ISSN: 000000000

CENTRE FOR WORKFORCE INTELLIGENCE

CfWI technical paper series no. 0011

Elicitation methods:
Applying elicitation methods to robust workforce planning

February 2015

www.cfwi.org.uk

Abstract

JEL Classification	**000, 000, 000, 000, 000**
Keywords	Elicitation methods, horizon scanning, system dynamics, scenario generation, CfWI, intervention, workforce planning
Author	Dr Graham Willis, CfWI
Reviewer	Joe Fernandez
Acknowledgements	Our thanks to Professor Tony O'Hagan for his helpful comments on this report.
Corresponding author	Dr Graham Willis, CfWI

Contents

1. Introduction

The Centre for Workforce Intelligence (CfWI) has investigated a range of methods for quantifying the values of critical and uncertain parameters for workforce modelling. Our robust workforce planning framework approach uses scenarios to think about the possible risks and opportunities of the future. Scenarios are challenging, yet plausible and consistent futures. Individually they are extremely unlikely to happen. However, it is not impossible that they might. They represent potential futures that workforce planners need to address.

We can evaluate our workforce policies against a set of these scenarios. Those policies that are the most effective – according to our measures of effectiveness – are taken to be the most ‘robust’ when compared against future uncertainty. Others may work badly – and would be subsequently rejected. There is also an in-between where a policy may work really well against all futures apart from one. We would then have to decide if this policy was worth pursuing. If we did, we might need to scan for the signs of this future in order to militate against it.

In order to model these scenarios we need to quantify the expected values for parameters that are inherently uncertain, since we may be looking 10, 15 or even 20 years into the future. Ideally, we would systematically review all the evidence and data. However, such data does not exist and we must obtain it in some way. We can project historical data forward, but the assumption that we can extrapolate the future from the past is problematic. An alternative approach is to ask experts their opinion and quantify these critical uncertainties. The following sections describe the context and background to expert elicitation; the development of a framework for elicitation; and our planned next steps.

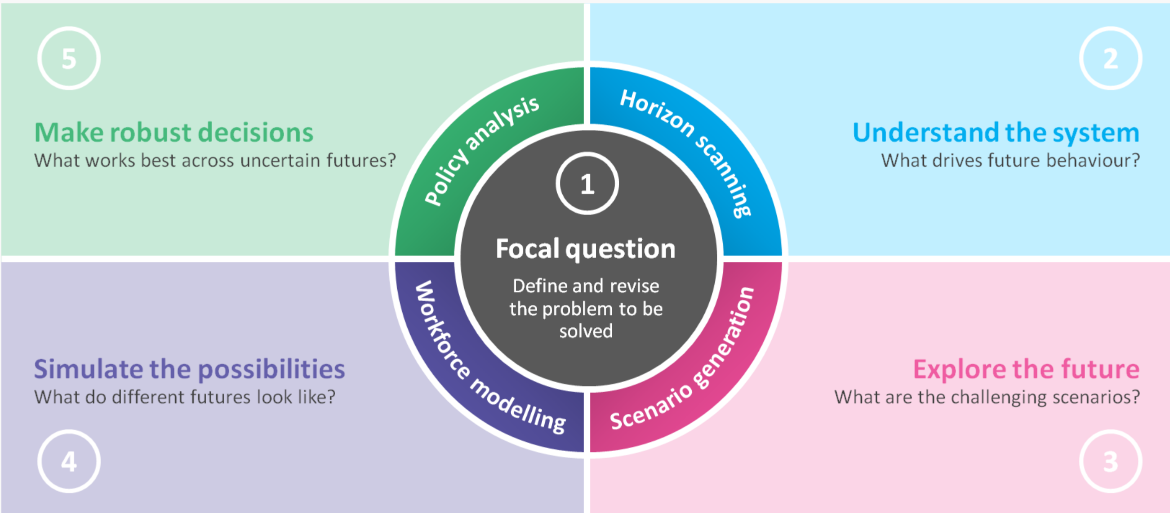
2. Context

2.1 Robust workforce planning

The CfWI use expert elicitation as part of our robust workforce planning framework, which we use in all workforce reviews. The purpose is to provide workforce planners with an understanding of how the future might evolve for health, public health and/or social care professions. It can also be used for groups of professions. This includes simulation of future supply and demand, in terms of workforce numbers and skills. It also includes an assessment of the effectiveness of different policy options, for example, increasing the intake to training, altering working patterns, or influencing the drivers of demand. Figure 1 illustrates the framework.

Figure 1: The CfWI Robust Workforce Planning Framework

The framework shows the core stages of **horizon scanning** to elicit the uncertain driving forces; **scenario generation** to produce challenging futures; **workforce modelling** to project workforce demand and supply across these futures; and **policy analysis** to determine which policies are the most robust across these uncertain futures.



