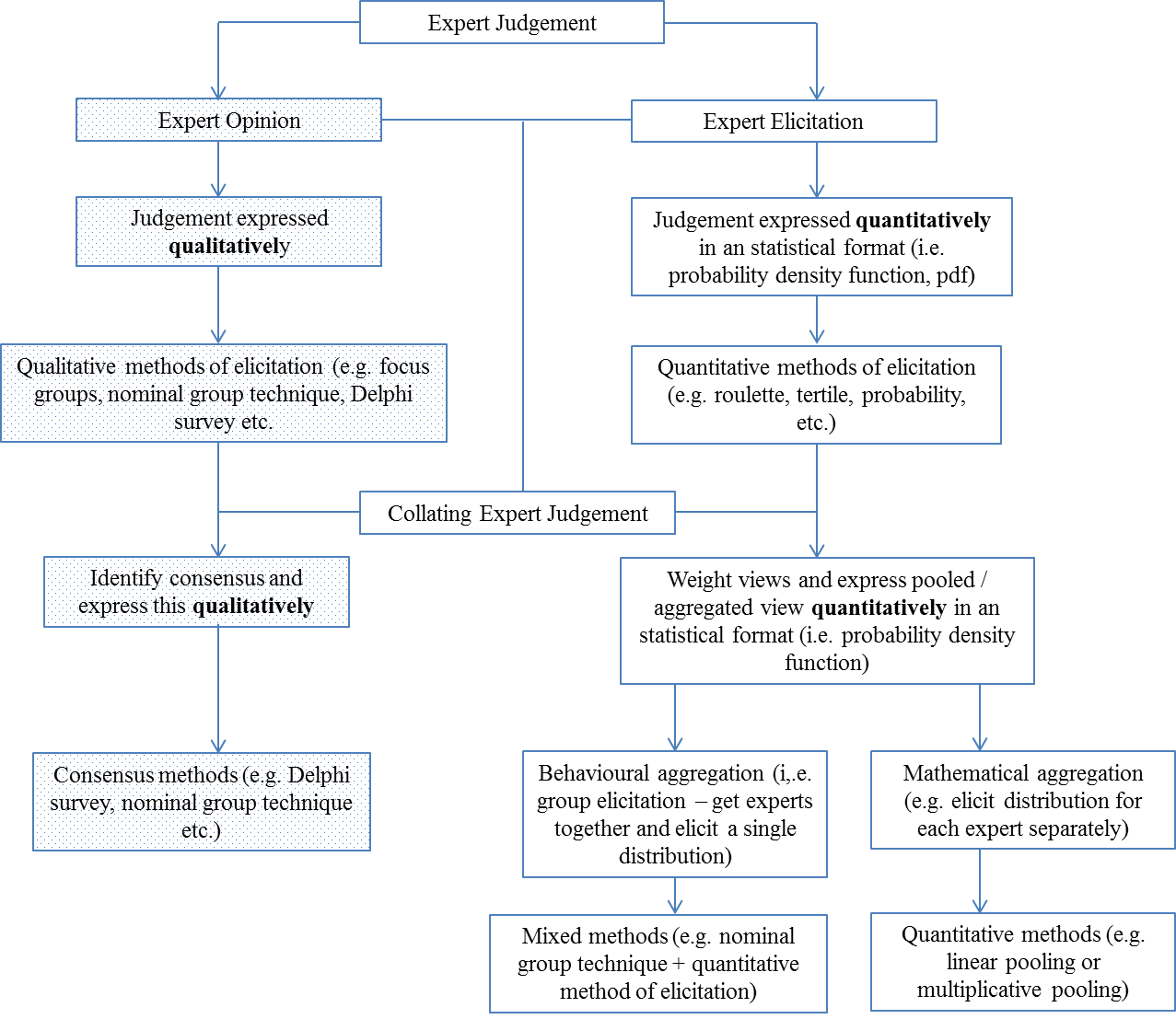
**Figure 1: Suggested nomenclature of expert judgement**

