Developing a reference protocol for structured expert elicitation in health-care decision-making: a mixed-methods study

Laura Bojke,[1*] Marta Soares,[1] Karl Claxton,[1] Abigail Colson,[2] Aimée Fox,[1] Christopher Jackson,[3] Dina Jankovic,[1] Alec Morton,[2] Linda Sharples[4] and Andrea Taylor[5]

[1]Centre for Health Economics, University of York, York, UK
[2]Department of Management Science, University of Strathclyde, Glasgow, UK
[3]MRC Biostatistics Unit, University of Cambridge, Cambridge, UK
[4]London School of Hygiene & Tropical Medicine, London, UK
[5]Leeds University Business School, Leeds, UK

*Corresponding author Laura.bojke@york.ac.uk

Declared competing interests of authors: Laura Bojke declares private consulting for BresMed (Sheffield, UK) and Roboleo & Co (Brackley, UK), not related to the current work. Alec Morton declares private consulting for AstraZeneca (Cambridge, UK) and Office of Health Economics (London, UK), not related to the current work. Alec Morton and Abigail Colson were previously supported by that European Union 7th Framework Programme that received in-kind support from pharmaceutical companies, including Astellas Pharma Inc. (Tokyo, Japan), Roche (Basel, Switzerland) and AstraZeneca.

Published June 2021
DOI: 10.3310/hta25370

Scientific summary

Structured expert elicitation in health-care decision-making

Health Technology Assessment 2021; Vol. 25: No. 37
DOI: 10.3310/hta25370

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

At the forefront of decisions in health care is the aim of maximising health, requiring judgements about interventions that may have higher health effects but potentially incur additional costs. The evidence used to establish cost-effectiveness is typically uncertain; for example, the evidence may not be on 'final' outcomes (e.g. cancer products licensed on evidence of progression-free survival), or the evidence base may not be well developed (e.g. in diagnostics, medical devices, early access to medicines scheme). It is important that the uncertainty in this evidence is characterised. If not, any analysis using this evidence may give decision-makers a misleading view of the risks associated with their decision.

In situations in which evidence is subject to uncertainty, the experience of experts may be essential. To ensure accountability in the decision, these expert judgements should be made explicit and incorporated transparently into the decision-making process. The process by which the beliefs of experts can be formally collected in a quantitative manner is structured expert elicitation. If conducted in an appropriate manner, structured expert elicitation can characterise uncertainties associated with the cost-effectiveness of competing interventions and assess the value of further evidence. This may be the approach best suited to a transparent decision-making process.

There is an increasing interest in structured expert elicitation, as new technologies are assessed progressively closer to their launch on the market. Structured expert elicitation is also valuable for 'early modelling' of new interventions or unknown diseases for which little or no evidence is available. A review of applied studies in health-care decision-making found heterogeneity in the methodology used and a lack of consideration for any existing guidance on the topic (Soares MO, Sharples L, Morton A, Claxton K, Bojke L. Experiences of structured elicitation for model-based cost-effectiveness analyses. *Value Health* 2018;**21**:715–23).

No standard guidelines exist to conduct expert elicitation in health technology assessments, but there are a number of generic guidance documents, some of which have been used in health technology assessment. The most notable of these are the Sheffield Elicitation Framework and Cooke's classical method. It is not clear if any of the existing guidelines, generic and domain specific, are appropriate for us in health-care decision-making.

Objectives

The overall aim of this report was to establish a reference protocol or guideline for the elicitation of experts judgements to inform health-care decision-making. To achieve this overall aim, the report focused on the following objectives:

1. Providing clarity on the methods for collecting and using experts judgements within an assessment of cost-effectiveness.
2. Exploring where alternative methodology may be required in particular context/constraints (e.g. time).
3. Establishing preferred approaches for elicitation for a range of parameters and a range of decision-making contexts.
4. Determining which elicitation methods allow experts to express parameter uncertainty, as opposed to variability.
5. Determining the applicability and usefulness of the reference protocol developed within a case study application.

