Generalized Beta Regression to Elicit Conditional Distributions of Medical Variables

Alessandro Magrini
Department of Economics
and Management
University of Pisa, Italy

Davide Luciani
IRCCS - Mario Negri Institute
for Pharmacological Research
Milan, Italy

Federico Mattia Stefanini
Department of Statistics,
Computer Science,
Applications
University of Florence, Italy

Abstract

Univariate conditional models are of core importance in supporting medical reasoning, as they allow to decompose a joint probability distribution using the chain rule. Although several methods are available for the elicitation of the joint prior distribution of parameters when the response is a medical categorical variable, the case of a medical continuous response is typically difficult to address, because its sample space is often bounded to an interval and its relationship with explanatory variables may be not linear. In these situations, the elicitation of an informative prior distribution on parameters of a univariate conditional model is challenging, because some level of statistical training is required to a medical expert for interpreting parameters and for retrieving appropriate quantitative information about them. The task can be eased and made efficient by recognizing that physicians typically distinguish among values involving medically normal and pathological patient conditions on the grounds of their personal clinical experience. In this paper, we propose a Generalized Beta regression where parameter elicitation is performed by establishing a correspondence among measured values expressed as relative positions within intervals with a clinical interpretation, regardless the original scales of variables. Software implementing the elicitation procedure is freely available.

Keywords: Bayesian elicitation, degree of belief, informative prior distribution, logistic function, rescaling procedure.

1. Introduction

Quantitative analysis is particularly challenging in medical problems, because very often a high number of variables is related into a highly structured stochastic system (see, for example, Luciani, Cavuto, Antiga, Miniati, Monti, Pistolesi, and Bertolini 2007; Luciani and Stefanini 2012; Magrini, Stefanini, and Luciani 2018). The study of such structures starts with the qualitative knowledge published in the specialized medical literature, and the process may benefit from the use of Directed Acyclic Graphs (DAGs), especially as a first step towards the development of a probabilistic network (Koller and Friedman 2009). In this case, the DAG describes a factorization of the joint probability distribution into the product of

univariate conditional models. Quantitative information required to specify a univariate conditional model is frequently scattered in many different sources and varies greatly in quality (Druzdzel and van der Gaag 2000), so that working with heterogeneous sources of information is often the only available option. On these grounds, the elicitation of experts' degree of belief is a valuable resource, because medical experts are typically able to retrieve appropriate quantitative information from the literature, or, if unavailable, to refer to their personal clinical experience. Nevertheless, judgements from experts may be subjected to several biases (Kahneman, Slovic, and Tversky 1982; Gigerenzer 2002; Kynn 2008), thus elicitation should be performed under a parametric model with a clear interpretation for the expert and/or making use of well characterized elicitation methods.

Canonical models (Díez and Druzdzel 2006) have been extensively used in medical problems involving categorical variables (Onisko, Druzdzel, and Wasyluk 2001; Díez, Mira, Iturralde, and Zubillaga 1997). In a canonical model, parameters represent conditional probabilities, thus elicitation is typically easy and reliable. However, continuous variables require to undergo discretization, thus eventually entailing loss of information.

Logistic regression models are very popular options when the response is a medical categorical variable and continuous explanatory variables are involved (Bagley, White, and Golom 2001). Logistic regression models properly deal with continuous explanatory variables, and parameters represent odds ratios instead of conditional probabilities. Several elicitation methods have been proposed for parameters of logistic regression models (see, for example, Chen, Ibrahim, and Yiannoutsos 1999).

Linear regression is widely used in medical problems when the response is a continuous variable (Schneider, Hommel, and Blettner 2010). The main advantage of linear regression is that parameters have a simple interpretation, that is they represent variations in the expected value of the response given a unitary variation of one explanatory variable at a time. Several elicitation methods have been proposed for the linear regression model (see, for example, Kadane, Dickey, Winkler, Smith, and Peters 1980; Kadane and Wolfson 1998).

However, linear regression is not a good solution for many medical applications, because it assumes that the relationship between the response and explanatory variables is linear and that errors have the same variance, although the scale of a medical continuous variable is typically bounded to an interval for living patients. Under these conditions, the regression model could predict values exceeding the plausible range of the response while disregarding that the error variance should decrease as predicted values approach the extremes of the scale. These problems could be addressed by working on a transformed scale (e.g., the logarithmic or the logit scale), by truncating the Gaussian distribution (DeMaris 2004, Chapter 9), or by assuming a distribution different from the Gaussian one (for instance the Beta distribution: Ferrari and Cribari-Neto 2004). Unfortunately, the interpretation of parameters changes and could become obscure to a physician, as proved by the fact that no relevant applications of these parametric models have been developed to elicit medical experts' degree of beliefs. An interesting proposal is the elicitation under piecewise-linear regression models (Garthwaite, Al-Awadhib, Fadlalla, and Jenkinsonc 2013), but this method is not specifically focused on the features which are relevant for physicians, thus additional training besides the medical standard one is required for the expert in order to obtain reliable assessments.

In our proposal, an informative prior distribution on parameters of a Generalized Beta regression is elicited after piecewise linear rescaling of continuous variables. Intervals in the piecewise linear rescaling gather medically normal and pathological values, as defined by the standard medical training. The rescaling procedure allows the expert to focus on clinically relevant concepts during parameter elicitation, like the correspondence among measured values expressed as relative positions within ranges of clinical interpretation, without considering the original scales of variables. Uncertainty is computed according to the number of patient cases on which each assessment is based. If available, evidence from a published clinical study on any parameter can be taken into account, and uncertainty is derived according to the number

of sampling units supporting the estimate of that parameter. As such, our proposal allows to circumvent the lack of quantitative information on pathophysiological and etiopathogenic relationships, which is mainly attested in animal, rather than human studies. Software implementing the elicitation procedure is available in the online supplementary materials.

This paper is structured as follows. In Section 2, the formal model for eliciting the dependence of a medical continuous response from each explanatory variable is described. In Section 3, the quantitative assessment is detailed. In Section 4, our elicitation method is applied to a variable (respiratory rate) included in a probabilistic network for the diagnosis of pulmonary embolism. Finally, our proposal is discussed in Section 5.

2. Model and notation

In this section, we detail the model proposed to elicit the dependence relationship between a medical continuous response, denoted as Y, and its explanatory variables (EVs), denoted as $X_1, \ldots, X_n$, where n is the number of EVs. Firstly, medical continuous scales are characterized on the basis of three different ranges of values, each involving medically normal, hypo-pathological and hyper-pathological patient conditions, respectively (Subsection 2.1). Secondly, the conditional expected value is defined as a symmetric logistic function after piecewise linear rescaling is applied to both the response and EVs (Subsection 2.2). Finally, the Generalized Beta family of distributions is assumed to complete the probabilistic specification of the response (Subsection 2.3).

2.1. Medical continuous scales

Medical continuous variables are typically bounded to an interval of values for living patients, that is values outside that interval cannot be observed until a patient is alive. The standard medical training and medical literature provide physicians the ability to properly recognize the extreme values of a medical continuous variable for a living patient, as well as to distinguish among values involving medically normal and pathological patient conditions (Jacobs, Oxley, and DeMott 2001; Irwin and Rippe 2011). On these grounds, the scale of a medical continuous variable V is partitioned into three intervals: a medical normal range (*n-range*), in which values are regarded as non-pathological, a hypo-pathological range (*lp-range*), including values lower than non-pathological ones, and a hyper-pathological range (*hp-range*), including values higher than non-pathological ones. Extreme values of the scale are denoted as v_{L2} and v_{H2}, while the cut points limiting *n-range* as v_{L1} and v_{H1}. The mid value of *n-range* is denoted as v_C, and the mid values of *lp-range* and *hp-range* are denoted as v_{LC} and v_{HC}, respectively. Value v_C is taken as a reference for all values representing a fully healthy patient condition, and is called *neutral value*. The partitioned scale of a medical continuous variable is shown in Figure 1. In the remainder, we will refer to *n-range*, *lp-range* and *hp-range* as *ranges of clinical interpretation*.

Special cases of continuous EVs are a *hyper-restricted* and a *hypo-restricted* medical scale, in which *lp-range* or *hp-range* are (or are supposed to be) of null size, respectively.

