Institution: University of Sheffield		
Unit of Assessment: B-10 Mathematical Sciences		
Title of case study: Expert knowledge elicitation in decision-making and risk analysis		
Period when the underpinning research was undertaken: 2001–2020		
Details of staff conducting the underpinning research from the submitting unit:		
Name(s):	**Role(s) (e.g. job title):**	**Period(s) employed by submitting HEI:**
Jeremy Oakley	Professor of Statistics	2002–present
Anthony O'Hagan	Professor of Statistics	1999–2012
Period when the claimed impact occurred: 2013–2020		
Is this case study continued from a case study submitted in 2014? N		

1. Summary of the impact (indicative maximum 100 words)

A protocol (the *Sheffield Elicitation Framework*: *SHELF*) was developed for eliciting expert knowledge about uncertain quantities, and representing uncertainty using probability distributions. Guidance and software ensured elicitation methods were accessible to diverse sectors. In parallel, "assurance" methods were developed for using elicitation to improve decision-making in clinical trial planning.

This research has changed practice in three areas. GlaxoSmithKline and Novartis now use the assurance method combined with SHELF when planning clinical trials, to improve productivity by reducing the risk of costly unsuccessful trials. The European Food Safety Authority now uses SHELF to improve risk assessments. The reinsurers Swiss RE now use SHELF to obtain more robust forecasts of life and health risks.

SHELF was used to help update the National Institute for Health and Care Excellence (NICE) guidelines on the use of surgical instruments in managing risks of CJD transmission: a policy decision affecting £13 million of NHS expenditure annually.

2. Underpinning research (indicative maximum 500 words)

In numerous decision-making and risk analysis problems, values of important quantities of interest are unknown, and suitable data to estimate these values are not available. Expert elicitation enables uncertainty about these quantities to be described using probability distributions, so that decision-makers can better understand the risks of proposed decision options, and consequently make decisions that are more robust to uncertainty.

There are three challenges in expert elicitation. Firstly, experts will typically *not* be experienced in making probability judgements required in most elicitation methods, and may find the process difficult. Secondly, the experts' judgements are, by definition, subjective, and may be biased; the judgements are not always reliable. Thirdly, different experts will have different opinions, and it is not always clear how to reconcile disagreements between experts.

In 2003-2006, O'Hagan led a team of statisticians and psychologists, funded by the NHS's Research Methodology Programme, researching best practice in expert elicitation. The team's research **[R1]** included the most comprehensive survey of elicitation methodology at the time

and was unique in drawing together perspectives from the fields of statistics and psychology. This multidisciplinary approach enabled different elicitation methods to be evaluated in terms of their feasibility and risk of bias, given the findings from the psychology literature.

This research **[R1]** underpinned the development of 'SHELF': the Sheffield Elicitation Framework (also known as 'the Sheffield Method'), designed by O'Hagan and Oakley. SHELF includes a new protocol for eliciting a single distribution from a group of experts, supporting software, training materials, and templates for reporting the results.

In addition to research on SHELF, methodology was developed to assist the planning of phase III randomised clinical trials **[R2, R3, R4]**. The standard practice in the pharmaceutical industry is to use 'power' calculations, where the number of patients to be recruited to a trial is chosen to achieve a desired probability of the trial producing a successful outcome. Critically, the effectiveness of the new drug to be tested is assumed to be its *desired* effectiveness. This can give the impression, sometimes misleading, that a successful trial can be 'guaranteed', as long as sufficiently many patients are recruited.

An alternative to a power calculation is the 'assurance' method: expert uncertainty about the performance of the drug is elicited, allowing a more realistic assessment of the probability of success if taken forward to phase III. Assurance methods for a range of different types of clinical trial data were developed at Sheffield **[R2, R3, R4]**. The key step in implementing the assurance method is the elicitation of the probability distribution for the effectiveness of the drug, and this is where the elicitation methodology of SHELF can be brought to bear. Uptake of assurance in the pharmaceutical industry depends on the availability and ease-of-use of suitable elicitation methods; for impact, it was necessary to research assurance and elicitation methods in parallel.

