Nigerian Journal of Mathematics and Applications
Volume **32**, (2022), 234 − 249. *Printed by Unilorin press*
©*Nig. J. Math. Appl.* *http://www.njmaman.com*

Incorporating Prior Beliefs into the Construction of Prior Distribution of the Parameters: A: A Survival Analysis Approach

Juliana I. Consul*

Abstract

Bayesian inference requires the combination of prior experience and the observed data (in the form of the likelihood). There is the need to discuss the specification of prior information about the model parameters. This research has built a structure for the joint prior distribution depending on the type of covariate in the context of a general linear model. Some issues associated with prior elicitation and some ways of incorporating our prior beliefs into the construction of the prior distribution of the parameters of the survival model are discussed in this research. This study also presents a way of specifying the covariance between parameters by thinking in terms of the coefficient of determination to find the correlation between the parameters. The research has given a direct and meaningful approach which reduces the burden on experts.

1. Introduction

Bayesian inference is increasingly popular in survival analysis. The introduction of Markov Chain Monte Carlo (MCMC) algorithm [23] has facilitated the spread of Bayesian methods in survival analysis. The beauty of Bayesian analysis lies in incorporating the prior information about the model parameters into the analysis

Received: 28/09/2022, Accepted: 19/10/2022, Revised: 25/10/2022. * Corresponding author.
2015 *Mathematics Subject Classification.* 62F15 & 62N03.
Keywords and phrases. Elicitation, Likelihood, Parameters, Variables, Survival
Department of Mathematics, Niger Delta University, Bayelsa State, Nigeria
E-mails of the corresponding author: iwori2001@yahoo.com
ORCID of the corresponding author: 0000-0002-8427-0483

(as prior distribution) in addition to the likelihood from the data. The prior distribution is an important issue in Bayesian inference because it is used in the calculation of the posterior distribution which brings the analyst closer to what is being modelled. Hence, the prior distribution is a key element of Bayesian inference which can be derived from either expert beliefs (subjective prior) or relevant empirical data (objective prior). However, most Bayesian analysts face the problem of choosing or deciding on the prior distribution to use.

The prior information is often an opinion or belief of an expert within the field of study from whom the prior is elicited. Prior elicitation thereby involves seeking an expert's opinion or knowledge about one or more uncertain parameters as an additional source of information. These elicited information are formalised as informative prior and expressed in the form of the prior distribution. Prior elicitation therefore helps the expert to express their current knowledge in a probabilistic form [12]. One challenge of an expert's elicitation is the problem of excessive overconfident or underconfident and the way by which the full probability distribution is constructed.

The Bayesian analyst is expected to find ways of using the limited specification of information from the experts when trying to construct a prior distribution for a parameter. One major benefit of considering the Bayesian approach is the ability of formally incorporating the expert's opinion a priori [5]. Although, the use of the elicited information from the informative prior from experts is usually seen as a way of introducing bias into the posterior inference. It will be reasonable to think of how to construct suitable prior distribution for parameters of interest in survival analysis. This research is very much interested in proposing and discussing how information are elicited from experts, incorporating the expert's opinion and thereby the construction of suitable prior distribution with application to survival analysis using a breast cancer data set as an example.

Many authors have discussed methods of prior elicitation. [16] incorporated the clinician's beliefs in estimating the treatment effects using Bayesian analysis. They also summarised some methods to elicit beliefs for Bayesian priors. [12] have also discussed some different methods of eliciting multivariate priors that have been proposed in literatures. [28] compared aggregated priors formed by equal weight aggregation, the classical method and the Sheffield elicitation using expert elicitation on a clinical trial. [1] used the Minimum Chi - square which is a classical estimation technique for elicitation of the parameters. [26] proposed a methodology for the construction of parametric prior on two treatment effect parameters based on graphical information (histogram) which were elicited from a group of experts. [2] combined expert prior by using elicited probabilities and elicited quantiles to construct a prior. [9] introduced a procedure where Beta hyperparameters were used to approximate the prior variance and situations where it is not possible to have an expert's opinion within the field of research, historical or published information about the parameters. [25] performed a Bayesian analysis of the Mycotic Ulcer Treatment Trial where an expert's opinion through an online questionnaire was used as a prior belief. [4] investigated Bayesian methods of prior elicitation methods with a focus on clinical trial research. [11] addressed the problem of quantifying expert opinion about a normal linear regression model. This research presents (with illustration) the construction of prior distributions of the parameters with application to survival analysis following the proportional hazard framework. In this research, the specification of prior distribution of the parameters in the model to survival analysis is discussed. The prior distribution of the baseline parameters of a survival model by supposing an exponential lifetime distribution and assuming a normal prior distribution for the logarithm of the hazards (log-hazard) will also be presented. This research also presents the construction of the prior distribution for the coefficient of a quantitative covariate and a 2 - Level factor with examples in survival models. An example on the construction of the prior distribution for unordered factors with application to survival model is given.