****

**Box 1: ‘Agreed’ definitions of core concepts**

|  |  |
| --- | --- |
| An **expert** is: | *‘someone who has in-depth knowledge of the topic of interest gained through their life experience, education or training’.* |
| A study designed to generate **expert parameter values** uses: | *‘a quantitative elicitation method to: derive point estimates and distributions for model input parameters; and/or pool expert judgement.’* |
| The purpose of an **expert elicitation study** is: | *‘to quantify a parameter value that appropriately represents the knowledge/judgement of the expert and the degree of uncertainty in that knowledge/judgement’.* |
| A study designed to collate **expert opinion** uses: | *‘a qualitative consensus method (e.g. a Delphi method or other approach) to collate views from experts to: frame the scope of the model-based economic evaluation; inform model conceptualisation; identify model face validity; quantify point estimates without specifying a distribution; pool elicitation results’.* |

|  |  |  |
| --- | --- | --- |
| **Table 1: Dictionary definitions of some core concepts** | | |
| Concept | Definition | Source |
| opinion | a view or judgement formed about something, not necessarily based on fact or knowledge; the beliefs or views of a group or majority of people - statement of advice by an expert on a professional matter | Oxford Dictionaries [35] |
| judgement | the ability to make considered decisions or come to sensible conclusions; an opinion or conclusion | Oxford Dictionaries [35] |
| belief | an acceptance that something exists or is true, especially one without proof; something one accepts as true or real; a firmly held opinion | Oxford Dictionaries [35] |
| knowledge | facts, information, and skills acquired through experience or education; the theoretical or practical understanding of a subject; the sum of what is known; information held on a computer system; true, justified belief; certain understanding, as opposed to opinion | Oxford Dictionaries [35] |
| expert | A) individual ‘whose knowledge can support informed judgement and prediction about the issues of interest’  B) someone who has knowledge of the subject of interest gained through their life experience, education or training’ | A)Morgan (2014) [9]  B) Garthwaite (2005). J Am Stat Assoc. [36] |
| elicitation | i) the act of getting information or a reaction from somebody, often with difficulty  ii) stimulation that calls up (draws forth) a particular class of behaviours;  iii) the process of making someone react in a particular way | i) Oxford Dictionary. [35]  ii) The Free Dictionary [37]  iii) Macmillan Dictionary [38] |

**Table 2: Rating scale**

|  |  |  |
| --- | --- | --- |
| **Score** | **Rating** | **Description** |
| *Five-point scale used to rate definitions\** | | |
| **1** | Strongly disagree | This means that you think the definition as written does NOT define the term and requires extensive modification |
| **2** | Disagree | This means that you think the definition as written does NOT define the term but only requires minor modification |
| **3** | Neither agree or disagree | This means that you do not have a strong opinion on the definition as written |
| **4** | Agree | This means that you think the definition as written DOES define the term but only requires some minor modification |
| **5** | Strongly agree | This means that you think the definition as written DOES define the term and requires no modification |
| *Five-point scale used to rate reporting criteria\** | | |
| **1** | definitely not required | This means that you think that the criteria should NOT be included in the reporting criteria. |
| **2** | possibly not required | This means that you think that the criteria could probably be omitted without any loss of key detail. |
| **3** | no strong opinion | This means that you do not have a strong opinion to indicate whether the criteria is, or is not, required. |
| **4** | possibly required | This means that the criteria could be included but it is not vital (it would be ‘nice to have’) |
| **5** | definitely required | This means you think that the criteria should be included in the reporting criteria otherwise key detail will not be reported. |