Binary EVs are another special case which is handled by assigning value 0 to the state involving non-pathological patient conditions, and value 1 to the state involving pathological patient conditions. Any polytomous EV is replaced by a set of dummy indicators, one for each non-neutral value. In the remainder, these dummy indicators will be considered as distinct EVs in order to keep the notation as simple as possible.

2.2. Dependence of a medical continuous response from its explanatory variables

Since physicians are typically able to assign a measured value of a medical continuous variable to one of the ranges of clinical interpretation, we believe that the most medical tasks, for
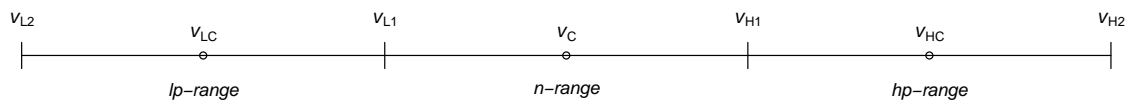
Figure 1: The scale of a medical continuous variable.

example medical diagnosis, widely depend on the relative position of measured values in those ranges, thus original scales become inessential for understanding changes in a patient's health. Similarly, we also believe that physicians are prepared to appreciate a change in the expected value of the response due to a change in the value of an EV without considering the original scales of variables, provided that changes involve an EV at a time while all other EVs take their respective neutral values, and that changes are still expressed as relative positions within the ranges of clinical interpretation. On these grounds, we propose to elicit the degree of dependence of the response from each EV after a rescaling procedure is applied (Algorithm 1) to make *n-range*, *lp-range* and *hp-range* of equal size and to map their mid values to 0, −1 and 1, respectively.

After rescaling, cut points v_{L2}, v_{L1}, v_{H1} and v_{H2} are mapped to −1.5, −0.5, 0.5 and 1.5, respectively (Figure 2). Computations are shown in Appendix. A rescaled variable or a value of a rescaled variable is indicated by the tilde symbol. For instance, $\widetilde{Y}$ denotes the rescaled response and $\widetilde{y}_{\mathrm{C}}$ denotes its neutral value. Similarly, $\widetilde{X}_i$ denotes a rescaled EV and $\widetilde{x}_{i,\mathrm{C}}$ denotes its neutral value.
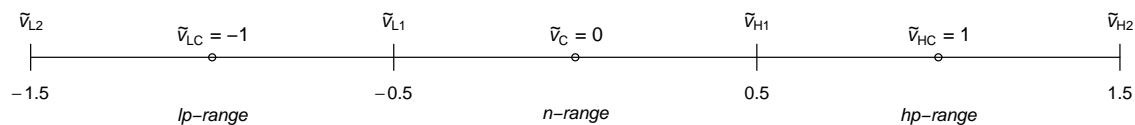

Figure 2: Cut points of a rescaled medical variable.

Algorithm 1 Rescaling procedure

For each variable, apply one of the following:

- if V is a continuous variable, then:

$$\widetilde{V} = \begin{cases} -1.5 + \frac{V - v_{\mathrm{L2}}}{v_{\mathrm{L1}} - v_{\mathrm{L2}}} & \text{if } V < v_{\mathrm{L1}} \\ -0.5 + \frac{V - v_{\mathrm{L1}}}{v_{\mathrm{H1}} - v_{\mathrm{L1}}} & \text{if } v_{\mathrm{L1}} \leq V < v_{\mathrm{H1}} \\ 0.5 + \frac{V - v_{\mathrm{H1}}}{v_{\mathrm{H2}} - v_{\mathrm{H1}}} & \text{if } V \geq v_{\mathrm{H1}} \end{cases}$$

- otherwise $\widetilde{V} = V$.

Our method to elicit the degree of dependence of the rescaled response $\widetilde{Y}$ from a rescaled EV $\widetilde{X}_i$ is based on a reference situation denoted as *fully-active*: changing $\widetilde{X}_i$ from value 0 to value −1 or 1 determines a change in the expected value of $\widetilde{Y}$ from value 0 to value −1 or 1. If a smaller change of $\widetilde{X}_i$ suffices to change the expected value of $\widetilde{Y}$ from value 0 to value −1 or 1, then we refer to a *hyper-active* relationship. If a change of $\widetilde{X}_i$ from value 0 to value −1 or 1 is insufficient to determine an analogous change in the expected value of Y, then we refer to a *hypo-active* relationship. Note that the definition of fully-, hypo- and hyper-active dependence relationship holds if the expected value of the rescaled response can be assumed equal to 0 when all rescaled EVs take value 0. We believe that this is often the case when no

relevant EV is omitted, thus we recommend particular care in the selection of EVs.

The expected value of the rescaled response $\widetilde{Y}$ given any configuration of rescaled EVs $\widetilde{X}_1 = \widetilde{x}_1, \ldots, \widetilde{X}_n = \widetilde{x}_n$ is assumed to be a member of the generalized logistic family of functions with horizontal asymptotes at $\widetilde{y}_{\mathrm{L2}} = -1.5$ and $\widetilde{y}_{\mathrm{H2}} = 1.5$:

$$
\mathrm{E}[\widetilde{Y} \mid \widetilde{X}_1 = \widetilde{x}_1, \ldots, \widetilde{X}_n = \widetilde{x}_n, \boldsymbol{\beta}] =
$$
$$
-1.5 + \frac{3}{1 + \exp\left\{-\log(5) \cdot (\nu(\widetilde{x}_1), \ldots, \nu(\widetilde{x}_n))' \boldsymbol{\beta}\right\}} \tag{1}
$$

where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)$ is a vector of parameters, hereinafter called *regression coefficients*, and:

$$
\nu(z) = \mathrm{sign}(z) \cdot |z|^r \qquad r = \frac{1}{\log(2)} \log\left(\frac{\log(5)}{\log(2)}\right) \approx 1.2153
$$

The function $\nu(\cdot)$ and the constant $\log(5)$ are introduced to ensure that, for any fully-active EV X_i, parameter β_i takes absolute value 1, and, given that all other EVs take their respective neutral values, the expected value of the rescaled response belongs to *n-range* if and only if X_i takes value in *n-range*. Computations are shown in Appendix.

Figure 3 displays the expected value function for different values of parameter β_i associated to EV X_i, given that all other EVs take their respective neutral values. The expected value function is defined only for values 0 and 1 of binary EVs, only for values not in *lp-range* for hyper-restricted EVs, and only for values not in *hp-range* for hypo-restricted EVs. If the response is a hyper- (or a hypo-) restricted variable, its expected value is left-bounded by $\widetilde{y}_{L1}$ (or right-bounded by $\widetilde{y}_{H1}$). A fully-active relationship is described by a regression coefficient with absolute value 1, while a hypo- (or hyper-) active relationship is described by a regression coefficient with absolute value smaller (or greater) than 1. Each regression coefficient has the same sign of the influence of the corresponding EV on the response.
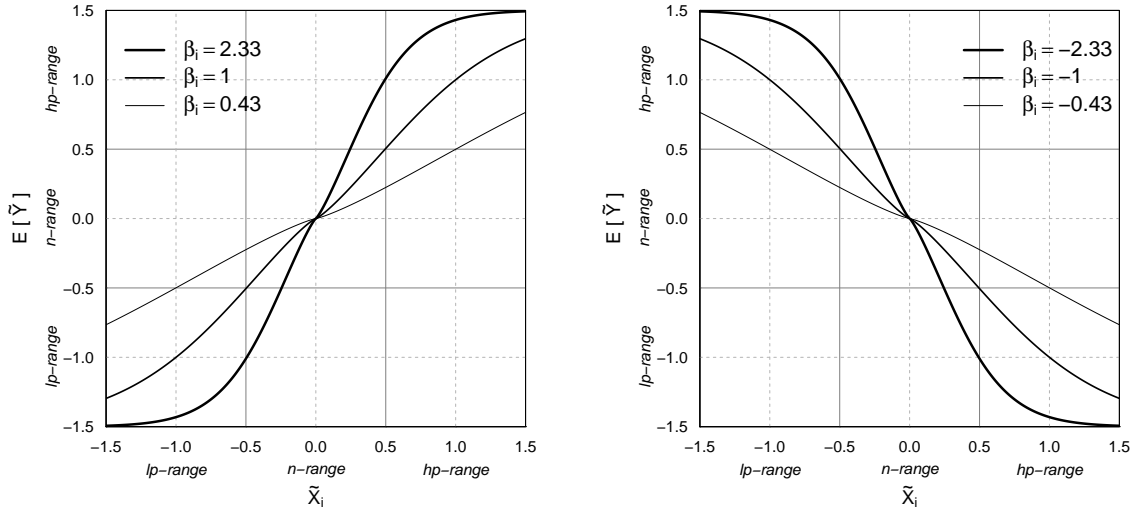

Figure 3: Expected value function for different values of parameter β_i associated to EV X_i, given that all other EVs take their respective neutral values. Left panel: monotonically increasing influence. Right panel: monotonically decreasing influence.