1.1. **Bayes' Theorem and Bayesian Inference.** Bayes' theorem can be used to show the relationship between two conditional probabilities that are the reverse of each other. In general, Bayes' theorem combines the prior experience (in the form of a prior probability) with observed data (in the form of a likelihood function) to interpret these data (in the form of a posterior distribution) in a process known as Bayesian inference [27]. The Bayes' theorem is therefore a probabilistic result which plays a central role in Bayesian inference.
A Bayesian analyst requires the specification of prior information about the model parameters by expressing beliefs about the parameters in the form of a probability distribution before the data is being observed. The prior distribution is expected to reflect information about the model parameters. The information used in the construction of the prior distribution is often an opinion or subjective belief of an "expert" within the field of investigation from whom information is being elicited. It is possible that experts provide incorrect information that reflect the beliefs of the expert which gives a different assessments of the prior distribution.

1.2. **Prior elicitation.** Prior elicitation is the process by which researchers attempt to construct the most suitable prior distribution for a parameter of interest in Bayesian Statistics. It is a pre-requisite to acquiring reliable expert information about the model parameters. In the process of prior elicitation, a person's knowledge or personal beliefs about some uncertain quantities are expressed in the form of a probability distribution of the quantities before the data is taken into account [4]. Prior elicitation helps expert to relate their thinking in statistical terms [12, 17]. It is therefore necessary that the Bayesian analyst find ways of

using the limited specification of beliefs from the experts in thinking and thereby fitting a prior distribution for the parameter.
In prior elicitation, there are cases where the expert gives a good reason to believe that the parameter is within some permissable range. In this case, the expert has substantial prior knowledge. In another case, it might be very difficult to properly elicit prior information about the unknown parameters and there may be no expert available to help guide the choice of the prior distribution. In this case, there is vague prior knowledge and it is assumed that all values of the parameter are equally likely used to represent complete prior ignorance. It is also usual that most Bayesian analyst choose prior distributions that may keep the mathematics simple when using the Bayes' theorem and hence they tend to assume a large or infinite variance for the parameter of interest.
Some authors agree that a good prior distribution will conform to the laws of probability and even if a prior distribution conforms to these laws of probability, some elicited distributions provide more useful inference than others. In this research, we will consider that we have substantial prior knowledge. [3] developed a simple online software to implement the methods of prior elicitation. It is also possible to have limited prior information about the parameter and a prior distribution which comes from the same family with the posterior distribution is chosen.
The process of prior elicitation is usually performed by asking the experts to report summaries about the covariate effects in terms of means, median, modes and percentiles of probability [4] and the analysts translates these percentiles into probability. It is generally easier to elicit the quantiles information about expert beliefs than the moments [13] as most authors investigate procedures to derive the model parameters of the distribution using mean and standard deviation [18, 19]. In the case of multivariate elicitation, experts elicit beliefs about two or more unknown parameters. If the parameters are independent, information about some of the parameters would not affect the beliefs about the other parameter and the joint probability density for these parameters will be the product of their marginal densities. In cases where the parameters are dependent, the dependence between the parameters can be expressed in terms of correlation by eliciting either the correlation coefficient or covariance. The correlation between parameters can be directly specified as between the values -1 and $+1$ [6]. [24] investigated different methods of eliciting prior correlations with applications to Bayes linear modelling of the reliability of defence systems.
The prior probability distributions are often classified into either informative or uninformative priors. A prior is informative when prior information about the parameter is available and included in the prior distribution. An expert's opinion is a form of an informative prior because the expert provides honest information of the present state of knowledge before updating with the new information [14].