For objective 4, statistical experiments were conducted. The aim of these experiments was threefold, to (1) evaluate alternative methods of elicitation and how they perform in representing parameter uncertainty; (2) explore individuals' ability to extrapolate from their knowledge base; and (3) explore how individuals revise their answers when presented with group summaries.

Methods

To achieve these objectives a mixed-methods approach was used, combining formal systematic review, targeted searches, experimental work and narrative synthesis. Specifically, first a systematic review of existing guidelines for formal elicitation, published in either the peer-reviewed or the grey literature, was conducted. This identified the approaches used in existing guidelines (the 'choices') and determined if dominant approaches evolve. Less formal targeted searches were also conducted to determine the state of the evidence on choices relating to the selection of experts, the level of elicitation, fitting and aggregation, assessing the expected accuracy of experts judgements, and heuristics and biases. The advantages and disadvantages of each available choice for these elements were extracted from the papers and potential constraints to their application in health-care decision-making determined.

Health-care decision-making is not a homogeneous domain, as different decision-makers face different constraints and this may have implications for expert elicitation methodology. The contexts in which structured expert elicitation in health-care decision-making may be conducted are therefore discussed in detail, as well as conclusions made regarding the use of a reference protocol for structured expert elicitation. Alongside this, a systematic review of structured expert elicitation applications in cost-effectiveness modelling was undertaken. This details the challenges that were reported by the authors conducting these analyses. When available, the basis for the methodological choices made in each application is extracted. This also provided a view of the current scope of the landscape with regards to applied structured expert elicitation in health-care decision-making.

When designing a structured expert elicitation, deciding what quantities to elicit is a major challenge. There is no guidance covering the spectrum of quantities that may be appropriate to elicit to inform health-care decision-making, including measures of treatment effects and baseline event rates. To address this lack of guidance, a review was undertaken of alternative quantities that can be elicited to inform the probability- or time-to-event-related parameters commonly used in health-care decision-making.

The statistical experiments, conducted to explore multiple uncertainties in structured expert elicitation methodology, utilised a simulated learning process (e.g. Wang H, Dash D, Druzdzel MJ. A method for evaluating elicitation schemes for probabilistic models. *IEEE Trans Syst Man Cybern B Cybern* 2002;**32**:38–43). Individuals' knowledge was determined by recorded observations. The 'data set' observed then determines participants' belief about the quantity of interest, from which accuracy can be measured. This approach allows the conditions of the experiment to be defined (e.g. equal vs. different knowledge base) and the isolation potential determinants (e.g. precision). Participants were shown random observations from a statistical model that represented an abstract medical problem. Following this, participants were asked to express their beliefs regarding treatment effectiveness. All participants ($n = 72$) were students at the University of York, the large majority of whom were undergoing clinical training. The exercises was delivered face to face and financial incentives were offered according to accuracy. The experiments measured:

- bias – difference in the means of the true and elicited (and fitted) distributions
- uncertainty – ratio of the standard deviations of the two distributions
- Kullback–Leibler divergence – information lost when one distribution is approximated by another (Soares MO, Sharples L, Morton A, Claxton K, Bojke L. Experiences of structured elicitation for model-based cost-effectiveness analyses. *Value Health* 2018;**21**:715–23)
- participants' preference for alternative methods.

Given the full range of evidence generated on which to base a reference protocol for structured expert elicitation in health-care decision-making, it was necessary to use this evidence to generate a set of principles that underpin the use of expert elicitation in health-care decision-making. Available choices, from the review of guidelines, are considered in the light of these principles and any empirical evidence available to support the choices. This informs the reference protocol by discounting or supporting particular choices.

The work also included an applied evaluation of the developed reference protocol. This uses an existing cost-effectiveness model, in which structured expert elicitation was used to generate initial estimates of uncertain parameters. In addition to demonstrating the usefulness of the reference protocol in navigating the structured expert elicitation process, the practicality of structured expert elicitation is determined using narrative feedback form experts and by generating estimates of resources required to design and conduct the structured expert elicitation.

Finally, a dissemination workshop was convened, which explored the usefulness and challenges in using structured expert elicitation in health-care decision-making. It was also used to refine, using discussion, a set of recommendations for further research.