Source: Centre for Workforce Intelligence

There are five parts to our approach:

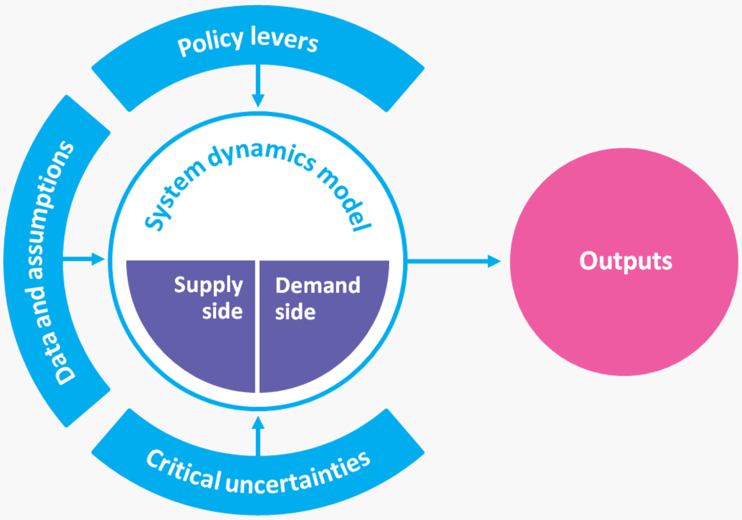
1. **Defining the focal question.** This establishes the context of the review, the workforce groups involved, and the timescale.
2. **Horizon scanning to understand the system and what drives future behaviour.** We use an online resource (www.horizonscanning.org.uk) to collect ideas about the future from stakeholders. These are then analysed using systems-thinking methods to determine what drives the future evolution of the workforce system. (CfWI, 2014a)

3. **Scenario generation to explore the future and generate challenging scenarios.** We use a workshop-based approach to scenario generation that tests the resulting scenarios for consistency using a form of cross-impact analysis (CfWI, 2014b). We then select the scenarios that are the most consistent, and quantify them for modelling using expert elicitation (as explored in this paper).
4. **Workforce modelling to simulate the different futures.** We use a tried and tested system dynamics modelling approach. The software allows us to compare what the different futures might look like in the absence of any policy actions. We can then test various policy options to see the difference they make. (CfWI, 2014c)
5. **Policy analysis to make robust decisions and decide which option works best.** We consider multiple future scenarios and test different policy options against each one. We then analyse which options are the most robust against the uncertainty of the future, i.e. perform best across all scenarios. (CfWI, 2014d)

Critical to this process is the ability to quantify parameters for modelling and simulation. Figure 2 illustrates this. Our models are driven by known data, reasonable assumptions we make where data is lacking or of low quality, controllable levers (like intake and retirement age), and intrinsic uncertainties.

Figure 2: Model inputs

The figure shows the different model input parameters. Some we know or can define, some we can control, and some are inherently unknowable because they exist in the future. It is the latter that we need to quantify using expert elicitation.



Source: Centre for Workforce Intelligence

2.2 Why do we need to use expert elicitation?

We need to use expert elicitation because the future is complex and uncertain. We cannot accurately forecast it. Many factors are uncertain and therefore open to challenge. For example, will students choose to train in a particular profession? Will people take any notice of public health advice? What will be the future health needs of the population? Will technology transform the way that we deliver services?

We might imagine a future where genomics and diagnostic technology combine to reduce demands on the health system, or in an alternative future, technology could present significant challenges to the health and care system and increase demand. Because of the complexity of factors and their interactions, the degree of uncertainty is not the same across all these futures.

In these different futures, some of the parameters that we use to model workforce demand and supply will vary. On the supply-side, this includes working hours and retirement age. On the demand-side, this includes the future need for health and care, the productivity of the delivery system, and the intensity of service delivery.

Although we have considerable historical data, forecasting the future by extrapolating the past can often lead to extreme results. Intuition and luck may lead to good forecasts, but there will be bad ones, and we cannot recognise good from bad until the future has happened. The absence of empirical data and the presence of considerable uncertainty suggests the need to use experts to inform decisions by characterising the uncertainty and filling the data gaps.

2.3 Scenarios and expert elicitation

There are many possible futures, which we capture as scenarios. A scenario has two parts: a narrative explaining how the driving forces, events and actions may unfold over time to lead to a future; and a set of quantified parameters that define this future.

The narrative of the scenario describes a chain of cause and effect; how events and actions unfold over time to reach the end state at a defined point in time. The richness of the narrative is important to paint a picture of a future that is challenging, compelling and consistent. Although this future is extremely unlikely, it is plausible; in other words, one cannot say that it could not happen.

To model any future, we first need to quantify key parameters. It would improve the accuracy of the modelling if we could capture how these parameters change over time, from today to the future. However, this is very difficult, so we usually make do with their value at the end state.

As the futures we are trying to model do not, and are unlikely to exist, we need to represent their inherent uncertainty. Giving a parameter a single fixed value without representing this uncertainty would be misleading. Parameters will vary in their certainty, depending on the future in question.

We can describe the uncertainty of a parameter on a likelihood scale, for example, from virtually certain to exceptionally unlikely. However, there are a number of problems with this approach. When asked for the numerical probability of terms like ‘probable’ or ‘possible’, people – even experts – differ widely in their interpretation (Wallsten et al, 1986). The same words can mean very different things to different people – or they could vary in meaning to the same person in different situations.

