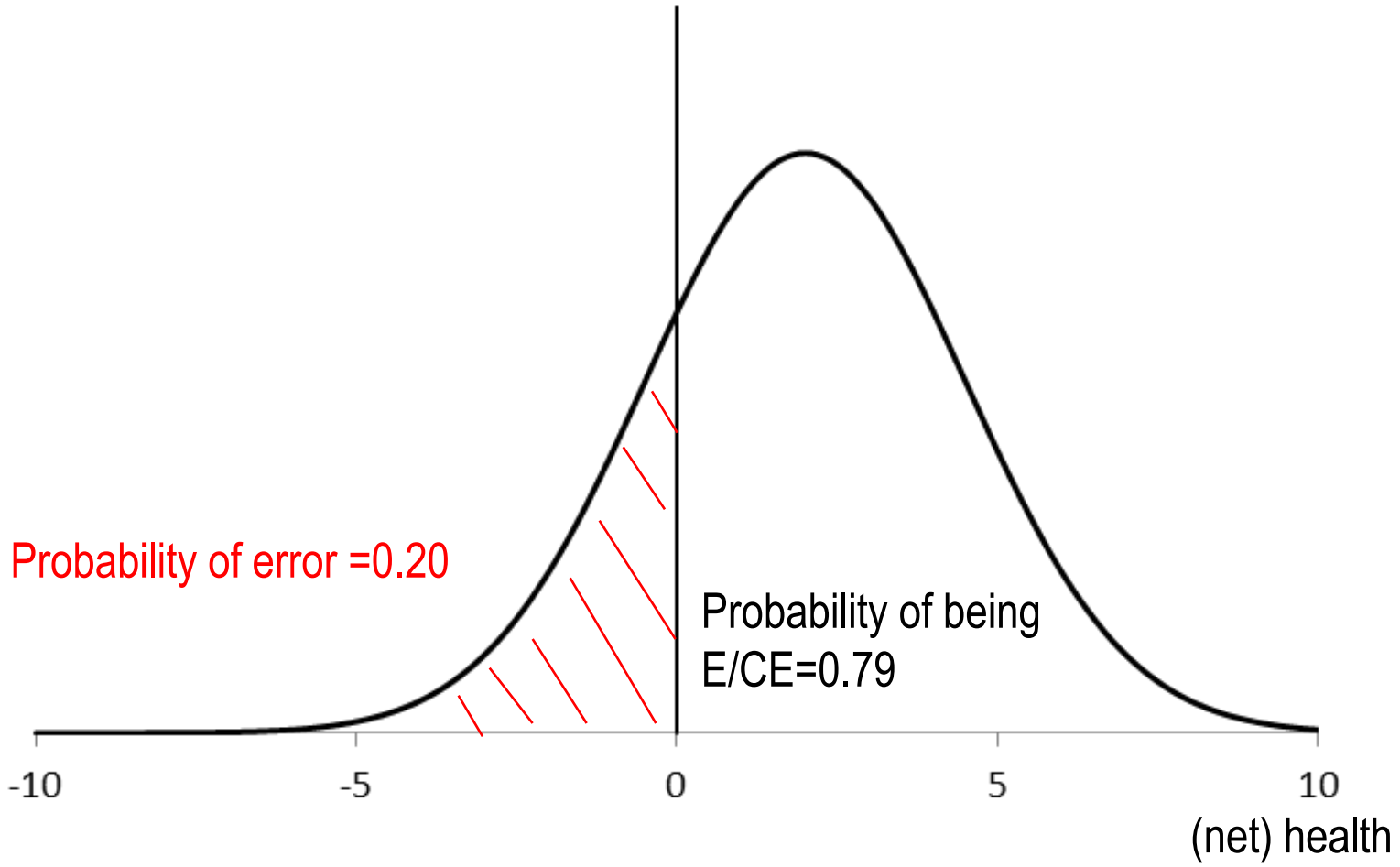
ADOPT



Uncertainty

- Uncertainty is ubiquitous, and decisions are often uncertain
- This means decisions made today may be wrong, other courses of action could potentially have been better in which case health would be lost
- Further research decreases uncertainty over decisions made today
- The value of research = value of avoiding the losses due to uncertainty