Results

A comprehensive list of elements and choices for structured expert elicitation was developed by reviewing existing protocols (work package 1). This covered the design, implementation and analysis stages of structured expert elicitation. The review showed that for many elements of the structured expert elicitation, there was a lack of consistency across the existing guidelines. Targeted searches also revealed that the majority of choices are not supported by any empirical evidence, both specific to health-care decision-making and more generally.

Empirical evidence generated by the experiments conducted here (work packages 2 and 3) determined that there is little difference between variable interval methods and fixed interval methods to encode judgements, in terms of procedural performance. Therefore, a decision-maker can consider either of these choices suitable. This experiment also determined that participants did not adjust uncertainty levels sufficiently to reflect differences in the underlying heterogeneity of the populations; in particular, uncertainty was consistently underestimated in the case of high heterogeneity. This case is frequently encountered in health-care settings. The experiments also sought to explore extrapolation beyond data observed and updating of priors after presentation of group summaries, issues which feed into multiple choices for structured expert elicitation. It was difficult to form definitive conclusions, given that the experiments were underpowered for these elements. The experiments did provide some evidence that experts changed their estimates in a rational way when provided with estimates from others, suggesting that group discussion or feedback may be useful. Extrapolation outside the observed sample does not seem to affect accuracy, suggesting that it is reasonable to ask experts about patients and practices of which they do not have direct clinical experience, or for whom there is no relevant literature.

In order to sift through the available choices, a set of principles that underpin the use of structured expert elicitation in health-care decision-making was defined using evidence generated from targeted searches, experimental evidence on methods to encode judgements and consideration of the constraints on the decision-making processes in health (work package 1). These nine principles are:

1. transparency
2. fitness for purpose
3. consistency, but respecting constraints of the decision-making context
4. reflecting uncertainty at the individual expert level
5. recognising and acting on biases

6. suitability for substantive experts, who are less likely to be normative
7. recognising where adaptive skills are required
8. recognising between-expert variation
9. promoting high performance.

Not all principles for structured expert elicitation in health-care decision-making were relevant for all elements. The most relevant principles for each element and components within structured expert elicitation were considered.

In almost all choices there is a lack of empirical evidence, and in some circumstances the principles are unable to provide sufficient justification for discounting particular choices (work package 1). It is, however, possible to define reference methods that could be used in a more narrowly defined area of health-care decision-making, namely health technology assessment. These include:

- Focus on gathering substantive expertise or experience. Normative skills can be developed during the training session as part of the structured expert elicitation.
- Simple observable quantities should be elicited when possible. Ratios or complex parameters, such as regression coefficients, should not be elicited directly.
- Minimise and record conflicts of interest among the experts. Include experts external to the structured expert elicitation task (i.e. not those involved in developing the task).
- Dependence between variables should be captured in structured expert elicitation. Expressing dependent variables in terms of independent variables is preferable when experts do not have strong normative skills.
- Use of either variable interval methods or fixed interval methods work well; however, decision-makers should aim for consistency across applications.
- Beliefs should be elicited from experts individually, even if a group interaction follows.
- Between-expert variation should be explored explicitly.
- Following fitting, a summary of the individual distributions should be obtained using linear pooling.
- Interaction should be face to face when possible, to allow a facilitator to deliver training to the expert.
- Training is crucial and should focus on avoiding bias and expressing uncertainty.
- All methodological choices for the structured expert elicitation must be documented and justified.

Additional considerations are required for decision-makers outside health technology assessment, for example at a local level, or for early technologies that have yet to progress through the regulatory process. Access to experts may be limited and in some circumstances group discussion may be needed to generate a distribution.

The application of the case study, a diagnostic model for asthma, explored practical issues. This highlighted sufficient information needs to be presented to the experts. The level of information presented to the experts and the wording of this information is paramount in ensuring that the quantity of interest is observable to the expert. When deciding on the information to provide to experts, it may be useful to consult existing policies. With regards to time constraints, the applied evaluation was undertaken over a 7-month period and involved three analysts in varying proportions. Overall, this equated to 5 months of full-time equivalent researcher time.