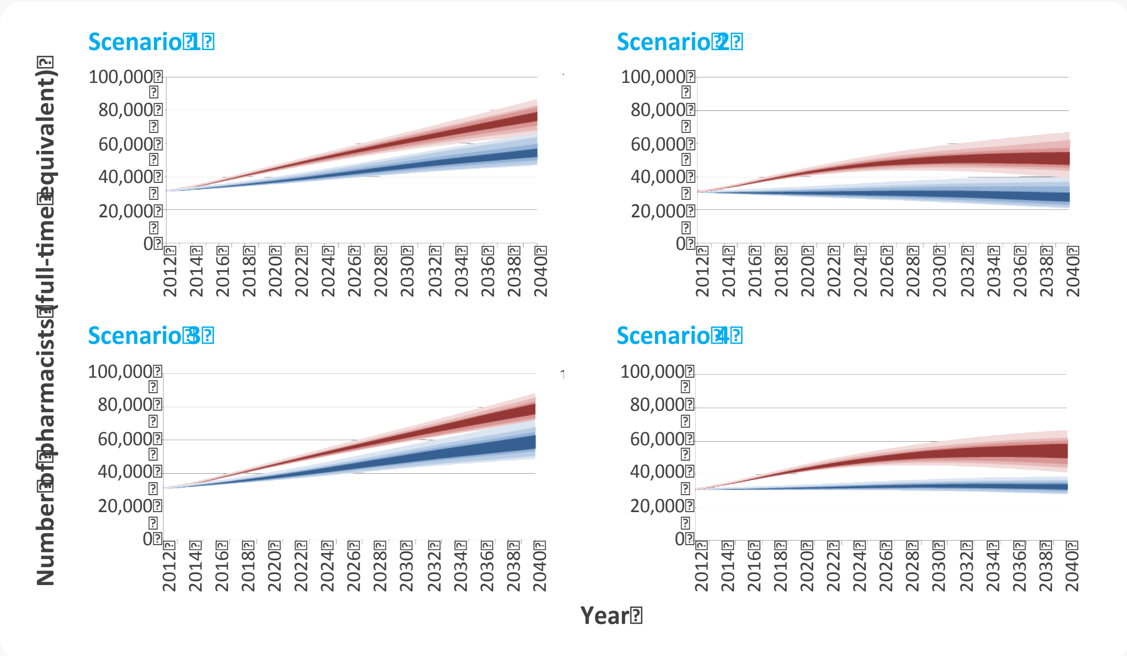
We cannot capture important differences in the judgment of experts and their reasoning with simple words. Representing differences of opinion using a simple scale is also difficult.

In our robust workforce planning framework, we often use Monte Carlo simulation software to sample the input parameters many times across their range of uncertainty. There may be several hundred or thousand iterations, building up a picture of how sensitive the outputs variables are to the uncertainty of the input variables. This allows us to see how the future defined by the scenario evolves over time, and the uncertainty of the outputs. Typically, the uncertainty increases the further that we look into the future.

We can display this as a graph with confidence bounds, as illustrated in Figure 3. In this example, the graphs show that the supply of pharmacists (the red bands) exceeds the demand for pharmacy services (the blue bands) in all four scenarios. It is reasonable to assume that, despite the uncertainty of the future, there is an over-supply and actions may need to be taken if this is not desirable.

Figure 3: Monte Carlo simulation

This shows the results of performing multiple simulations where parameters are sampled across their probability distribution. The resulting graphs – often called fan charts – show the confidence bounds. The graphs are from the CfWI strategic review of the future pharmacist workforce.



Source: Centre for Workforce Intelligence (CfWI, 2013)

3. Background to expert elicitation

3.1 Definition

Expert elicitation can be defined as follows:

'Expert knowledge elicitation (EKE) refers to the drawing out of knowledge from one or more experts. Experts can be asked for specific information (facts, data, sources, requirements, etc.) or for judgements about things (preferences, utilities, probabilities, estimates, etc.). Elicitation of specific information is relatively simple; the expert either does or does not know the answer, so two experts who both know the answer should give exactly the same values. Eliciting judgements is more challenging as we wish the expert to use his or her expertise to, for instance, make an estimate of an uncertain quantity; now it is clear that different experts can give different answers. In the case of eliciting information we must select experts who know the right answer, whereas in eliciting judgements we need experts who not only have skill at estimation, but who can also give realistic judgements as to the accuracy of their estimates.' (European Food Safety Authority, 2014, p.7)

In the case of the CfWI robust workforce planning framework, we are interested in judgements about the future values of parameters, expressed as a probability distribution.

3.2 Representing uncertainty

We can represent the uncertainty of each parameter as a probability. Although these are subjective judgments, it is important that they behave as probabilities. If we are certain of an outcome occurring this would be given a probability of 1. This states that the outcome is 100 per cent certain. If we assign a probability of 0 we are saying that the outcome has a 0 per cent chance of occurring. Consequently, the probability of a given outcome will be between 0.0 and 1.0.

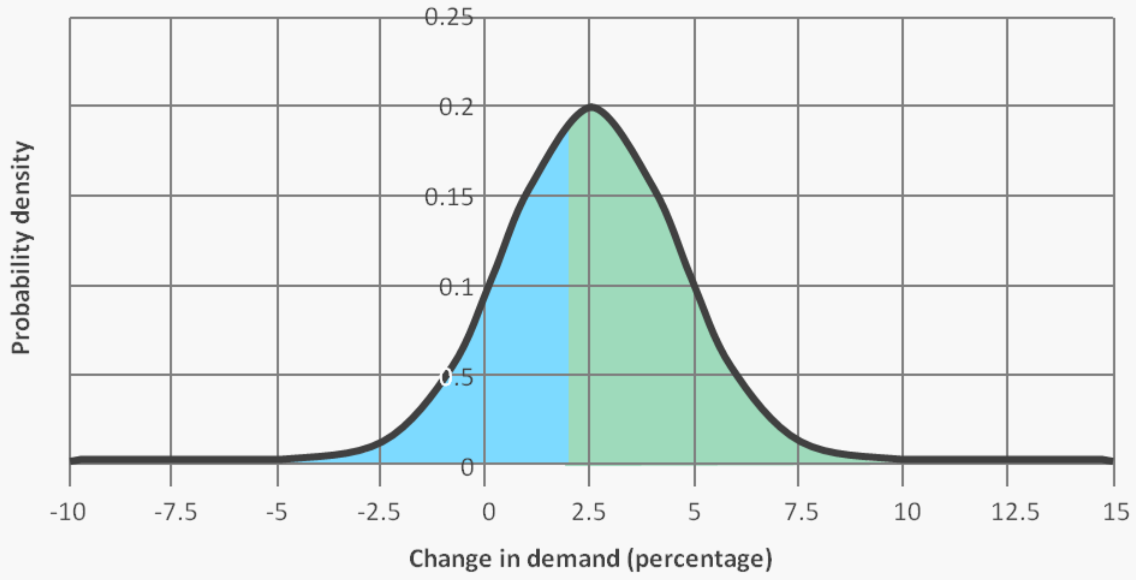
For workforce planning, we might say that the probability of demand increasing by 2 per cent is 0.6 or 60 per cent. Unfortunately, this would be completely wrong. There is not a probability of 0.6 or even 0.0001 that demand increases by *exactly* 2 per cent. More realistically, we might say that there is a probability that demand increases by 2 per cent *or more*. This is useful but it does not provide all the information we need. For instance, it does not give the probability that demand increases by 3 per cent or more, or that it increases by between 1 and 2 per cent.

