Use of Expert Judgement Across NICE Guidance-Making Programmes: A Review of Current Processes and Suitability of Existing Tools to Support the Use of Expert Elicitation

Alison Peel¹ · Michelle Jenks¹ · Moni Choudhury² · Rosemary Lovett² · Juan Carlos Rejon-Parrilla³ · Andrew Sims^{4,5} · Joyce Craig¹

Published online: 3 August 2018
© The Author(s) 2018

Abstract

Objectives This study aimed to review current use of experts within National Institute for Health and Care Excellence (NICE) guidance-making programmes, identify improvements in use of expert judgement, and to assess tools and protocols to support the elicitation of information from experts for use by NICE.

Methods The study comprised a review of NICE process guides; semi-structured interviews with individuals representing each NICE guidance-making programme and a comparison of the suitability of currently available tools and protocols for expert elicitation to the requirements of NICE programmes.

Results Information elicited from experts and the way in which it is used varies across NICE guidance-making programmes. Experts' involvement can be as intensive as being a member of a committee and thus having direct influence on recommendations or limited one-off consultations on specific parameters. We identified 16 tools for expert elicitation that were potentially relevant. None fully met the requirements of NICE, although an existing tool could be potentially adapted. Ongoing research to develop a reference protocol for expert elicitation in healthcare decision making may be of use in future.

Conclusions NICE uses expert judgement across all its guidance-making programmes, but its uses vary considerably. There is no currently available tool for expert elicitation suitable for use by NICE. However, adaptation of an existing tool or ongoing research in the area could address this in the future.

Key Points for Decision Makers

Use of expert judgement is widespread in NICE guidance making, but there is no standard approach across programmes and little use of published tools or protocols for expert elicitation.

Current practice could be improved by using a tool or protocol to facilitate more formal methods of expert elicitation. The tool or protocol would need to meet a number of requirements in order to be adopted by NICE, including fitting into the time and resource constraints of the guidance-making process.

None of the tools and protocols identified in this study were entirely appropriate for use by NICE. However, information on the available tools and protocols could be of use to international HTA agencies and other groups interested in conducting expert elicitation to inform healthcare decision making.

✉ Alison Peel
alison.peel@york.ac.uk

¹ York Health Economics Consortium, Enterprise House, Innovation Way, University of York, York, UK

² National Institute of Health and Care Excellence, London, UK

³ National Institute of Health and Care Excellence, Manchester, UK

⁴ Newcastle upon Tyne Hospitals National Health Service Foundation Trust, Freeman Hospital, Newcastle upon Tyne, UK

⁵ Faculty of Medical Sciences, Institute of Cellular Medicine, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, UK

1 Background

Expert judgement has many uses in the context of health technology assessment (HTA) and economic evaluation, and can be broadly categorised into *expert opinion* and *expert elicitation* (Box 1) [1]. *Expert opinion* involves qualitative expression of an individual's judgement (e.g. asking an expert to define the natural history of a disease). *Expert elicitation* involves collecting quantitative information from experts (e.g. using elicitation to obtain estimates of model inputs, and often their distributions, as well as associated uncertainty). Although generally acknowledged to be a poor quality of evidence [2], experts are routinely relied upon to fill gaps in higher quality evidence during development of evidence-based recommendations.

There is no universal agreement on the most appropriate technique for expert elicitation, although a number of protocols and methodologies are available [3–6]. Several digital tools and software to support expert elicitation exist [7], some of which are described in further detail in this paper.

1.1 Use of Expert Judgement in HTA and Guideline Production

The National Institute for Health and Care Excellence (NICE) uses information from both professional and lay experts in the production of guidance [8]. Professional experts are individuals with knowledge of the experience and views of practitioners, government and policy, or research and practice. Lay experts are those with experience of using health or care services, carers, advocates, or members or officers of a voluntary or community organisation potentially affected by NICE guidance. This article focuses predominantly on professional experts, referred to simply as 'experts' in some instances. Points referring to lay experts are clearly indicated.

There are several programmes within NICE which produce guidance on a variety of technologies, interventions, and services [9]. Within each programme, one or more

committees are responsible for considering the available evidence and making recommendations for guidance.

1.2 Objectives

The aims of this study were:

- To review current use of experts across NICE guidance-making programmes and to identify how the use of expert judgement could be improved during HTA and guidance development;
- To assess tools and protocols that could assist with the elicitation of information from experts for use by NICE and in HTA worldwide. We considered a 'tool' to be an interactive digital resource used to assist a user with expert elicitation, and a 'protocol' as a guide that provided methodology on how to design, conduct or best report expert elicitation.

2 Methods

2.1 Review of Current Use of Expert Judgement in NICE Guidance

We obtained methods and process manuals (March 2017) for each NICE guidance-making programme from the NICE website and extracted information on the use of professional and lay experts and/or expert judgement.

In March and April 2017, we conducted interviews with NICE staff members, NICE committee chairs, and individuals with experience of undertaking contracted work for NICE programmes (specifically in those programmes where external groups undertake expert elicitation). We used semi-structured interview guides ("Appendix 1"). Questions were adapted as necessary to reflect the differing roles of interviewees representing the various programmes, lay experts and those working outside of NICE. The interviews aimed to capture how NICE uses experts during guidance making, including numbers of experts consulted, their identification,

Expert elicitation:

This describes obtaining *quantitative* values from experts for use in decision modelling. For example, experts may be asked to provide an estimate of the proportion of patients who would receive a particular treatment, and define the distribution and uncertainty associated with their estimate.

Expert opinion:

This describes obtaining *qualitative* information from experts. For example, experts may be asked to define the place of a new therapy in the care pathway.

Box 1 Definitions of expert elicitation and expert opinion used in this study [1]

the process of gathering expert judgement, use of elicited information by NICE programmes, and the opportunities/challenges associated with current processes. We conducted interviews with representatives from the following guidance-making programmes:

- NICE guidelines.
- Diagnostics guidance.
- Highly specialised technologies guidance.
- Interventional procedures guidance.
- Medical technologies guidance.
- Technology appraisals guidance.

We also interviewed a representative from the Public Involvement Programme in order to understand the use of lay experts across all guidance-making programmes.

Interviewees were informed that their full responses were confidential, with themes and examples being reported rather than full interview transcripts. This allowed for more open discussion during interviews.

We identified key issues emerging from interviews and the review of manuals to determine a set of requirements for a tool or protocol to assist with expert elicitation at NICE. Although we did not ask interviewees directly for a list of requirements, their responses fed into the list presented in this paper and, thus, we judge this list to be representative of the requirements of NICE programmes using expert elicitation. Furthermore, four authors of this paper have experience conducting informal expert elicitation for the development of NICE guidance and so could provide insight into the issues emerging from interviews and the relevance of the identified requirements.