Limitations

The major limitation of the work conducted here lies not in the methods employed but in the evidence available from the wider literature on which to base the set of choices and determine how appropriate these are. Concluding on the suitability of the choices available from the existing guidelines is challenging owing to the lack of empirical evidence to support specific choices. Instead, it was necessary to develop

principles for structured expert elicitation in health-care decision-making, using the sources of evidence as described above and published guidelines for good structured expert elicitation. Using only the principles, in the absence of empirical evidence, meant that it was not always possible to give definitive conclusions on choices.

Areas for further research

In considering the appropriateness of choices for structured expert elicitation in health-care decision-making and exploring how these choices may be affected by the context in which the structured expert elicitation is applied, there are several areas in which further research is required before definitive statements can be made regarding their appropriateness for a reference protocol. Researchable questions in these areas include the following:

- Which methods for expert recruitment are most practical and what are the challenges?
- What training strategies can be used to minimise bias?
- Which methods for eliciting dependent quantities work best for non-normative experts?
- Which consensus approach works best in health-care decision-making in practice and for which types of quantities and decision-makers?
- Should individual priors be combined when there is significant expert variation? If so, how?

At the dissemination workshop, participants were asked to discuss areas for further research, specifically considering what decision-makers in health-care decision-making may require when determining a reference protocol for structured expert elicitation for use within their setting. Participants were not asked to define which research topics are highest priority for their setting. Selecting experts, minimising bias, adaptation to specific setting in which structured expert elicitation may be applied (e.g. choosing individual or group elicitation), appropriate wording of questions, methods for multivariate elicitation and what information should be presented to the experts to help them formulate their beliefs. Some of these topics would benefit from empirical research and others may be resolved though application of the proposed reference protocol to health-care decision-making, including in settings with a range of constraints.

Conclusions

Structured expert elicitation can offer opportunities in health-care decision-making, particularly reimbursement decisions supported by model-based economic evaluation. Structured expert elicitation allows the uncertainty in the evidence used to populate these models to be characterised, or, when evidence is completely lacking, provides additional information needed to reach a decision.

The work described in this report has attempted to generate evidence which is useful for analysts and decision-makers in health-care decision-making. Structured expert elicitation conducted in this context to date has not used a set of consistent methods and, above all, has not considered the implications of the choices made when designing and conducting a structured expert elicitation. To improve the accountability of health-care decision-making, the procedure used to derive expert judgements should be transparent.

The reference protocol presented here is intended to serve as a guide to good practice and reporting, and is flexible in many choices rather than being prescriptive regarding methods. It can therefore be thought of as a reference guide. This was necessary owing to the lack of empirical data specific to health-care decision-making and more generally to structured expert elicitation. This may be a useful characteristic, as it is possible to apply this reference protocol across different settings.

Funding

This project was funded by the NIHR Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 25, No. 37. See the NIHR Journals Library website for further project information. This work was also funded by the Medical Research Council (reference MR/N028511/1).

Health Technology Assessment

HTA MRP–MRC

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 3.370

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, the Cochrane Library and Clarivate Analytics Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme or, commissioned/managed through the Methodology research programme (MRP), and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

This report

This issue of the Health Technology Assessment journal series contains a project commissioned/managed by the Methodology research programme (MRP). The Medical Research Council (MRC) is working with NIHR to deliver the single joint health strategy and the MRP was launched in 2008 as part of the delivery model. MRC is lead funding partner for MRP and part of this programme is the joint MRC–NIHR funding panel 'The Methodology Research Programme Panel'.

To strengthen the evidence base for health research, the MRP oversees and implements the evolving strategy for high-quality methodological research. In addition to the MRC and NIHR funding partners, the MRP takes into account the needs of other stakeholders including the devolved administrations, industry R&D, and regulatory/advisory agencies and other public bodies. The MRP funds investigator-led and needs-led research proposals from across the UK. In addition to the standard MRC and RCUK terms and conditions, projects commissioned/managed by the MRP are expected to provide a detailed report on the research findings and may publish the findings in the HTA journal, if supported by NIHR funds.

The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded under a MRC–NIHR partnership. The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the MRC, NETSCC, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, the MRC, NETSCC, the HTA programme or the Department of Health and Social Care.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).