On quantifying expert opinion about multinomial models that contain covariates

Journal Item

On Quantifying Expert Opinion about Multinomial Models that Contain Covariates

Fadlalla G. Elfadaly

The Open University, Milton Keynes, UK
Cairo University, Giza, Egypt

Paul H. Garthwaite

The Open University, Milton Keynes, UK

Summary. This paper addresses the task of forming a prior distribution to represent expert opinion about a multinomial model that contains covariates. The task has not previously been addressed. We suppose the sampling model is a multinomial logistic regression and represent expert opinion about the regression coefficients by a multivariate normal distribution. This logistic-normal model gives a flexible prior distribution that can capture a broad variety of expert opinion. The challenge is to (i) find meaningful assessment tasks that an expert can perform and which should yield appropriate information to determine the values of parameters in the prior distribution, and (ii) develop theory for determining the parameter values from the assessments. A method is proposed that meets this challenge.
The method is implemented in interactive user-friendly software that is freely available. It provides a graphical interface that the expert uses to assess quartiles of sets of proportions and the method determines a mean vector and a positive-definite covariance matrix to represent the expert's opinions. The chosen assessment tasks yield parameter values that satisfy the usual laws of probability without the expert being aware of the constraints this imposes. Special attention is given to feedback that encourages the expert to consider his/her opinions from a different perspective. The method is illustrated in an example that shows its viability and usefulness.

Keywords: Elicitation method; Interactive graphical software; Logistic normal prior; Multinomial logit model; Multinomial logistic model; Prior distribution

1. Introduction

The purpose of an elicitation method is to help an expert quantify his or her opinions in a mathematically useful form. The expert performs meaningful assessment tasks and a probability distribution is formed from the elicited assessments. In a Bayesian analysis this distribution would be the prior distribution, but the distribution might also be used to express uncertainty about parameters in a decision analysis, or to communicate an expert's opinion succinctly. Different elicitation methods are required for different situations depending, in particular, on the sampling model and the model used to represent opinion. Here we consider the important case where the sampling model is a multinomial model with covariates.

Multinomial models arise when there is a set of complementary and mutually exclusive categories and each observation falls into one of these categories. In the simplest case, the probability of falling into any specified category is the same for each observation and observations are independent of each other. Then observations follow a multinomial distribution with, say, probability p_i that an observation falls in the ith category. A number of papers have addressed the task of quantifying an expert's opinions about the p_i so as to form a prior distribution for this situation. Early work focused on representing expert opinion by a Dirichlet distribution, which is the natural conjugate prior distribution for multinomial sampling; reviews of this work may be found in Garthwaite *et al.* (2005) and O'Hagan *et al.* (2006). Modelling opinion by a Dirichlet distribution has continued to attract attention (Zapata-Vázquez *et al.*, 2014; Evans *et al.*, 2017), but methods for eliciting more flexible prior distributions have also been proposed. Elfadaly and Garthwaite (2013) give a method for quantifying expert opinion as both a Dirichlet distribution and a Connor-Mosimann distribution, and give an example to illustrate the greater flexibility of the latter. More flexible models have also been obtained through the use of copulas. Elfadaly and Garthwaite (2017) use a Gaussian copula prior distribution to model expert opinion and Wilson (2018) uses a D-vine copula. With the latter copula, specification of the multivariate prior distribution can be separated into assessments about univariate marginal distributions and unconditional bivariate copulas, thus simplifying the assessment tasks that the expert performs (Wilson, 2018). Recent applications in which experts have assessed prior distributions for a multinomial model include Németh *et al.* (2017), who quantify the opinions of three experts about treatments for schizophrenia, representing their opinions by Dirichlet distributions, and Wilson *et al.* (2017), who elicited expert opinion about disease progression of untreated melanoma, representing opinions by Connor-Mosimann distributions.

Previous work has not addressed the task of eliciting a prior distribution for the common situation where covariates influence the values of the p_i (the multinomial probabilities), other than in the special case where there are only two categories and the sampling model reduces to logistic regression (Al-Labadi *et al.*, 2018; Bedrick *et al.*, 1996, 1997; Garthwaite *et al.*, 2013; Tsutakawa and Lin, 1986). Here we develop an elicitation method for quantifying opinion about a multinomial model when there are covariates and any number of categories. We take a multinomial logistic regression as the sampling model, which is the most common form of multinomial model with covariates. A linear regression links the covariates to functions of the multinomial probabilities (Agresti, 2002).

While the Dirichlet distribution is the standard prior distribution for a multinomial sampling model when there are no covariates, using the logistic normal distribution as the prior distribution has also been considered (see, for example, O'Hagan and Forster (2004), Sections 12.14 to 12.19). However, ways of choosing its parameters to model an expert's opinions have not been proposed and in this paper we first develop an elicitation method to rectify this deficiency. It does not seem to have been suggested before, but it is easy to expand the scope of the logistic normal prior so that it is a suitable prior distribution for a multinomial logistic regression model. We extend our elicitation method so it elicits the parameters of the resulting logistic normal prior.

Eliciting parameters of multivariate distributions is not, in general, an easy task, especially if variates are not independent (O'Hagan *et al.*, 2006). The assessed subjective distribution must satisfy the usual laws of probability and, with multinomial models, these include additional requirements because the probabilities of the different categories must sum to one. The elicitation method proposed here uses assessment tasks and a task structure that leads to coherent assessments without the expert having to be conscious of coherence constraints. The method has been implemented in interactive graphical user-friendly software that aims to help the expert quantify his/her opinion effectively and as painlessly as possible. This software is free and is available on the web at `http://statistics.open.ac.uk/elicitation`. It has an online *Help* facility and its output includes files for running prior-posterior analysis in standard Bayesian packages.

Use of the elicitation method is illustrated through an application that concerns the potential cost to Malta of a new drug for diabetes. In Malta, the standard treatments for diabetes could be classified into four groups, as Metformin, Sulfonylureas, Metformin plus Sulfonylureas, and Insulin. A fifth new treatment, Gliptins, is to be introduced. However, this will increase the overall cost of treating diabetes, as the new treatment is more expensive per patient. In order to estimate the future cost, an expert quantified his opinion about the proportions of patients in Malta who would be on each treatment. The diagrams we use in describing the new elicitation method are screen shots from this example.

In Section 2 we consider multinomial models that do not contain covariates; we define the logistic normal prior model and consider its assumptions. The required assessments and their use to elicit the prior and obtain feedback is given in Section 3. In Section 4 we consider the case where the sampling model does contain covariates and extend our elicitation method to model opinion about a logistic normal distribution. Section 5 illustrates application of the method through the example concerning diabetes treatments. Some concluding comments are made in Section 6.

2. Elicitation for a multinomial sampling model

The flexibility of a logistic normal distribution makes it an attractive means of representing expert opinion about a multinomial sampling model. Here we quantify opinion as an *additive* logistic normal prior distribution, which is the form most widely used [see, for example, Aitchison (1986)]. We suppose there are k categories and that each observation in a sample belongs to exactly one category. Under a multinomial sampling model, observations are independent of each other and the probability that an observation belongs to any specified category is the same for each observation.

Let p_i denote the probability that an observation is in the ith category ($i = 1, \ldots, k$) and put $\mathbf{p} = (p_1, \ldots, p_k)$. For $i = 2, \ldots, k$, define r_i and Y_i by

$$r_i = p_i/p_1 \tag{1}$$

and

$$Y_i = \log(p_i/p_1), \tag{2}$$

so r_i is the probability ratio of the ith category relative to the first category and Y_i is the log probability ratio. From equation (2),

$$p_1 = \frac{1}{1+\sum_{j=2}^{k}\exp(Y_j)}, \qquad p_i = \frac{\exp(Y_i)}{1+\sum_{j=2}^{k}\exp(Y_j)}, \quad i = 2,\ldots,k. \tag{3}$$

Let $\boldsymbol{Y}_{k/1} = (Y_2,\ldots,Y_k)'$, where the first category is suppressed. Then equation (3) defines the *additive logistic transformation* from $\boldsymbol{Y}_{k/1}$ to $\mathbf{p}$. Also, from equation (3), it follows that $\sum_{i=1}^{k} p_i = 1$, which will be referred to as the *unit sum constraint*. The vector $\mathbf{p}$ is said to have a logistic normal distribution if

$$\boldsymbol{Y}_{k/1} \sim \text{MVN}(\boldsymbol{\mu}_{k/1}, \boldsymbol{\Sigma}_{k/1}) \tag{4}$$

and $\boldsymbol{\Sigma}_{k/1}$ is positive-definite. We assume the prior distribution takes this form.

In the above formulae, p_1 seems to be treated differently to the other p_i variables. However, the additive logistic normal distribution is permutation invariant. That is, whatever the ordering of the elements of the vector $\mathbf{p}$, the density function given above is invariant. For a theoretical proof of this property see Aitchison (1986). Hence, any order of the elements of $\mathbf{p}$ can be considered. With the representation given in equations (2)–(3), p_1 is referred to as the fill-up variable and the first category as the fill-up category. (In multinomial logistic regression, the standard choice for the fill-up category is the first or last category.) However, while in principle the choice of fill-up category is arbitrary and the assessed prior distribution should be invariant to its choice, in practice the assessed prior distribution would typically vary if the expert repeated the elicitation process but with a different category chosen as the fill-up category. We favour choosing the most common category as the fill-up (first) category, because we believe that choice makes the assessment tasks easier for the expert.

O'Hagan and Forster (2004, Sections 12.15 to 12.18) give the prior-posterior analysis for a $(k-1)\times 1$ vector Φ whose components are a set of log contrasts. Their results can be applied to the prior distribution that we elicit by equating Φ to $\boldsymbol{Y}_{k/1}$ and taking the distribution of $\boldsymbol{Y}_{k/1}$ as the prior distribution of Φ.