In the absence of an expert's knowledge, prior information may be obtained from earlier studies. We could also have an informative prior when the posterior distribution of the previous model which is similar to the form of a present model is used as the prior distribution of the present. In this case, the present model is not starting from scratch but based not only on the present data but the cumulative effects of past and present data are taken into account [28].

2. Materials and methods

2.1. **Structures of prior distribution.** When we have several parameters in a model, it is usually best to build a structure for their joint prior distribution. The structures of the prior distribution are constructed depending on the type of the covariate. Our discussion in this research will be in the context of a general linear model. We discuss the structures of prior distribution for coefficient of binary factor or 2 - level, coefficient of categorical factors, correlation or covariance between parameters and parameters representing a set of proportion. It is advisable and reasonable to center quantitative covariates so that sensible prior is chosen.

2.1.1. *Centering of covariates.* Quantitative variables can sometimes be standardized or centered (i.e subtracting a central value such as the mean). It is good and important to center the covariates because it makes it easier to choose a sensible prior. The constant (intercept) will correspond to a plausible case thereby making elicitation more realistic. Furthermore, it will then be more reasonable for the intercept and the coefficient of the covariate to be independent in the prior. For instance, the age of a group of breast cancer patients can be centered by subtracting the estimated sample mean age of the patients from the covariate age.

2.1.2. *Structure of prior distribution for the coefficient of a 2 - Level or binary factor.* We will suppose that we have a $2-$ Level categorical factor or covariate and that the model also contains an intercept then we would have only 1 degree of freedom. We could have only a single contrast with coefficients -1 and 1 [21]. It is possible to think of constructing the priors for the coefficient of a $2-$ Level factor by representing the covariate values as $-1, 1$. This would just be the same as a quantitative covariate.

2.1.3. *Structure of prior distribution for the coefficients of categorical (unordered) factors.* We will suppose that we want to construct prior for a categorical factor, for example four categorical factor. It is possible that we might choose Level 1 as the baseline (known) and consider the other levels of the categorical factor as in the scheme in Table 1.

Table 1. Scheme for construction of prior for four categorical factor with Level 1 as baseline

Level	covariate		
1	0	0	0
2	1	0	0
3	0	1	0
4	0	0	1

The scheme in Table 1 gives a different prior variance to Level 1 compared to the other three levels. In reality, we might not have greater prior knowledge about Level 1. There are some possible ways that not having greater prior knowledge about the baseline is overcome. One way is by constraining the parameters to sum to zero and making the parameters exchangeable [15]. That is, the parameters could all have the same means, variances and each pair of parameters would also have the same covariance. We illustrate constraining the parameters to sum to zero by supposing that we have a factor with four levels. The following scheme of orthogonal contrasts [21] given in Table 2 could be used to constrain parameters to sum to zero.

Table 2. Scheme of orthogonal contrasts constrain to sum to zero

Level	contrast		
1	1	1	1
2	1	-1	-1
3	-1	1	-1
4	-1	-1	1

When the number of levels is not a power of 2 it is more complicated to construct a scheme of constrasts. The following scheme given in Table 3 could be used for a general case of m levels.

Table 3. A scheme when the number of levels is not a power of 2

Level	contrast				
1	-1	-1	-1	....	-1
2	1	-1	-1	....	-1
3	0	2	-1	....	-1
4	0	0	3	....	-1
$\vdots$	$\vdots$	$\vdots$	....	$\vdots$	
m	0	0	0	....	m - 1