**\***There was also an option for the respondent to indicate if ‘they do not know’ the answer

**Table 3: Reporting criteria for an expert elicitation study**

|  |  |  |
| --- | --- | --- |
| **Criterion** | **Description** | **Note** |
| Research rationale | The need for using an expert elicitation exercise should be described | This should ideally include some reference to the design and conduct of systematic reviews to identify key input parameters for the decision analytic model and a statement confirming that these reviews did not identify data relevant for the model-based economic analysis as specified |
| Research problem | All uncertain quantities (model input parameters) that will be elicited should be described | In some instances, there may be a substantial number of uncertain quantities required, and a degree of 'pre-selection' will have occurred to identify a relevant sub-set. Clear justification for model parameters identified as key for the decision problem needs to be provided. |
| Measurement type of uncertain quantities | The rationale for the measure type of each uncertain quantity elicited should be described | The measurement type of uncertain quantities can be (but not limited to): scalar quantities (i.e. numbers); proportions (e.g. probabilities); ratios (e.g. odds, hazard); risk (e.g. relative); rate (e.g. mortality), etc. Some measures are easier to understand and elicit than others thus it is important to fully justify the selection of any measurement type. |
| Definition of an expert | The nature of the expert population should be described to clearly state what topic of expertise they represent and why | It is unlikely that a single expert will be sufficient and it is generally necessary to elicit judgement from a group of experts that were selected to represent the views of a larger population |
| Number of experts | The selection criteria and final number of experts recruited to provide expert judgement should be reported. | Selection criteria need to be described in detail. There should be clear and specific pre-defined criteria used to identify how experts were selected and if/how their elicited quantities were used. |
| Preparation | There should be clear reference made to a protocol that describes the design and conduct of the elicitation exercise. | None |
| Piloting | It should be clearly reported if the elicitation exercise process was piloted and a summary of any modifications made. | The selection and number of experts used in the piloting process should be reported. Key aspects that may have required modification include: selection of experts; measure type and number of uncertain quantities to be elicited; training exercise; framing of the elicitation question; method of aggregation. |
| Data collection | The approach to collect the data should be reported. | Data can be collected from individual experts or a group/s of experts. Collecting data from individual experts means that a mathematical aggregation process may need to be used. Collecting data from a group/s of experts means that behavioural aggregation methods may be used. |
| Administration | The mode of administering the elicitation exercise should be reported. | Elicitation exercises can be conducted face-to-face or via the telephone and/or computer. In a limited number of situations it may be feasible to collect the data using a self-administered online or postal survey but this is unlikely to be successful in most instances. Both face-to-face and telephone data collection is likely to be supported by using a computer. |
| Training | The use of training materials should be reported and made available. | This may include background training materials sent to the experts and/or training in the use of probabilities and nature of distributions. This document need to provide explanation of efforts made to prevent influencing experts' knowledge and judgement. In practice, this recommendation will require a copy of the elicitation exercise to be included, which is likely to be presented as electronic supplementary material |
| The exercise | The number and framing of questions used in the exercise should be reported and made available. | This will require a copy of the elicitation exercise to be included, which is likely to be presented as electronic supplementary material. |
| Data aggregation | The type of aggregation method (mathematical or behavioural) should be reported together with a description of the method or process used to aggregate the data. | Mathematical aggregation (relevant when data were collected from multiple individual experts) can be conducted using a range of methods, for example: Bayesian methods; opinion pooling; Cooke's method. Behavioural aggregation (relevant when data were collected from group/s of experts) can be conducted using processes such as, for example: Delphi or Nominal Group technique. |
| Measures of performance for data aggregation | The processes followed to estimate measures of performance (calibration/information) for data aggregation need to be fully described | Calibration is the process of measuring the performance of experts by comparing their judgement with a 'seed parameter' (parameter whose true values are known or can be found within the duration of a study). Calibration scores represent the probability that any differences between expert's probabilities and observed values of 'seed parameters' might have arisen by chance. Information represents the degree to which an expert's distribution is concentrated, relative to some user-selected background measure. |
| Ethical issues | The ethical issues for the expert sample and research community should be described. | The use of expert elicitation should acknowledge the issues of ethical responsibility, anonymity, reliability, and validity in an ongoing manner throughout the data collection and aggregation process. |
| Presentation of results | The individual, and aggregated, point estimate(s) and distribution for each uncertain quantity (quantities) should be presented. | **T**he units of measurement should be clear and attention should be paid to the style of presentation that may benefit from the use of figures rather than relying on a tabular format. |
| Interpretation of results | The interpretation of uncertain quantities elicited should be presented together with a description of how the results will be used in the model-based economic analysis. | This should include an explanation of how the reader should interpret the results. It should be recognised that the number and type of experts used will affect the results obtained. The interpretation of results should comment on the degree of uncertainty observed. |