3. Elicitation method without covariates

An expert should be asked only meaningful questions, so he will not be asked directly about the $Y_i = \log(r_i) = \log(p_i/p_1)$, as the Y_i are not easily interpretable. (We will use he/his rather than she/her because the expert in the example we give is male.) However, his assessments must provide information about the Y_i. Our approach is to ask him to focus on two categories at a time, one of which is the fill-up category. Thus, for the ith category we ask the expert to consider all the items that fall in the ith category or the first category. The expert is then asked about the proportion of those items that fall in the ith category relative to the proportion that falls in the

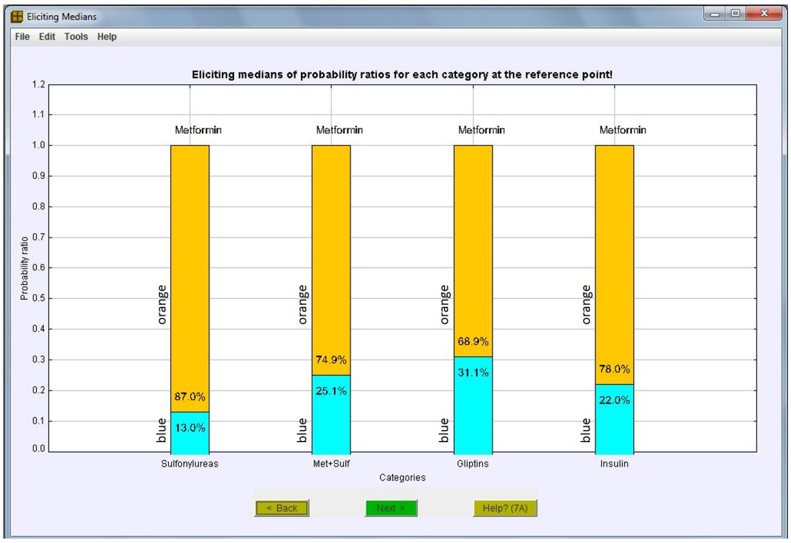

Fig. 1. Assessing medians of probability ratios. The proportion of a bar that is blue is the assessed median of r_i $(i = 2, \ldots, 5)$.

first category. He makes his assessments on a bar chart using a graphical interface (see Fig. 1) and may think about this proportion as the volume in the lower (blue) part of a box relative to the volume in the upper (orange) part of the box. (Labels naming colours have been added to figures in this paper to assist understanding with black-and-white copies.) That is, he focuses on r_i. Alternatively, if he finds it easier he can consider $p_i/(p_1 + p_i)$, the proportion of the complete bar that is blue. The expert assesses medians and quartiles (medians are being assessed in Fig. 1) and in both cases we treat his assessments as values of r_i.

As the log transformation from r_i to Y_i is strictly monotonic increasing, quantile assessments of r_i provide the corresponding quantiles of Y_i. We will make substantial use of this approach and have found that it gives a viable means of eliciting opinion about Y_i. Where possible, feedback is also given to the expert in the form of unconditional marginal quantiles of each probability p_i $(i = 1, \ldots, k)$, so as to improve the quality of the elicitation method (*cf.* Section 3.3).

3.1. Eliciting the mean vector $\boldsymbol{\mu}_{k/1}$ using median assessments

For $i = 2, \ldots, k$, let m_i^* denote the median of $r_i = p_i/p_1$ and let m_i denote the median of Y_i. (With this notation, deletion of an asterisk 'transforms' a median of r_i to a median of Y_i.) We use $C(\,.\,)$ to denote the centre (median) of the quantity in braces.

Here we relate $\boldsymbol{\mu}_{k/1} = (\mu_2, \ldots, \mu_k)'$ to the m_i^*. Using the symmetry of the distribution of Y_i $(i = 2, \ldots, k)$, together with equations (1)–(4), we have that $\mu_i = E(Y_i) = C(Y_i) = C[\log(r_i)]$. Hence, the components of $\boldsymbol{\mu}_{k/1}$ are given by

$$\mu_i = \log(m_i^*) \quad \text{for } i = 2, \ldots, k. \tag{5}$$

Using this approach, $\boldsymbol{\mu}_{k/1}$ is assessed in a simple direct way. The expert assesses unconditional medians, m_i^*, of r_i for each category. These assessments do not need to be constrained at all. However, based on these assessments, unconditional marginal medians of p_i that sum to 1 are approximated and presented as feedback to the expert who is invited to revise them until they form an acceptable representation of his opinion. More than one iteration is seldom needed. Our approach to approximate the marginal medians of p_i $(i = 1, \ldots, k)$ is discussed in Section 3.3.

The expert uses interactive graphics to give his median assessments, m_i^*, of the probability ratio r_i. The screen shot given earlier in Fig. 1 illustrates the process. The expert has assessed his median of the probability ratio for each of the four categories relative to the first category. He did this by clicking the mouse-pointer on each (blue/orange) stacked bar to divide it in two according to his median assessment for the ratio of p_i, represented by the lower (blue) area of each stacked bar, to p_1, which is represented by the upper (orange) area of each bar. Fig. 1 and all other figures in this paper are from the example reported in Section 5, in which an expert quantified his opinion about the potential cost of a new drug in Malta.

3.2. *Eliciting the variance matrix $\boldsymbol{\Sigma}_{k/1}$ using conditional assessments*

Kadane *et al.* (1980) give a method of eliciting the "spread" matrix of a vector-variate that has a multivariate-t distribution. We adopt their method with minor modifications and with the degrees of freedom (df) set equal to infinity, so that the multivariate-t distribution becomes an MVN distribution and the spread matrix becomes a variance-covariance matrix. Kadane *et al.* (1980) use $S(.)$ to denote the spread of a t distribution but here $S(.)$ will be the variance of a normal distribution.

Our procedure for determining $\boldsymbol{\Sigma}_{k/1} = S(\boldsymbol{Y}_{k/1})$ requires the following quantities.

1. $C(Y_i)$ for $i = 2, \ldots, k$.
2. $S(Y_2)$ and $S(Y_{i+1} \,|\, Y_2 = y_2^\diamond, \ldots, Y_i = y_i^\diamond)$ for $i = 2, \ldots, k-1$.
3. $C(Y_j \,|\, Y_2 = y_2^\diamond, \ldots, Y_{i-1} = y_{i-1}^\diamond, Y_i = y_i^0)$ for $i = 2, \ldots, k-1$; $j = i+1, \ldots, k$, where $y_2^\diamond, \ldots, y_{i-1}^\diamond$ and y_i^0 are chosen as detailed below.

The quantities in 3 provide information about the off-diagonal elements of $\boldsymbol{\Sigma}_{k/1}$ – the degree to which the conditions change the median assessment of Y_j reflects the correlation between Y_j and the conditioning variables.

The purpose of Section 3.1 was to elicit $\boldsymbol{\mu}_{k/1}$, whose ith component is $C(Y_i)$. Hence the quantities in 1 have already been assessed. The processes for assessing the other quantities are described next.

3.2.1. *Assessing $S(Y_2)$ and $S(Y_{i+1} \,|\, Y_2 = y_2^\diamond, \ldots, Y_i = y_i^\diamond)$*

To obtain $S(Y_2)$ the expert is asked to assess the lower and upper quartiles for the probability ratio $r_2 = p_2/p_1$, which we denote by L_2^* and U_2^*. Then the upper and lower quartiles of Y_2 are

$$L_2 = \log(L_2^*) \quad \text{and} \quad U_2 = \log(U_2^*), \tag{6}$$

since the log transformation from r_2 to Y_2 is monotonic strictly increasing. As the interquartile range of a standard normal distribution is 1.349, we put

$$S(Y_2) = \left[\frac{U_2 - L_2}{1.349}\right]^2. \tag{7}$$

So as to define the assessment tasks involving conditional spreads, values must be specified for $y_2^\diamond, \ldots, y_{k-1}^\diamond$. Kadane *et al.* (1980) set them equal to the expert's assessment of upper quartiles. We prefer to give them values from the middle of the expert's subjective distribution, when possible, because of the unit sum constraint. Otherwise, conditioning can imply that some categories have a tiny probability of occurring, even though the expert initially thought their occurrence is quite likely. For this reason we choose each $y_i^\diamond$ so that the corresponding condition on r_i is $r_i = m_i^*$. Hence, $y_i^\diamond = m_i$ for $i = 2, \ldots, k-1$. To shorten notation, we replace $S(Y_{i+1} \,|\, Y_2 = m_2, \ldots, Y_i = m_i)$ by $S(Y_{i+1} \,|\, Y_2, \ldots, Y_i)$ as, for any MVN distribution, the values actually specified for $Y_2, \ldots, Y_i$ do not affect $S(Y_{i+1} \,|\, Y_2, \ldots, Y_i)$.

To obtain $S(Y_{i+1} \,|\, Y_2, \ldots, Y_i)$ the expert assesses quartiles of r_{i+1} conditional on $r_j = m_j^*$ for $j = 2, \ldots, i$. Let L_{i+1}^* and U_{i+1}^* denote the expert's lower and upper quartile assessments. The log transformation from r_{i+1} to Y_{i+1} is monotonic strictly increasing. Also, $r_2, \ldots, r_i$ determine $Y_2, \ldots, Y_i$. Hence,

$$L_{i+1} = \log(L_{i+1}^*) \quad \text{and} \quad U_{i+1} = \log(U_{i+1}^*) \tag{8}$$

are the lower and upper quartiles of $Y_{i+1} \,|\, Y_2, \ldots, Y_i$. We put,

$$S(Y_{i+1} \,|\, Y_2, \ldots, Y_i) = \left[\frac{U_{i+1} - L_{i+1}}{1.349}\right]^2, \quad \text{for } i = 2, \ldots, k-1. \tag{9}$$

The interactive graphical interface is used to elicit the required quartile assessments. *Help* messages describe the meaning of quartiles to the expert and explain how to use the method of bisection to assess them. The expert first clicks the computer's mouse on the left-most stacked bar to assess L_2^* and U_2^*. Then, for each remaining r_{i+1} ($i = 2, \ldots, k-1$), he assesses quartiles L_{i+1}^* and U_{i+1}^* given that $r_2 = m_2^*, \ldots, r_i = m_i^*$. This is illustrated in Fig. 2, where the expert has assessed the two quartiles of r_4, represented by the two short (dark blue) horizontal lines on the third (blue/orange) bar, conditional on r_2 and r_3 having the median values given by the two leftmost (purple/orange) bars. To assist the expert, the software presents a separate graph showing the pdf curve of the lognormal distribution of $(r_{i+1}|r_2 = m_2^*, \ldots, r_i = m_i^*)$, for $i = 2, \cdots, k-1$; see Fig. 2. The expert can change his assessed conditional quartiles of r_{i+1} until the conditional pdf curve forms an acceptable representation of his opinion.