2.3. The Generalized Beta family of distributions

After defining the conditional expected value of the rescaled response for a given configuration of EVs (Equation 1), we provide a full probabilistic specification of the rescaled response by assuming that it belongs to the *Generalized Beta* family (Johnson, Kotz, and Balakrishnan

1995, Vol. 2). The standard formulation of the Generalized Beta density is:

$$p(\widetilde{y} \mid a_1, a_2, \widetilde{y}_{\mathrm{L2}}, \widetilde{y}_{\mathrm{H2}}) = \frac{\Gamma(a_1 + a_2)}{(\widetilde{y}_{\mathrm{H2}} - \widetilde{y}_{\mathrm{L2}})^{a_1 + a_2 - 1} \Gamma(a_1) \Gamma(a_2)} (\widetilde{y} - \widetilde{y}_{\mathrm{L2}})^{a_1 - 1} (\widetilde{y}_{\mathrm{H2}} - \widetilde{y})^{a_2 - 1} \quad (2)$$

with $a_1 > 0$ and $a_2 > 0$. A parameterization based on the relative mean $0 < \delta = \frac{\mathrm{E}[\widetilde{Y}] - \widetilde{y}_{\mathrm{L2}}}{\widetilde{y}_{\mathrm{H2}} - \widetilde{y}_{\mathrm{L2}}} < 1$ and on a precision parameter $\tau > 0$ (Ferrari and Cribari-Neto 2004) is considered here:

$$\begin{cases} \delta = \frac{a_1}{a_1 + a_2} \\ \tau = a_1 + a_2 \end{cases} \quad (3)$$

The resulting probability density function is:

$$\mathrm{GenBeta}(\widetilde{y} \mid \delta, \tau, -1.5, 1.5) = \frac{\Gamma(\tau)}{3^{\tau - 1} \Gamma(\tau \delta) \Gamma(\tau(1 - \delta))} \; (\widetilde{y} + 1.5)^{\tau \delta - 1} (1.5 - \widetilde{y})^{\tau(1 - \delta) - 1} \quad (4)$$

After reparameterization, the expected value and variance are:

$$\begin{aligned} \mathrm{E}[\widetilde{Y}] &= -1.5 + 3\delta \\ \mathrm{Var}[\widetilde{Y}] &= \frac{9(1 - \delta)\delta}{1 + \tau} \end{aligned} \quad (5)$$

where explicit conditioning of parameters is omitted to shorten the notation. The precision parameter τ regulates heteroscedasticity so that the variance of the rescaled response takes the maximum value $2.25/(1 + \tau)$ when the expected value is equal to the neutral value. The variance has limiting value 0 as the expected value approaches the extreme values of the scale (Figure 4).
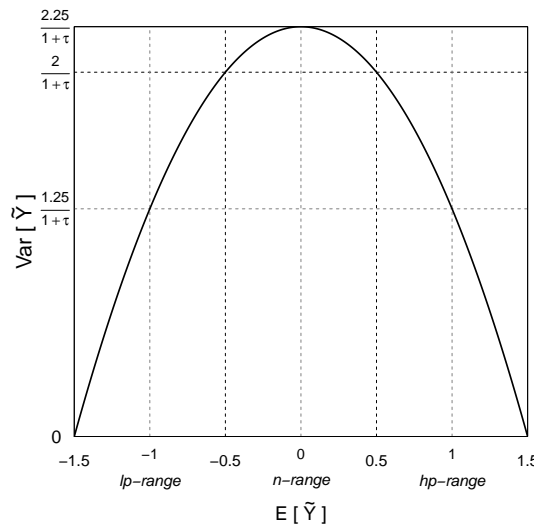

Figure 4: Conditional variance of the rescaled response as a function of different values of the conditional expected value.

3. Quantitative assessment

The quantitative assessment of model parameters is performed by means of Algorithm 2. Afterwards, the elicited joint prior distribution is revised by providing predictive distributions as feedback to the expert. In the next subsections, Algorithm 2 is illustrated and the revision procedure is detailed.

Algorithm 2 Parameter elicitation

---

1) Repeat for $i = 1, \dots, n$:

    i) if X_i is a continuous variable, then:

        - ask the expert for a range, which is not *n-range*, and for a relative position π_i within that range matching a value of X_i for which he/she is able to assess the corresponding expected value of the response, given that all other EVs take their respective neutral values. Compute:

$$\underline{\widetilde{x}_i} = \begin{cases} -1.5 + \pi_i, & \text{if } \textit{lp-range} \text{ is chosen} \\ 0.5 + \pi_i, & \text{if } \textit{hp-range} \text{ is chosen} \end{cases}$$

    otherwise:

        - set $\underline{\widetilde{x}_i} = 1$;

    ii) ask the expert for a range, which is not *n-range*, and for a relative position ϖ_i within that range, matching the expected value of the response given that X_i takes the value chosen in step 1.i and all other EVs take their respective neutral values. Compute:

$$\widetilde{m}_{i,\underline{\widetilde{x}_i}} = \begin{cases} -1.5 + \varpi_i, & \text{if } \textit{lp-range} \text{ is chosen} \\ 0.5 + \varpi_i, & \text{if } \textit{hp-range} \text{ is chosen} \end{cases}$$

    iii) ask the expert for the number of patient cases q_i on which the assessment $\widetilde{m}_{i,\underline{\widetilde{x}_i}}$ is based;

    iv) compute:

$$\widehat{\beta}_i = \frac{1}{\log(5) \cdot \nu(\underline{\widetilde{x}_i})} \cdot \log\left(\frac{1.5 + \widetilde{m}_{i,\underline{\widetilde{x}_i}}}{1.5 - \widetilde{m}_{i,\underline{\widetilde{x}_i}}} \right)$$

---

2) Ask the expert for sets of interacting EVs. Go to step 3 if no set is provided, otherwise denote the sets as $\boldsymbol{X}_{\mathcal{S}} = \{X_{\mathcal{S}_1}, \dots, X_{\mathcal{S}_r}\}$, where r is the number of sets of interacting EVs. Repeat for $j = 1, \dots, r$:

   i) initialize $\widetilde{\boldsymbol{x}}_{\mathcal{S}_j}$ as an empty vector. For each $k \in \mathcal{S}_j$, insert value $\widetilde{\underline{x}}_k$ into $\widetilde{\boldsymbol{x}}_{\mathcal{S}_j}$;

   ii) compute the expected value of the response given that variables in $\mathcal{S}_j$ take the values chosen by the expert in step 1.i and all other EVs take their respective neutral values:

   $$\widetilde{m}_{\mathcal{S}_j, \widetilde{\underline{\boldsymbol{x}}}_{\mathcal{S}_j}} = \frac{1}{1 + \exp\left\{ -\log(5) \cdot \nu(\widetilde{\boldsymbol{x}}_{\mathcal{S}_j})' \{\widehat{\beta}_k : \ k \in \mathcal{S}_j\} \right\}}$$

   iii) feedback such value to the expert as a relative position within one among *lp-range*, *n-range*, or *hp-range*. If the expert agrees with such value, remove $X_{\mathcal{S}_j}$ from $\boldsymbol{X}_{\mathcal{S}}$. Otherwise:

   - ask him/her for a new range and/or a new relative position $\varpi_{\mathcal{S}_j}$ within that range. Compute:

   $$\widetilde{m}_{\mathcal{S}_j, \widetilde{\underline{\boldsymbol{x}}}_{\mathcal{S}_j}} = \begin{cases} -1.5 + \varpi_{\mathcal{S}_j}, & \text{if } \textit{lp-range} \text{ is chosen} \\ -0.5 + \varpi_{\mathcal{S}_j}, & \text{if } \textit{n-range} \text{ is chosen} \\ 0.5 + \varpi_{\mathcal{S}_j}, & \text{if } \textit{hp-range} \text{ is chosen} \end{cases}$$