To get a proper, comprehensive, description of uncertainty about a parameter like future demand for a particular service, we use a probability density function or PDF. Figure 4 provides a simple PDF. Parameter values where the PDF is high are judged to be the most probable. Even though an individual point has a zero probability (as explained below), the probability of being near a high probability density is much higher than the probability of being near one with a very low density. The area under the curve between two points represents the probability of a change in demand between these values. For example, the green area from 0 to 2 represents a probability of 0.3 or 30 per cent that demand is *between* 0 and 2 per cent. Similarly, the blue area from 2 and above represents a probability of 0.6 or 60 per cent that demand increases by 2 per cent *or more*. We can examine the graph to provide other probabilities, for example 0.4 or 40 per cent where demand increases by 3 per cent or more, and 0.175 or 17.5 per cent for an increase of between 1 and 2 per cent.

However, it is worth noting that the closer our two points are together, the smaller the area under the curve and thus the smaller the probability of a variable being between these points. So in any probability distribution over a continuous variable, the probability of any exact point value is zero.

Figure 4: Probability density function (PDF)

PDF representing the relative likelihood of an individual value. Note that the curve is symmetrical, which is unusual in most natural systems.



Source: Centre for Workforce Intelligence

3.3 Approaches to expert elicitation

We could ask experts for their estimate of the probability distribution of key parameters and by asking many experts, we could average their distributions in some way, or use their individual distributions to understand their spread of uncertainty.

If the experts shared and discussed their answers we might try to get a group consensus; the best estimate of the group. However, while these approaches may seem appealing they miss out much expert knowledge. Our experts may be making guesses and assumptions, or basing their answers on data or research known only to them.

One always needs to consider key questions in eliciting expert input such as:

- What assumptions have they made?
- What data and research have they used and where is it from?
- Is it peer-reviewed?
- Are there dissenting views and debates that we are not recording?

We cannot capture all this information in a single distribution, or even a set of distributions. However, expert elicitation is concerned with getting knowledge from one or more experts. This could be asking for information that we know to exist but perhaps we cannot easily obtain, for example, numbers that we could verify but only with a lot of effort.

Alternatively, we could be asking experts to make judgements about quantities that are uncertain and whose values cannot be verified. It is the latter that we are interested in. In the context of our robust workforce planning framework we will be eliciting judgements about:

- the expected future
- a set of challenging futures described by scenarios.

We do not use the expected future as a forecast, but as a baseline or reference for eliciting values across the set of scenarios.

There are many approaches to expert elicitation. Most follow a well-defined process. For example, the US Environmental Protection Agency (2014) describes the following steps as good practice:

1. Clear problem definition.
2. Appropriate structuring of the problem.
3. Appropriate staffing to conduct expert elicitation and select experts.
4. Protocol development and training, including the consideration of group processes and methods to combine judgement, if appropriate.
5. Procedures to verify expert judgements.
6. Clear and transparent documentation.
7. Appropriate peer review for the situation.

Elicitation can involve individuals or groups. Individual elicitation can be helpful in exploring the factors contributing to the overall uncertainty, but cannot provide a consensus view. Although using an individual avoids group bias, people may still exhibit their own biases. A set of individual elicitations can build a picture of the different perspectives on uncertainty, but without debate, people cannot challenge or modify the views of others.

In the group approach, we combine expert judgements in some way to obtain the ‘best’ estimate. We can aggregate these judgements in two basic ways:

- Mathematic aggregation, where we combine expert judgements mathematically. The aggregation is done *after* the elicitation process. It is not necessary for the experts to interact, and may be preferable that they do not, to reduce bias.
- Behavioural aggregation, where we assume that we can reach a more informed evaluation by having the experts interact in some way. The aggregation is done *during* the elicitation session, not after. Behavioural aggregation requires the use of a facilitator and relies on the experts being willing to listen to differing opinions and to compromise.

3.4 Avoiding bias

It is important for the elicitation to reflect as accurately as possible the real knowledge and uncertainty of experts in the field. Any deviation of the elicited probability distribution from this ideal can be seen as a bias, and can lead to erroneous advice and decisions in the planning of future resources. Avoidance of bias is therefore a prime consideration in devising and carrying out an elicitation protocol. Bias can arise in many ways. Psychologists have shown that some kinds of short cuts in making judgements, known as heuristics, can lead to biases. Certain heuristics, such as availability bias or anchoring and adjustment, are common sources of bias in elicitation.