2.1.4. *Structure of prior correlation or covariance between parameters.* In multivariate elicitation, if the parameters are dependent, the degree of dependence between the two parameters can be expressed in terms of correlation. The correlation coefficient or covariances between the parameters can be elicited. Eliciting covariance between parameters directly could be difficult as experts do not think in terms of covariance. [24] discussed four different methods of determining the dependency between variables. These methods include direct calculation, direct elicitation of correlation, adjusted expectation and adjusted uncertainty.
We will want to specify the covariance between pairs of parameters by incorporating our prior beliefs into the parameterisation and construction of the covariance matrix. We could make the covariance between parameters zero if we reasonably think that the parameters of a model are independent of each other in our beliefs. One way of thinking of how to get actual values for the prior covariance between any two parameters is to think of the correlation between the parameters. It is also possible to think in terms of the coefficient of determination which is the square of the correlation coefficient between the parameter. The coefficient of determination can be expressed as the proportion of the variance of one of the parameter can be lost if we knew the value of the other parameter. The covariance will be zero if knowing the value of one parameter does not affect our belief about another parameter. A covariance of zero between parameters would mean that one parameter is unable to explain some of the uncertainty by knowing the other parameter.
For instance, we will suppose that if we knew one parameter then 50% of the proportion of the variance of the other parameter will be lost. Then, the coefficient of determination, r^2 is 0.50 and the correlation r, will be the square root of 0.50 which is 0.7. It is therefore expected that the analysts decides on the sign of the correlation by thinking if the parameter is bigger in the same direction with the other parameter in which case we have a positive sign.
The covariance between the parameters X and Y, $Covar(X, Y)$ can then be calculated using

$$Covar(X, Y) = r\sigma_X\sigma_Y$$

where σ_X and σ_Y are the standard deviation of X and Y respectively and r is the correlation between the parameters.

2.1.5. *Structure of prior for parameters representing a set of proportion.* The Dirichlet distribution could be a choice of expressing our beliefs about a set of proportions because of the convenience of being a conjugate family to the multinomial likelihood in Bayesian analysis (see [10]). Suppose we wish to elicit expert beliefs about a set of uncertain proportion $\vartheta = (\vartheta_1, \ldots, \vartheta_k)$ of k categories where $\vartheta_i \geq 0$ for $i = 1, 2, \ldots, k$ and $\sum_{i=1}^{k} \vartheta_i = 1$. The Dirichlet distribution [7] has a probability density function which is given by

$$f(\theta; a_1, a_2) = \frac{\Gamma(a_1 + a_2)}{\Gamma(a_1)\Gamma(a_2)}\theta^{(a_1-1)}(1-\theta)^{(a_2-1)} \text{ for } 0 \leq \theta \leq 1$$

and parameters $a_1, \ldots, a_k$ where $\sum_{i=1}^{k} a_i = n$ and $\vartheta_i = a_i/n$. We can elicit the ϑ_i's as the probability that we think that a single observation is in category i. [30] suggested thinking in terms of a judgement of the expected value of each ϑ_i together with judgement concerned with uncertainty to identify n. Some recent work on elicitation of parameters of Dirichlet distribution are given by [10].

3. Application to survival analysis

We applied the discussion on the elicitation and construction of prior distribution to the parameters in the model of a survival model using a breast cancer data set from from the University of Illorin teaching hospital, Illorin, Nigeria for a period of five years [22]. The data set consists of the length of stay and the status (dead or alive) after treatment from year 2011 to 2016. The covariates used in this breast data set are given as follows:
Age: This is the age (in years) of the patient.
Sex: This is the sex of the patient. Female was indicated as "1" while male as "2".
Mode of diagnosis (mode): This is the mode of diagnosis of the cancer. Cytological was indicated as "1" while Histological was indicated as "2".
Location of breast cancer (location): This indicates the location of the breast cancer on the survivability of the breast cancer patients. Left breast was indicated as "1", right breast was indicated as "2" and both breast was indicated as "3".