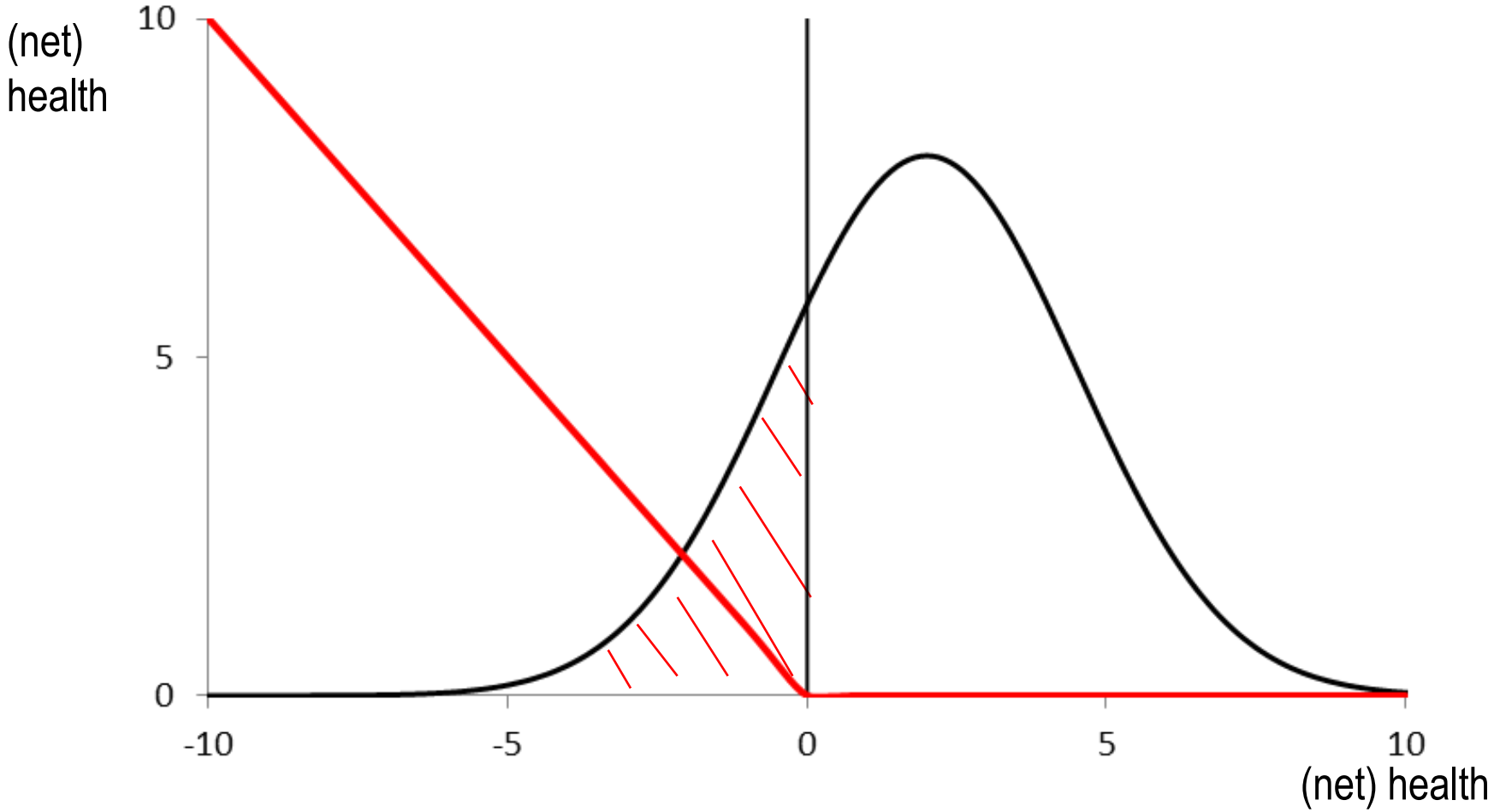
Uncertainty over value



Consequences of uncertainty

loss function

Probability of error =0.20
Expected losses = 0.3 health units

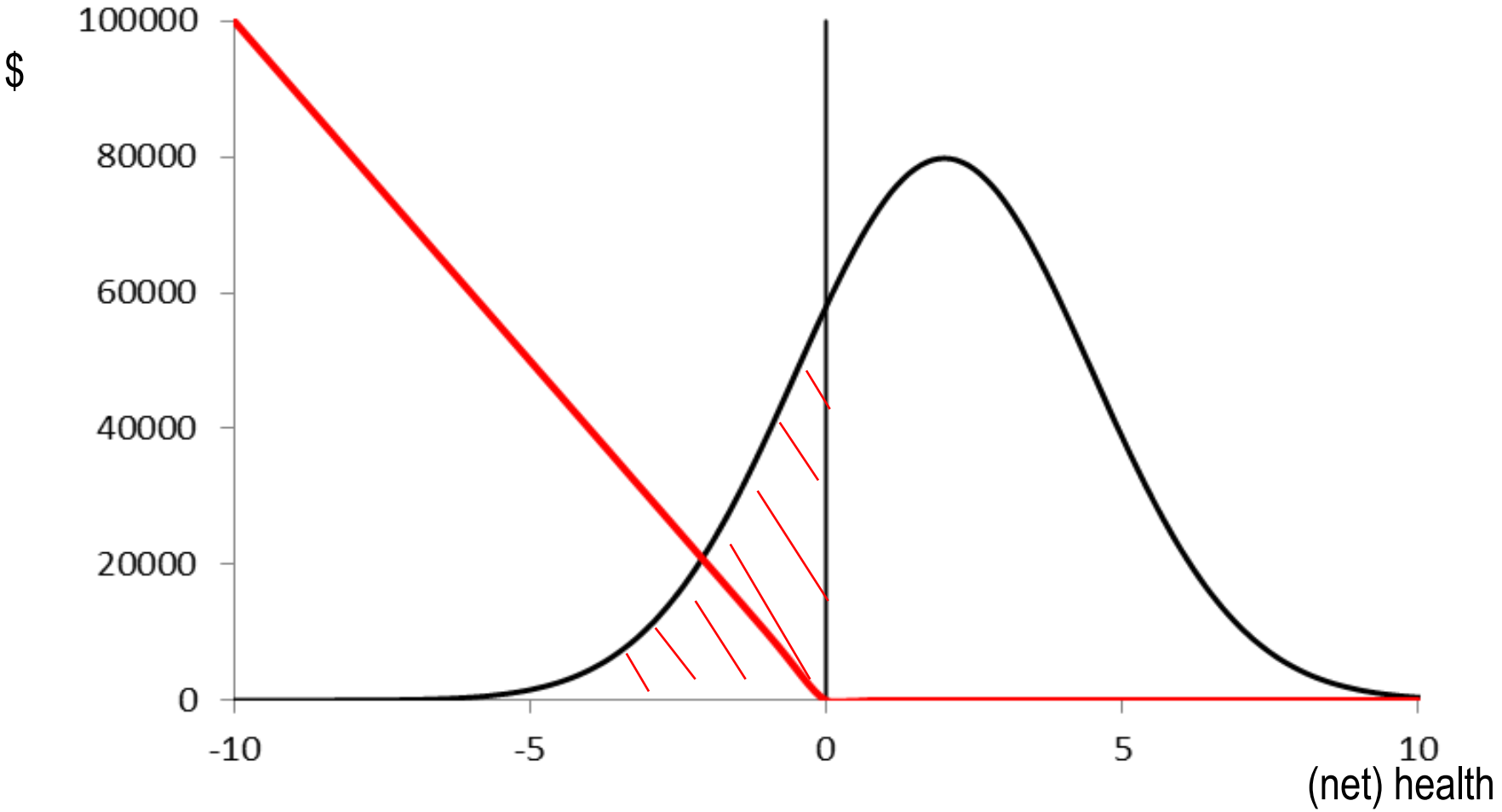


Consequences of uncertainty

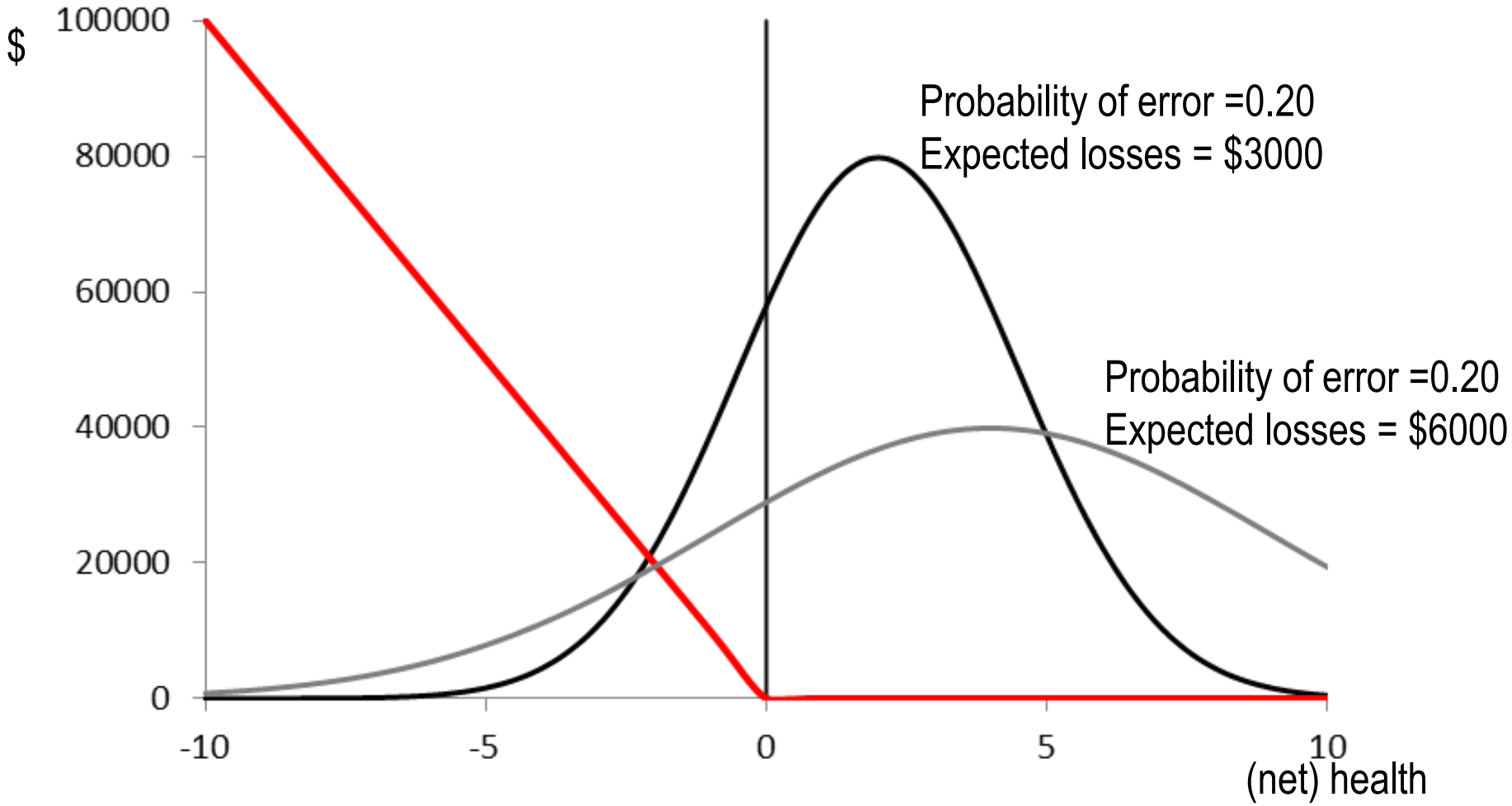
loss function: $10000/health unit

Probability of error =0.20

Expected losses = EVPI = 0.3 health units = $3000



Consequences of uncertainty



Is further evidence worthwhile?

- EVPI Value of eliminating uncertainty in all parameters = maximum return to research

$$EVPI = E_{\theta} \max_j NB(j, \theta) - \max_j E_{\theta} NB(j, \theta) = EVPI$$

- Expected value of information in a subset of input parameters

$$EVPI_{\theta 1} = E_{\theta 1} \max_j E_{\theta 2|\theta 1} NB(j, \theta 1, \theta 2) - \max_j E_{\theta} NB(j, \theta)$$

θ { $\theta 1$ = parameter of interest; $\theta 2$ = other uncertainties

- Expected value of sample information (EVSI)

$$EVSI = E_{\theta 1} E_{D|\theta 2} \max_j E_{\theta 1,(\theta 2|D)} NB(j, \theta 1, \theta 2) - \max_j E_{\theta} NB(j, \theta 1, \theta 2)$$

What is expert elicitation?

- A process that aids experts to formulate a quantitative judgement based on their own beliefs for a specific quantity

  an elicitation is intended to link an expert's beliefs to an expression of these in a statistical (numerical) form – basically getting them down on paper.

- Although formal elicitation techniques have been seldom used, expert opinion is commonly asked for informally.