Availability bias arises when an expert gives undue weight to more recent events, or events that happened to the expert personally, when making judgements. This is simply because they are more memorable. To avoid availability bias it is important to make sure that an expert reviews all relevant evidence so that it is equally ‘available’ and memorable.

Anchoring and adjustment arises when an expert uses an earlier judgement, consciously or unconsciously, as a starting point for a subsequent judgement. The psychological heuristic in this case is that the expert evaluates the second judgement by moving from the starting point of the first value, and, in practice, they do not move far enough. The first value is seen as an ‘anchor’ holding the expert back from fully evaluating the second. To avoid the anchoring bias it is important to devise questions so as not to put values in the discussion that might serve as anchors.

Bias can also arise if we allow the facilitator’s opinions or preconceptions to influence the experts. The facilitator’s role is to manage the elicitation process, and care should be taken not to bias the outcome.

4. Developing an elicitation framework

4.1 Experiences with the Delphi method

The Delphi method is a systematic consensus process for collecting and refining the knowledge of a group of experts (Linstone & Turoff, 2002). It is an iterative process where experts provide their individual estimates for the parameters under consideration, together with their reasoning. The facilitator then anonymously shares these answers and their reasoning with all the other participants in the exercise. Participants then have the opportunity to revise their estimates over two or more rounds.

The CfWI has considerable experience in the use of online questionnaires to conduct Delphi exercises. We have successfully used this approach in a number of projects, including the review of medical and dental student intakes (MDSI) (Department of Health, 2012). For MDSI, we used the Delphi method to obtain the values of key parameters for modelling, including future health demand, workforce retirement age, and the ratio of full to part-time working. Delphi participants estimated these parameters across four future scenarios.

However, on other occasions, we have needed to reappraise the value of Delphi because the number of required values has been high. For example, where we have segmented the workforce by gender and professional groupings, the number of Delphi questions could exceed one hundred. This is too many. It risks experts disengaging, forgetting their previous responses, and the quality of their answers suffering as a result.

There are other concerns with the Delphi method, including:

- The time taken for the process can be considerable, from initially sending out the questionnaire to having it completed over several rounds.
- Running multiple rounds increases the risk of experts not responding and therefore dropping out. This affects the quality of the consensus reached.
- There is no open conversation or debate, as the group never meets, so it is impossible to freely share ideas and assess people's expertise. However, this may be an advantage, as powerful individuals cannot dominate the group.
- Expert responses are limited to short statements, resulting in loss of complexity of the issues, themes and concepts.
- The anonymity of the process means it is difficult to find out about dissenting answers.
- Pooling of single opinions means that their significance is lost, and some may be highly significant.
- Experts may not reach consensus in the second round, or their consensus may be artificial, as they may want to end the process.
- The classical Delphi method underestimates the uncertainty, as we are only looking at the distribution of the medians.
- If not conducted carefully, the expert panel may raise criticisms of the overall Delphi process, therefore losing faith in the method.

4.2 Developing a new approach

Having examined different approaches towards elicitation, the CfWI looked at two particular methods in more detail. These were the SHeffield ELicitation Framework (SHELF) and Cooke's method (EFSA, 2014

Cooke's method of elicitation relies on seed questions (or questions that the facilitator knows the answers to, but the experts do not). When the experts give probability judgements about the seed variables, two things are tested. The first of these is how well these judgements fit to reality, in the sense that one half of all true values (known to the facilitator) fall below the experts' lower quartile, one quarter between the lower quartile and the median, and so on.

We can score the experts on how closely they fit this ideal of calibration. We can also score them on how informative their judgements are. The closer together the quartiles are the more informative is the judgement. The final weighting of experts uses both these scores, although calibration is more important than 'informativeness'.

There are two difficulties with Cooke's method. Firstly, it is very difficult to find seed questions where we know the answer *and* that are sufficiently close to the elicitation questions, where we do not know the answers. If we ask experts questions that are substantially different to the elicitation questions, the weightings will not be valid.

Secondly, as the CfWI operates in the highly political environment of health, public health and social care, and workshop participants are selected to ensure the proceedings are seen as credible, ranking experts would not be politically acceptable to either the experts or the wider stakeholder group. For these reasons, Cooke's method for elicitation was not tested.

The other approach we investigated was SHELF, a package of materials for elicitation, created by Tony O'Hagan and Jeremy Oakley at the University of Sheffield (see www.tonyohagan.co.uk/shelf/ for further information).

It is designed for eliciting the knowledge of a group of experts in a face-to-face workshop, and capturing a probability distribution to represent their judgements. The method uses behavioural aggregation, where the experts interact to achieve consensus. This requires a well-defined protocol for managing the interaction to ensure that individual and group biases do not detract from the benefits of pooling knowledge and sharing multiple perspectives. The presence of a facilitator is essential, and it is helpful to have a recorder to support them. The process of holding a SHELF elicitation workshop is described below.

Preparation for the workshop

Good workshop preparation is critical to success. It includes creating clear definitions, for not only the parameters that we require elicited, but also for the selection of the group of experts. We need to make a decision on the number of workshops, since experts are asked to discuss their judgements fully and this can take a considerable time. Consequently, only a small number of parameters can be elicited in a single workshop. A workshop may take one or more days; three is the effective limit before tiredness sets in. The number of experts should be no more than six to eight. Fewer experts risks having insufficient knowledge to pool, but more experts risks discussions taking longer than necessary for the process. The experts are briefed prior to the workshop, specifying the problem and sharing any relevant information.