2.2 Review of Tools and Protocols for Expert Elicitation

We conducted a targeted literature search to identify tools and protocols designed to support expert elicitation. We included resources meeting the following criteria:

- Tool or software designed and used for eliciting data from experts published or updated since 2005;
- Protocol (i.e. a guide) describing the elicitation of data from experts published since 2005.

Only tools and protocols published or updated since 2005 were included, as those developed before the publication of a seminal text in the field [4], were judged less relevant. We did not consider tools and protocols designed for related purposes, such as decision analysis. Further, we excluded case studies reporting expert elicitation in a single or small number of instances (as opposed to guides reporting how expert elicitation *should* be undertaken).

First, we undertook an Ovid MEDLINE search in March 2017 to find systematic reviews of methods for eliciting probability distributions from experts in the context of HTA and economic evaluation (search strategy provided in “Appendix 2”). We screened records to identify eligible systematic reviews and assessed studies included in reviews for relevance. Second, we conducted targeted searching of Google scholar for published work and Google for grey literature and online tools. We consulted websites describing tools for expert elicitation and identified ongoing work by contacting UK-based researchers known to have published in the area and checking funding body webpages. Grey literature searches were not limited to tools used in the context of HTA and economic evaluation.

A single reviewer (MJ) conducted study selection and a second reviewer (JC) confirmed the relevance of included results. We extracted relevant information on tools and protocols and, where required, contacted tool developers to request additional information. We assessed the suitability of tools and protocols first against the criteria emerging from interviews and second through consultations with NICE staff or stakeholders involved in expert elicitation (for example, a model developer from NICE guidelines who is required to elicit model input parameters from expert committee members).

Our review focused on tools and protocols designed specifically for *expert elicitation*. However, we also assessed a generic online survey tool suitable for the collection of *expert opinion* following discussion with NICE staff, who judged that a tool for obtaining qualitative information from experts could be helpful.

3 Results

3.1 Review of Current Use of Expert Judgement in NICE Guidance

From March to June 2017, we conducted 12 semi-structured interviews with NICE staff. Further interviewees included two NICE committee chairs, two health economists contracted to work on NICE guidelines, and a professor who led an independent external group contracted to critically appraise evidence for guidance.

Table 1 summarises the methods we used to review the use of expert judgement and the principal ways in which experts are involved in guidance making in each of the NICE programmes (i.e. if they are full committee members, external advisers or both).

Sections 3.1.1 and 3.1.2 describe the ways in which *expert opinion* and *expert elicitation* are used in guidance making. It is worth noting that although we present these activities separately they are not typically discrete processes. For example, in the questionnaires used by the interventional

Table 1 Summary of use of experts by NICE guidance-making programmes

Guidance-making programme	NICE staff member interviewed?	Advice sought from independent external group member?	Process and methods manual reviewed?	Topic specific experts are committee members?	Topic specific experts provide advice (without being formal committee member)?
NICE guidelines	✓	✓	✓	✓	✓
Diagnostics guidance	✓	✓	✓	✓	✓
Highly specialised technologies guidance	✓	✓	✓	X	✓
Interventional procedures guidance	✓	X	✓	X	✓
Medical technologies guidance	✓	✓	✓	X	✓
Technology appraisals guidance	✓	✓	✓	X	✓
Quality standards	X	X	X	N/A	N/A

✓ yes, X no, N/A unknown as the manual was not reviewed and interviews did not take place due to programme restructuring at the time of this study

procedures programme there are questions around both expert elicitation *and* expert opinion.

3.1.1 Use of Expert Opinion

Qualitative expert opinion plays an important role in developing NICE guidance and is collected at various stages across all NICE programmes. Box 2 summarises the use of expert opinion, based on both the review of manuals and the interviews conducted.

Generally, expert opinion is sought in an unstructured way, during committee meetings (in person), by telephone or email. This process is flexible, and experts are asked anything that committee members and chairs deem appropriate. Many of the programmes also use standardised questionnaires to gather expert opinion, particularly at the scoping stage. In some cases, an interactive approach may be used. For example, staff working on NICE guidelines have

facilitated workshops with professional experts for ‘conceptual mapping’ of clinical pathways in complex disease areas.

The use of opinion from lay experts is more limited. Their input generally relates to their own areas of expertise, notably patient experience, quality of life, and any undocumented side effects or benefits of the technology/service under evaluation.

3.1.2 Use of Expert Elicitation

All NICE programmes use quantitative information elicited from experts in some circumstances. However, this is less common than the use of expert opinion and is generally restricted to cases where there are significant gaps in the evidence base. Where expert elicitation is used, information is usually elicited from professional experts. The input of lay experts is sought where there are questions related to their experience of a technology, service or disease area that the committee must take into account to make an informed

Professional expert opinion is sought for the following:

- Background information on the disease area, current NHS practice, and how this might alter through use of a new technology, service, or intervention.
- Interpreting evidence and its generalisability to the current NHS.
- Defining the decision problem, e.g. identifying relevant comparators and the patient population or subgroups most likely to benefit from the new technology.
- Sense-checking data used in clinical and cost-effectiveness reviews and economic modelling (including the plausibility of model structures, inputs, and results).
- Formal consultation on draft guidance

Box 2 Use of expert opinion at NICE

recommendation (although this typically involves opinion gathering rather than elicitation of quantitative values).

Four of the programmes producing guidance (technology appraisal guidance, highly specialised technologies guidance, medical technologies guidance, and diagnostics guidance) contract independent external groups to review evidence. In these programmes, it is more likely that the external group (or company submitting evidence) carries out expert elicitation, rather than staff from the programme itself. Hence, for all four programmes, we sought advice from these external groups. Interviewees from the technology appraisal programme advised that external evidence review groups used professional experts to estimate quantitative information for parameterising economic models. Based on the experience of the interviewee, these inputs would normally be parameters such as resource use data, rather than the efficacy of a treatment. Typically, experts are asked for a central estimate and some measure of uncertainty around this, for example, minimum and maximum plausible values. In a previous appraisal, experts were asked to plot data on a histogram in order to elicit the distribution of the input value. Information from experts was aggregated by linear pooling and a mean value, distribution and associated uncertainty were calculated [10].

Within the NICE guidelines programme, interviewees described a variety of uses and techniques for expert elicitation. Elicitation ranged from structured approaches with a large group of experts (for example, the full committee) to less formal exercises with smaller numbers of experts. Experts were typically asked for means and ranges of quantitative estimates and means across experts were calculated. One health economist with experience of working on NICE clinical guidelines described using a modified Delphi technique via an online survey tool to elicit information from professional experts. This involved multiple rounds of elicitation, with responses shared after each round, and experts allowed to refine their answers. Other structured approaches to elicitation included using online questionnaires with a subgroup of topic-specific committee members in order to obtain initial estimates of model inputs, which were then presented to the full committee.