3.2.2. Assessing $C(Y_j \,|\, Y_2 = y_2^\diamond, \ldots, Y_{i-1} = y_{i-1}^\diamond, Y_i = y_i^0)$

The value we give to y_i^0 must cause the expert to revise his opinions, so it must differ from m_i by a reasonable amount. In particular y_i^0 must not equal $y_i^\diamond = m_i$, as in that case the conditional centre would exactly equal m_i. We set $y_i^0 = L_i$ for

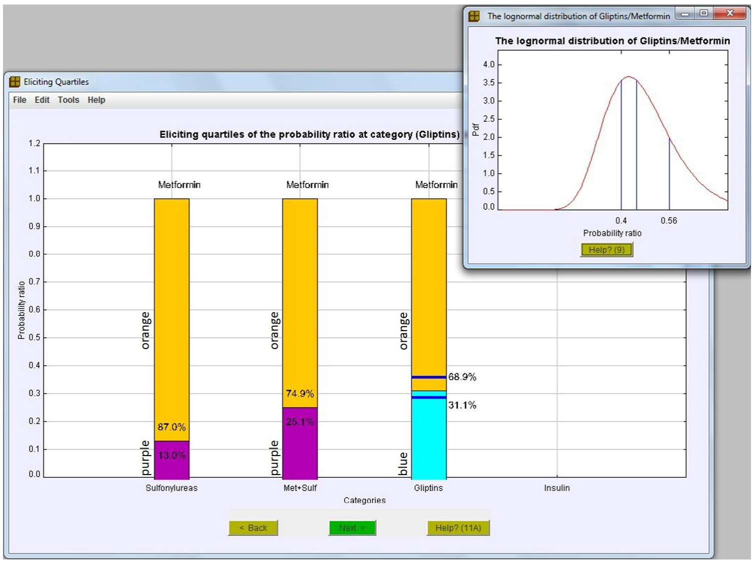

Fig. 2. Assessing conditional quartiles of probability ratios. The expert is asked to assume that the proportions in the first two (purple/orange) bars are correct. These give r_2 and r_3. Conditional on this assumption, the expert assesses quartiles L_4^* and U_4^* for the proportion in the third (blue/orange) bar.

$i = 2, \ldots, k-1$. The condition $Y_2 = L_2$ is equivalent to $r_2 = L_2^*$ and, the condition $Y_i = L_i$ is equivalent to $r_i = L_i^*$. The expert has earlier assessed $L_2^*, \ldots, L_{k-1}^*$, giving $L_2, \ldots, L_{k-1}$ from equations (6) and (8). Here the expert first assesses the median of $(r_j \,|\, r_2 = L_2^*)$ for $j = 3, \ldots, k$, and we let m_{2j}^* denote the assessments. He next assesses the median of $(r_j \,|\, r_2 = m_2^*, \ldots, r_{i-1} = m_{i-1}^*, r_i = L_i^*)$ for $i = 3, \ldots, k-1$, $j = i+1, \ldots, k$ and we let m_{ij}^* denote the assessments.

Let $C(Y_j \,|y_2^0)$ denote $C(Y_j \,|Y_2 = y_2^0)$ and let $C(Y_j \,|\, y_2^\diamond, \ldots, y_{i-1}^\diamond, y_i^0)$ denote $C(Y_j \,|\, Y_2 = y_2^\diamond, \ldots, Y_{i-1} = y_{i-1}^\diamond, Y_i = y_i^0)$. We have

$$C(Y_j \,|\, y_2^0) = \log(m_{2j}^*) \quad \text{for } j = 3, \ldots, k \tag{10}$$

and

$$C(Y_j \,|y_2^\diamond, \ldots, y_{i-1}^\diamond, y_i^0) = \log(m_{ij}^*) \quad \text{for } i = 3, \ldots, k-1;\; j = i+1, \ldots, k. \tag{11}$$

To implement this part of the elicitation process, the interactive program displays a bar-chart with a pair of stacked bars for each category (except for the first category), as illustrated in Fig. 3. The left-hand stacked bars show the expert's unconditional median assessments; the right-hand bars relate to the conditional assessments, some giving conditions and the remainder showing the expert's conditional assessments. The expert is first asked to assume that L_2^* is the actual value of r_2. This condition is specified by the right-hand stacked bar in the first pair of bars. Taking this information into account, he clicks the mouse on the right-hand bar in each remaining pair to re-assesses his medians of $r_3, \ldots, r_k$, giving $m_{22}^*, \ldots, m_{2k}^*$.

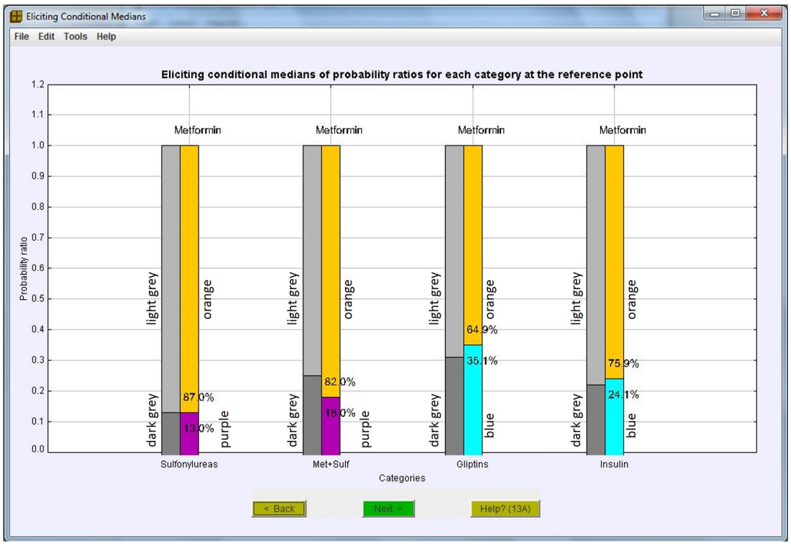

Fig. 3. Assessing conditional medians. The expert is asked to assume that the proportions in the first two (purple/orange) bars are correct; only the second of these differs from his unconditional median assessment. Conditional on these assumptions, the expert assesses medians m^*_{34} and m^*_{35} for the proportions in the third and fourth (blue/orange) bars.

These conditional median assessments, as with the unconditional median assessments m^*_i $(i = 2, \ldots, k)$, are not constrained and may take any positive value.

Then, in each successive step i (for $i = 3, \cdots, k-1$), the expert is asked to suppose that the true values of $r_2, \ldots, r_{i-1}$ and r_i are $m^*_2, \ldots, m^*_{i-1}$ and L^*_i, respectively. In Fig. 3, these conditions are shown by the right-hand (purple/orange) stacked bars in the first two categories. Given this information, the expert is asked to revise his earlier unconditional assessments of the medians of $r_{i+1}, \ldots, r_k$. His new medians are m^*_{ij} $(j = i+1, \ldots, k)$. The left-hand (dark-grey/light-grey) bars in Fig. 3 are his unconditional median assessments from Fig. 3 and they are displayed to aid the expert.

3.2.3. *Determining* $\boldsymbol{\Sigma}_{k/1}$

Conformably partition $\boldsymbol{\Sigma}_{k/1} = S(\mathbf{Y}_{k/1})$ as

$$\boldsymbol{\Sigma}_{k/1} = \begin{bmatrix} \boldsymbol{\Sigma}_{(k-1)/1} & \boldsymbol{\zeta} \\ \boldsymbol{\zeta}' & \sigma^2_k \end{bmatrix}, \tag{12}$$

where $\boldsymbol{\Sigma}_{(k-1)/1} = S(Y_2, \ldots, Y_{k-1})$. A slight variant of the method of Kadane *et al.* (1980) is used to estimate all the elements of $\boldsymbol{\Sigma}_{k/1}$. Differences between our method and that of Kadane *et al.* (1980) arise from the choices of $y^\diamond_2, \ldots, y^\diamond_{k-1}$. The required spreads and centres are obtained from the elicited assessments, using equations (7) and (9)–(11). Our procedure gives the desired property that $\boldsymbol{\Sigma}_{k/1}$ is certain to be positive-definite. Details are given in on-line supplementary materials.

3.3. *Feedback using marginal quartiles*

After initial hyperparameter values of the logistic normal distribution have been assessed, the expert is given feedback about his prior distribution through a bar-chart. This displays the unconditional median and unconditional quartiles of p_j ($j = 1, \ldots, k$) that are implied by his prior distribution. This form of feedback has two clear benefits.

(i) It encourages the expert to examine his assessed opinions from a different perspective. The feedback requires the expert to consider the probabilities p_j, rather than the probability ratios $r_j = p_j/p_1$.

(ii) Conditional assessment tasks are generally harder to perform than unconditional tasks. Other than for the probability ratio r_2, all assessments of upper and lower quartiles have been conditional assessments for the remaining probability ratios. The feedback asks the expert to consider *un*conditional assessments.

An example of the bar-chart is shown in Fig. 4, where the left-hand (grey) bar in each pair and the associated short (black) horizontal lines show the median and quartile values given as feedback. When the feedback is displayed, the expert is invited to change the median and/or quartiles of any categories he wants by adjusting the right-hand (blue) bars and the short (dark blue) horizontal lines.