Context

SHELF has a set of forms to guide the elicitor through a well-structured protocol. The SHELF1 form sets the context and captures details of the session including:

- elicitation title, session, date and start time
- participants and their roles
- purpose of the elicitation
- formal statement of the protocol
- orientation and training (only completed after training has been conducted)
- participants' expertise and declaration of interests
- strengths and weaknesses of the group
- review of evidence
- structuring of the elicitation questions – review and opportunity for the experts to revise the questions
- definitions
- end time and supporting documents.

Training may be given before the workshop, or as part of the workshop. The training question follows the same protocol as that used for the formal elicitation, and is as described in the next section.

Elicitation

The elicitation process uses the SHELF2 form for each parameter of interest. A two-stage process is used. First, the experts are asked to make individual judgements, and then these are discussed to provide an aggregate probability distribution. The steps in the process are as follows:

- elicitation title, session, date and start time (same as for SHELF1)
- definition of the parameter to be elicited (referred to as X)
- evidence
- plausible range
- median
- upper and lower quartiles
- fitting
- group elicitation
- fitting and feedback
- chosen distribution
- discussion
- end time and supporting documents (or next elicitation).

The detailed process is as follows. Experts are first asked for the plausible upper (U) and lower (L) limits of X. These are the limits that would leave the experts extremely surprised if X were not somewhere between the limits. The experts first write down their own limits, without discussion. The elicitor then suggests setting the plausible range to the largest U value and lowest L value. The experts' individual values are otherwise not revealed. The range is discussed and may be revised, for example, if it is unfeasibly wide. This discussion of the range is done first to avoid anchoring on a central value.

The next steps (median, and upper and lower quartiles) are done individually without discussion. The first step is to ask for the median (M), giving careful guidance if the group is not experienced to avoid simply setting M equidistant between L and U. Experts are then asked to make individual judgements of the lower (Q1) and upper quartiles (Q3). Again, it is important to give guidance. For example, Q1 and Q3 should in general be closer to M (the most likely value) than L or U (the least likely).

Once the experts have written down M, Q1 and Q3, they are presented (anonymously) to the facilitator and recorded on the SHELF2 form. Individual probability distributions are then fitted to the experts' judgements.

The group elicitation starts with sharing and discussing the individual distributions. This is a critical part of the process. The experts need to discuss their differences, understand the various interpretations of the evidence, and exchange information. Careful guidance is needed by the facilitator to ensure that all views are shared, and that the conversation is not dominated by a few individuals. This discussion is recorded. During the discussion some experts will change their opinions (which is necessary for consensus) but this should be as a result of gaining greater understanding and insight, rather than being pressurised by others.

Once the discussion is complete (no more useful points are being made), the facilitator asks the group for joint values of M, Q1 and Q3. These are the values that an intelligent and impartial person would give, having listened to the discussions and different opinions and arguments. In law, this would be the 'fair-minded and informed observer'. The facilitator then fits a distribution to these numbers, and the closeness to the individual distributions is discussed. It is important to discuss where individual distributions differ from the aggregated distribution, to confirm that the respective experts are happy that this is what an impartial observer would say, reflecting on all opinions and the evidence. Equally, the facilitator also has to be happy with the final distribution.

4.3 Expert elicitation framework

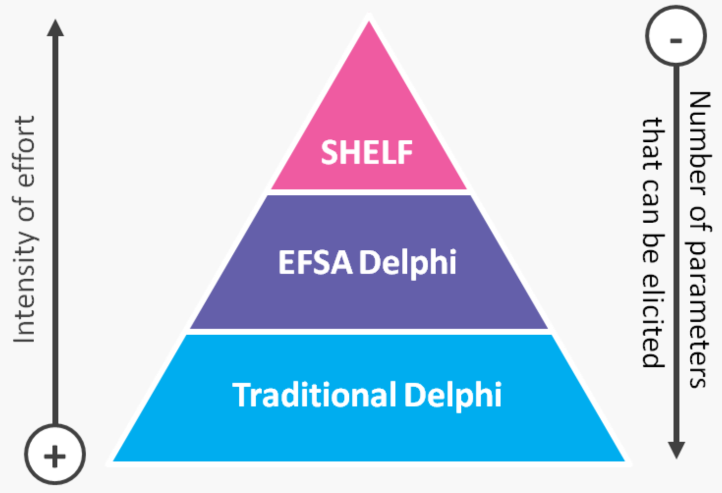
The CfWI has used the SHELF method for several elicitation exercises. The experts involved have commented favourably on the process. It has made them aware that quantifying the future values of parameters is very difficult, but that a carefully designed, formal and documented elicitation method is the best approach.

The SHELF approach was therefore selected as the ideal way to quantify uncertainty around parameter values in CfWI analyses, but it is costly. A single elicitation workshop requires the attendance of typically four to eight experts for a full day, and it is only feasible to elicit probability distributions for two or three uncertain quantities in one workshop. A CfWI analysis with a range of scenarios to quantify may involve dozens or even hundreds of uncertain quantities. It is therefore not practical to use the 'gold standard' SHELF method for all of these.

We therefore adopt a framework in which three different levels of intensity are applied in three layers. This is described in Figure 5. It is important to recognise the compromises that have necessarily been made in this framework and are detailed further in this paper.

Figure 5: Expert elicitation framework

The three levels of the framework increase upwards in intensity of effort, and downwards in the number of parameters that can be elicited.



Source: Centre for Workforce Intelligence

The top layer is SHELF elicitation, involving the highest level of resource intensity. This is used for just a small number of quantities that have strategic importance in the model. After which, the experts who have participated in these workshops, and so have been trained to make the necessary probability judgements, are sent a questionnaire in which they are given a range of additional quantities to assess using the middle layer method.