   - ask the expert for the number of patient cases $q_{\mathcal{S}_j}$ on which the assessment $\widetilde{m}_{\mathcal{S}_j, \widetilde{\underline{\boldsymbol{x}}}_{\mathcal{S}_j}}$ is based;
   - compute:

   $$\widehat{\beta}_{\mathcal{S}_j} = \frac{1}{\log(5) \cdot \prod_{k \in \mathcal{S}_j} \nu(\widetilde{\underline{x}}_k)} \cdot \log\left(\frac{1.5 + \widetilde{m}_{\mathcal{S}_j, \widetilde{\underline{\boldsymbol{x}}}_{\mathcal{S}_j}}}{1.5 - \widetilde{m}_{\mathcal{S}_j, \widetilde{\underline{\boldsymbol{x}}}_{\mathcal{S}_j}}} \right) - \sum_{k \in \mathcal{S}_j} \widehat{\beta}_k$$

   (continued on next page)

3) Ask the expert for a range and for a relative position ϱ within that range matching the first quartile of the response when the conditional expected value is equal to the neutral value y_{C} (whichever the configuration of EVs). Compute:

$$
\widetilde{y}_{0.25} = \begin{cases} -1.5 + \varrho, & \text{if } \textit{lp-range} \text{ is chosen} \\ -0.5 + \varrho, & \text{if } \textit{n-range} \text{ is chosen} \\ 0.5 + \varrho, & \text{if } \textit{hp-range} \text{ is chosen} \end{cases}
$$

Compute the corresponding value of the precision parameter τ:

$$
\widehat{\tau} = \arg_{\tau} \left| 0.25 - \int_{\widetilde{y}_{\mathrm{L2}}}^{\widetilde{y}_{0.25}} \mathrm{GenBeta}(z, 0.5, \tau, \widetilde{y}_{\mathrm{L2}}, \widetilde{y}_{\mathrm{H2}}) \mathrm{d}z = 0 \right|
$$

4) Initialize Λ as an empty matrix.

   i) Repeat s times, where s is the number of bootstrap replications:
   - initialize $\boldsymbol{z}$ as an empty vector;
   - for each row g_l of model matrix $\mathbf{G}$:
     · sample a value z_l such that:

$$
p(z_l) = \mathrm{GenBeta}\left(z_l, \frac{1}{1 + \exp\left\{ -\log(5) \cdot \nu(g_l)'\widehat{\boldsymbol{\beta}} \right\}}, \widehat{\tau}, 0, 1 \right)
$$

   (a simple Beta distribution);
     · insert value z_l into $\boldsymbol{z}$;
   - compute maximum likelihood estimation for $\boldsymbol{\beta}$ and τ, say $(\boldsymbol{\beta}_{\mathrm{ML}}, \tau_{\mathrm{ML}})$, given the model:

$$
p(\boldsymbol{z} \mid \boldsymbol{\beta}, \tau) = \prod_l \mathrm{GenBeta}\left(z_l, \frac{1}{1 + \exp\left\{ -\nu(g_l)'\boldsymbol{\beta} \right\}}, \tau, 0, 1 \right)
$$

   (the Beta regression model proposed by Ferrari and Cribari-Neto 2004 without intercept, where function $\nu(\cdot)$ is applied to the values of EVs). Quasi-Newton optimization with analytic first derivatives provided by Ferrari and Cribari-Neto (2004) may be used;
   - set $\boldsymbol{\beta}_{\mathrm{REP}} = \boldsymbol{\beta}_{\mathrm{ML}} / \log(5)$ and $\tau_{\mathrm{REP}} = \tau_{\mathrm{ML}}$;
   - insert $(\boldsymbol{\beta}_{\mathrm{REP}}, \log(\tau_{\mathrm{REP}}))$ into Λ as a row.
   ii) Compute Σ as the empirical covariance matrix of Λ.

5) Return:

$$
\begin{pmatrix} \boldsymbol{\beta} \\ \log(\tau) \end{pmatrix} \sim \mathrm{MVN}_{n+m+1}\left(\begin{pmatrix} \widehat{\boldsymbol{\beta}} \\ \log(\widehat{\tau}) \end{pmatrix}, \Sigma \right)
$$

3.1. Parameter elicitation

We now illustrate how Algorithm 2 works. In step 1, the expert is asked to consider one configuration of EVs at a time, such that each EV takes a value $\widetilde{x}_i$ outside *n-range* at his/her choice, while all others take their respective neutral values. It is important that the chosen value is outside *n-range* because it will be compared to the neutral one. For each of the considered configurations of EVs, the expert chooses the value $\widetilde{x}_i$, assesses the expected value of the response, say $\widetilde{m}_{i,\widetilde{x}_i}$, and declares the number of patient cases on which such assessment is based, say q_i. For a binary EV, it is meant that $\widetilde{x}_i=1$. During the elicitation, the expert assesses relative positions within the ranges of clinical interpretation, without considering the original scales of variables.

In step 2, the expert is asked to individuate sets of interacting EVs, then, for each one, the expected value of the response in the absence of interaction is returned and eventually adjusted. Evidence from a clinical study published in the literature can be taken into account by setting values $\widetilde{x}_i$ and $\widetilde{m}_{i,\widetilde{x}_i}$ coherently with the ones reported. Analogously, q_i can be set equal to the number of sampling units in the study where $\widetilde{X}_i = \underline{x}_i$. Once values $\widetilde{m}_{i,\widetilde{x}_i}$ are elicited for each configuration of EVs, the implied value of regression coefficients $\boldsymbol{\beta}$, say $\widehat{\boldsymbol{\beta}}$, is computed by inverting Equation 1.

In step 3, the expert assesses the first quartile of the response when the conditional expected value is equal to the neutral value y_{C}, whichever the configuration of EVs, and the implied value of the precision parameter τ is computed, say $\widehat{\tau}$. Again, the expert assesses relative positions within the ranges of clinical interpretation, without considering the original scale of the response.

Note that steps 1 and 2 define a virtual experiment with model matrix:

$$\mathbf{G} = \begin{pmatrix} \mathbf{1}_{q_1}\,\widetilde{x}_1 & \ldots & 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \mathbf{1}_{q_n}\,\widetilde{x}_n & 0 & \ldots & 0 \\ \mathbf{1}_{q_1}\,\widetilde{x}_1\,\mathrm{I}(1 \in \mathcal{S}_1) & \ldots & \mathbf{1}_{q_n}\,\widetilde{x}_n\,\mathrm{I}(n \in \mathcal{S}_1) & \mathbf{1}_{q_{\mathcal{S}_1}}\prod_{k\in\mathcal{S}_1}\widetilde{x}_k & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \mathbf{1}_{q_1}\,\widetilde{x}_1\,\mathrm{I}(1 \in \mathcal{S}_r) & \ldots & \mathbf{1}_{q_n}\,\widetilde{x}_n\,\mathrm{I}(n \in \mathcal{S}_r) & 0 & \ldots & \mathbf{1}_{q_{\mathcal{S}_r}}\prod_{k\in\mathcal{S}_r}\widetilde{x}_k \end{pmatrix} \quad (6)$$

where $\mathbf{1}_{q_i}$ is the unitary column vector of q_i elements and $\mathrm{I}(\cdot)$ is the indicator function. The model matrix is composed of one row for each assessment and one column for each EV or set of interacting EVs for which the assessment is performed.

In step 4, the joint prior distribution on $\begin{pmatrix} \boldsymbol{\beta} \\ \log(\tau) \end{pmatrix}$ is assumed to be well represented by a Multivariate Gaussian distribution with mean equal to $\begin{pmatrix} \widehat{\boldsymbol{\beta}} \\ \log(\widehat{\tau}) \end{pmatrix}$ and covariance matrix equal to the covariance matrix of the maximum likelihood estimator of $\begin{pmatrix} \boldsymbol{\beta} \\ \log(\tau) \end{pmatrix}$ in the virtual experiment, which is computed by parametric bootstrap simulation (Hastie, Tibshirani, and Friedman 2009, Section 7.11).

3.2. Revision

The joint prior distribution on model parameters resulting from quantitative assessment does not necessarily correspond to the expert's actual degree of belief, thus we propose a revision

procedure providing feedback to the expert on the implied predictive distributions for some configurations of EVs at his/her choice. The revision procedure exploits the original scale of the response in order to bring the potential bias due to rescaling procedure into consideration.