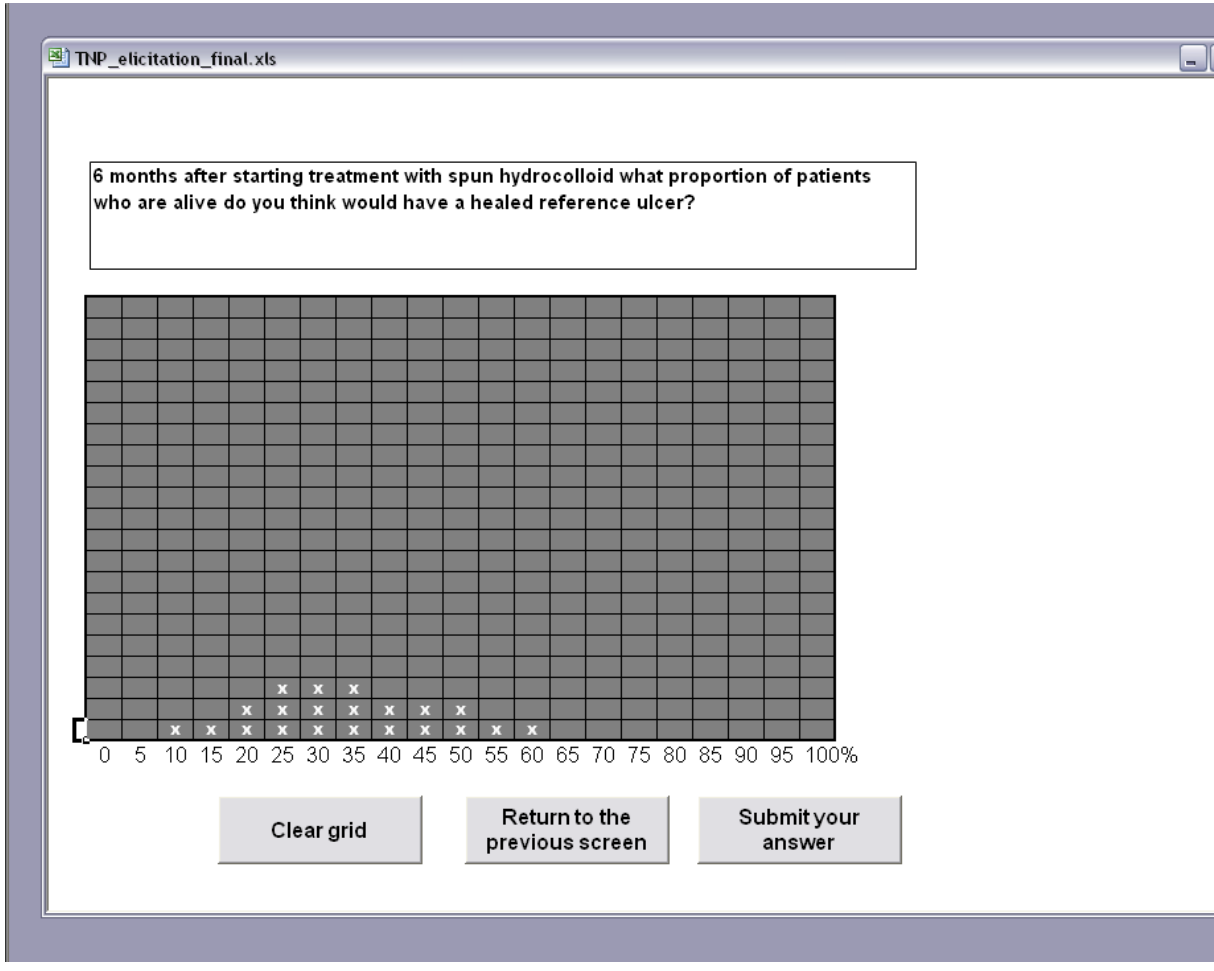
Epistemic uncertainty

- Experts not expected to know for sure the exact answer
  - If unsure the expert should still answer the question
  - Express how uncertain (or certain) the expert is

- Uncertainty in elicitation for HTA
  - Subjective (personal) probability
    - degree of belief in an uncertain proposition
    - reflect epistemic uncertainties (imperfect knowledge)
    - Do not reflect aleatory uncertainty (variability)
  - Good elicitation should eradicate bias, heuristics, irrationality…
  - Inevitably, probabilities elicited are personal and inaccurate

Histogram method

- Histogram or grid method
  - 21 crosses need to be placed in a grid
  - Expert expresses degree of certainty

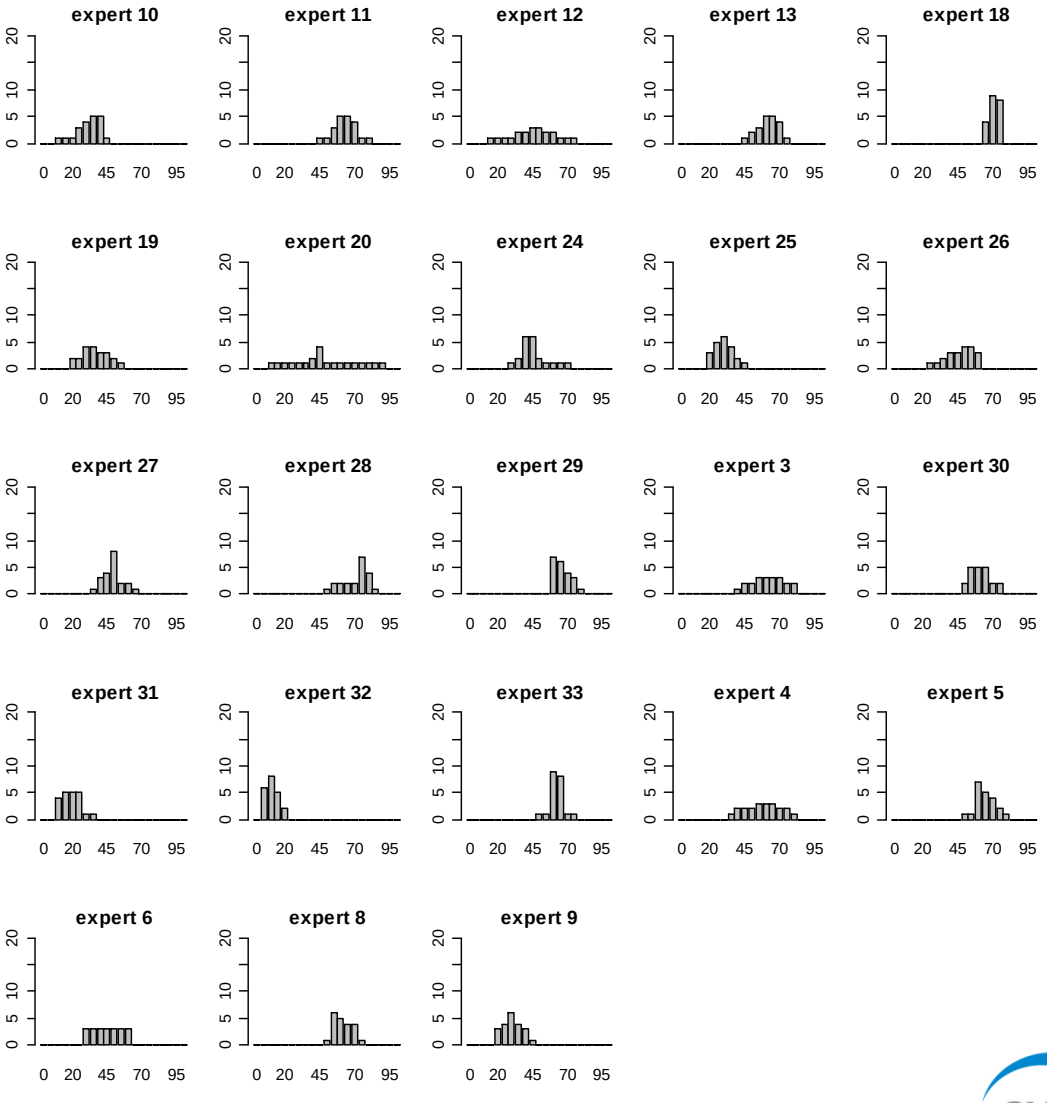


THE UNIVERSITY of York

Histogram method

- Example reply

What proportion of patients do you think would have a grade 3 reference ulcer (rather than a grade 4 reference ulcer)?