The middle layer is an adaptation of the Delphi approach which is described in EFSA (2014). Whereas the traditional Delphi asks experts only to provide an estimate of the uncertain quantity, this method, which we refer to as EFSA Delphi, asks them for the same judgements as in SHELF. That is, they provide for each quantity lower and upper credible bounds, their median and quartiles.

This corresponds to the first phase of a SHELF elicitation, the assessment of individual judgements. In place of the second phase – the group judgements – we have a Delphi iteration process. The experts' judgements and rationales are relayed anonymously back to the experts and they are asked to provide revised judgements. After one or two Delphi iteration rounds, the experts' individual probability distributions are averaged to provide the final aggregate distribution.

EFSA Delphi is a feasible approach because it is used only after the experts have learnt how to make the required judgements in a full SHELF workshop. In practice, they will typically be asked in the same questionnaire to make simple estimates of the remaining quantities, thereby combining the second and third layers. However, EFSA Delphi is less desirable than the SHELF approach because it does not allow the same degree of expert interaction, and because it is not moderated by a facilitator who can intercept unrealistic judgements and correct misunderstandings. It is important that the quantities chosen for EFSA Delphi are representative of the full range of quantities to be estimated using traditional Delphi.

EFSA Delphi is still a more intensive exercise than traditional Delphi, and would still be impractical to use for large numbers of quantities. Accordingly, the bottom layer of the framework is a traditional Delphi method, used for the remaining quantities. The traditional Delphi suffers from the same difficulties already mentioned, particularly the fact that it does not elicit any measure of experts' confidence or uncertainty. This is addressed in the expert elicitation framework by assuming that the uncertainty around these quantities is similar to that which the experts specify in the SHELF and EFSA Delphi methods for comparable quantities.

4.4 The framework in practice

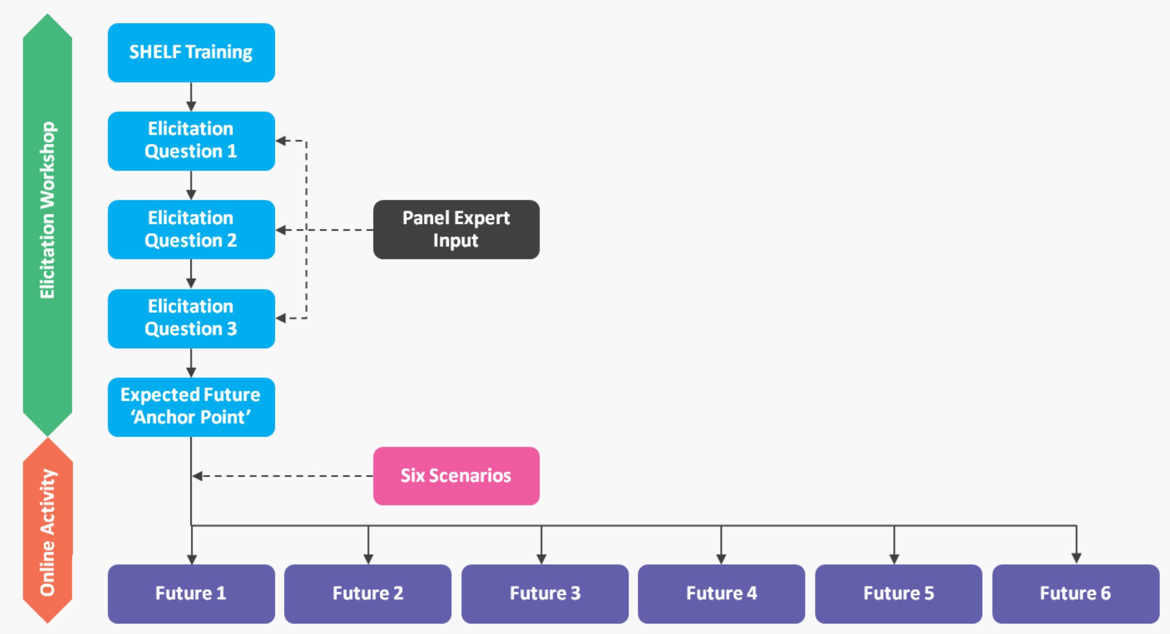
Recent SHELF exercises have elicited the following distributions:

- expected future workforce productivity
- expected workforce effort to meet demand (excluding population growth) as a result of long-term physical conditions
- expected workforce effort to meet demand (excluding population growth) as a result of long-term mental health conditions

The distributions elicited were for the *expected* future. We know from our work on scenarios that the most likely future will almost certainly never happen. However, the expected future is useful for reducing bias when eliciting values for the scenarios, as experts can be asked to give estimates relative to the expected future. This avoids direct comparisons between scenarios. Figure 6 describes the process used.

Figure 6: Elicitation framework in practice

The expected future is used as a reference point for eliciting values for the six scenarios below, using EFSA Delphi. Note that only the top two layers of the framework are used.



Source: Centre for Workforce Intelligence

5. Next steps

5.1 Alternative SHELF model

We have been using the SHELF quartile method in all elicitations so far. SHELF offers two other principal methods, the tertile (dividing the distribution into three equal parts) and roulette (dividing the distribution into a number of equally sized bins) methods. It would be relatively easy to adapt the existing approach, including the EFSA Delphi method. We initially plan to test the tertile method, and review the findings before deciding if we should investigate the roulette method.

5.2 Correlations

Eliciting beliefs about quantities that are not independent is a long-standing challenge in elicitation methodology. Unfortunately, correlated quantities can make material differences when input to models, compared with assuming independence. This is something that we need to address in order to meet our aspirations for a robust methodology.