3. References to the research (indicative maximum of six references)

R1. **O' Hagan, A.,** Buck, C. E., Daneshkhah, A., Eiser, J. E., Garthwaite, P. H., Jenkinson, D. J., **Oakley, J. E**. and Rakow, T. (2006). *Uncertain Judgements: Eliciting Expert Probabilities*. Chichester: Wiley. https://doi.org/10.1002/0470033312 (1,611 citations)

R2. **O'Hagan,** A. & Stevens, J. W. (2001). Bayesian assessment of sample size for clinical trials of cost effectiveness. *Medical Decision Making 21*(3), 219-230. https://doi.org/10.1177/0272989X0102100307(118 citations)

R3. **O'Hagan,** A., Stevens, J. W., & Campbell, M. J. (2005). Assurance in clinical trial design. *Pharmaceutical Statistics, 4*(3), 187–201. https://doi.org/10.1002/pst.175 (203 citations)

R4. Ren, S. & **Oakley, J. E**. (2014). Assurance calculations for planning clinical trials with time-to-event outcomes. *Statistics in Medicine 33*(1), 31-45. https://doi.org/10.1002/sim.5916 (22 citations)

4. Details of the impact (indicative maximum 750 words)

[Text removed for publication]: changing clinical trials planning to improve productivity

[Text removed for publication].

Two pharmaceutical companies (both in the world's top ten by revenue) use assurance as a matter of policy, regularly combined with SHELF. [Text removed for publication]

[Text removed for publication]

National Institute for Health and Care Excellence (NICE): updated guidance on reducing the risk of transmission of Creutzfeldt–Jakob disease (CJD) from surgery

In 2020 NICE issued updated guidance **[S3]** on surgical procedures considered at risk of transmitting CJD. The guidance was based on a decision-modelling exercise **[S4]**, in which elicitation with SHELF was essential to quantify uncertainty in the model predictions. The decision-modelling showed that a precautionary approach of switching to single-use instruments would not be cost-effective, even accounting for the uncertainty quantified by SHELF. Based on the annual number of operations and the costs of single-use surgical instruments, the cost of implementing this precautionary approach in England was estimated to be approximately £13 million every year; such costs would have had significant implications for patients elsewhere in the healthcare system **[S4]**.

European Food Safety Authority (EFSA): changing the use of expert knowledge in food safety risk assessments

In 2012, SHELF was one of three methods recommended by EFSA's new guidance on best practice in expert elicitation. In 2015, 50 EFSA staff and 17 EFSA experts were trained in the use of SHELF **[S5]**. EFSA have since used SHELF for elicitation in risk assessments in areas to provide quantitative data where qualitative was only possible before, [Text removed for publication] **[S6]**.

[Text removed for publication].

Swiss Re: change to the practice of forecasting life and health risks

Swiss Re, the world's second largest reinsurers, have changed their process for forecasting life and health risks, incorporating the use of SHELF in 2018. [Text removed for publication].

SHELF was used to elicit distributions for the size of the market for critical illness cover in the UK in five years, and changes in long-term mortality improvements (MI). [Text removed for publication].

5. Sources to corroborate the impact (indicative maximum of 10 references)

S1. [Text removed for publication].

S2. [Text removed for publication].

S3. Reducing the risk of transmission of Creutzfeldt–Jakob disease (CJD) from surgical instruments used for interventional procedures on high-risk tissues Interventional

procedures guidance-IPG666. National Institute of Health and Care Excellence. Published: 22 January 2020. https://www.nice.org.uk/guidance/ipg666

S4. Stevenson, M., Uttley, L., Oakley, J. E., Carroll, C., Chick, S. E., & Wong, R. (2020). Interventions to reduce the risk of surgically transmitted Creutzfeldt–Jakob disease: a cost-effective modelling review. *Health Technology Assessment, 24*(11), 1–150. https://doi.org/10.3310/hta24110.

S5. Hart, A., O'Hagan, A., Quigley, J., & Bolger, F. (2016). Training Course on Steering an Expert Knowledge Elicitation. EFSA Supporting Publications, 13(5), 1009E. https://doi.org/10.2903/sp.efsa.2016.en-1009

S6. [Text removed for publication].

S7. [Text removed for publication].