The revision starts by displaying to the expert the probability density of the predictive distribution implied by the first configuration of EVs, approximated via Monte Carlo simulation. If the median of the predictive distribution determines surprise or disbelief in the expert, then the prior mean of regression coefficients is multiplied by a positive real value $g \in \mathbb{R}$ chosen by the expert. If this is the case, the joint prior distribution on model parameters is recomputed, the probability density of the predictive distribution is approximated, and its median is inspected again. When the median of the predictive distribution appears adequate to the expert, the interquartile range is inspected. If the interquartile range appears too tight (or wide) to the expert, then he/she is invited to decrease (or to increase) the value of the first quartile when the conditional expected value is equal to the neutral value y_C. If this is the case, the joint prior distribution on model parameters is recomputed, the probability density of the predictive distribution is approximated, and its interquartile range is inspected again. When the interquartile range of the predictive distribution does not raise disbelief anymore, the procedure is repeated for each predictive distribution implied by the other configurations of EVs. The revision stops when no predictive distribution raises disbelief anymore.

4. A worked example

Our elicitation procedure was applied to the medical variable 'respiratory rate' included in BayPAD, a probabilistic network for the diagnosis of pulmonary embolism (Magrini *et al.* 2018). The variable is defined as the number of breaths per minute for a patient. The second author played the role of medical expert in the elicitation task. The source of information for each assessment is explicitly reported.

Scale of the response

Since 0 breaths per minute cannot be reported in living patients, v_L2 was set at 0, corresponding to a state of apnea (no breathing). Respiratory rate values higher than 40 breaths per minute are instead observed in experimental settings only, where they also fail to increase ventilation (Braun 1990), thus v_H2 was set at 40. The medical normal range varies with age, going from 15 to 25 breaths per minute in healthy adults (Banner, Kirby, Kirton, DeHaven, and Blanch 1995). In summary, the elicited cut points for the response were: $y_\mathrm{L2} = 0$, $y_\mathrm{L1} = 15$, $y_\mathrm{H1} = 25$, $y_\mathrm{H2} = 40$.

Selection of EVs

Altered respiratory rate values often result from an abnormal oxygen and carbon dioxide content in arterial blood, which is signalled to special chemo-receptors, in turn affecting the neurological drive controlling respiratory muscles and ventilation. Alterations of the neurological drive are often distinguished into two main general categories, i.e., increased intra-pulmonary shunt and increased dead space (Hamid, Shannon, and Martin 2005). Shunt and dead space are respectively measured as percentage of total cardiac output and total ventilation failing to reach the alveolar space. When a pulmonary region is characterized by a pathological value of shunt (or dead space), other pulmonary regions may react by increasing their dead space (or shunt), a compensatory phenomenon known as mismatching. On these grounds, the expert considered two EVs representing intra-pulmonary shunt and dead space to explain altered respiratory rate values in the presence of mismatching. In order to explain altered respiratory rate values without mismatching, namely when pulmonary

circulation is bypassed before reaching the lungs, the expert considered two further EVs representing reduced alveolar space and extra-pulmonary shunt. Furthermore, conditions affecting the neurological drive and respiratory muscles sustaining ventilation, here denoted as neuromuscular conditions, were considered to explain low respiratory rate values (Hamid *et al.* 2005, Chapter 18). Non-medical conditions like panic disorders were also considered as a cause of increasing respiratory rate when there is no respiratory gas exchange to restore (Papp, Martinez, Klein, Coplan, Norman, Cole, de Jesus, Ross, Goetz, and Gorman 1997).

Scale of continuous EVs

Intra-pulmonary shunt, dead space, extra-pulmonary shunt and reduced alveolar space represent percentage, thus their extreme values were set equal to 0 and 100, respectively. Values higher than 5 percent for shunt and reduced alveolar space, as well as values higher than 30 percent for dead space, are regarded as pathological (Rhoades and Bell 2012, Chapter 20). No relevant clinical effects are associated with low values of dead space, extra-pulmonary shunt and reduced alveolar space, thus their scale was considered as hyper-restricted. Conversely, low values of intra-pulmonary shunt may decrease respiratory rate: by applying the ventilation-perfusion equation (Hamid *et al.* 2005, Chapter 18) to yield an unsustainable oxygen saturation of 90 percent (Papiris, Kotanidou, Malagari, and Roussos 2002) in the case of mismatching, the expert found that values lower than 2 percent are pathological. In summary, elicited cut points were $(0, 2, 5, 100)$ for intra-pulmonary shunt, $(0, 0, 30, 100)$ for dead space, $(0, 0, 5, 100)$ for both extra-pulmonary shunt and reduced alveolar space.

Quantitative assessment

In step 1 of Algorithm 2, the expert performed the quantitative elicitation of the influence of each EV at a time. For what concerns panic disorders, the expert retrieved from published evidence a mean respiratory rate of 28.77 (relative position equal to 0.25 within *hp-range*) among 25 patients (Papp *et al.* 1997). For what concerns neuromuscular conditions, the mean respiratory rate depends on disease severity. Since the most common scenario pertains to patients exposed to respiratory muscle fatigue after prolonged and extremely demanding breathing, the definition of neuromuscular conditions was limited to those cases requiring mechanical ventilation, where a low respiratory rate often precedes an apnea event (Braun 1990; Cretikos, Bellomo, Hillman, Chen, Finfer, and Flabouris 2008). On these grounds, the expert assessed a relative position equal to 0.6 within *hp-range* for the conditional expected value of respiratory rate, based on 100 patient cases. No clinical study was found addressing the influence of intra-pulmonary shunt, dead space, extra-pulmonary shunt and reduced alveolar space on respiratory rate, thus the expert assessed a fully active relationship between the response and each of these EVs based on his personal experience (5 patient cases). In step 2 of Algorithm 2, the expert individuated intra-pulmonary shunt and dead space as a set of interacting EVs. We returned the current expected value of respiratory rate when both the two EVs take the mid value of *hp-range* and all other EVs take their respective neutral values, that is relative position 0.8846 within *hp-range*. The expert wanted to increase such relative position to 0.9, based on his personal experience (5 patient cases). In step 3 of Algorithm 2, the expert was asked for the first quartile of respiratory rate when the conditional mean is equal to the neutral value and provided a relative position equal to 0.3 within *n-range*. Expert's assessments are summarized in Table 1.

In step 4 of Algorithm 2, we computed the resulting joint prior distribution on parameters (50000 bootstrap replications):

$$\begin{pmatrix} \boldsymbol{\beta} \\ \log(\tau) \end{pmatrix} \sim \text{MVN}_8(\mu, \Sigma)$$

Table 1: Expert assessments during step 1 of the elicitation procedure. In step 2, the expert assessed $\varrho=0.3000$ (*n-range*).

	π_i	ϖ_i	q_i	$\widetilde{\underline{x}}_i$	$\widetilde{m}_{i,\underline{x}_i}$
Intra-pulmonary shunt	0.50 (*hp-range*)	0.50 (*hp-range*)	5	1.0000	1.0000
Dead space	0.50 (*hp-range*)	0.50 (*hp-range*)	5	1.0000	1.0000
Extra-pulmonary shunt	0.50 (*hp-range*)	0.50 (*hp-range*)	5	1.0000	1.0000
Reduced alveolar space	0.50 (*hp-range*)	0.50 (*hp-range*)	5	1.0000	1.0000
Panic disorder	–	0.25 (*hp-range*)	25	1.0000	0.7500
Neuromuscular conditions	–	0.60 (*lp-range*)	100	1.0000	-0.9000
Intra-pulmonary shunt : Dead space	0.50 (*hp-range*); 0.50 (*hp-range*)	0.90 (*hp-range*)	5	1.0000	1.4000

where:

$$\mu = \begin{pmatrix} 1.0000 \\ 1.0000 \\ 1.0000 \\ 1.0000 \\ 0.6826 \\ -0.8614 \\ 0.0922 \\ 3.2526 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} 0.019510 & 0.000161 & 0.000018 & -0.000055 & -0.000042 & 0.000072 & -0.019699 & 0.000962 \\ 0.000161 & 0.019380 & 0.000325 & -0.000051 & 0.000084 & -0.000008 & -0.019871 & 0.000870 \\ 0.000018 & 0.000325 & 0.019357 & -0.000062 & -0.000141 & -0.000120 & 0.000733 & 0.000629 \\ -0.000055 & -0.000051 & -0.000062 & 0.019043 & -0.000070 & 0.000048 & 0.000430 & 0.000509 \\ -0.000042 & 0.000084 & -0.000141 & -0.000070 & 0.002950 & 0.000017 & 0.000464 & 0.000366 \\ 0.000072 & -0.000008 & -0.000120 & 0.000048 & 0.000017 & 0.000881 & -0.000208 & -0.000583 \\ -0.019699 & -0.019871 & 0.000733 & 0.000430 & 0.000464 & -0.000208 & 0.092830 & 0.000884 \\ 0.000962 & 0.000870 & 0.000629 & 0.000509 & 0.000366 & -0.000583 & 0.000884 & 0.014011 \end{pmatrix}$$

Marginal quantile summaries are provided in Table 2.