Consider two quantities, X and Y. We can always elicit marginal distributions for these by the existing method(s). The question is: how to get a measure of correlation, and then how to turn that into a suitable joint distribution. There are a number of possible additional questions that we could ask the experts in order to get a handle on correlation but little experimental evidence about whether, and if so how well, they work. We plan to investigate this area further.

5.3 The three layers

We have a framework that comprises three layers – SHELF, EFSA Delphi and traditional Delphi. An issue that we have not fully considered is how to turn traditional Delphi estimates into elicited distributions. One approach would be to use a suitable comparable quantity whose distribution has been elicited by either SHELF or EFSA Delphi. However, there may be several candidates for this comparable quantity, so we need to look at some examples from the work so far in order to evaluate whether there is a problem, and if so how to address it.

6. References

CfWI (2013). *A strategic review of the future pharmacist workforce: informing pharmacist student intakes.* Centre for Workforce Intelligence. [online]. Available: http://www.cfwi.org.uk/publications/a-strategic-review-of-the-future-pharmacist-workforce [Accessed October 2014].

CfWI (2014a). *Horizon scanning: Analysis of key forces and factors*. Centre for Workforce Intelligence. Technical Paper Series No. 0006. [online]. Available at: http://www.cfwi.org.uk/our-work/research-development/cfwi-technical-paper-series [Accessed October 2014].

CfWI (2014b). *Scenario generation: Enhancing scenario generation and quantification.* Centre for Workforce Intelligence. Technical Paper Series No. 0007. [online]. Available at: http://www.cfwi.org.uk/our-work/research-development/cfwi-technical-paper-series [Accessed October 2014].

CfWI (2014c). *Developing robust system dynamics based workforce models: A best practice approach*. Centre for Workforce Intelligence. Technical Paper Series No. 0008. [online]. Available at: http://www.cfwi.org.uk/our-work/research-development/cfwi-technical-paper-series [Accessed October 2014].

CfWI (2014d). *Policy analysis: Applying robust decision-making to the workforce planning framework.* Centre for Workforce Intelligence. Technical Paper Series No.0009 [online]. Available at http://www.cfwi.org.uk/our-work/research-development/cfwi-technical-paper-series [Accessed October 2014].

Department of Health (2012). *Review of Medical and Dental Student Intakes in England.* [online]. Available at: https://www.gov.uk/government/publications/planning-medical-and-dental-school-intakes-for-the-future-needs-of-the-nhs [Accessed October 2014]

EFSA (European Food Safety Authority) (2014). *Guidance on Expert Knowledge Elicitation in Food and Feed Safety Risk Assessment*. EFSA Journal 2014, Vol.12, No.6: 3734.

Linstone, H.A. and Turoff, M. (2002). *The Delphi Method: Techniques and Applications*. [online]. Available at: http://is.njit.edu/pubs/delphibook/ [Accessed October 2014].

US Environmental Protection Agency (2011). *Expert Elicitation Task Force White Paper*. [online]. Available at: http://www.epa.gov/stpc/pdfs/ee-white-paper-final.pdf [Accessed October 2014]

Wallsten, T.S., Budescu, D.V., Rapoport, A., Zwick, R., and Forsyth, B.H. (1986). *Measuring the vague meanings of probability terms*. Journal of Experimental Psychology: General, Vol.115, pp.348-365.

Disclaimer

The Centre for Workforce Intelligence (**CfWI**) is an independent agency working on specific projects for the Department of Health and is an operating unit within Mouchel Management Consulting Limited.

This report is prepared solely for the Department of Health by Mouchel Management Consulting Limited, in its role as operator of the CfWI, for the purpose identified in the report. It may not be used or relied on by any other person, or by the Department of Health in relation to any other matters not covered specifically by the scope of this report.

Mouchel Management Consulting Ltd has exercised reasonable skill, care and diligence in the compilation of the report and Mouchel Management Consulting Ltd only liability shall be to the Department of Health and only to the extent that it has failed to exercise reasonable skill, care and diligence. Any publication or public dissemination of this report, including the publication of the report on the CfWI website or otherwise, is for information purposes only and cannot be relied upon by any other person.

In producing the report, Mouchel Management Consulting Ltd obtains and uses information and data from third party sources and cannot guarantee the accuracy of such data. The report also contains projections, which are subjective in nature and constitute Mouchel Management Consulting Ltd's opinion as to likely future trends or events based on i) the information known to Mouchel Management Consulting Ltd at the time the report was prepared; and ii) the data that it has collected from third parties.

Other than exercising reasonable skill, care and diligence in the preparation of this report, Mouchel Management Consulting Ltd does not provide any other warranty whatsoever in relation to the report, whether express or implied, including in relation to the accuracy of any third party data used by Mouchel Management Consulting Ltd in the report and in relation to the accuracy, completeness or fitness for any particular purposes of any projections contained within the report.

Mouchel Management Consulting Ltd shall not be liable to any person in contract, tort (including negligence), or otherwise for any damage or loss whatsoever which may arise either directly or indirectly, including in relation to any errors in forecasts, speculations or analyses, or in relation to the use of third party information or data in this report. For the avoidance of doubt, nothing in this disclaimer shall be construed so as to exclude Mouchel Management Consulting Ltd’s liability for fraud or fraudulent misrepresentation.

The CfWI produces quality intelligence to inform better workforce planning that improves people's lives

Centre for Workforce Intelligence
209-215 Blackfriars Road
London SE1 8NL United Kingdom

T +44 (0)20 7803 2707
E enquiries@cfwi.org.uk

www.cfwi.org.uk