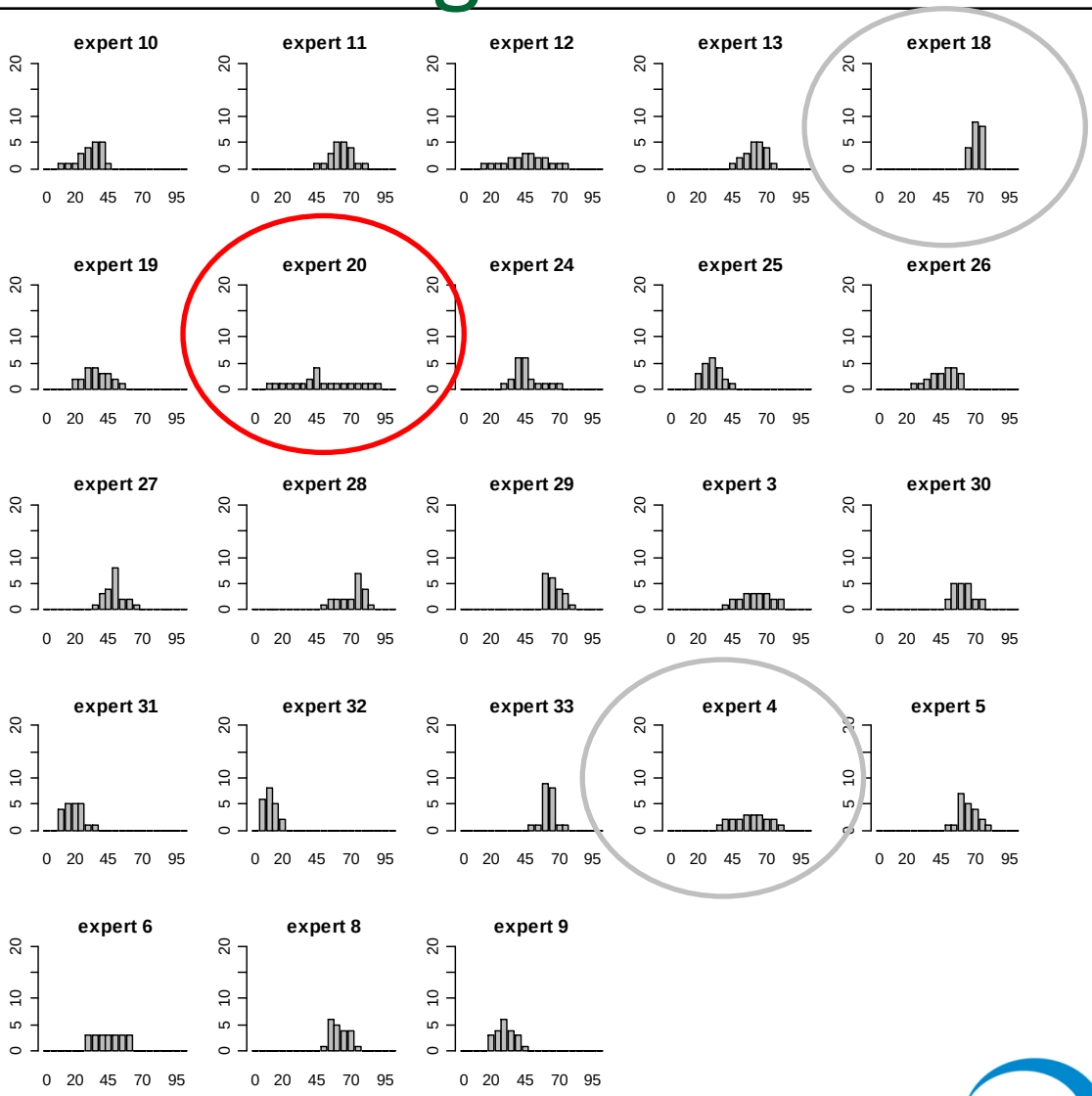
CHE
Centre For Health Economics

THE UNIVERSITY of York

Histogram method

- Example reply

What proportion of patients do you think would have a grade 3 reference ulcer (rather than a grade 4 reference ulcer)?



- Quite uncertain – most likely value is 45%

Quite certain – most likely value is 70% (range 65 to 75%)

In-between – most likely value is 60% (range 30 to 85%)

CHE
Centre For Health Economics

THE UNIVERSITY of York

Histogram method

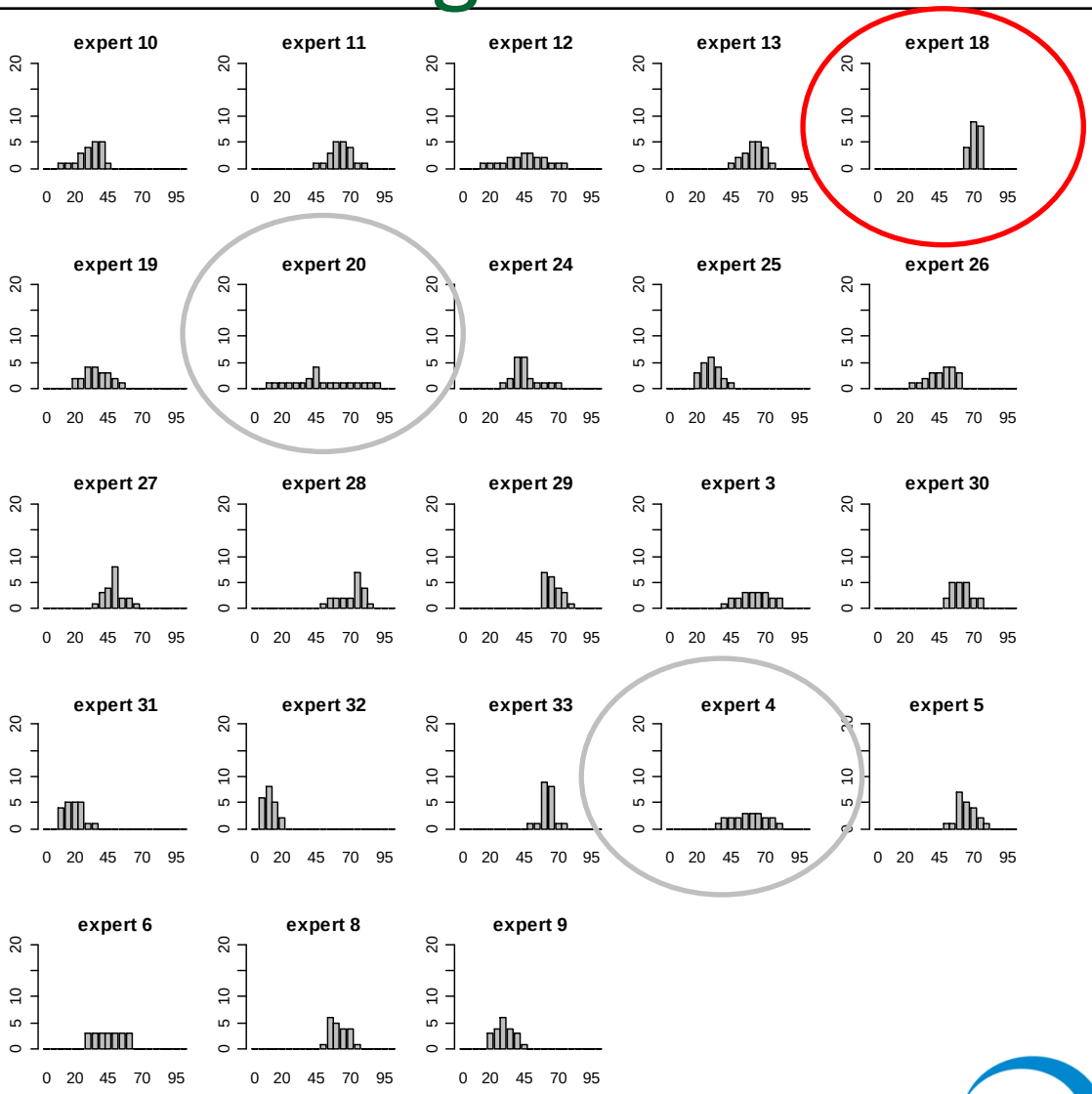
- Example reply

What proportion of patients do you think would have a grade 3 reference ulcer (rather than a grade 4 reference ulcer)?