We will follow the classic proportional hazard framework in modelling the hazard rate of survival times in terms of the covariates [20].
The proportional hazard model is the most commonly used method to relate the hazard function to the covariate values for a subject using the proportionality assumption [8]. Suppose we have S covariates for $s = 1, 2, \cdots, S$ and n subjects for $i = 1, 2, \cdots, n$. We denote the covariate vector for a subject i by $\underline{X}_i = (1, x_{i,1}, x_{i,2}, \cdots, x_{i,S})$. These covariates may be continuous or quantitative, categorical or even indicator variables (equal to 1 if present and 0 if absent). The proportional hazard model assumes that any two subjects i and j with the hazard function $h_i(t)$ and $h_j(t)$ at time t and covariate vectors $\underline{X}_i = (1, x_{i,1}, x_{i,2}, \cdots, x_{i,S})$ and $\underline{X}_j = (1, x_{j,1}, x_{j,2}, \cdots, x_{j,S})$, have their hazards related by

$$h_i(t) = \lambda_{i,j} \times h_j(t)$$

We have that $\lambda_{i,j}$ is a constant and does not depend on t. The proportional hazard model can also be written as:

$$h_i(t) = \lambda_i \times h_0(t) \tag{1}$$

where $h_0(t)$ is the baseline hazard function which is a function of time t but does not involve the covariates $\underline{X}_i = (1, x_{i,1}, x_{i,2}, \cdots, x_{i,S})$. The quantity λ_i is the hazard multiplier which depends on (or is related to) the covariates of the subject i but not on the time variable t. We must have that $\lambda_i \geq 0$. This is usually done using a logarithmic link function to a linear predictor η_i. So,

$$\log \lambda_i = \eta_i = \beta_0 + \sum_{s=1}^{S} \beta_s x_{i,s} \tag{2}$$

where $x_{i,s}$ is the value of covariate s for subject i and β_0 is the baseline parameter. The linear predictor can then be used as a prognostic index.

We will suppose a baseline hazard using the exponential distribution. We assume that the baseline hazard using the exponential distribution which has constant hazard given by

$$h(t) = \lambda.$$

The survivor function is

$$S(t) = \exp\{-\lambda t\}$$

and the density is

$$f(t) = \lambda \exp\{-\lambda t\}$$

3.1. **Construction of prior distribution for the coefficients of a linear predictor.** We will construct the joint prior distribution for the coefficients in a linear predictor using a multivariate normal distribution. In the case of an exponential lifetime distribution so that there are no additional parameters unlike the case of a Weibull distribution where there is, in addition, a shape parameter. The construction of the prior distribution will often involve elicitation of prior beliefs from a person other than the analyst. This person will often have expert knowledge of the subject matter of the study. The expert may not have great expertise in probability and statistics. Translating the beliefs of the expert into a statistical form to suit the analysis might be challenging. The expert may not be able to provide probability distributions for the parameters of interest directly but rather the researcher is required to find appropriate questions to ask the expert in order to determine the probability distribution of the parameters. It is generally considered that the expert should be asked questions about observable quantities.

The details of the elicitation method may vary depending on the type of covariates involved. We will give examples of the construction of priors with covariates of different types. We consider eliciting judgements about the survival probability at some specified time and beliefs about the effects of several covariates.

3.2. **Identifying a "baseline" case.** We first identify a suitable "baseline" case. This is a hypothetical patient with "typical" or "central" covariate values. We will then consider the effects of changing covariate values away from this baseline case by varying one covariate at a time from the baseline values. By choosing such a central case as a baseline, it becomes reasonable to suppose that the coefficients of covariates are independent of the baseline logarithm of the hazard (log - hazard) in the prior distribution and may be elicited separately. It does not follow that the coefficients of covariates are mutually independent but this may be a reasonable assumption in many cases in practice. It would be possible to extend the method to allow dependence between coefficients by considering the effects of varying more than one covariate.

3.3. **Eliciting a baseline patient and median survival time of the baseline patient.** We will elicit a baseline patient from an expert by asking the following questions.

Question 1: *"Please can you identify a set of covariate values for a "baseline" patient. This should be a "typical" case so that it is relatively easy to use your experience to express judgements about this case. The baseline case should be the case about which you have the least uncertainty. Making this case "central" or "average" enables us to ask about judgements for this case separately and then to ask other questions about the effects of changing covariate values from this case."*

Response to Question 1: *"I can identify the values for a "baseline" patient as one which gives the choice of central value of the continuous covariate, "Age=45 ", "Sex= Female", " Mode of diagnosis = Cytological" and "Location of breast cancer = both breast"*

We would want to ask about the survival probability at a specified time and it will be necessary to identify a suitable choice of time to specify. The next question is to find the median survival time for a baseline patient.