Expert elicitation may also be conducted via questionnaires in other programmes. An interviewee from the interventional procedures programme reported asking experts for quantitative information in the questionnaires they complete prior to committee meetings. Some of these questions may include pre-defined ranges for quantitative estimates. For example: “Please estimate the proportion of doctors in your specialty who are doing this procedure (choose one): (a) More than 50% of specialists engaged in this area of work; (b) 10–50% of specialists engaged in this area of work; (c) Fewer than 10% of specialists engaged in this area of work; (d) Cannot give an estimate”. The interviewee reported that

a single round of expert elicitation was conducted, and individual answers were not shared across experts. In general, it appeared to be rare for multiple rounds of elicitation to be considered necessary for the production of guidance.

Interviewees indicated that in many cases elicitation is informal and may be conducted over the telephone or by email with individual experts, or with groups or individuals during committee meetings. For example, during a meeting, the committee may ask experts about the mean duration of a procedure where this has not been reported in the evidence base. In these instances, it is unlikely that the uncertainty around experts’ estimates would be quantified as it may be judged unnecessary or impractical to do so.

3.1.3 Potential for Improvements in the Use of Expert Opinion and Expert Elicitation

The most frequently reported challenge for guidance programmes was the limited time available for obtaining and using expert judgement (i.e. expert opinion or expert elicitation). Recruiting and retaining experts can be challenging and time-consuming and interviewees suggested that the use of online surveys could help to improve engagement, particularly in the case of lay experts who may be recruited through their clinicians and asked to return a paper questionnaire. In other cases, use of survey software for elicitation or gathering opinions may save time because experts would not need to attend face-to-face meetings. Use of software that automatically collated experts’ responses would reduce administration time. However, it was noted that NICE staff and committee members often prefer to speak with experts in person in order to obtain the most appropriate information possible for the guidance (via expert opinion).

Interviewees with experience of conducting *expert elicitation* thought that this process could be improved if a protocol was available that provided advice on the best approach to elicitation and how to avoid bias. A tool or protocol would avoid the need to begin afresh when designing each elicitation. If a tool or protocol led to time savings, expert elicitation could be used more frequently and robustly across all programmes. Another suggestion was to provide training to experts prior to elicitation, which would allow them to provide information better tailored to the evidence gaps. However, this might have recruitment and resource implications.

A tool or protocol to support expert elicitation would be of particular benefit for internal and external modellers working on NICE guidelines, for external academic groups contracted by the diagnostics, medical technologies, technology appraisals, and highly specialised technologies programmes, and for companies submitting to NICE.

3.2 NICE Requirements from a Tool or Protocol for Expert Elicitation

Our findings suggest that any single existing protocol or tool is unlikely to meet all of the needs of the different NICE programmes. However, resources that do not fulfil all requirements of all programmes could still be of use in guidance making.

A tool to support the use of expert elicitation in NICE guidance making would ideally be flexible, affordable, user-friendly, and be used within current timeframes for guidance making, as well as meeting a number of technical requirements. Specific requirements are described in more detail in Table 2, alongside an analysis of the extent to which the two most suitable tools met these criteria.

The majority of requirements that emerged from our review of manuals and interview data referred to *tools* for expert elicitation (or possibly a tool-protocol combination, i.e. a tool where a protocol was available to use alongside it). However, we also identified some key features of expert elicitation for which an ideal protocol would provide information on best practice: selection of experts (including numbers of experts to include); avoidance of bias (including during recruitment); weighting of expert judgements according to the quality of responses; reporting of expert elicitation exercises; and use of elicitation outputs in decision making.

3.3 Review of Protocols and Online Tools for Expert Elicitation

Targeted, pragmatic searches identified 1017 records, of which 16 tools [3, 11–25] and four protocols [1, 26–28] for expert elicitation were deemed suitable for inclusion.

Further details of the tools identified and an analysis of their potential for use by NICE are presented in “Appendix 3”. Although no tool met all of the requirements identified in our review, five were considered further. In the case of these tools, it was judged either that they were likely to be of use for expert elicitation by NICE, or that insufficient information was available in the public domain (i.e. in published studies or on the tools’ websites) to be able to exclude them. The tools’ developers were contacted to obtain further information and the suitability of the tools was analysed in light of this information.

Two tools, Excalibur [11] and PEGS (prior elicitation of graphical software) [24], were subsequently excluded because they required elicitation to be conducted during face-to-face interviews with experts and there were concerns around how user-friendly the software was, particularly because associated training was not available for either. The remaining three tools were judged to be potentially more useful to NICE. These were MATCH [17], SHELF [3] (these

two tools use much of the same functionality, as described in Sect. 3.3.1), and ExpertLens [21].

3.3.1 MATCH and SHELF

MATCH [17] is an online interface for the SHELF package, which is considered to represent good practice in expert elicitation [29]. The SHELF package includes a series of documents that provide extensive information on designing and conducting expert elicitation, which could be considered a protocol [3]. A potential advantage of using MATCH is that it provides a user-friendly online platform and, unlike SHELF, does not require use of R statistical software. For these reasons, we assessed MATCH in favour of SHELF for use by the NICE guidance-making programmes. All of the information provided with SHELF is applicable to MATCH and is freely available online.

MATCH fulfils some of NICE’s requirements from a tool. However, it does not allow for pooling of estimates and must be used with one expert at a time, or by obtaining consensus opinions from multiple experts, face-to-face. In addition, facilitators would require training in the use of the SHELF methodology to use MATCH. Training is available but could be costly should a large number of staff require training.

3.3.2 ExpertLens

ExpertLens [21] is an online tool for eliciting information from a large number of experts. It uses a modified Delphi technique, and elicitation occurs online in a series of rounds. The tool fulfils some of NICE’s requirements. However, it does not capture the uncertainty associated with elicited values and encourages experts to agree estimates by consensus. Therefore, the tool would not be appropriate as it currently stands. At present, the tool is best suited to the collection of expert opinion. However, with the addition of a bolt-on module to enable elicitation of input values and associated uncertainty, this tool could meet the technical requirements for use in NICE guidance making.

Table 2 summarises the extent to which MATCH and ExpertLens met NICE’s requirements (considering that MATCH employs SHELF functionality but with a more user-friendly interface, and so the latter was not considered further).