Quite uncertain – most likely value is 45%

- Quite certain – most likely value is 70% (range 65 to 75%)

In-between – most likely value is 60% (range 30 to 85%)



CHE
Centre For Health Economics

THE UNIVERSITY of York

Histogram method

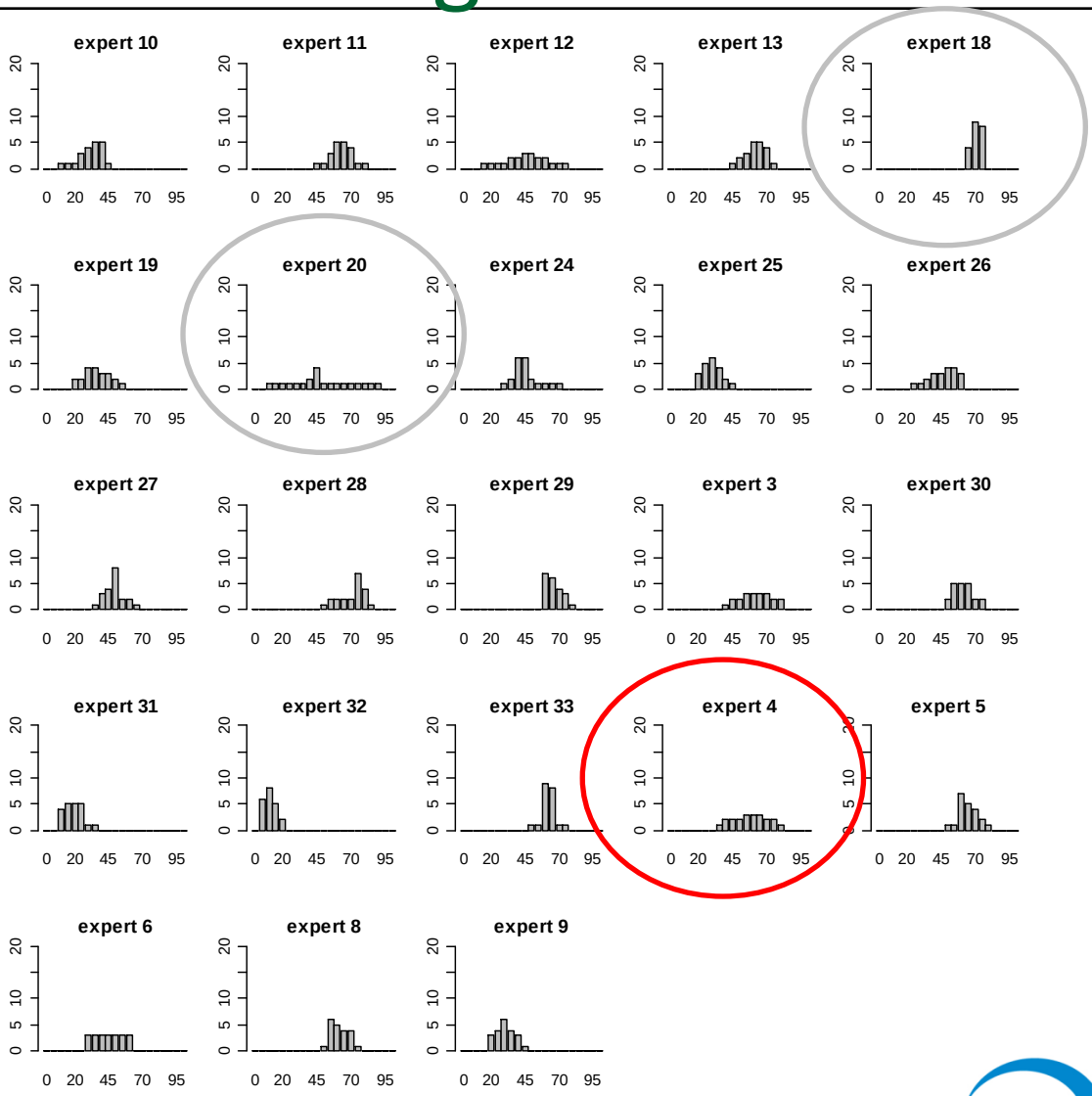
- Example reply

What proportion of patients do you think would have a grade 3 reference ulcer (rather than a grade 4 reference ulcer)?

Quite uncertain – most likely value is 45%

Quite certain – most likely value is 70% (range 65 to 75%)

- In-between – most likely value is 60% (range 30 to 85%)



CHE
Centre For Health Economics

Conduct

Conduct of the exercise

- Face to face, computer based (Excel VBA) exercise
- Experienced facilitator + tutors
- More than 30 questions, 18 uncertain (*not shown here)*
- Multiple experts (N=23), individual elicitation
- Extensive training over
  - The concept of uncertainty
  - Impact of bias
  - The computer instrument
  - The method of eliciting distributions

Eliciting relative effectiveness

Relative effectiveness on healing for foam dressings

Your strongest belief was that <<ref>>% of patients had a healed ulcer 6 months after starting hydrocolloid. Assume that this value is true.

6 months after starting treatment with foam dressings what proportion of patients who are alive do you think would have a healed reference ulcer?

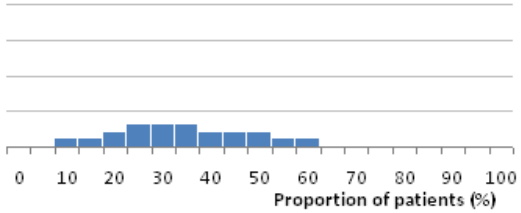
THE UNIVERSITY of York

Eliciting relative effectiveness

Section 2 - Healing (1/1)

Think of UK patients with at least one debrided grade 3 or 4 pressure ulcer (greater than 5 cm^2 in area). Patients start treatment with a **non-silver spun hydrocolloid/hydrofibre dressing** as the primary contact layer. If patients have multiple grade 3 or 4 ulcers, assume that you are treating the deepest ulcer (we will refer to this as the reference ulcer).

On a previous screen you recorded the proportion of living patients you thought had a healed reference ulcer 6 months after starting this spun hydrocolloid treatment. This answer is shown again in the plot on the right.