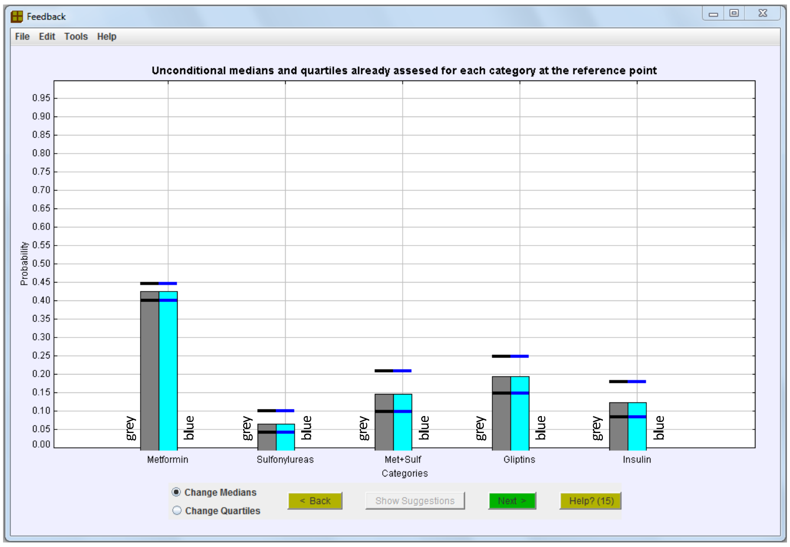

Fig. 4. Feedback giving marginal medians and quartiles of the p_j. The right-hand (blue) bars and the short (dark blue) horizontal lines should be adjusted by the expert if the feedback values in the left-hand (grey) bars are not a satisfactory representation of his opinions.

To add this feedback to the software, a method is required for estimating unconditional quartiles from the elicited hyperparameters $\boldsymbol{\mu}_{k/1}$ and $\boldsymbol{\Sigma}_{k/1}$. Unfortunately, a closed-form method for estimating the unconditional moments, or quartiles, of the logistic normal distribution does not exist. This is only a minor inconvenience though – we simply generate a large sample of, say, 100 000 random observations of $\boldsymbol{Y}_{k/1}$ from MVN$(\boldsymbol{\mu}_{k/1}, \boldsymbol{\Sigma}_{k/1})$. Each observation gives a value of $(p_1, \ldots, p_k)$, so the generated observations yield a random sample of each p_i $(i = 1, \ldots, k)$, from which quantiles of each p_i's marginal distribution can be calculated. Further detail is given in the supplementary materials.

If the expert revises the quantiles of the p_i, we also require a method of determining $\boldsymbol{\mu}_{k/1}$ and $\boldsymbol{\Sigma}_{k/1}$ from the new quantiles. This is a trickier problem and the procedure we construct is also described in the supplementary materials. An important feature of the procedure is that the different categories are treated symmetrically. In particular, no category has the distinction of being classified as the 'fill-up category'. To achieve this we define the $k \times (k-1)$ matrix $\mathbf{H}$ by

$$\mathbf{H} = \begin{bmatrix} -1/k & -1/k & -1/k & \ldots & -1/k \\ (k-1)/k & -1/k & -1/k & \ldots & -1/k \\ -1/k & (k-1)/k & -1/k & \ldots & -1/k \\ \vdots & & & & \\ -1/k & -1/k & -1/k & \ldots & (k-1)/k \end{bmatrix}.$$

That is, each element of the top row of $\mathbf{H}$ is $-1/k$, while its other $k-1$ rows form a square matrix whose diagonal elements are $(k-1)/k$ and whose non-diagonal elements are $-1/k$, We transform from $\boldsymbol{Y}_{k/1}$ to the $k \times 1$ vector $\boldsymbol{T} = (T_1, \ldots, T_k)'$ by putting $\boldsymbol{T} = \mathbf{H}\boldsymbol{Y}_{k/1}$. Then

$$T_i = \log(p_i) - \frac{1}{k}\sum_{j=1}^{k}\log(p_j) \qquad \text{for } i = 1, \ldots, k. \tag{13}$$

The distribution of $\boldsymbol{T}$ is a singular normal distribution $(\sum_{j=1}^{k} T_i = 0)$. We let $\boldsymbol{\Gamma} = (\gamma_1, \ldots, \gamma_k)' = \mathbf{H}\boldsymbol{\mu}_{k/1}$ denote its mean and write its variance matrix, $\mathbf{H}\boldsymbol{\Sigma}_{k/1}\mathbf{H}'$, as $\mathbf{A\Omega A}$, where $\mathbf{A}$ is a diagonal matrix with ith diagonal element $a_i = \{\mathrm{Var}(T_i)\}^{1/2}$ and $\boldsymbol{\Omega}$ is the correlation matrix of $\boldsymbol{T}$. Aitchison (1986) refers to $\mathbf{A\Omega A}$ as the *centred logratio covariance matrix.*

In modifying the prior distribution to reflect an expert's re-assessments of quantiles of $p_1, \ldots, p_k$, we treat $\boldsymbol{\Omega}$ as fixed and only revise the estimates of $\mathbf{A}$ and $\boldsymbol{\Gamma}$. It is because we want to take the correlation matrix as fixed that we must transform to variables that treat the different categories symmetrically. From equation (13), revision of the quantiles of p_i primarily affect γ_i and a_i.

If the expert revises quantiles of $p_1, \ldots, p_k$, we determine $\boldsymbol{\Gamma}$ and $\mathbf{A}$ (using the method described in the supplementary materials), re-estimate $\boldsymbol{\mu}_{k/1}$ and $\boldsymbol{\Sigma}_{k/1}$, and generate a large sample of observations from MVN$(\boldsymbol{\mu}_{k/1}, \boldsymbol{\Sigma}_{k/1})$. From the sample, we calculate quantiles of the p_i that mimic any asymmetry in the expert's previous quantile assessments. The quantiles are unlikely to exactly match the expert's previous assessments, as the expert gave more assessments than the number of parameters, so his assessments will seldom satisfy the requirements for statistical coherence. The expert is invited to revise the suggested quantiles and the cycle is repeated until the suggested quantiles form an acceptable representation of his assessments. On completion of the elicitation, the software outputs the elicited hyperparameters of the logistic normal prior distribution, $\boldsymbol{\mu}_{k/1}$ and $\boldsymbol{\Sigma}_{k/1}$, in a suitable format for further Bayesian analysis.

4. Elicitation method with covariates

We now turn to the situation that motivated the work in this paper, where the multinomial sampling model contains one or more covariates that influence the membership probabilities of different categories. We suppose the sampling model is a multinomial logistic regression, which is a generalisation of logistic regression. In Section 4.1 we define the multinomial logistic model and also the form of our prior distribution for its parameters, which is a multivariate normal distribution.

In the context of ordinary logistic regression (when there are only two categories), there are disadvantages in using diffuse normal distributions as the prior distributions for the regression coefficients (see, for example, Al-Labadi *et al.*, 2018). In particular, the probabilities of each category are usually the quantities of primary interest and the resulting marginal prior distributions of these probabilities will be concentrated at 0 and 1. This would also happen with a multinomial logistic model if diffuse priors were used, giving further motivation for eliciting an informative prior distribution for its regression coefficients.

Our method of quantifying opinion to obtain the parameters of the prior distribution is given in Sections 4.2 and 4.3. The method makes extensive use of the elicitation procedure developed in Section 3, repeatedly using the procedure to quantify opinion when the covariates take specified values, for a variety of values.

4.1. Sampling and prior models

If a covariate is a factor with i levels, we assume that it has been re-expressed as $i-1$ dummy 0/1 variables: these all equal 0 for the first level of the factor and a different one of them equals 1 for each other factor level. We suppose the dummy variables and continuous covariates give m variables in total. Let $\mathbf{x} = (x_1, \ldots, x_m)'$ be the vector that they form and let $p_i(\mathbf{x})$ denote the probability that an observation with covariate values $\mathbf{x}$ falls in the ith category ($i = 1, \ldots, k$). The sampling model is obtained by setting Y_i equal to $\alpha_i + \mathbf{x}'\boldsymbol{\beta}_i$ in equation (3), where α_i and $\boldsymbol{\beta}_i = (\beta_{1,i}, \ldots, \beta_{m,i})'$ are the constant and vector of regression coefficients for the ith category ($i = 2, \ldots, k$). This gives

$$p_i(\mathbf{x}) = \begin{cases} \dfrac{1}{1 + \sum_{j=2}^{k} \exp(\alpha_j + \mathbf{x}'\boldsymbol{\beta}_j)}, & i = 1 \\ \\ \dfrac{\exp(\alpha_i + \mathbf{x}'\boldsymbol{\beta}_i)}{1 + \sum_{j=2}^{k} \exp(\alpha_j + \mathbf{x}'\boldsymbol{\beta}_j)}, & i = 2, \ldots, k, \end{cases} \tag{14}$$

which defines the multinomial logistic (logit) model. To simplify the notation, we henceforth denote $p_i(\mathbf{x})$ by p_i. As defined earlier, $\boldsymbol{Y}_{k/1} = (Y_2, \ldots, Y_k)'$, where $Y_i = \log(p_i/p_1)$ for $i = 2, \ldots, k$. While the prior distribution for the α_i and $\boldsymbol{\beta}_i$ $(i = 2, \ldots, k)$ is a multivariate normal distribution, the prior model for $(p_1, \ldots, p_k)$ is referred to as a logistic normal distribution.

It is convenient to rearrange the regression coefficients into a matrix, say $\mathbf{B}$, of the form

$$\mathbf{B} = \left[\begin{pmatrix} \alpha_2 \\ \boldsymbol{\beta}_2 \end{pmatrix}, \cdots, \begin{pmatrix} \alpha_k \\ \boldsymbol{\beta}_k \end{pmatrix} \right], \tag{15}$$

because we will be focusing on one row of $\mathbf{B}$ at a time. We define the new set of vectors $\boldsymbol{\alpha}$, $\boldsymbol{\beta}_{(h)}$, for $h = 1, \ldots, m$, as the rows of $\mathbf{B}$, so that

$$\boldsymbol{\alpha} = (\alpha_2,\, \alpha_3, \cdots,\, \alpha_k)', \tag{16}$$

$$\boldsymbol{\beta}_{(h)} = (\beta_{h,2},\, \beta_{h,3}, \cdots,\, \beta_{h,k})'. \tag{17}$$

We assume that the prior distribution for $(\boldsymbol{\alpha}', \boldsymbol{\beta}'_{(1)}, \ldots, \boldsymbol{\beta}'_{(m)})'$ is a multivariate normal distribution, $\mathrm{MVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. The task is to elicit $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$.