Table 2: Marginal quantile summaries of the elicited joint prior distribution.

	0.025	0.25	0.5	0.75	0.975
Intra-pulmonary shunt	0.6402	0.7262	1.0000	1.2738	1.3598
Dead space	0.6414	0.7271	1.0000	1.2729	1.3586
Extra-pulmonary shunt	0.6416	0.7273	1.0000	1.2727	1.3584
Reduced alveolar space	0.6445	0.7295	1.0000	1.2705	1.3555
Panic disorder	0.5427	0.5762	0.6826	0.7891	0.8225
Neuromuscular conditions	-0.9378	-0.9195	-0.8614	-0.8032	-0.7849
Intra-pulmonary shunt : Dead space	-0.6926	-0.5049	0.0922	0.6894	0.8770
Precision (τ)	19.0613	20.5027	25.8563	32.6079	35.0737

Revision

The expert chose four configurations of EVs to check the consistency of the elicited prior distribution on parameters. Each configuration consisted of dead space and reduced alveolar space taking the minimum value of *hp-range*, combined with:

1. absence of panic disorder, absence of neuromuscular conditions;

2. absence of panic disorder, presence of neuromuscular conditions;

3. presence of panic disorder, absence of neuromuscular conditions;

4. presence of panic disorder, presence of neuromuscular conditions.

The first configuration was chosen to test the joint influence of two continuous non-interacting EVs on the response. Since the four EVs involve fully-active dependence relationships with

the response, it is expected that the predictive distribution is mainly located in *hp-range* when they all involve weakly hyper-pathological conditions. The second configuration was chosen to test if the (negative) influence neuromuscular conditions is able to neutralize the joint influence of two fully-active EVs involving weakly hyper-pathological conditions. The third configuration was chosen to explore the addition of the influence of panic disorder, and the fourth to evaluate to which extent neuromuscular conditions can counterbalance the result.

Probability density of the inspected predictive distributions are shown in Figure 5 (Monte Carlo approximation with 50000 replications). Their median and interquartile range (within brackets) appeared coherent with the expert's expectations, thus revision was not necessary: 27.62 (5.63) breaths per minute for the first configuration, 16.67 (4.46) breaths per minute for the second configuration, 35.26 (3.84) breaths per minute for the third configuration, 24.86 (5.18) breaths per minute for the fourth configuration.

Configuration 1

N = 50000 (bw: 0.387658)   Median: 27.62019   IQR: 5.634964

Configuration 2

N = 50000 (bw: 0.344346)   Median: 16.675139   IQR: 4.463249

Configuration 3

N = 50000 (bw: 0.296528)   Median: 35.264059   IQR: 3.843451

Configuration 4

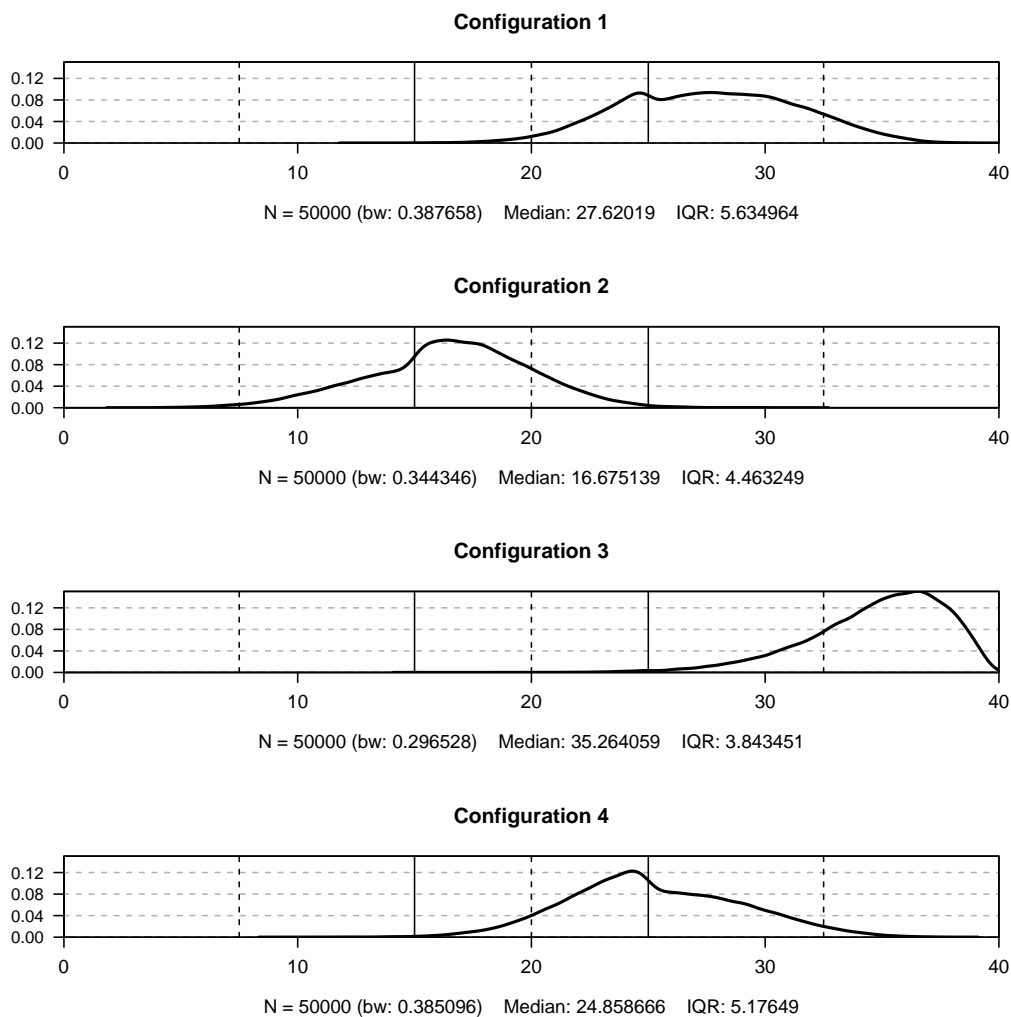N = 50000 (bw: 0.385096)   Median: 24.858666   IQR: 5.17649

Figure 5: Probability density of the predictive distributions inspected during the revision of the joint prior distribution elicited in the worked example (Monte Carlo approximation with 50000 replications). Each configuration of EVs consists of dead space and reduced alveolar space taking the minimum value of *hp-range*, combined with: absence of panic disorder, absence of neuromuscular conditions (Configuration 1); absence of panic disorder, presence of neuromuscular conditions (Configuration 2); presence of panic disorder, absence of neuromuscular conditions (Configuration 3); presence of panic disorder, presence of neuromuscular conditions (Configuration 4). The original scale of the response (respiratory rate) is used. Vertical straight lines denote the cut points: $y_{L2} = 0$, $y_{L1} = 15$, $y_{H1} = 25$, $y_{H2} = 40$.

5. Discussion

In this paper, we addressed the elicitation of an informative prior distribution on parameters of a univariate conditional model when the response is a medical continuous variable. Such task is typically difficult because some level of statistical training is required to a medical expert for interpreting parameters and for retrieving appropriate quantitative information about them. Our method is founded on reliable information contained in the medical literature, that is the correspondence among measured values expressed as relative positions within ranges of clinical interpretation, regardless the original scales of variables. As such, it can be eventually applied to elicit an informative prior distribution on the parameters of a probabilistic network.