3.3.3 Protocols for Expert Elicitation

Of the four included protocols, none met all the requirements of the NICE programmes. However, as the requirements that emerged from our review referred largely to characteristics of a tool or a combined tool and protocol (see Sect. 3.2), it was anticipated that this would be the case. Nevertheless, each protocol presented some information that could be of

Table 2 NICE requirements from a tool and the extent to which two tools met these criteria

NICE Requirement	MATCH	ExpertLens
Pre-elicitation material		
Provide training materials for experts to go alongside tool (e.g. on probabilities and uncertainty)	● Materials are not provided, although some materials for use with SHELF methodology could be used.	● Materials are not provided.
Administering the Elicitation Tool or Software		
Allow for flexibility in the manner in which elicitation is conducted (e.g. via the internet, in person or via telephone).	● Designed to be conducted in person.	● Designed to be conducted via the internet but could be used in person or via telephone.
Allow multiple parameters to be elicited during an elicitation session (both in permitting an analysis of covariance and to capture information about multiple independent parameters).	● Tool allows for this.	● Tool allows for this.
Be user friendly, allowing non-experts (either in statistics or the programme behind the tool/software) to use the programme and generate results.	● Training required.	● Involvement from tool's developers required.
Be affordable and allow elicitation to be completed within given timeframes.	● The tool is free but expenses would be incurred for the necessary training and there would be issues with conducting elicitation within given timeframes.	● There is potential for the tool to be costly and there would be issues with conducting elicitation within given timeframes.
Analysis of Elicited Data		
Capture uncertainty in estimates elicited from experts and variability between elicited values, allowing distributions and ranges to be elicited to thus inform probabilistic sensitivity analysis.	● Captures uncertainty in estimates, but values from multiple experts cannot be combined.	● Does not capture uncertainty and values from multiple experts cannot be combined.
Allow for external evidence (e.g. from published literature) to be synthesised alongside expert elicitation.	● The tool does not allow for this unless captured within a consensus.	● The tool does not allow for this unless captured within a consensus.
Provide summary outputs (statistics or diagrams) for use in reporting to Committees.	● These are provided, but only based on outputs from one expert or a consensus output from a group.	● These are provided.
KEY: ● Does not meet requirements; ● Partially meets requirements; ● Meets requirements in full.		

use in expert elicitation to inform guidance making. We extracted this information and present this below.

With regard to the selection of experts, Durbach et al. [26] state that the number of experts will vary depending on factors including time and resource constraints. Hoffmann et al. [28] recommend use of snowball sampling and sample size calculations to select and determine numbers of experts; however, this may not be feasible within current NICE timeframes. These protocols support that efforts should be made to prevent bias or overconfidence in experts' judgements and experts should be asked to capture uncertainty associated with their judgements by providing highest and lowest plausible values in addition to a central estimate. The US Environmental Protection Agency [27] suggest various methods to reduce bias, including providing background briefing information, training experts, piloting, and documenting all stages of elicitation. In addition, they describe various methods for weighting and combining expert responses. Durbach et al. [26] recommend using linear pooling to combine expert judgements and assigning weightings to each expert, although no information on how weightings should be calculated is given. Iglesias et al. [1] present detailed reporting guidelines for expert elicitation exercises and recommend that full results should be presented and interpreted, ensuring that uncertainty is described, in order to inform decision making.

3.3.4 Ongoing Research

A project funded by the Medical Research Council is currently being conducted at the University of York and aims to develop a reference protocol for elicitation of quantitative values from experts [30]. The protocol will be designed for use in healthcare decision making and it is anticipated that it will be suitable for use by model developers involved in the production of NICE guidance. It is likely to be available in early 2019. The project does not aim to develop a tool for use alongside the protocol, although it is possible that this will occur as follow-up work.

3.4 Online Survey Tools for Collating Expert Opinion

A standard tool for gathering *expert opinion* might also be of use to the NICE guidance-making programmes. At present, many of the programmes ask experts to provide opinions using paper questionnaires and synthesising this information can be time-consuming. Therefore, we assessed a generic online survey tool (Snap Surveys) for suitability and consulted staff members across guidance-making programmes as to whether use of the tool would be beneficial.

Overall, responses suggest that an online survey tool could represent an efficient method for obtaining information

from experts. In some programmes, use of the tool could allow for a greater number of experts to be engaged at lower cost, or for experts to be consulted at more time points during the guidance-making process. A representative from the interventional procedures programme advised that they are already piloting use of electronic surveys to gather information from experts. However, online surveys could not completely replace expert opinions provided at committee meetings because it is not known in advance what questions will be asked.

An online survey tool could also have the potential to be used for expert elicitation if clear guidelines were given around structuring questions, recruitment, informing experts of the purpose of the exercise, analysis of responses and presentation of results. Alternatively, if a bolt-on module was developed for ExpertLens as described above, this tool could also be used for both collecting expert opinion and expert elicitation.

4 Discussion

All NICE guidance-making programmes use expert judgement to inform decision making and, especially in cases where there is a paucity of evidence, this judgement can be highly influential in the development of guidance. Experts' roles vary depending on the NICE programme of interest, the aims of that programme, and the way in which expert judgement is incorporated. The role of professional experts is two-fold. First, they provide expert opinion, including background information and sense checking clinical and economic evidence. Second, quantitative information is elicited to fill gaps in the evidence base. Roles for lay experts are more limited and do not involve elicitation of quantitative information.

Currently, there is little use of published tools or protocols for expert elicitation across NICE, with no standard approach specified in the methods and process manuals. Further, existing timelines in certain programmes are often a constraint to conducting formal elicitation. Interviewees judged that current practice could be improved by using more formal methods for expert elicitation guided by a tool or protocol. This would improve the reliability of elicited data, the sample sizes used in elicitations, and the information generated about uncertainty. However, time and resource constraints mean that a tool or protocol for use by NICE would need to adopt a streamlined process of expert elicitation. In a paper published since our review, the authors compared remote elicitation with face-to-face elicitation and reported that issues with recruitment, expert attrition and the time needed to obtain data remained prevalent when remote elicitation was used [31]. A tool or protocol may also require adaptation for the use of lay experts, although a

survey tool for completing questionnaires online and synthesising answers could significantly cut-down administration time and may improve engagement.

None of the tools and protocols for expert elicitation identified in the pragmatic review were entirely appropriate for use in NICE guidance making. Two tools for expert elicitation, ExpertLens and MATCH, do not currently meet NICE's requirements but could potentially be adapted for use in guidance making through work with the tools' developers. In particular, development of a module for ExpertLens could allow it to be adapted for expert elicitation of input values for decision models. Ongoing research that aims to develop a reference protocol for expert elicitation in healthcare decision making could also be of use. However, because the protocol is still in development it was not possible to assess it against NICE's requirements.

Online survey tools for gathering expert opinion could be utilised by NICE to reach a wider pool of both professional and lay experts in a cost-effective manner. However, use of surveys could not entirely replace conversations with experts. Rather, a survey tool may be useful for collecting and analysing information typically obtained through questionnaires, or for gaining additional information.

4.1 Limitations

Due to the pragmatic nature of this review (including primarily searching for systematic reviews only), we did not identify every available tool or protocol to support expert elicitation. Further, by focusing primarily on tools for expert elicitation in the context of HTA, we may have missed relevant knowledge and resources from other sectors. The small sample size of individuals from each of the NICE guidance-making programmes and external evidence review groups is a further limitation to this research, especially considering the apparent diversity in the use of expert judgement across programmes. The processes and opinions described were typically based upon the views of one NICE staff member and either a committee chair or an individual from an external organisation involved in guidance development.