To make the elicitation problem manageable we shall assume that, given the value of $\boldsymbol{\alpha}$, the vectors $\boldsymbol{\beta}_{(h)}$ and $\boldsymbol{\beta}_{(g)}$ are *a priori* conditionally independent for all h and g $(h \neq g)$. Thus, $\boldsymbol{\Sigma}_{|\alpha} = S(\boldsymbol{\beta}'_{(1)}, \ldots, \boldsymbol{\beta}'_{(m)} \,|\, \boldsymbol{\alpha})$ is a block-diagonal matrix:

$$\boldsymbol{\Sigma}_{|\alpha} = \begin{pmatrix} \boldsymbol{\Sigma}_{\beta,1|\alpha} & 0 & \cdots & 0 \\ 0 & \boldsymbol{\Sigma}_{\beta,2|\alpha} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \boldsymbol{\Sigma}_{\beta,m|\alpha} \end{pmatrix}, \tag{18}$$

where $\boldsymbol{\Sigma}_{\beta,h|\alpha} = S(\boldsymbol{\beta}_{(h)}|\boldsymbol{\alpha})$. We do not make further independence assumptions for the following reasons.

(i) As in ordinary linear regression, changing the origins of the x-variables will change the correlation between α_i and $\boldsymbol{\beta}_i$ $(i = 1, \ldots, k)$. Hence, as there is no reason to believe that the x-variables each have a natural origin, it is unreasonable to assume that α_i and $\boldsymbol{\beta}_i$ are uncorrelated in the prior model.

(ii) If we only consider items that have some specified covariate values, $\mathbf{x}^*$ say, then these items follow a simple multinomial distribution with membership probabilities $p_1, \ldots, p_k$. Thus, if changing the value of the h^*th component of $\mathbf{x}$ increased one of these probabilities, then it must decrease one or more of the other probabilities, because of the unit sum constraint. It follows that the components of $\boldsymbol{\beta}_{(h^*)}$ cannot be independent of each other.

In the next subsections we give procedures for eliciting:

1. $E(\boldsymbol{\alpha})$ and $S(\boldsymbol{\alpha})$.
2. $E(\boldsymbol{\beta}_{(h)})$ and $\boldsymbol{\Sigma}_{\beta,h|\alpha}$ for $h = 1, \ldots, m$.

3. $\mathbf{\Sigma}_{\alpha,\beta,h}$ for $h = 1, \ldots, m$, where $\mathbf{\Sigma}_{\alpha,\beta,h}$ is the covariance matrix for the covariances between $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}_{(h)}$.

Details for determining $\boldsymbol{\mu}$ and $\mathbf{\Sigma}$ from the quantities in 1, 2 and 3 are given in the supplementary material.

4.2. Assessment of $E(\boldsymbol{\alpha})$ and $S(\boldsymbol{\alpha})$

To make the assessment tasks easier for the expert, each covariate is given a reference value and, in the elicitation procedure, only one covariate at a time is varied. All other covariates are assumed to be at their reference values/levels. By doing this for each covariate in turn, the expert can concentrate on revising his assessments as a result of the change in just one variable. For notational convenience, we assume that the scale of each continuous covariate has been centred so that 0 is its reference value. For a factor, its first level is taken as its reference level, so that its dummy variables each equal 0 at the reference level. Let $\mathbf{x}_0$ be the $m \times 1$ vector of 0s. When $\mathbf{x} = \mathbf{x}_0$, all variables take their reference value (0), so we refer to $\mathbf{x}_0$ as the *reference point* and the *reference subpopulation* consists of those items whose covariate values equal $\mathbf{x}_0$.

To elicit $E(\boldsymbol{\alpha})$ and $S(\boldsymbol{\alpha})$, the expert is asked to restrict attention to items in the reference subpopulation. Considering just this subpopulation, he performs the same assessment tasks as in Section 3. In that section, the expert's assessments yielded $\boldsymbol{\mu}_{k/1} = E(\boldsymbol{Y}_{k/1})$ and $\mathbf{\Sigma}_{k/1} = S(\boldsymbol{Y}_{k/1})$. Now they yield $E(\boldsymbol{Y}_{k/1} \,|\, \mathbf{x}_0)$ and $S(\boldsymbol{Y}_{k/1} \,|\, \mathbf{x}_0)$, which are the quantities $E(\boldsymbol{\alpha})$ and $S(\boldsymbol{\alpha})$, since $\mathbf{x}_0$ is a vector of 0s.

4.3. Assessment of $E(\boldsymbol{\beta}_{(h)})$ and $\mathbf{\Sigma}_{\beta,h|\alpha}$

In this phase of the elicitation process, one covariate at a time is varied from its reference value/level and the remaining covariates are set equal to their reference values, so that only one component of $\mathbf{x}$ is non-zero. For $h = 1, \ldots, m$, let $\mathbf{x}_h^*$ denote a $m \times 1$ vector whose elements are 0 apart from its hth element, which equals some value that we specify, x_h^* say. (For a factor level, x_h^* is set equal to 1.)

For each value of h, $E(\boldsymbol{\beta}_{(h)})$ and $\mathbf{\Sigma}_{\beta,h|\alpha}$ can be elicited through full assessment or through a short-cut option. With either choice, the expert is asked to make assessments for the subpopulation of items for which $\mathbf{x} = \mathbf{x}_h^*$. The expert may choose to use the full-assessment procedure for some values of h and the short-cut option for others.

Full assessment procedure

The expert performs the same assessment tasks as in Section 3 (and Section 4.2) except that:

(i) He is asked to restrict attention to items in the subpopulation with covariate values $\mathbf{x}_h^*$.

(ii) In making assessments he should primarily consider the difference between the medians of the subpopulation of current interest and the medians of the reference subpopulation. When making assessments about probability ratios, he is

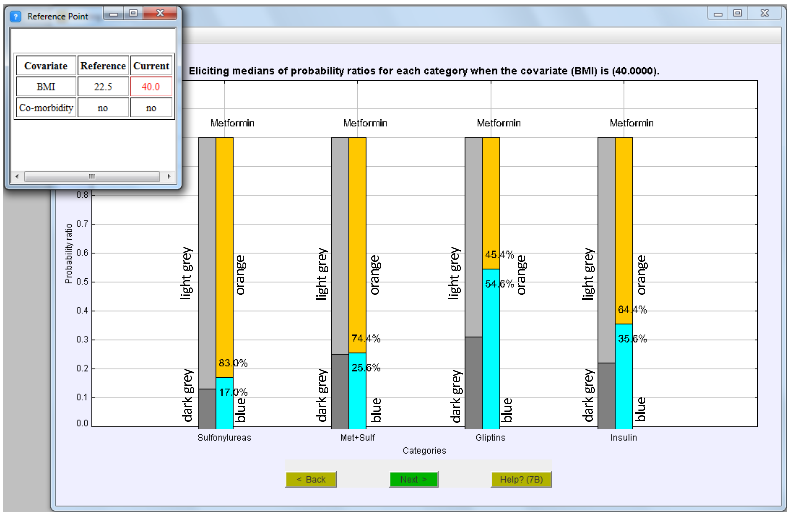

Fig. 5. Assessing medians of probability ratios for specified covariate values. The expert is asked to assume that the left-hand (light-grey/dark-grey) bar in each pair gives the actual probability ratio for the reference subpopulation. Conditional on this assumption, the expert assesses medians for the subpopulation with covariate values $\mathbf{x}_h^*$. These assessments are the blue/orange bars on the right of each pair.

also asked to assume that his assessed medians for the reference subpopulation are the true probability ratios for that subpopulation (which fixes the value of $\boldsymbol{\alpha}$). For the feedback screen showing probabilities, he is asked to assume that his median assessments equaled the true probabilities for the reference subpopulation (which also fixes the value of $\boldsymbol{\alpha}$).

The assessment tasks that the expert performs are illustrated in Fig. 5, Fig. 6 and two figures given in the supplementary material, Figs. S1 and S2. Each figure shows a bar chart with pairs of bars. The bars to the left in each pair (the light-grey/dark-grey bars) show the expert's medians for the probability ratios (or, in Fig. 6, the probabilities) for the reference subpopulation. In other respects, the information contained/elicited in Figs. 5, S1, S2 and 6 corresponds to the information in Figs. 1, 2, 3 and 4, respectively.

A table showing levels/values for the reference subpopulation and the subpopulation of current interest is displayed in all assessment screens but can be re-positioned or closed. To avoid obscuring the main figure, it is only shown in Figs. 5 and S1.