Three critical phases can be highlighted in our proposal. The first is the selection of explanatory variables, which was supported by several medical arguments in the worked example. Although physicians agree on what the potential explanations of a response are, their selection typically depends on the granularity of the representation required by the medical context of interest, represented by the diagnosis of pulmonary embolism (Luciani, Marchesi, and Bertolini 2003; Luciani *et al.* 2007; Magrini *et al.* 2018) in the worked example.

The second critical phase is the elicitation of cut points of the original scales, which are assumed to be known without uncertainty. In the worked example, cut points were elicited according to variations of clinical relevance with reference to arterial oxygen saturation. This is a reasonable assumption as long as conclusive evidence is retrievable from the literature. The choice of cut points may also influence the shape of conditional predictive distributions on the original scale, in particular the probability density function may show sharp changes at the cut points.

The third critical phase is the assessment of uncertainty. In our method, the variance of the joint prior distribution of parameters is computed according to the number of patient cases on which each assessment is based. In the worked example, the expert paid attention to assess a number of patient cases substantially lower than the one typically employed by clinical studies. Such approach fulfills the recommendations of the Evidence-Based Medicine (Guyatt, Cairns, and Churchill 1992) paradigm for the clinical practice, inviting physicians to comply with the results of clinical studies as long as they are available, and only upon unavailability to base judgements on personal clinical experience.

Our proposal is a combination of the CMP (Bedrick, Christensen, and Johnson 1996) and the *g-prior* (Geweke 1986; Bové and Held 2011). Similarly to the CMP, we aim to elicit independent conditional means, but uncertainty on parameters is derived from the number of patient cases on which each assessment is based, rather than from several elicited quantiles. In the *g-prior*, the model matrix of the virtual experiment is taken from data in order to achieve conjugacy. We instead characterize the virtual experiment according to the elicited information, in order to obtain a prior distribution which is independent of the data at hand and closer to the expert's degree of belief.

The proposed piecewise linear rescaling is expected to enrich the class of relationships approximately described by a symmetric logistic function, but its adequacy could be empirically checked by assessing the conditional expected value for several values of the same explanatory variable (overfitting) (O'Hagan, Buck, Daneshkhah, Eiser, Garthwaite, Jenkinson, Oakley, and Rakow 2006, Chapter 5.4). Similarly, the adequacy of the conditional variance function could be investigated by assessing the first quartile of the response for several conditional expected values. Furthermore, the dependence of the response from each explanatory variable was assumed to be monotonic, a case that covers a relevant amount of useful applications, but non-monotonic relationships could be allowed.

Several studies caution against the negative effect of implicit heuristics during the elicitation of beliefs (Kahneman *et al.* 1982; Gigerenzer 2002; Kynn 2008). We believe that our proposal is more robust to implicit heuristics than existing methods, because assessments can be directly based on the evidence from published clinical studies, or, if unavailable, the expert can focus on clinically relevant concepts, like the correspondence among measured values expressed as

relative positions within ranges of clinical interpretation.

Future work could be directed towards the evaluation of the validity of our proposal. On one hand, the effective availability of the information required to perform the elicitation can be confirmed by assessing practical difficulties encountered by experts with different clinical experience or belonging to different medical specialties. On the other hand, the extent to which the model developed by an expert is accepted by another expert can be evaluated. Comparative studies with elicitation procedures based on the original scales of variables may be also addressed, where real patient data could be employed to evaluate the reliability of predictive distributions obtained with each procedure.

Software implementing the elicitation procedure is available in the online supplementary materials and, besides efficiently performing the elicitation-revision cycle, it can be used to detect critical issues emerging from the practical application of our method.

Acknowledgements

References

Bagley SC, White H, Golom BA (2001). "Logistic Regression in the Medical Literature: Standards for Use and Reporting, with Particular Attention to One Medical Domain." *Journal of Clinical Epidemiology*, **54**(10), 979–985.

Banner MJ, Kirby RR, Kirton OC, DeHaven CB, Blanch PB (1995). "Breathing Frequency and Pattern are Poor Predictors of Work of Breathing in Patients Receiving Pressure Support Ventilation." *Chest*, **108**(5), 1338–1344.

Bedrick EJ, Christensen R, Johnson W (1996). "A New Perspective on Priors for Generalized Linear Models." *Journal of the American Statistical Association*, **91**, 1450–1460.

Bové DS, Held L (2011). "Hyper g-Priors for Generalized Linear Models." *Bayesian Analysis*, **6**(3), 387–410.

Braun SR (1990). "Respiratory Rate and Pattern." In HK Walker, WD Hall, JW Hurst (eds.), *Clinical Methods: The History, Physical, and Laboratory Examinations*, 3rd edition. Chapter 43, Butterworths, Boston, US-MA.

Chen MH, Ibrahim JG, Yiannoutsos C (1999). "Prior Elicitation, Variable Selection and Bayesian Computation for Logistic Regression Models." *Journal of the Royal Statistical Society. Series B (Methodological)*, **61**, 223–242.

Cretikos MA, Bellomo R, Hillman K, Chen J, Finfer S, Flabouris A (2008). "Respiratory Rate: the Neglected Vital Sign." *Medical Journal of Australia*, **188**(11), 657–659.

DeMaris A (2004). *Regression with Social Data: Modeling Continuous and Limited Response Variables*. John Wiley & Sons, Hoboken, US-NJ.

Díez FJ, Druzdzel MJ (2006). "Canonical Probabilistic Models for Knowledge Engineering." *Technical Report CISIAD-06-01*, UNED, Madrid, E.

Díez FJ, Mira J, Iturralde E, Zubillaga S (1997). "DIAVAL, a Bayesian Expert System for Echocardiography." *Artificial Intelligence in Medicine*, **10**, 59–73.

Druzdzel MJ, van der Gaag LC (2000). "Building Probabilistic Networks: Where Do the Numbers Come From?" *IEEE Transactions on Knowledge and Data Engineering*, **12**(4), 481–486.

Ferrari SLP, Cribari-Neto F (2004). "Beta Regression for Modelling Rates and Proportions." *Journal of Applied Statistics*, **31**(7), 799–815.

Garthwaite PH, Al-Awadhib SA, Fadlalla GE, Jenkinsonc DJ (2013). "Prior Distribution Elicitation for Generalized Linear and Piecewise-Linear Models." *Journal of Applied Statistics*, **40**(1), 59–75.

Geweke J (1986). "On Assessing Prior Distributions and Bayesian Regression Analysis with g-Prior Distributions." In DA Berry, KM Chaloner, JK Geweke (eds.), *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti*, pp. 233–243. Elsevier, Amsterdam, NL.

Gigerenzer G (2002). *Calculated Risks: How to Know When Numbers Deceive You.* Simon & Schuster, New York, NY.

Guyatt G, Cairns J, Churchill D (1992). "Evidence-Based Medicine. A New Approach to Teaching the Practice of Medicine." *Journal of the American Medical Association*, **268**(17), 2420–2425.

Hamid Q, Shannon J, Martin J (2005). *Physiological Basis of Respiratory Disease.* BC Decker Inc., Hamilton, CN-ON.

Hastie T, Tibshirani R, Friedman J (2009). *The Elements of Statistical Learning - Data Mining, Inference, and Prediction.* 2nd edition. Springer, Stanford, US-CA.

Irwin RS, Rippe JM (2011). *Irwin and Rippe's Intensive Care Medicine.* 7nd edition. Lippincott Williams & Wilkins, Philadelphia, US-PA.

Jacobs DS, Oxley DK, DeMott WR (2001). *Jacobs & DeMott Laboratory Test Handbook.* 5nd edition. Lexi-Comp, Cleveland, US-OH.

Johnson NL, Kotz S, Balakrishnan N (1995). *Continuous Univariate Distributions.* 2nd edition. Wiley, Hoboken, US-NJ.

Kadane JB, Dickey JM, Winkler RL, Smith WS, Peters SC (1980). "Interactive Elicitation of Opinion for a Normal Linear Model." *Journal of the American Statistical Association*, **75**, 845–854.

Kadane JB, Wolfson LJ (1998). "Experiences in Elicitation." *The Statistician*, **47**, 3–19.

Kahneman D, Slovic P, Tversky A (1982). *Judgement under Uncertainty: Heuristics and Biases.* Cambridge Univeristy Press, Cambridge, UK.