Question 2: *"Please can you identify a survival time t such that you feel that it is equally likely that a baseline patient would die before or after this time."*

Response to Question 2: *"I can identify a survival time to be $t = 3$ years. "*

Let the reference time elicited in response to Question 2, or some rounded value close to it, be t_0, although, in subsequent questions, the actual value would be given rather than the symbol "t_0".
We suppose that given a sample of baseline patients, the number who survive beyond t_0 has a binomial distribution. We wish to elicit beliefs about the parameter of this binomial distribution. [29] discusses a number of methods for eliciting

prior distribution for binomial parameters. These methods may be used but we might also use a more direct approach with questions such as the following.

Question 3: *"Please can you think about the proportion of baseline patients who would survive beyond t_0. Please give a value p such that you think that it is equally likely that the proportion is less than or greater than p."* Let the given value be p_2.

Response to Question 3: *"I think that the proportion is about* 0.5 *of the baseline patients."*

Question 4: *"If you are still thinking about the proportion of baseline patients who would survive beyond t_0, please can you give a value p such that you think that it is equally likely that the proportion is between p_2 and p and that it is greater than p."* Let the given value be p_3.

Response to Question 4: *"I think that the proportion is about* 0.4 *of the baseline patients."*

Question 5: *"Again, if you are still thinking about the proportion of baseline patients who would survive beyond t_0, please can you give a value p such that you think that it is equally likely that the proportion is between p and p_2 and that it is less than p."* Let the given value be p_1.

Response to Question 5: *"I think that the proportion is about* 0.6 *of the baseline patients. "*

We have thus elicited the three quartiles, p_1, p_2, p_3, of the prior distribution for the survival probability at time t_0. These quartiles may be transformed to give the quartiles of the prior distribution for the baseline log-hazard β_0. Since the survival probability is $\exp(-\lambda t_0)$ where $\lambda = \exp(\beta_0)$, the three quartiles for β_0 become

$$b_{0,h} = \ln\left\{-\frac{\ln p_{4-h}}{t_0}\right\} \qquad (3)$$

for $h = 1, 2, 3$.

Using a normal prior distribution, we can determine the mean as $(b_{0,1} + b_{0,3})/2$ and the standard deviation as $(b_{0,3} - b_{0,1})/1.349$. The additional value, $b_{0,2}$ can be used as a check of normality, or, at least, symmetry, since we should have $b_{0,2} \approx (b_{0,1} + b_{0,3})/2$.

The reference time elicited in response to Question Q2 is $t_0 = 3$. The response to Question Q3, Q4 and Q5 are 0.5, 0.4 and 0.6 respectively. Thus, the three quartiles p_1, p_2 and p_3 at time t_0 as 0.6, 0.5 and 0.4 and we follow (3) to get the quartiles of β_0: $\beta_{0,1} = -1.77$, $\beta_{0,2} = -1.5$ and $\beta_{0,3} = -1.23$. Using a normal prior distribution, the mean and standard deviation of β_0 are given as $\mu_0 = -1.5$ and standard deviation $\sigma_0 = 0.4$.

3.4. **Eliciting of prior distributions for other parameters in the log-hazard for non-baseline patients.** We will consider the elicitation of prior distributions for the effect of a quantitative covariate x_j. We introduce a second hypothetical patient. We will also suppose that the covariate values for this patient are the same as those for the baseline patient except in the case of x_j where the baseline value $x_{j,0}$ is replaced by $x_{j,0}+\delta_j$. A suitable value for δ_j should be determined in consultation with the expert by asking the following questions.

Question 6: *"We wish to consider the effect of a change in the value of (covariate x_j). Please can you suggest a value, different to the baseline value but where you feel able to give judgements based on your experience which may be used for this purpose. The value should not be too close to the baseline value but should be within the range of your experience."*

Response to Question 6: *"I would think of the quantitative covariate "Age" and suppose the age of "55" years as the age that is not too close to the baseline value. "*

Again we use the same method as in discussion on the construction of the prior distribution for the baseline parameter. We therefore elicit a normal distribution for the log-hazard in this case, which is $\beta_0 + \beta_j\delta_j$.