Short-cut option

Using the full procedure requires quite a number of assessments if there are several covariates. As a short-cut, for any $\mathbf{x}_h^*$ (but not for $\mathbf{x}_0$) the expert may choose to skip the assessment of probability ratios (the tasks illustrated in Figs. 5, S1 and S2) and just give assessments for the feedback screen. The feedback screen contains pairs of bars (cf. Fig. 6). The left-hand bars show the expert's unconditional median probabilities when all the covariates are at their reference values, $\mathbf{x}_0$. Conditional

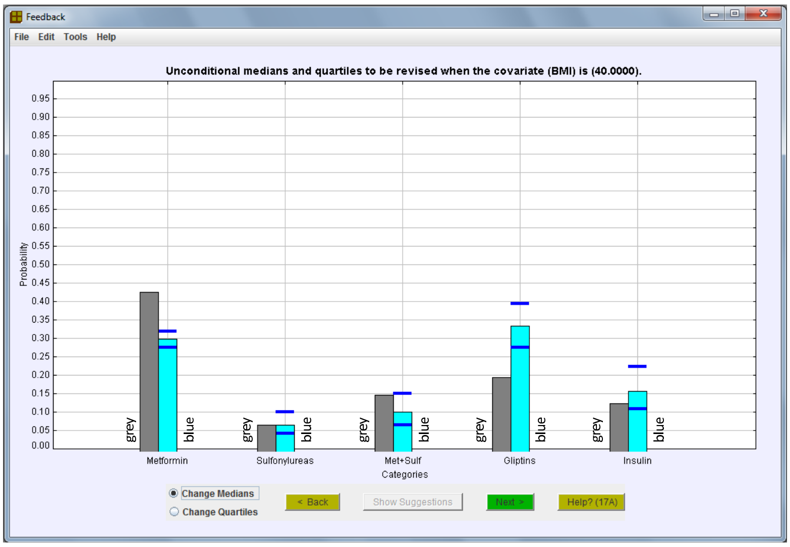

Fig. 6. Feedback screen for both the full procedure and the short-cut option. The left-hand (dark-grey) bars show the expert's unconditional median probabilities when all the covariates are at their reference values, x_0. The expert is asked to assume that these are the true membership probabilities for the reference subpopulation. Under that assumption, the expert assesses medians and quartiles for the membership probabilities of the subpopulation whose covariate values are x_h^* (rather than x_0). These assessments form the right-hand (sky-blue) bars and short (dark-blue) horizontal lines.

on these being the correct values, in the right-hand bars the expert assesses medians and quartiles for membership probabilities when the values of the covariates are $\mathbf{x}_h^*$. In making these assessments, the expert should focus on the difference between the left-hand and right-hand bars in a pair, as his assessments should reflect the change in probabilities as the covariate values change from $\mathbf{x}_0$ to $\mathbf{x}_h^*$.

The short-cut option can shorten the elicitation procedure significantly, but requires the assumption that prior knowledge gives the same correlation structure to changes in membership probabilities as to the membership probabilities themselves. With both the full procedure and the short-cut option, assessments for the subpopulation with covariate values $\mathbf{x}_h^*$ yield $E(\boldsymbol{Y}_{k/1} \,|\, \mathbf{x}_h^*)$ and $S(\boldsymbol{Y}_{k/1} \,|\, \mathbf{x}_h^*, \boldsymbol{\alpha})$. In the supplementary material we show that

$$E(\boldsymbol{\beta}_{(h)}) = \{E(\boldsymbol{Y}_{k/1} \,|\, \mathbf{x}_h^*) - E(\boldsymbol{\alpha})\}/x_h \tag{19}$$

and

$$\boldsymbol{\Sigma}_{\beta,h|\alpha} = S(\boldsymbol{Y}_{k/1} \,|\, \mathbf{x}_h^*, \boldsymbol{\alpha})/x_h^2 \tag{20}$$

where, as defined earlier, x_h is the only non-zero component of $\mathbf{x}_h^*$.

4.4. Assessment of $\boldsymbol{\Sigma}_{\alpha,\beta,h}$

After the assessments in Section 4.3 have been elicited for the subpopulation with covariate values $\mathbf{x}_h^*$, the elicitation procedure continues to focus on this subpopu-

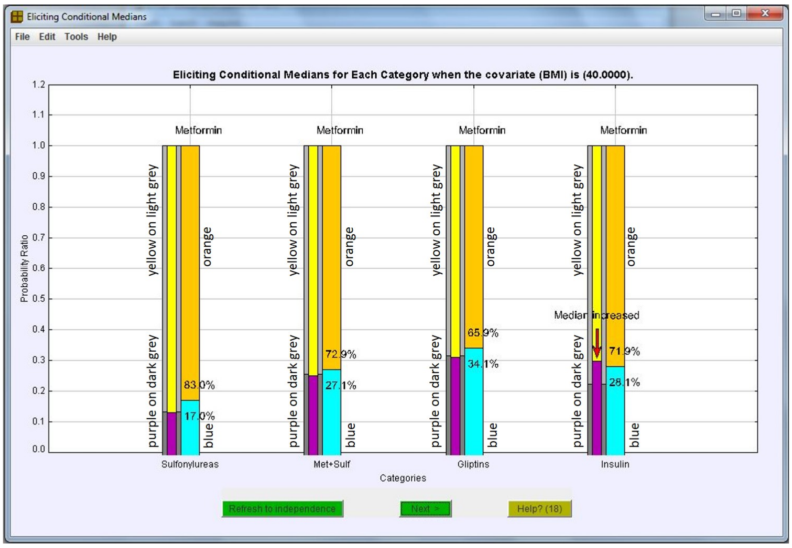

Fig. 7. Assessing correlations between $\boldsymbol{\alpha}$ and $\beta_{(h)}$. The left-hand (light-grey/dark-grey) bars have narrow inner (purple/yellow) bars within them. The expert is asked to assume that the inner bars give the correct proportions for the reference population while the bars behind them show the expert's medians for those proportions. The largest difference between the inner and outer bars is marked with a red arrow (fourth bar). The expert gives his medians for the subpopulation with covariate values $\mathbf{x}_h^*$ in the right-hand bars.

lation and the covariance between $\boldsymbol{\alpha}$ and its β-coefficients ($\boldsymbol{\beta}_{(h)}$) are elicited. The strategy for this task is to choose $k-1$ different values for $\boldsymbol{\alpha}^*$, say $\boldsymbol{\alpha}_2^*, \ldots, \boldsymbol{\alpha}_k^*$ and, for each $\boldsymbol{\alpha}_i^*$, obtain assessments from the expert that determine $E(\boldsymbol{\beta}_{(h)}|\boldsymbol{\alpha} = \boldsymbol{\alpha}_i^*)$. The covariance between $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}_{(h)}$ is reflected in the changes in $\boldsymbol{\beta}_{(h)}$ as $\boldsymbol{\alpha}^*$ varies.

The values chosen for $\boldsymbol{\alpha}_2^*, \ldots, \boldsymbol{\alpha}_k^*$ are derived from $E(\boldsymbol{\alpha})$ and $S(\boldsymbol{\alpha})$, the prior mean and prior variance of $\boldsymbol{\alpha}$, which were elicited in Section 4.2. To form $\boldsymbol{\alpha}_i^*$ ($i = 2, \ldots, k$), each of its components is set equal to the corresponding component of $E(\boldsymbol{\alpha})$, apart from its ith component, which is set equal to the upper quartile of the marginal distribution of α_{i+1}. (The ith component of $\boldsymbol{\alpha}$ is α_{i+1}.) This upper quartile is equated to $\mu_{\alpha,i+1} + 0.674\,\xi_{i+1}$, where $\mu_{\alpha,i+1}$ and ξ_{i+1}^2 denote the ith component of $E(\boldsymbol{\alpha})$ and the ith diagonal element of $S(\boldsymbol{\alpha})$, respectively. (0.674 is the upper quartile of the standard normal distribution.)

Under the condition that $\boldsymbol{\alpha} = \boldsymbol{\alpha}_i^*$, the probability ratios for the reference population are

$$r_j(\mathbf{x}_0) = \exp(\alpha_{i,j}^*) \quad \text{for } j = 2, \ldots, k, \tag{21}$$

where $\alpha_{i,j}^*$ is the jth component of $\boldsymbol{\alpha}_i^*$. The expert is asked to assume that these are the true probability ratios for items whose covariate values are at their reference values. Conditional on this assumption, he then assesses the medians of $r_2, \ldots, r_k$ for the subpopulation of current interest. These give $E(\boldsymbol{Y}_{k/1}|\mathbf{x}_h^*, \boldsymbol{\alpha} = \boldsymbol{\alpha}_i^*)$, from which $E(\boldsymbol{\beta}_{(h)}|\boldsymbol{\alpha} = \boldsymbol{\alpha}_i^*)$ is obtained.

Fig. 7 illustrates the assessment tasks that the expert performs. Within each of

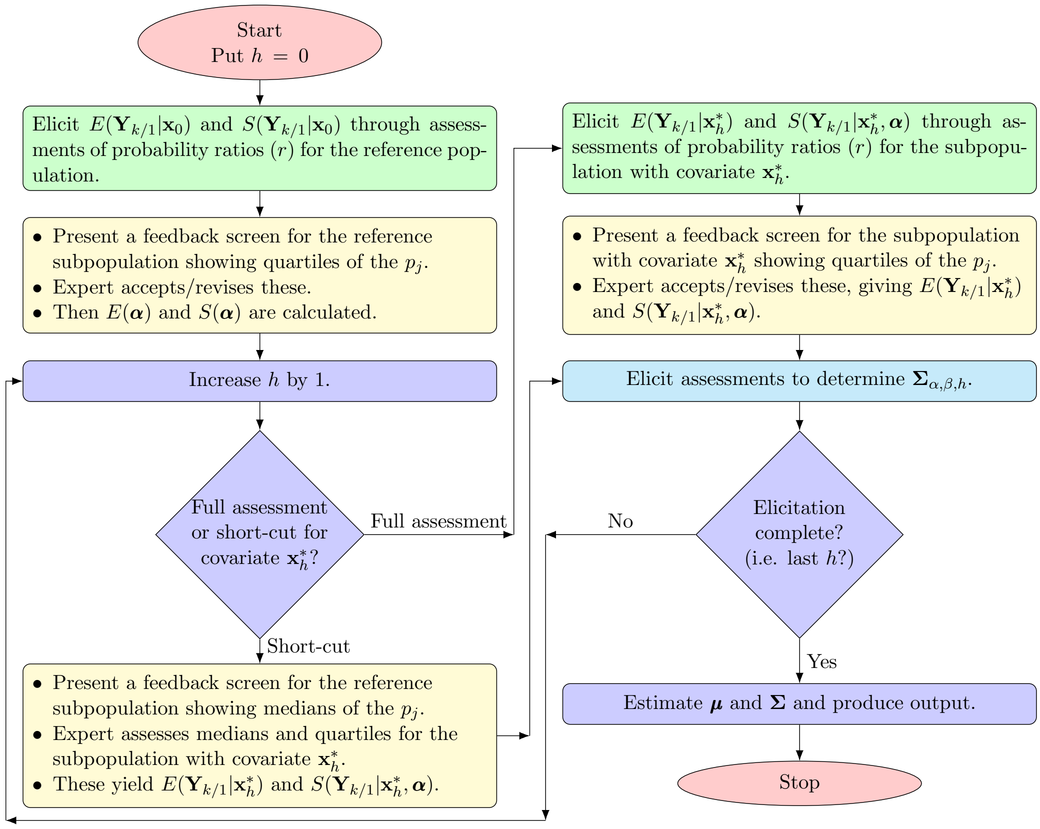

Fig. 8. Flow diagram showing the main steps of the elicitation method when the sampling model has covariates.

the left-hand bars is a narrower (purple/yellow) bar. The expert is asked to assume that these narrower bars give the true probability ratios for items with covariates at their reference value. These 'true' ratios correspond to the condition $\boldsymbol{\alpha} = \boldsymbol{\alpha}_i^*$. The right-hand (blue/orange) bars are the expert's conditional median assessments of probability ratios (given $\boldsymbol{\alpha} = \boldsymbol{\alpha}_i^*$) for items with covariate values $\mathbf{x} = \mathbf{x}_h^*$. To help the expert with this assessment task, we suggest median values for the conditional probability ratios that the expert may accept or change. The suggested values are calculated under the assumption that $\boldsymbol{\beta}_{(h)}$ and $\boldsymbol{\alpha}$ are independent *a priori*.

