Structured expert elicitation

1. Description of elicited parameters

We required outcomes for patients with Hospital Acquired Pneumonia (HAP), Ventilator Associated Pneumonia (VAP), and complicated urinary tract infections (cUTIs) caused by carbapenem-resistant gram negative bacteria. We were only interested in outcomes following microbiology-directed treatment for patients with an infection caused by *Enterobacterales* with an OXA-48 or MBL resistance mechanism, or *Pseudomonas* with a MBL resistance mechanism.

Outcomes were elicited depending on whether the infectious pathogen is susceptible to treatment. Therefore outcomes only depend on whether a patient is susceptible to treatment or not, and not to the specific treatment given. The outcomes we were interested in were 30-day mortality, length of stay in hospital, and the type of ward these patients would stay on in hospital.

As background information we provided experts with several related studies (see appendix 10). In these studies, infecting pathogens were not confirmed to be susceptible to the antibiotics administered (cefiderocol or CAZ-AVI); however, in our assessment, they are likely to have been susceptible.

For HAP, VAP and cUTI, both for susceptible and not-susceptable patients, the following questions were asked of experts:

Question 1. In this patient population, what proportion of patients will still be alive 30 days after starting microbiology directed treatment?

Question 2. In the patient population described at the top of the page, what will be the average length of stay?

Question 3. In the patient population described at the top of the page, what proportion of hospital stay would be spent on each of the following wards? This number should represent the average for all such patients, regardless of their outcome.

## Protocol for elicitation

The following sections describe the details of the elicitation exercise, according to the elements as described in the MRC elicitation guidance.

* 1. Selecting the quantities (preparation and design stage)

The choice of quantity considered the following three objectives:52 fitness for purpose; directly observable and homogeneity in the quantities elicited. Eliciting the same summaries throughout will reduce the burden of training.201

For question 1 the quantities elicited relate to the *proportion of patients with an event at a certain time.*  Question 2 relates to a continuous outcome, length of stay (LOS), which, in principle, can take values up to ∞. Question 3 relates to the proportionate split of LOS between the three types of ward – general ward, HDU and ICU. As the total proportion must sum to 100, these quantities were not elicited with uncertainty, and instead a mean proportion elicited.

* 1. Methods to encode judgements (preparation and design stage)

Either the Chips and Bins method or a Bisection method have been shown to work equally well in health care elicitation. The Chips and Bins approach however, is viewed as less complex and easier to complete by health care professionals, and so this method is used here.

Experts were first asked to express the range for their beliefs, the minimum, which is the value such that the experts believes that there is a 1% probability that the proportion is less than that value, and the maximum, a value, such that the experts believe that there is a 1% probability that the proportion is more than that value. Grids were then generated based on this range and experts were asked to place ‘chips’ on this grid to represent their beliefs.

* 1. Validation (preparation and design stage)

At the end of each task, experts were given a qualitative summary of their responses. If experts felt that these did not represent their views they were encouraged to revise their responses. Experts also had an opportunity to revise their responses following the feedback round (see below).

* 1. Selecting experts (preparation and design stage)

The models developed for this project span across HAP, VAP and cUTI and also relate to likely outcomes depending on susceptibility to treatment. Therefore there are multiple types of experts relevant for this task. Here we have included hospital consultants, microbiologists and pharmacists as experts. As part of the task, experts were asked to identify which of these disciplines they worked in. Experts were not expected to have any normative skills. Experts were recruited using recommendation from peers.

### Pilot exercise (preparation and design stage)

The wording of the questions was piloted for clarity and adequacy. The draft exercise was sent to a lead clinician and feedback sought. Following feedback the questions were modified, specifically the wording of the questions.

* 1. Training and preparation for experts (preparation and design stage)

A narrated power-point training session was delivered to experts prior to the task. The training session described the objectives of the elicitation exercise, clarified concepts such as uncertainty, familiarised the experts with the quantities elicited, described and explained the impact of bias and heuristics, and trained experts on the methods of elicitation used. A recorded version of the training slides was also sent to the experts following the session and also key details from this repeated in the task itself.

Experts were also reminded throughout the SEE that they were to elicit uncertainty on their estimate rather than thinking about variability across this heterogeneous group of patients

* 1. Level of elicitation (elicitation stage)

Each expert elicited their judgements individually without interaction with other experts. Eliciting judgements individually reduced the risk of estimates being biased by a subset of experts. In the SEE elicitation literature, there are concerns that experts may not feel confident in eliciting judgements individually, however, the experts in this SEE process elicited their beliefs on a condition that they encounter regularly in general practice. Concerns regarding individual level elicitation and lower confidence amongst experts generally arises when dealing with problems/technologies or conditions that are new or unknown to the experts.

* 1. Mode of administration (elicitation stage)

The elicitation exercise was administered via an application in SHINY. The task was delivered remotely, due to current restrictions on face to face meetings. Experts were offered the opportunity to complete the exercise remotely alongside one of the team. Email contacts were given to provide any support needed.

* 1. Feedback to experts and revision (elicitation stage)

Once experts expressed their beliefs and completed each question, they were presented with graphical feedback of what their estimates looked like. Experts were able to see how the grid looked once they have placed all of their chips on it. In addition, once experts had completed the grid, a summary of their answers was relayed to them. This provided the following information:

Your answers imply that (example quantities given)

* There is a 17% probability that the proportion of patients is between 19 and 20%
* There is a 50% probability that the proportion of patients is between 20 and 21%
* There is a 33% probability that the proportion of patients is between 21 and 22%

Following the individual elicitation beliefs were then aggregated using linear opinion pooling. This overall distribution was then relayed back to experts and they were given the opportunity to revise their own beliefs on the histograms they previously completed. This approach has been show to generated less biased parameters when the quantities elicited are unknown to the experts. Following this revision, expert’s beliefs were aggregated using the same approach, linear opinion pooling, and the final parameter values determined.

* 1. Opportunity for interaction (elicitation stage)

Given the individual level of elicitation that was chosen, there was no opportunity for interaction between the experts. The revision stage was done remotely so experts did not interact with each other.

* 1. Feedback from experts on process (elicitation stage)

Qualitative feedback on the elicitation process was collected from the experts, including rationales for their responses. This was collected during the task using free text boxes. This form of validation helps to highlight if experts understood the task and responded as best they could.

* 1. If/how to aggregate (aggregation, analysis and post-elicitation)

As an individual level of elicitation was chosen, mathematical aggregation was applied to generate the distributions, specifically linear opinion pooling using equal weighting of experts. First a probability distribution was fitted to each expert’s beliefs from the histogram and then these were pooled, assuming that each expert contributed equally to the group overall distribution.

This overall distribution was then relayed back to experts and they were given the opportunity to revise their own beliefs. Following this revision, expert’s beliefs were aggregated using the same approach, linear opinion pooling, and the final parameter values determined.

* 1. Fit to distribution (aggregation, analysis and post-elicitation)

A Beta distribution was fitted to expert’s distributions for question 1 as these relate to proportions. For question 2 a lognormal distribution was fitted. Question 3 only asked for point estimates so not fitting was required.

* 1. Data Protection and Anonymity (aggregation, analysis and post-elicitation)

Experts were asked to give their opinions individually (not in groups). The information provided, including personal details, is kept anonymous and confidential, stored securely and only accessed by those carrying out the study.