**Table 4: Reporting criteria for a Delphi study designed to identify expert opinions**

|  |  |  |
| --- | --- | --- |
| **Criterion** | **Description** | **Note** |
| Research problem | The research problem should be clearly defined and ideally framed explicitly as a research question to be addressed. | When clarifying the research problem, remember the Delphi technique is a group facilitation technique and as such only lends itself to group involvement. |
| Research rationale | The topic and use of the Delphi method should be justified. | The Delphi is best used when the research requires anonymity to avoid dominance of one opinion. It should also be remembered that the strength of the Delphi method lies in the use of iteration in which the process of gaining opinion occurs in rounds to allow individuals to change their opinion. |
| Literature review | The rationale for using the Delphi method must be informed by a clear description of the evidence base for the topic of the study. | The focus of using the Delphi method should be where unanimity of opinion does not exist owing to a poor evidence base. This section should also describe the process of determining the most important issues to refer to in the design of the initial round of the Delphi. |
| Data collection | This should include a clear explanation of the Delphi method employed. | This should be sufficiently detailed for a reader to be able to duplicate the process of conducting the Delphi method. This includes a description of the types of questions used (qualitative or quantitative and ranking, rating or scoring scale used). This section should describe which medium was used to collect the data (electronic or written communication). This section should also describe how results from previous rounds were fed back to the experts and whether feedback is given to the group and/or individual response. |
| The survey | A copy of each round of the survey used in the Delphi method should be presented. | The use of journal supplementary appendices should be exploited to allow the reader access to a full copy of the survey used for each round of the Delphi. |
| Rounds | This should state the number of rounds planned and used together with the plans for moving from one round to the next. | The structure of the initial round (either qualitative or quantitative) should be decided from the protocol stage of the study together with the number of rounds to be used. |
| The sample | The sample or 'expert' panel should be described in terms of the definition of an expert in the context of the study and the selection and composition of the panel including how it was formed from a sampling frame and response rate achieved. | It should be noted that the composition of the panel will affect the results obtained from the Delphi method. Careful thought should be given to the criteria employed to define an expert, the justification of a participant as an 'expert' and the use of non-probability sampling techniques (such as purposive or criterion methods). |
| Ethical issues | The ethical issues for the expert sample and research community should be described. | The use of the Delphi method should acknowledge the issues of ethical responsibility, anonymity, reliability, and validity in an ongoing manner throughout the data collection and analysis process. |
| Data analysis | The management of opinions, analysis and handling of both qualitative and quantitative data should be described. | As with any other survey-based approach, a pre-specified data analysis plan should be prepared. This should include a clear description of the meaning of 'consensus' in relation to the stated aim of the study and how 'agreement' is defined. This should also take account of reliability and validity issues identified. |
| Presentation of results | The results for each round, and final round, should be presented clearly while taking account of the audience of the study findings. | The response rate for each round should be stated. Careful consideration should be paid on how to present the interim (between round) and final results in either graphical and/or statistical representations. In round 1, a summary of the total number of issues generated should be presented. In the final round, the strength of overall consensus should be summarised. Reporting data from quantitative questions should acknowledge the limitations associated with eliciting point estimates (e.g. no indication of uncertainty). |
| Interpretation of results | The interpretation of consensus (not) gained should be presented together with the meaning of the results and direction of further research needed, | This should include an explanation of how the reader should interpret the results, and how to digest the findings in relation to the emphasis being placed upon them. It should be recognised that the composition of the panel will affect the results obtained. The interpretation of results should state whether 'outliers' to the overall consensus were asked for the reasons for their answers. |