In supplementary materials, detail is given for calculating $\boldsymbol{\Sigma}_{\alpha,\beta,h}$ from $E(\boldsymbol{\beta}_{(h)}|\boldsymbol{\alpha} = \boldsymbol{\alpha}_2^*), \ldots, E(\boldsymbol{\beta}_{(h)}|\boldsymbol{\alpha} = \boldsymbol{\alpha}_k^*)$ and quantities obtained in Sections 4.2 and 4.3. After all the required assessments have been elicited for the population with covariate values $\mathbf{x}_h^*$, the procedure moves on to the next subpopulation. Fig. 8 gives a flow diagram that summarises the complete elicitation process.

5. Example: Potential cost of a new drug

As noted in the introduction, in Malta there were four standard treatments for diabetes – Metformin (Mtf), Sulfonylureas (Sul), Metformin plus Sulfonylureas (M+S), Insulin (Ins) – and a fifth treatment, Gliptins (Glip), is to be introduced. The new treatment is more expensive per patient, so the overall cost to Malta of treating diabetes will increase, which is a concern to Malta's Ministry for Health and also to its Ministry of Finance. In order to estimate the future cost, an expert quantified his opinion about the proportions of patients in Malta who would be on each treatment. The expert, Dr Neville Calleja, is Director of Health Information and Research in Malta's Ministry for Health. Dr Calleja is both a medical doctor and has a PhD in statistics that was supervised by one of this paper's authors. Consequently, Dr Calleja has a high level of expertise in statistics and was enthusiastic about the concept of quantifying an expert's background knowledge as a prior distribution.

The elicitation session was conducted with both authors acting as facilitators. In preparatory discussion, the expert described his problem of interest and the facilitators gave an overview of the elicitation procedure. The expert was also given some practice at assessing the median and quartiles of a scalar quantity (such as the distance from Manchester to Edinburgh). The elicitation program was then loaded and the expert input his answers and assessments to the computer, first defining the sampling model. He thought that body-mass index (BMI) and co-morbidity would affect which treatment a diabetic patient was likely to take, where co-morbidity is a binary yes/no variable, with 'yes' meaning the patient also suffers from hypertension or has a history of heart disease. Thus BMI and co-morbidity are the covariates in this example. The categories in our multinomial model are the five diabetes treatments Mtf, Sul, M+S, Glip, and Ins, with membership probabilities $p_1, \ldots, p_5$, respectively. The task is to quantify the expert's opinions about the membership probabilities for these five categories and how the probabilities are affected by the covariates. We take Mtf as the fill-up category. The following describes the assessments that the expert made.

5.1. Assessments at the reference point

As a reference point ($\mathbf{x}_0$), the expert chose patients who had a BMI of 22.5 and no co-morbidity diseases, and he was asked to consider the subpopulation of patients with these characteristics. Unconditional median assessments of the probability ratios for each of the last four categories were first elicited. Fig. 1 shows these assessments. He then assessed upper and lower quartiles of probability ratios (i) unconditionally for the second category, and (ii) conditionally for the third, fourth and fifth categories, under the assumptions that the true values of some probability ratios actually equalled their median assessments. Table 1 summarizes this set of assessments. A screen shot showing assessments for the fourth category is given in Fig. 2.

Conditional median assessments were elicited next, under the assumption that the true values of some probability ratios equalled their assessed lower quartiles. These sets of conditional median assessments are summarized in Table 2. Fig. 3

Table 1. Median and conditional quartile assessments at the reference point

	$\frac{p_2}{p_1+p_2}$	$\frac{p_3}{p_1+p_3}$	$\frac{p_4}{p_1+p_4}$	$\frac{p_5}{p_1+p_5}$
Upper quartile	0.18	0.31[1]	0.36[2]	0.25[3]
Median	0.13	0.25	0.31	0.22
Lower quartile	0.08	0.18[1]	0.29[2]	0.15[3]

[1] Given $p_2/(p_1+p_2) = 0.13$.
[2] Given $p_2/(p_1+p_2) = 0.13$ and $p_3/(p_1+p_3) = 0.25$.
[3] Given $p_2/(p_1+p_2) = 0.13$, $p_3/(p_1+p_3) = 0.25$ and $p_4/(p_1+p_4) = 0.31$.

Table 2. Conditional median assessments at the reference point[1]

	$\frac{p_2}{p_1+p_2}$	$\frac{p_3}{p_1+p_3}$	$\frac{p_4}{p_1+p_4}$	$\frac{p_5}{p_1+p_5}$
Given $p_2/(p_1+p_2)$	$\underline{0.08}$	0.27	0.33	0.22
Given $p_2/(p_1+p_2)$ and $p_3/(p_1+p_3)$	$\underline{0.13}$	$\underline{0.18}$	0.35	0.24
Given $p_2/(p_1+p_2)$, $p_3/(p_1+p_3)$ and $p_4/(p_1+p_4)$	$\underline{0.13}$	$\underline{0.25}$	$\underline{0.29}$	0.25

[1] Underlined values are the conditions; assessments are not underlined.

is a screen shot taken as the expert assessed medians for the last two categories, conditional on the second and third categories having probability ratios ($p_2/(p_1+p_2)$ and $p_3/(p_1+p_3)$) of 0.13 and 0.18.

The expert was then shown the feedback graph in Fig. 4, which displays the unconditional median and quartile values that were calculated by the software based on his assessments at the reference point. He accepted the suggested unconditional medians and quartiles as a reasonable representation of his opinions. This completed the assessments that determine $E(\boldsymbol{\alpha})$ and $S(\boldsymbol{\alpha})$.

5.2. Assessments at different covariate values

During the rest of this elicitation process, for each covariate in turn, the expert was asked to consider people who had a specified value for one covariate (a value not equal to its reference value), while the other covariate was at its reference value/level.

The expert gave 40 as a value of BMI for which he could make reasonable assessments and which differed markedly from the reference BMI value of 22.5. To quantify his opinions about $E(\boldsymbol{\beta}_{(1)})$ and $\boldsymbol{\Sigma}_{\beta,1\,|\alpha}$, the expert made modifications to the feedback screen that had resulted from his assessments at the reference point. He was asked to suppose that his median assessments at the reference point were correct and to assess medians and quartiles for a person who had a BMI of 40 instead of the reference value of 22.5 (but who still had no co-morbidity). This step is illustrated in Fig. 6 – his new median assessments were 0.29, 0.06, 0.10, 0.35 and 0.16. The computer then suggested values for the unconditional quartiles and the expert found them an acceptable representation of his opinions.

To obtain $\boldsymbol{\Sigma}_{\alpha,\beta,1}$, the expert again assessed probability ratios for just the subpopulation with BMI=40 and the other covariate at its reference level (i.e. no co-morbidity). These assessments were conditional assessments, where the conditions specified the probability ratios at the reference point. Most of the specified values equalled his assessed medians at the reference point, but for one category at a time (starting with the second category) a slightly larger value was specified. This information could change the expert opinions about the subpopulation with BMI=40, and he was invited to revise his median assessments of the probability ratios for that subpopulation. He only wanted to make changes when the fourth or fifth category had the larger value and gave the following as his median probability ratios for the subpopulation with BMI=40 and no co-morbidity:

	$\frac{p_2}{p_1+p_2}$	$\frac{p_3}{p_1+p_3}$	$\frac{p_4}{p_1+p_4}$	$\frac{p_5}{p_1+p_5}$
α_4 increased	0.23	0.37	0.56	0.36
α_5 increased	0.15	0.27	0.34	0.28

Fig. 7 is a screen shot of the assessments when a value larger than the expert's median assessment was specified for α_5 (the probability ratio at the reference point for the fifth category).

The assessment tasks for the subpopulation with BMI=40 and co-morbidity at its reference level (i.e. no co-morbidity) were repeated for the subpopulation that *had* co-morbidity and had BMI at the reference level of 22.5. Conditional on the probability ratios at the reference point equalling the expert's median assessments, the expert gave 0.14, 0.11, 0.10, 0.40 and 0.20 as his medians for this new subpopulation. All other values suggested by the computer were accepted by the expert as a fair reflection of his opinions, except when conditioning specified that the fourth category had a larger value at the reference point than his median assessment. This larger value affected the expert's opinions about some of the probability ratios of this new subpopulation – he set the conditional medians for the fourth and fifth categories to 0.64 and 0.46.

5.3. Results

The output from the software gives the hyperparameters ($\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$) of a multivariate normal distribution that corresponds to the expert's assessments. These hyperparameters are given in the supplementary material, Tables S1 and S2. The output also presents the correlations corresponding to the off-diagonal elements of $\boldsymbol{\Sigma}$. Fig. S3 in the supplementary material shows the correlations of the hyperparameters given by $\boldsymbol{\Sigma}$ in a graphical display developed by Murdoch and Chow (1996).