4.2 Conclusions

This review of current NICE processes for using expert judgement and suitability of tools, software and protocols to assist with this process identifies valuable information relating to potential areas for improvement and further research

at NICE. We outline the specific requirements of NICE from a tool or protocol for expert elicitation and explore how existing tools/protocols could be adapted to meet these requirements. These findings could inform international HTA agencies and other groups interested in conducting expert elicitation to inform healthcare decision making.

Data Availability Statement Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Acknowledgements The authors are extremely grateful to all of the individuals who took part in the interviews for this project. Further, the authors thank the experts researching this area who provided details of their upcoming research, the tool and software developers who provided further information, and the NICE staff who took part in appraisals of tools and software. Katy Harrison and Luke Cowie of NICE provided helpful feedback on the draft manuscript and Mick Arber of York Health Economics Consortium undertook the literature search informing the review.

Author Contributions All authors helped to conceptualise the research, reviewed the results, helped to write and revise the manuscript, and approved the manuscript submitted for publication. AP and MJ led the writing of the manuscript.

Compliance with Ethical Standards

This research was funded by the National Institute for Health and Care Excellence (NICE).

Conflict of interest Newcastle upon Tyne Hospitals and York Health Economics Consortium are funded by NICE to support the Medical Technologies Evaluation Programme as an External Assessment Centre (EAC). Four of the authors (AP, MJ, JC, AS) work for the EAC but otherwise have no conflicts of interest. MC, JCRP and RL are employees of NICE.

Open Access This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix 1: The guides used during semi-structured interviews

N.B. A follow-up question about requirements from a tool or protocol was also asked.

Figure 1.1: General interview guide

- 1: Setting up the elicitation exercise:**
 - When in the process of guidance production does the programme seek expert advice? *(If there are multiple points, provide answers to each of the following questions for each point of expert input e.g. topic selection/guidance production)?*
 - How are experts recruited and selected? Within this, please explain the take up of involvement and also what influences experts dropping out;
 - What determines how many experts are used?
 - Is expert advice from patients also incorporated? If so, please explain where in the process this is incorporated and how patients are recruited/selected and their number determined;
 - Are experts (including patient experts) provided with any training?
 - Are experts asked to declare potential conflicts of interest?
- 2: The expert elicitation itself:**
 - How are the questions determined and structured?
 - Are the questions piloted?
 - What background information is provided?
 - How is elicitation conducted (i.e. in person, via email etc.);
 - What evidence is elicited (e.g. quantitative/qualitative etc.);
 - If quantitative, what are the typical values requested (e.g. mean, 95% confidence intervals, interquartile ranges, median, mode, minimum, maximum)?
 - Are multiple rounds of elicitation conducted?
 - Are results from first round shared and experts given the opportunity to refine answers in light of the pooled responses?
- 3: Converting expert elicitation into something useful for guidance:**
 - How is the information used in decision making, including:
 - How are data from experts analysed (i.e. are values weighted)?
 - Is uncertainty in individual experts' responses captured? If so, are decision makers advised of the uncertainty, and if yes, how is this done?
 - Is variability between experts captured? If so, are decision makers advised of the uncertainty, and if yes, how is this done?
 - How are any values generated used (e.g. are they combined with other evidence in a meta-analysis or are they presented alongside clinical evidence)?
 - How much does expert opinion influence the recommendations made?
 - Do you have examples of information elicited from experts on previous guidelines?
- 4: Overarching questions:**
 - Considering the current process, what are the opportunities and challenges associated with this?
 - Do you consider the current process produces reliable, valid and unbiased answers?
 - Are you happy to be involved in the judgment of any existing tools identified for their suitability to your guidance-making programme?

Figure 1.2: Interview guide for NICE lay expert staff**1: Setting up the elicitation exercise:**

- When in the process of guidance production does the programme seek patient expert advice? *(If there are multiple points, provide answers to each of the following questions for each point of expert input e.g. topic selection/guidance production)?*
- How are patient experts recruited and selected? Within this, please explain the take up of involvement and also what influences experts dropping out;
- What determines how many experts are used?
- Are patient experts provided with any training?
- Are patients asked to declare potential conflicts of interest?

2: The expert elicitation itself:

- How are the questions determined and structured?
- Are the questions piloted?
- What background information is provided?
- How is elicitation conducted (i.e. in person, via email etc.);
- What evidence is elicited (e.g. quantitative/qualitative etc.);
- If quantitative, what are the typical values requested (e.g. mean, 95% confidence intervals, interquartile ranges, median, mode, minimum, maximum)?
- Are multiple rounds of elicitation conducted?
- Are results from first round shared and patients given opportunity to refine answers in light of the pooled responses?

3: Converting expert elicitation into something useful for guidance:

- How is the information used in decision making:
 - How are data from patient experts analysed (i.e. are values weighted)?
 - Is uncertainty in individual patient experts' responses captured? If so, are decision makers advised of the uncertainty, and if yes, how is this done?
 - Is variability between patient experts captured? If so, are decision makers advised of the uncertainty, and if yes, how is this done?
 - How are any values generated used (e.g. are they combined with other evidence in a meta-analysis or are they presented alongside clinical evidence)?
 - How much does patient expert opinion influence the recommendations made?
- Do you have examples of information elicited from patient experts on previous guidelines?

4: Overarching questions:

- Considering the current process, what are the opportunities and challenges associated with this?
- Do you consider the current process produces reliable, valid and unbiased answers?

Figure 1.3: Interview guide for NICE patient expert staff

- 1: Setting up the elicitation exercise:**
- What is your role in guidance development?
 - When in the process of guidance production do you seek expert advice? *(If there are multiple points, provide answers to each of the following questions for each point of expert input e.g. topic selection/guidance production)?*
 - Are all experts identified through NICE? If not, please explain how experts are recruited and selected;
 - What determines how many experts are used?
 - Is expert advice from patients also incorporated? If so, please explain where in the process this is incorporated and if patients are identified through NICE;
 - Do you provide experts (including patient experts) with any training?
 - Do you ask experts to declare any potential conflicts of interest?
- 2: The expert elicitation itself:**
- How are the questions determined and structured?
 - Are the questions piloted?
 - What background information is provided?
 - How is elicitation conducted (i.e. in person, via email etc.);
 - What evidence is elicited (e.g. quantitative/qualitative etc.);
 - If quantitative, what are the typical values requested (e.g. mean, 95% confidence intervals, interquartile ranges, median, mode, minimum, maximum)?
 - Are multiple rounds of elicitation conducted?
 - Are results from first round shared and experts given opportunity to refine answers in light of the pooled responses?
- 3: Converting expert elicitation into something useful for guidance:**
- How is the information used in decision making, including:
 - How are data from experts analysed (i.e. are values weighted)?
 - Is uncertainty in individual experts' responses captured? If so, are decision makers advised of the uncertainty, and if yes, how is this done?
 - Is variability between experts captured? If so, are decision makers advised of the uncertainty, and if yes, how is this done?
 - How are any values generated used (e.g. are they combined with other evidence in a meta-analysis or are they presented alongside clinical evidence)?
 - How much does expert opinion influence the recommendations made?
 - Do you have examples of information elicited from experts on previous guidelines on which you have worked?
- 4: Overarching questions:**
- Considering the current process, what are the opportunities and challenges associated with this?
 - Do you consider the current process produces reliable, valid and unbiased answers?