Koller D, Friedman N (2009). *Probabilistic Graphical Models. Principles and Techniques.* The MIT Press, Cambridge, US-MA.

Kynn M (2008). "The 'Heuristics and Biases' Bias in Expert Elicitation." *Journal of the Royal Statistical Society ser. A*, **171**, 239–264.

Luciani D, Cavuto S, Antiga L, Miniati M, Monti S, Pistolesi M, Bertolini G (2007). "Bayes Pulmonary Embolism Assisted Diagnosis: A New Expert System for Clinical Use." *Journal of Emergency Medicine*, **24**(3), 157–164.

Luciani D, Marchesi M, Bertolini G (2003). "The Role of Bayesian Networks in the Diagnosis of Pulmonary Embolism." *Journal of Thrombosis and Haemostasis*, **1**(4), 698–707.

Luciani D, Stefanini FM (2012). "Automated Interviews on Clinical Case Reports to Elicit Directed Acyclic Graphs." *Artificial Intelligence in Medicine*, **55**(1), 1–11.

Magrini A, Stefanini FM, Luciani D (2018). " A Probabilistic Network for the Diagnosis of Acute Cardiopulmonary Diseases." *Biometrical Journal*, **60**(1), 174–195.

O'Hagan A, Buck CE, Daneshkhah A, Eiser JR, Garthwaite PH, Jenkinson DJ, Oakley JF, Rakow T (2006). *Uncertain Judgements: Eliciting Experts' Probabilities*. John Wiley & Sons, Chichester, UK.

Onisko A, Druzdzel MJ, Wasyluk H (2001). "Learning Bayesian Networks Parameters from Small Data Sets: Applications of Noisy-OR Gates." *International Journal of Approximate Reasoning*, **27**, 165–182.

Papiris S, Kotanidou A, Malagari K, Roussos C (2002). "Clinical review: Severe Asthma." *Critical Care*, **6**, 30–44.

Papp LA, Martinez JM, Klein DF, Coplan JD, Norman RG, Cole R, de Jesus MJ, Ross D, Goetz R, Gorman JM (1997). "Respiratory Psychophysiology of Panic Disorder: Three Respiratory Challenges in 98 Subjects." *American Journal of Psychiatry*, **154**, 1557–1565.

Rhoades R, Bell DR (2012). *Medical Physiology: Principles for Clinical Medicine*. 3rd edition. Lippincott Williams & Wilkins, Philadelphia, US-PA.

Schneider A, Hommel G, Blettner M (2010). "Linear Regression Analysis: Part 14 of a Series on Evaluation of Scientific Publications." *Deutsches Arzteblatt International*, **107**(44), 776–782.

Appendix

Cut points after rescaling It is required that *n-range*, *lp-range* and *hp-range* are of equal size and their mid values are mapped to 0, −1 and 1, respectively.

The constraints are:

$$
\begin{cases}
\frac{\widetilde{v}_{\text{L2}}+\widetilde{v}_{\text{L1}}}{2} = -1 \\
\frac{\widetilde{v}_{\text{L1}}+\widetilde{v}_{\text{H1}}}{2} = 0 \\
\frac{\widetilde{v}_{\text{H1}}+\widetilde{v}_{\text{H2}}}{2} = 1 \\
\widetilde{v}_{\text{L1}} - \widetilde{v}_{\text{L2}} = \widetilde{v}_{\text{H1}} - \widetilde{v}_{\text{L1}} = \widetilde{v}_{\text{H2}} - \widetilde{v}_{\text{H1}}
\end{cases}
$$

which imply:

$$
\begin{cases}
\widetilde{v}_{\text{L2}} = -1.5 \\
\widetilde{v}_{\text{L1}} = -0.5 \\
\widetilde{v}_{\text{H1}} = 0.5 \\
\widetilde{v}_{\text{H2}} = 1.5
\end{cases}
$$

Conditional expected value function It is required that, for any fully-active EV X_i, (1) parameter β_i takes absolute value 1; (2) given that all other EVs take their respective neutral values, the expected value of the rescaled response belongs to *n-range* if and only if X_i takes value in *n-range*.

The first feature is satisfied if:

$$
\begin{cases}
\text{E}[\widetilde{Y} \mid \widetilde{X}_1 = 0, \ldots, \widetilde{X}_{i-1} = 0, \widetilde{X}_i = 1, \widetilde{X}_{i+1} = 0, \ldots, \widetilde{X}_n = 0; \beta_i = 1] = 1 \\
\text{E}[\widetilde{Y} \mid \widetilde{X}_1 = 0, \ldots, \widetilde{X}_{i-1} = 0, \widetilde{X}_i = -1, \widetilde{X}_{i+1} = 0, \ldots, \widetilde{X}_n = 0; \beta_i = 1] = -1 \\
\text{E}[\widetilde{Y} \mid \widetilde{X}_1 = 0, \ldots, \widetilde{X}_{i-1} = 0, \widetilde{X}_i = 1, \widetilde{X}_{i+1} = 0, \ldots, \widetilde{X}_n = 0; \beta_i = -1] = -1 \\
\text{E}[\widetilde{Y} \mid \widetilde{X}_1 = 0, \ldots, \widetilde{X}_{i-1} = 0, \widetilde{X}_i = -1, \widetilde{X}_{i+1} = 0, \ldots, \widetilde{X}_n = 0; \beta_i = -1] = 1
\end{cases}
$$

and the second feature is satisfied if:

$$
\begin{cases}
\mathrm{E}[\widetilde{Y} \mid \widetilde{X}_1 = 0, \ldots, \widetilde{X}_{i-1} = 0, \widetilde{X}_i = 0.5, \widetilde{X}_{i+1} = 0, \ldots, \widetilde{X}_n = 0; \beta_i = 1] = 0.5 \\
\mathrm{E}[\widetilde{Y} \mid \widetilde{X}_1 = 0, \ldots, \widetilde{X}_{i-1} = 0, \widetilde{X}_i = -0.5, \widetilde{X}_{i+1} = 0, \ldots, \widetilde{X}_n = 0; \beta_i = 1] = -0.5 \\
\mathrm{E}[\widetilde{Y} \mid \widetilde{X}_1 = 0, \ldots, \widetilde{X}_{i-1} = 0, \widetilde{X}_i = 0.5, \widetilde{X}_{i+1} = 0, \ldots, \widetilde{X}_n = 0; \beta_i = -1] = -0.5 \\
\mathrm{E}[\widetilde{Y} \mid \widetilde{X}_1 = 0, \ldots, \widetilde{X}_{i-1} = 0, \widetilde{X}_i = -0.5, \widetilde{X}_{i+1} = 0, \ldots, \widetilde{X}_n = 0; \beta_i = -1] = 0.5
\end{cases}
$$

In summary, the constraints are:

$$
\begin{cases}
1 = -1.5 + \frac{3}{1+\exp\{-h \cdot \nu(1) \cdot 1\}} \\
-1 = -1.5 + \frac{3}{1+\exp\{-h \cdot \nu(-1) \cdot 1\}} \\
-1 = -1.5 + \frac{3}{1+\exp\{-h \cdot \nu(1) \cdot (-1)\}} \\
1 = -1.5 + \frac{3}{1+\exp\{-h \cdot \nu(-1) \cdot (-1)\}} \\
0.5 = -1.5 + \frac{3}{1+\exp\{-h \cdot \nu(0.5) \cdot 1\}} \\
-0.5 = -1.5 + \frac{3}{1+\exp\{-h \cdot \nu(-0.5) \cdot 1\}} \\
-0.5 = -1.5 + \frac{3}{1+\exp\{-h \cdot \nu(0.5) \cdot (-1)\}} \\
0.5 = -1.5 + \frac{3}{1+\exp\{-h \cdot \nu(-0.5) \cdot (-1)\}}
\end{cases}
$$

By choosing $\nu(z) = \mathrm{sign}(z) \cdot |z|^r$, the solution is unique:

$$
\begin{cases}
h = \log(5) \\
r = \frac{1}{\log(2)} \log\left(\frac{\log(5)}{\log(2)}\right) \approx 1.2153
\end{cases}
$$

Affiliation:

Alessandro Magrini
Department of Economics and Management
University of Pisa
I-56124 Pisa, Italy
E-mail: alessandro.magrini@for.unipi.it
ORCID iD: http://orcid.org/0000-0002-7278-5332