The reason for quantifying the expert's opinion was to estimate the cost of treating diabetes in Malta when Gliptins can be prescribed through its national health service. To this end, the expert's opinions were combined with data from a survey that had been conducted in Malta as part of the European Health Interview Survey, 2008 (Ministry for Social Policy, 2008). The survey was conducted on a sample of 5 500 individuals aged 15 years or over. It was drawn from a population register for the Maltese islands and was stratified by age, gender and locality. The survey gave complete data on the required covariates (BMI and co-morbidity status) for 217 diabetics, so for each of these individuals we have the vector $\mathbf{x}$ in equation (14). A summary of these data is given in Table 3. The other information that was needed were the average monthly costs per patient of each treatment. To the nearest five euros, these were Mtf: 65; Sul: 40; M+S: 85; Glip: 230 and Ins: 30.

Table 3. BMI and co-morbidity status of 217 diabetics from the European Health Interview Survey

BMI	Co-morbidity		Total
	Yes	No	
<20.0	4	3	7
20.0 $<$ 22.5	3	0	3
22.5 $<$ 25.0	20	14	34
25.0 $<$ 30.0	56	38	94
30.0 $<$ 40.0	50	22	72
$\geqslant 40$	6	1	7
Total	139	78	217

A vector was randomly generated from the multivariate normal distribution that represented the expert's opinions. This vector gives the values of α_j and $\boldsymbol{\beta}_j$, say $\alpha_j^{\#}$ and $\boldsymbol{\beta}_j^{\#}$ $(j = 2, \ldots, 5)$. Applying equation (14), we determined $p_1(\mathbf{x}), \ldots, p_5(\mathbf{x})$ for each of the diabetics with covariate data. These probabilities were multiplied by the costs of each treatment to give the expected cost for that person, and hence the average cost per patient for the people surveyed could be determined. The number of people with diabetes in Malta is known reasonably accurately (approximately 20 000 adult diabetics), so scaling up gave the cost of diabetes treatments in Malta when $\alpha_j = \alpha_j^{\#}$ and $\boldsymbol{\beta}_j = \boldsymbol{\beta}_j^{\#}$ for $j = 2, \ldots, 5$.

The calculation was repeated for 10 000 randomly selected vectors from the expert's multivariate normal distribution, each giving an estimate of the cost of diabetes treatments in Malta. Their average was 2.38 million euros per month and, based on $250th$ smallest and $250th$ largest estimates, a 95% credible interval for the monthly cost is (1.29, 3.77). R code that performed these calculations is given in supplementary material. Without introduction of the new drug treatment, the corresponding cost is 1.26 million euros per month, based on estimates of the proportions of diabetics in Malta on the other four treatments (Mtf, Sul, M+S and Ins) at the time the expert assessed his prior distribution.

6. Concluding comments

Different sampling models require different forms of prior distribution and, until now, no prior distribution had been proposed for a multinomial sampling model that contains covariates. However, in this context a logistic normal distribution has commonly been used as the sampling model, making it a natural choice as a prior model for representing expert opinion. The distribution contains a large number of parameters and we made reasonable independence assumptions to reduce their number and make the elicitation problem manageable. Nevertheless, quantifying expert opinion as a logistic normal distribution would be well nigh impossible without the use of interactive computing, but the use of interactive graphics leads to a viable elicitation method.

In most of the elicitation tasks that the expert performs, assessments are required of a probability ratio: the proportion of items that fall in a specified category relative to the proportion that fall in the fill-up category. It may seem simpler to ask the expert about the actual proportion of items that fall in the category (i.e. ask about p_i rather than p_i/p_1). Indeed, that was the assessment task we tried to use in a prototype version of the elicitation method developed here. Unfortunately, it led to assessment tasks in which the expert was asked to assume that p_j took a specified value and his opinion about the distribution of $Y_i \mid p_j$ was elicited. This caused problems, because the value of p_j does not determine the value of Y_j, so while the distribution of $Y_i \mid Y_j$ would be a normal distribution, $Y_i \mid p_j$ does not follow a normal distribution, making mathematics intractable. The problem does not arise when a value of $r_j = p_j/p_1$ is specified, as r_j does determine Y_j (we have $Y_j = \log(r_j)$), so that $Y_i \mid r_j$ follows a normal distribution.

The people intending to use the prior distribution that is assessed by an expert will typically include one or more people with good knowledge of statistics. If the expert's background in statistics is limited, then a facilitator who has good knowledge of statistics should guide the expert through the elicitation process. Features from the Sheffield elicitation framework (SHELF) might also form part of the elicitation process if a transparent record is required. SHELF contains a protocol for capturing information about an elicitation exercise including an expert's backgrounds and potential conflicts of interest and any reasoning or key sources of information that underpin the expert's judgements (Gosling, 2018). The elicitation method has been implemented in interactive software that is freely available on the web at `http://statistics.open.ac.uk/elicitation`. All screens include *help* buttons to assist the expert and facilitator. There is also a *User Guide* that includes some training material to give the expert practice at quantifying his/her opinion. The output available from the software includes files that can be run in WinBUGS (Lunn *et al.*, 2000), OpenBUGS (Lunn *et al.*, 2009) or rJAGS (Plummer, 2015), making it straightforward to combine the prior distribution with data. The output also gives a file that documents the assessments elicited at each step of the process and, using a second file that is also provided, the software can sequentially reproduce screen shots of the elicitation process. As the example in Section 5 indicates, the elicitation software provides a practical means of quantifying expert opinion about a multinomial model that includes covariates, and the assessed prior distribution may be readily exploited to yield useful inferences.

Acknowledgement

The work reported here was supported by a studentship from the Open University. We are very grateful to Dr Neville Calleja whose opinion is quantified in the example. We are grateful to two reviewers and the editor for constructive comments that led to clear improvements in this paper and the work it reports.

References

Agresti, A. (2002). *Categorical Data Analysis*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., New Jersey, second edition.

Aitchison, J. (1986). *The Statistical Analysis of Compositional Data*. Chapman and Hall, London.

Al-Labadi, L., Baskurt, Z., and Evans, M. (2018). Statistical reasoning: Choosing and checking the ingredients, inferences based on a measure of statistical evidence with some applications. *Entropy*, **20:289**, doi:10.3390/e20040289.

Bedrick, E. J., Christensen, R., and Johnson, W. (1996). A new perspective on priors for generalized linear models. *Journal of the American Statistical Association*, **91**, 1450–1460.

Bedrick, E. J., Christensen, R., and Johnson, W. (1997). Bayesian binomial regression: Predicting survival at a trauma center. *American Statisticion*, **51**, 211–218.

Elfadaly, F. G. and Garthwaite, P. H. (2013). Eliciting Dirichlet and Connor–Mosimann prior distributions for multinomial models. *Test*, **22**, 628–646.

Elfadaly, F. G. and Garthwaite, P. H. (2017). Eliciting Dirichlet and Gaussian copula prior distributions for multinomial models. *Statistics and Computing*, **27**, 449–467.

Evans, M., Guttman, I., and Li, P. (2017). Prior elicitation, assessment and inference with a Dirichlet prior. *Entropy*, **19:564**, doi:10.3390/e19100564.

Garthwaite, P. H., Kadane, J. B., and O'Hagan, A. (2005). Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association*, **100**, 680–701.

Garthwaite, P. H., Al-Awadhi, S. A., Elfadaly, F. G., and Jenkinson, D. J. (2013). Prior distribution elicitation for generalized linear and piecewise-linear models. *Journal of Applied Statistics*, **40**, 59–75.

Gosling, J. P. (2018). SHELF: the sheffield elicitation framework. In L. C. Dias, A. Morton, and J. Quigley, editors, *Elicitation: The Science and Art of Structuring Judgement*. Springer, New York.

Kadane, J. B., Dickey, J. M., Winkler, R., Smith, W., and Peters, S. (1980). Interactive elicitation of opinion for a normal linear model. *Journal of the American Statistical Association*, **75**, 845–854.

Lunn, D., Thomas, A., Best, N., and Spiegelhalter, D. (2000). WinBUGS - a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing*, **10**, 325–337.

Lunn, D., Spiegelhalter, D., Thomas, A., and Best, N. (2009). The BUGS project: Evolution, critique and future directions (with discussion). *Statistics in Medicine*, **28**, 3049–3082.

Ministry for Social Policy (2008). *European Health Interview Survey 2008*. Department of Health Information and Research, Malta.

Murdoch, D. J. and Chow, E. D. (1996). A graphical display of large correlation matrices. *The American Statistician*, **50**, 178–180.

Németh, B., Molnár, A., Akehurst, R., Horváth, M., Kóczián, K., Németh, G., Götze, A., and Vokó, Z. (2017). Quality-adjusted life year difference in patients with predominant negative symptoms of schizophrenia treated with cariprazine and risperidone. *Journal of Comparative Effectiveness Research*, **6**, 639–648.

O'Hagan, A. and Forster, J. (2004). B*ayesian Inference*, volume 2B of *Kendall's Advanced Theory of Statistics*. Arnold, London, second edition.

O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., Oakley, J. E., and Rakow, T. (2006). *Uncertain Judgements: Eliciting Experts' Probabilities*. John Wiley, Chichester.

Plummer, M. (2015). rjags: Bayesian graphical models using MCMC. R package version 3-15.

Tsutakawa, R. K. and Lin, H. Y. (1986). Bayesian estimation of item response curves. *Psychometrika*, **51**, 251–267.

Wilson, E. C. F., Usher-Smith, J. A., Emery, J., Corrie, P. G., and Walter, F. M. (2017). Expert elicitation of multinomial probabilities for decision-analytic modeling: an application to rates of disease progression in undiagnosed and untreated melanoma. *Value in Health*. dx.doi.org/10.1016/j.jval.2017.10.009.

Wilson, K. J. (2018). Specification of informative prior distributions for multinomial models using vine copulas. *Bayesian Analysis*, **13**, 749–766.

Zapata-Vázquez, R. E., O'Hagan, A., and Bastos, L. S. (2014). Eliciting expert judgements about a set of proportions. *Journal of Applied Statistics*, **41**, 1919–1933.