Appendix 2: Search strategy for MEDLINE (Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to Present)

1	Expert Testimony/ and elicit\$.ti,ab,kf. (72)
2	(expert\$ or stakeholder\$1 or advisory group\$1 or advisory panel\$1).ti,ab,kf. (167178)
3	elicitation\$1.ab,kf. or elicit\$.ti. or elicit\$.ab. /freq=2 (50682)
4	2 and 3 (592)
5	(Technology Assessment, Biomedical/ or Cost-Benefit Analysis/ or exp Models, Economic/) and (elicitation\$1.ab,kf. or elicit\$.ti.) (170)
6	(health technology assessment reports or health technology assessment rockville md or health technology assessment winchester england or "international journal of technology assessment in health care").jn. and elicit\$.ti,ab,kf. (58)
7	(Technology Assessment, Biomedical/ or Cost-Benefit Analysis/ or exp Models, Economic/) and ((expert\$ or stakeholder\$1 or clinician\$1) adj (knowledge or view or views or viewpoint\$1 or advice or testimony or testimonies or input or inputs or contribution\$1 or subjective or estimate or estimates or opinion\$1 or judgement\$1)).ti,ab,kf. (533)
8	(health technology assessment reports or health technology assessment rockville md or health technology assessment winchester england or "international journal of technology assessment in health care").jn. and ((expert\$ or stakeholder\$1 or clinician\$1) adj (knowledge or view or views or viewpoint\$1 or advice or testimony or testimonies or input or inputs or contribution\$1 or subjective or estimate or estimates or opinion\$1 or judgement\$1)).ti,ab,kf. (52)
9	(technology assessment\$1 or technology appraisal\$1 or hta or htas or model\$ or economic\$ or cost).ti,kf. and (elicitation\$1.ab. or elicit\$.ti,kf.) (657)
10	((expert\$ or stakeholder\$1 or clinician\$1) adj10 elicit\$.ti,ab,kf. (925)
11	((advisory group\$1 or advisory panel\$) adj10 elicit\$.ti,ab,kf. (3)
12	((expert\$ or stakeholder\$1 or clinician\$1) adj3 (prior or belief\$1)).ti,ab,kf. (680)
13	(elicit\$ adj5 (distribution\$1 or belief\$1)).ti,ab,kf. (534)
14	(probability distribution\$1 or prior distribution\$1).ti,ab,kf. (7054)
15	subjectivist.ti,ab,kf. (56)
16	(subjective adj3 probabilit\$).ti,ab,kf. (334)
17	or/1,4-16 (10829)
18	Expert Testimony/ or (expert\$ or stakeholder\$1 or clinician\$1).ti,ab,kf. (335263)
19	(advisory group\$1 or advisory panel\$).ti,ab,kf. (2307)
20	(prior adj (distribution\$1 or knowledge or belief\$1 or view or views or viewpoint\$1 or advice or testimony or testimonies or input or inputs or contribution\$1 or subjective or estimate or estimates or opinion\$1 or probabil\$ or judgement\$1)).ti,ab,kf. (8136)
21	priors.ti,ab,kf. (2348)
22	(parameter\$1 adj3 (estimate or estimates)).ti,ab,kf. (11372)
23	(bayes\$ and (probabilit\$ or probabilistic\$ or uncertain\$ or subjective or distribution\$1 or prior or posterior\$1)).ti,ab,kf. (17290)
24	(probabilistic adj5 sensitivity analys\$).ti,ab,kf. (2013)
25	or/20-24 (37423)
26	(18 or 19) and 25 (1565)
27	((expert\$ or stakeholder\$1 or clinician\$1) adj3 (knowledge or view or views or viewpoint\$1 or advice or testimony or testimonies or input or inputs or contribution\$1 or subjective or estimate or estimates or opinion\$1 or judgement\$1)).ti,ab,kf. (29284)
28	((expert\$ or advisory) adj2 (panel\$1 or group\$1)).ti,ab,kf. (16536)
29	(elicit\$ adj3 (knowledge or view or views or viewpoint\$1 or advice or testimony or testimonies or input or inputs or contribution\$1 or subjective or estimate or estimates or opinion\$1 or judgement\$1)).ti,ab,kf. (1704)
30	((parameter\$1 or input or inputs) adj3 (knowledge or belief\$1 or view or views or viewpoint\$1 or advice or testimony or testimonies or contribution\$1 or subjective or estimate or estimates or opinion\$1 or judgement\$1)).ti,ab,kf. (15209)
31	or/27-30 (61203)

32	Bayes Theorem/ (25424)
33	bayes\$.ti,ab,kf. (37419)
34	Probability/ (53252)
35	(probabilit\$ or probabilistic\$.ti,ab,kf. (179453)
36	(prior\$1 or posterior\$1).ti,ab,kf. (735227)
37	uncertainty/ (9130)
38	uncertain\$.ti,ab,kf. (130238)
39	or/32-38 (1082549)
40	Technology Assessment, Biomedical/ (8963)
41	(health technology assessment reports or health technology assessment rockville md or health technology assessment winchester england or "international journal of technology assessment in health care").jn. (3355)
42	(technology assessment\$1 or technology appraisal\$1 or hta or htas).ti,ab,kf. (6668)
43	exp Models, Economic/ (12626)
44	models, theoretical/ or models, statistical/ or Computer Simulation/ or monte carlo method/ or markov chains/ (362180)
45	Decision Trees/ (10085)
46	(model\$ or decision tree\$1 or decision analytic\$ or decision analys\$ or decision-making or markov or cohort simulation\$1 or agent-based or event-based or event simulation\$1 or micro-simulation\$1 or microsimulation\$1 or state-transition or monte carlo simulation\$ or mcmc).ti,ab,kf. (2370213)
47	or/40-46 (2539786)
48	Economics/ (27063)
49	exp "costs and cost analysis"/ (208808)
50	Economics, Dental/ (1894)
51	exp economics, hospital/ (22309)
52	Economics, Medical/ (9037)
53	Economics, Nursing/ (3982)
54	Economics, Pharmaceutical/ (2732)
55	(economic\$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic\$.ti,ab,kf. (632970)
56	(expenditure\$ not energy).ti,ab,kf. (24031)
57	value for money.ti,ab,kf. (1325)
58	budget\$.ti,ab,kf. (24232)
59	or/48-58 (761406)
60	((energy or oxygen) adj cost).ti,ab,kf. (3553)
61	(metabolic adj cost).ti,ab,kf. (1138)
62	((energy or oxygen) adj expenditure).ti,ab,kf. (21398)
63	or/60-62 (25184)
64	59 not 63 (755657)
65	64 or ec.fs. (921400)
66	31 and 39 and (47 or 65) (4796)
67	17 or 26 or 66 (16085)
68	systematic\$ review\$.ti,ab. (102893)
69	meta-analysis as topic/ (15730)
70	meta-analytic\$.ti,ab. (5262)
71	meta-analysis.ti,ab,pt. (115640)
72	metanalysis.ti,ab. (156)
73	metaanalysis.ti,ab. (1384)
74	meta analysis.ti,ab. (93758)
75	meta-synthesis.ti,ab. (525)
76	metasynthesis.ti,ab. (229)