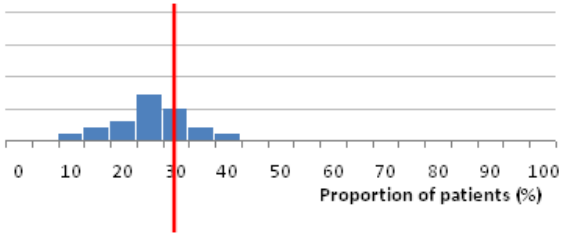
Your strongest belief was that 30% of patients had a healed ulcer 6 months after starting a spun hydrocolloid. Assume that this value is true.

We have marked this value on the next plots and grids with a red bar.

We now want to know what proportion of patients would have their reference ulcer healed if treatment started with foam, alginates or TNP.

6 months after starting treatment with **non-silver FOAM dressings** what proportion of patients who are alive do you think would have a healed reference ulcer?

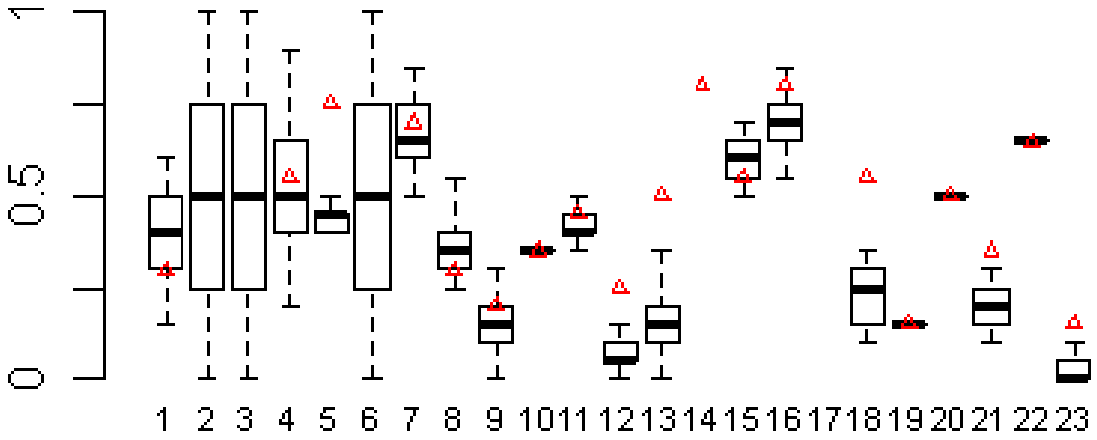
Click here to answer



Continue to exercise menu

Heterogeneity

Proportion of patients healed with F compared to HC



Further transformation to Log HR (d)

Normal, mean=-0.96, CI= [-6.32 to 4.40]

Existing and elicited

Impact of elicited data over effectiveness

Existing data for F was uninformative*

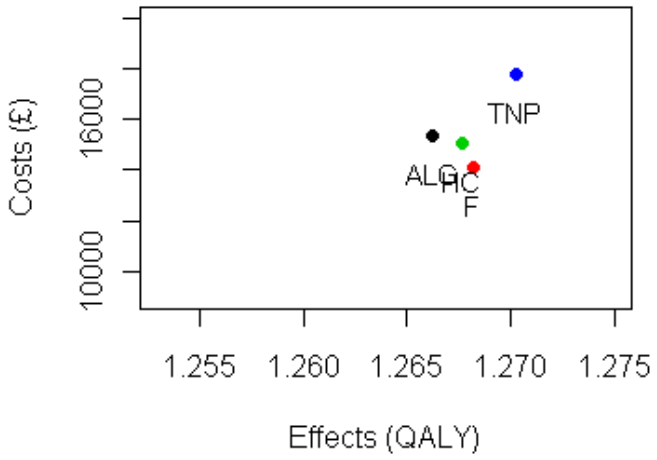
F assumed to be informed by an 'average' effect of all dressing treatments*

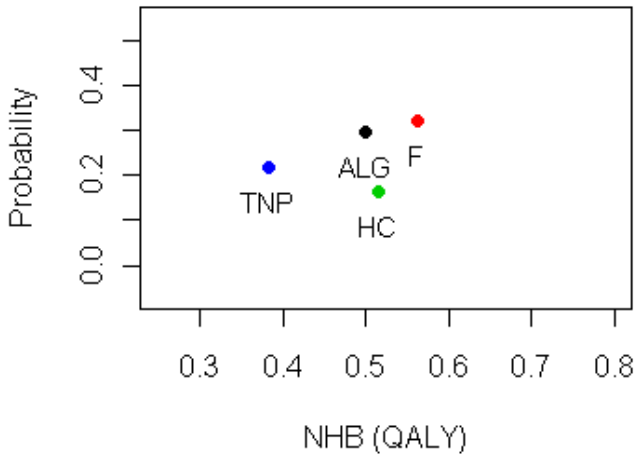
	Existing evidence	Elicited evidence	Existing and elicited evidence collated
d[F]	0.03 [-1.97 to 1.86]	-0.96 [-6.32 to 4.40]	-0.91 [-2.14 to 0.21]
d[ALG]	-0.19 [-1.76 to 1.13]	0.003 [-0.63 to 0.64]	-0.27 [-2.12 to 1.57]
d[TNP]	0.18 [-2.17 to 2.63]	0.45 [-0.66 to 1.56]	0.47 [-1.18 to 2.10]

Impact on decision

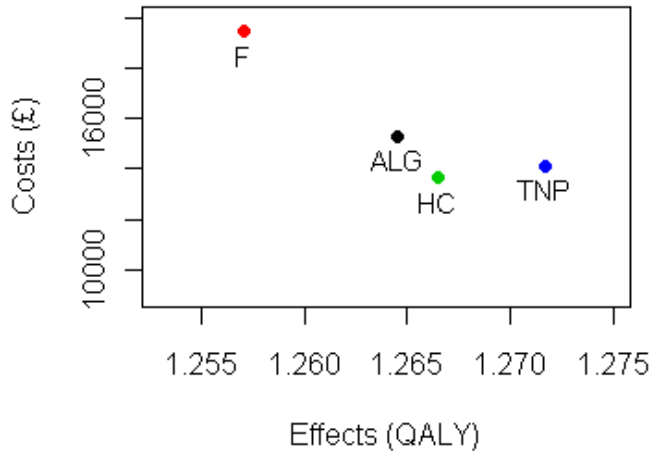
Impact of elicited data over cost effectiveness