77	meta synthesis.ti,ab. (525)
78	meta-regression.ti,ab. (4516)
79	metaregression.ti,ab. (436)
80	meta regression.ti,ab. (4516)
81	(synthes\$ adj3 literature).ti,ab. (2205)
82	(synthes\$ adj3 evidence).ti,ab. (6459)
83	integrative review.ti,ab. (1682)
84	data synthesis.ti,ab. (9757)
85	(research synthesis or narrative synthesis).ti,ab. (1569)
86	(systematic study or systematic studies).ti,ab. (9788)
87	(systematic comparison\$ or systematic overview\$).ti,ab. (2584)
88	evidence based review.ti,ab. (1693)
89	comprehensive review.ti,ab. (10381)
90	critical review.ti,ab. (13422)
91	quantitative review.ti,ab. (581)
92	structured review.ti,ab. (639)
93	realist review.ti,ab. (156)
94	realist synthesis.ti,ab. (115)
95	or/68-94 (237875)
96	review.pt. (2258402)
97	medline.ab. (84210)
98	pubmed.ab. (64924)
99	cochrane.ab. (51588)
100	embase.ab. (53920)
101	cinahl.ab. (17171)
102	psyc?lit.ab. (936)
103	psyc?info.ab. (16894)
104	(literature adj3 search\$).ab. (41203)
105	(database\$ adj3 search\$).ab. (39342)
106	(bibliographic adj3 search\$).ab. (1810)
107	(electronic adj3 search\$).ab. (14671)
108	(electronic adj3 database\$).ab. (18390)
109	(computer?ed adj3 search\$).ab. (3221)
110	(internet adj3 search\$).ab. (2457)
111	included studies.ab. (13473)
112	(inclusion adj3 studies).ab. (10769)
113	inclusion criteria.ab. (57176)
114	selection criteria.ab. (25362)
115	predefined criteria.ab. (1529)
116	predetermined criteria.ab. (903)
117	(assess\$ adj3 (quality or validity)).ab. (58174)
118	(select\$ adj3 (study or studies)).ab. (51559)
119	(data adj3 extract\$).ab. (43484)
120	extracted data.ab. (10011)
121	(data adj2 abstracted).ab. (4225)
122	(data adj3 abstraction).ab. (1220)
123	published intervention\$.ab. (142)
124	((study or studies) adj2 evaluat\$).ab. (144740)
125	(intervention\$ adj2 evaluat\$).ab. (8525)
126	confidence interval\$.ab. (312859)
127	heterogeneity.ab. (124790)
128	pooled.ab. (65103)

```

129 pooling.ab. (9843)
130 odds ratio$.ab. (204985)
131 (Jadad or coding).ab. (149946)
132 or/97-131 (1100238)
133 96 and 132 (178456)
134 review.ti. (353285)
135 134 and 132 (85111)
136 (review$ adj4 (papers or trials or studies or evidence or intervention$ or evaluation$)).ti,ab.
    (142261)
137 95 or 133 or 135 or 136 (416274)
138 67 and 137 (876)
139 exp animals/ not humans/ (4360649)
140 (news or comment or editorial or letter or case reports).pt. or case report.ti. (3444719)
141 138 not (139 or 140) (859)
142 limit 141 to (english language and yr="2010 -Current") (542)
143 remove duplicates from 142 (523)

```

Key to Ovid symbols and commands

\$	Unlimited truncation
\$N	Limited truncation - restricts the number of characters following the word to N
?	Wildcard symbol
ti,ab,kf,	Searches are restricted to the Title, Abstract, Keyword Heading Word fields
/	Searches are restricted to the Subject Heading field
exp	The subject heading is exploded
*	The subject heading is searched as a major descriptor only
pt.	Search is restricted to the publication type field
ab. /freq=2	Search is restricted to records where the terms occur at least twice in the abstract
or/1-3	Combines sets 1 to 3 using OR

Appendix 3: The 16 tools for expert elicitation initially included in our review and their potential usability by NICE

Name of tool and if considered further (bold)	Analysis of potential for use by NICE
@RISK 7.5 [18] No	Tool uses Monte Carlo simulation to analyse risk (commonly used in context of finance). A separate protocol for eliciting experts' judgements would be required, no protocols for expert elicitation are provided and elicitation is not a primary aim of the tool
AgenaRisk [13] No	Need to know how to use the software, thus facilitator would need to be trained in this. Tool appears to be designed to be used in person
ElicitN [14] No	For face-to-face use and requires use of R, which are both barriers to use
Elicitor (v. 2010) [15] No	For face-to-face use and requires use of R, which are both barriers to use
Excalibur [11] Yes	For face-to-face use and programme may not be user-friendly, which could be barriers to use. Uses seed questions with known answers to weight experts' judgements. The program was considered further following a discussion between the authors, as relatively little information was available online about the software
ExpertLens [21] Yes	Uses an online platform and seems useful for a large group of experts (there are also case studies of use with smaller groups, e.g. 4/5 people). Although time may be a constraint with a large group, this could be streamlined with fewer experts. A cost is likely to be incurred each time it is used by NICE
MATCH [17] Yes	A more useable interface for SHELF (see below) and so was explored further. Users may still require training in how to use the tool. Pooling of experts' responses is not possible within the tool, meaning that users would also need the skills and knowledge to do this, and additional software may be required
Modelrisk 3.0 [25] No	There is a potential cost associated and experts would need to provide a consensus. Also more suited for sensitivity analysis than expert elicitation. All of these form barriers to use
PEGS (Prior elicitation of graphical software) [24] Yes	May be issues with usability as the interface does not look user friendly and experts are expected to input data themselves. Also, may not combine multiple expert options. The program was considered further following a discussion between the authors, as relatively little information was available online about the software