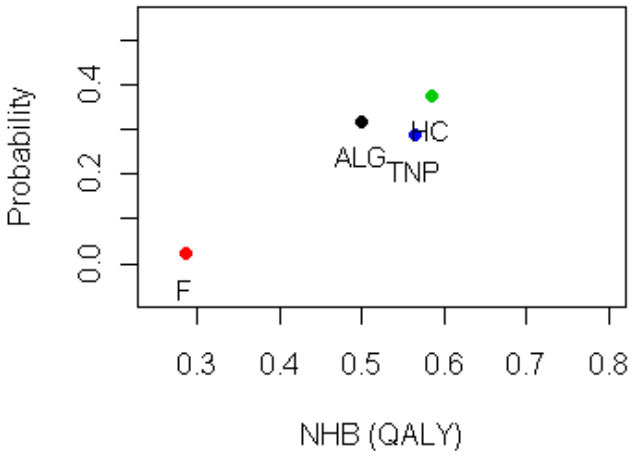
Existing data





Existing + elicited data





Impact on research decision

Table 6 Estimates of the Value of Further Research

	Individual Value: NHB, QALY (NMB, £)	Population Values[a]: NHB, QALY (NMB, £)
EVPI	0.114 (£2273)	4888 (£97.8 million)
EVPI for transition parameters		
Absolute healing parameters and death rates	0 (£0)	0 (£0)
Relative treatment effects	0.101 (£2010)	4327 (£87 million)
F v. HC	0.001 (£13)	29 (£0.6 million)
ALG v. HC	0.041 (£817)	1757 (£35.1 million)
NPWT v. HC	0.056 (£1114)	2395 (£47.9 million)
EVPI for other relevant events and payoffs		
Related events (surgery and complications)	0 (£0)	0 (£0)
Costs, discontinuation, and number of dressing changes	0.018 (£363)	780 (£15.6 million)
Utilities	0 (£0)	0 (£0)

Note: NPWT = negative-pressure wound therapy; HC = spun hydrocolloid; ALG = alginate; F = foam; NHB = net health benefit; NMB = net monetary benefit; EVPI = expected value of perfect information; QALY = quality-adjusted life year.
a. Benefits from research are assumed to sustain for 10 years.

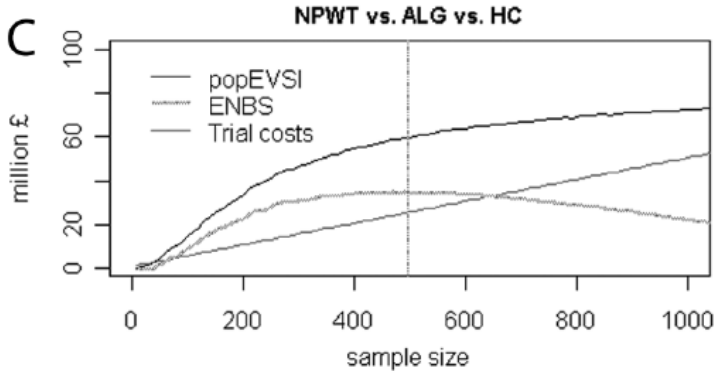
Impact on research decision

- For NPWT, optimal design was a 3 arm trial with longer follow-up with approx 400 patients

Table 8 Optimal Sample Size and ENBS for Alternative Designs of Further Research

	NPWT v. Spun Hydrocolloid		NPWT v. Alginate		NPWT v. Spun Hydrocolloid v. Alginate	
Follow-up Time	Maximum ENBS	Optimal Sample Size, N*	Maximum ENBS	Optimal Sample Size, N*	Maximum ENBS	Optimal Sample Size, N*
0.5 years	—	—	£12.3 million	272	£154,028	403
1 year	£14.0 million	476	£27.2 million	306	£34.7 million	497
2 years	£27.1 million	389	£35.2 million	234	£54.6 million	411

Note: The maximum ENBS was calculated from smoothed ENBS functions using a polynomial function of degree 5. Smoothing did not provide a good fit in one scenario (3-arm trial following up patients for 0.5 years), in which case the observed maximum ENBS and correspondent sample size are presented. NPWT = negative-pressure wound therapy; ENBS = expected net benefit of sampling.



Discussion

- Advantages of elicitation in HTA
  - Appropriately represent epistemic uncertainties
  - Transparent, timely and defensible decisions
  - Elicited information is relatively cheap
- Difficulties of elicitation in HTA
  - Representing epistemic uncertainty
  - Complex exercises with multiple quantities
  - Substantive vs. normative experts
  - Heterogeneity
  - Lack of guidance and standardised procedures
- Issues common to other areas
  - What is an expert and how many experts are sufficient?
  - Calibration and differential weighting

THE UNIVERSITY of York

Research Article

Statistics in Medicine

Received 19 October 2010, Accepted 11 April 2011 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.4288

Methods to elicit experts' beliefs over uncertain quantities: application to a cost effectiveness transition model of negative pressure wound therapy for severe pressure ulceration

Marta O. Soares,[a*†] Laura Bojke,[a] Jo Dumville,[b] Cynthia Iglesias,[b] Nicky Cullum[b] and Karl Claxton[a]

We can use decision models to estimate cost effectiveness, quantify uncertainty regarding the adoption decision and provide estimates of the value of further research. In many cases, the existence of only limited data with

Journal of the Royal Statistical Society
SERIES A Statistics in Society

J. R. Statist. Soc. A (2014)

Treatment comparisons for decision making: facing the problems of sparse and few data

Marta O. Soares and Jo C. Dumville

University of York, UK

and A. E. Ades and Nicky J. Welton

University of Bristol,UK

[Received November 2011. Final revision December 2012]

Summary. Advanced evidence synthesis techniques such as indirect or mixed treatment comparisons provide powerful analytic tools to inform decision making. In some cases, however, existing research is limited in quantity and/or existing research data are 'sparse'. We demon-

Methods to Assess Cost-Effectiveness and Value of Further Research When Data Are Sparse: Negative-Pressure Wound Therapy for Severe Pressure Ulcers

Marta O. Soares, MSc, Jo C. Dumville, PhD, Rebecca L. Ashby, PhD, Cynthia P. Iglesias, PhD, Laura Bojke, PhD, Una Adderley, MSc, Elizabeth McGinnis, PhD, Nikki Stubbs, MSc, David J. Torgerson, PhD, Karl Claxton, PhD, Nicky Cullum, PhD

Health care resources are scarce, and decisions have to be made about how to allocate funds. Often, these decisions are based on sparse or imperfect evidence. One such example is negative-pressure wound therapy (NPWT), which is a widely used treatment for severe pressure ulcers; how-

the impact of each on cost-effectiveness was evaluated. An analysis of the value of further information indicated that a randomized controlled trial may be worthwhile in reducing decision uncertainty, where from a set of alternative designs, a 3-arm trial with longer follow-up was esti-

THE UNIVERSITY of York

marta.soares@york.ac.uk

CHE
Centre For Health Economics