Name of tool and if considered further (bold)	Analysis of potential for use by NICE
RAMAS Constructor [19] No	High cost and the fact that the tool is not specifically designed for expert elicitation and cannot combine multiple expert opinions are barriers to use
RAMAS Risk Calc [20] No	The cost and the fact that experts would need to provide a consensus or probability distributions form barriers to use. Also more suited for sensitivity analysis than expert elicitation
RiskSim (now SimVoi) [16] No	A separate programme/protocol for eliciting experts' judgements would be required to allow these values to be fed into the risk analyses, no protocols for expert elicitation are provided and it is not a primary aim of the tool
SHELF 3.0 [32] Yes	SHELF is designed for use in a health-care setting. However, the user needs knowledge of R and, therefore, the cost of training staff in using R and considerations around whether NICE staff would want to act in this facilitator role may prohibit use. It is also designed to be used in a workshop setting. MATCH (see above) is a more accessible version of SHELF The program was considered further following discussion between the authors, as it is one of the most established and developed methodologies for expert elicitation
UncertWeb-Elicitor tool [12] No	Likely to require expertise in R. Also may no longer be available (email delivery failed when contacting developer for further information). Therefore, the tool was not considered further
Unicorn [22] No	Not useful across all types of expert elicitation required in guidance making: can only be used for eliciting certain parameter types. Therefore, the tool was not considered further
Variogram [23] No	Designed for one very specific type of elicitation ('Kriging' in geostatistics), which is not applicable to health care decision making

References

1. Iglesias CP, Thompson A, Rogowski WH, Payne K. Reporting guidelines for the use of expert judgement in model-based economic evaluations. *Pharmacoeconomics*. 2016;34(11):1161–72.
2. Evans D. Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions. *J Clin Nurs*. 2003;12(1):77–84.
3. O'Hagan A, Oakley J. SHELF: the Sheffield Elicitation Framework v3.0. 2016. Available from: <http://www.tonyohagan.co.uk/shelf/>.

4. O'Hagan A, Buck CE, Daneshkhan AJ, Eiser R, Garthwaite PH, Jenkinson DJ, et al. *Uncertain judgements: eliciting experts' probabilities*. New York: Wiley; 2006.
5. Garthwaite PH, Kadane JB, O'Hagan A. Statistical methods for eliciting probability distributions. *J Am Stat Assoc*. 2005;100(470):680–701.
6. Cooke RM. *Experts in uncertainty: opinion and subjective probability in science*. Oxford: Oxford University Press; 1991.
7. Devilee J, Knol A. Software to support expert elicitation: an exploratory study of existing software packages. RIVM Letter report 630003001/2011. National Institute for Public Health and the Environment of the Netherlands. 2012. Available from: <https://www.rivm.nl/bibliotheek/rapporten/630003001.pdf>.
8. Sharma T, Choudhury M, Kaur B, Naidoo B, Garner S, Littlejohns P, et al. Evidence informed decision making: the use of “colloquial evidence” at NICE. *Int J Technol Assess Health Care*. 2015;31(3):138–46.
9. National Institute for Health and Care Excellence. NICE Guidance. 2017. [cited 17th July 2017]. Available from: <https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance>.
10. Speight P, Palmer S, Moles D, Downer M, Smith D. The cost-effectiveness of screening for oral cancer in primary care. *Health Technol Assess*. 2006;10(14):1–144, iii–iv.
11. Ababei D. EXCALIBUR. Available from: <http://www.lighttwist.net/wp/excalibur>.
12. Bigagli L. UncertWeb. 2013. Available from: <https://wiki.aston.ac.uk/foswiki/bin/view/UncertWeb/UncertWebSoftware>.
13. Fenton N, Neil M. AgenaRisk 7.0. 2016. Available from: <http://www.agenarisk.com/>.
14. Fisher R, O'Leary R. ElicitN. 2011. Available from: <http://www.downloadcollection.com/elicitn.htm>.
15. James A, Low Choy S, Mengerson K. Elicitor (v. 2010). 2010.
16. Middleton M. RiskSim. 2013. Available from: <http://www.decisiontoolworks.com/>.
17. Morris D, Oakley J, Crowe J. MATCH. 2014. Available from: <http://optics.eee.nottingham.ac.uk/match/uncertainty.php>.
18. Palisade. @Risk 7.5. [program] UK; 2016. Available from: <http://www.palisade.com/>.
19. RAMAS. RAMAS Constructor. 2016. Available from: <http://www.ramas.com/constructor-1>.
20. RAMAS. RAMAS Risk Calc. 2016. Available from: <http://www.ramas.com/riskcalc>.
21. RAND. ExpertLens. 2011. Available from: <http://www.rand.org/pubs/tools/expertlens.html>.
22. Technical University of Delft. Unicorn. 2007. Available from: <http://www.lighttwist.net/wp/unicorn-download>.
23. Truong P, Heuvelink G, Gosling J. VARIOGRAM. 2013. Available from: <http://www.variogramelicitation.org/>.
24. Open University. PEGS (Prior elicitation of graphical software). Available from: <http://statistics.open.ac.uk/elicitation>.
25. Vose Software. ModelRisk 3.0. 2008. Available from: <http://vosesoftware.com/knowledgebase/newsletters/newsletters/06/index.html>.
26. Durbach I, Mervin B, McCall B. Expert elicitation of auto-correlated time series with application to e3 (energy-environment-economic) forecasting models. *Environ Model Softw*. 2017;88:93–105.
27. Expert Elicitation Task Force. Expert Elicitation Task Force White Paper. 2009. Available from: [https://yosemite.epa.gov/sab/sabproduct.nsf/fedrgstr_activites/F4ACE05D0975F8C68525719200598BC7/\\$File/Expert_Elicitation_White_Paper-January_06_2009.pdf](https://yosemite.epa.gov/sab/sabproduct.nsf/fedrgstr_activites/F4ACE05D0975F8C68525719200598BC7/$File/Expert_Elicitation_White_Paper-January_06_2009.pdf).
28. Hoffmann S, Fischbeck P, Krupnick A, McWilliams M. Elicitation from large, heterogeneous expert panels: using multiple uncertainty measures to characterize information quality for decision analysis. *Decis Anal*. 2007;4(2):91–109.
29. Sperber D, Mortimer D, Lorgelly P, Berlowitz D. An expert on every street corner? Methods for eliciting distributions in geographically dispersed opinion pools. *Value Health*. 2013;16(2):434–7.
30. Medical Research Council. Successfully funded proposals. 2017. [cited 24th July 2017]. Available from: <https://www.mrc.ac.uk/funding/how-we-fund-research/highlight-notices/methodology-for-eliciting-expert-opinion/successfully-funded-proposals/>.
31. Grigore B, Peters J, Hyde C, Stein K. EXPLICIT: a feasibility study of remote expert elicitation in health technology assessment. *BMC Med Inform Decis Mak*. 2017;17:131.
32. O'Hagan A, Oakley J. SHELF: the Sheffield Elicitation Framework v3.0. [program]. 2016. <http://www.tonyohagan.co.uk/shelf/>.