Given the underlying linear model and assuming that the value of β_j is not known with certainty, the variance of this distribution should be greater than that of β_0. If we find that it is not then the elicitation process may be repeated to try to overcome this problem. Alternatively, if the variance given in this second case is less than that for the baseline, the elicitation might be repeated with this case as the baseline. If the two variances are equal then increasing the value of $|\delta_j|$ might solve the problem.

We assume that the value given for $\mathrm{Var}(\beta_0+\beta_j\delta_j)$ is greater than that for $\mathrm{Var}(\beta_0)$ and that β_0 and β_j may be considered to be independent, we have

$$E(\beta_j) = \frac{1}{\delta_j}\{E(\beta_0 + \beta_j\delta_j) - E(\beta_0)\} \tag{4}$$

and

$$\mathrm{Var}(\beta_j) = \frac{1}{\delta_j^2}\{\mathrm{Var}(\beta_0 + \beta_j\delta_j) - \mathrm{Var}(\beta_0)\}. \tag{5}$$

The prior distribution for the coefficient of a quantitative variable is constructed using the covariate "*Age*". The response to Question Q6 is 55 years and thus $\delta_j = 10$. We elicit a normal distribution for the log-hazard of $\beta_0 + \beta_j\delta_j$. Thus, we have $E(\beta_0 + \beta_j\delta_j) = -1.1$ and $\mathrm{Var}(\beta_0 + \beta_j\delta_j) = 0.238$. We follow (4) and (5) to get the prior distribution for β_{age} as mean $m = 0.04$ and standard deviation $\sigma = 0.028$.

The prior distribution of a 2 - Level factor or a binary covariate by choosing the so-called "corner constraint" and setting the effect of the baseline level of the factor to zero. Then a prior distribution for the effect of the other level of the factor can be obtained in the same way as for a quantitative covariate, with $\delta_j = 1$.
The prior distribution for the parameter of sex, β_{sex} is constructed by thinking that it is a quantitative covariate with $\delta_j = 1$. We get the prior distribution for β_{sex} as mean $m = 0.046$ and standard deviation $\sigma = 0.15$. We perform a similar process for the other coefficients of the 2 - level factor covariates such as the mode of diagnosis (mode).

We might want to consider the case of unordered factors with $K > 2$ levels such as "Location of breast cancer (location)". We might think of constructing the prior distribution for the effects of the levels by choosing a baseline or reference level (see Scheme in Table 1). In our case, we will overcome the problem of having greater prior knowledge about Level 1 by constraining the parameters to sum to zero (see Scheme in Table 2) and making it exchangeable. We would choose having the same means, varainces and each pair of parameters would have the same covariance. We construct the prior distribution for the coefficient of "Location of breast cancer (location)" using the judgements about survival at three years so $t_0 = 3$.

Conclusion. The Bayesian paradigm provides a way of combining the prior information with data to get a posterior distribution which brings the analyst closer to what is being modeled. It is known that Bayesian analysis does not tell you how to select a prior distribution but it requires skills and techniques to translate subjective prior beliefs into a mathematically formulated prior. The choice of the prior distribution is important for obtaining useful posterior inferences. This research has therefore discussed how to incorporate prior beliefs from past experience or elicitation from expert's opinion and the construction of suitable prior distribution of the model parameters with application to survival analysis using a breast cancer data set. The Bayesian analyst takes the information about the unknown parameters from experts in the field of study and translates it into a probability distribution. The posterior distribution represents the updated belief or probability distribution, given the data (likelihood) and the prior distribution. This research discussed the structures of prior distribution for coefficient of binary factor or 2-level, coefficient of categorical factors, correlation or covariance between parameters and parameters representing a set of proportion. The prior beliefs and information were elicited from experts in the field of study by asking questions about observable quantities and responses to questions were given. The construction of a good prior distribution requires a supportive process from the experts. In practical, the implementation of an prior elicitation process and construction of prior distribution requires ensuring that the correct experts are chosen. The prior distribution was constructed for the baseline and non-baseline parameters using the breast cancer data set and assuming an exponential lifetime distribution.

The methods of prior elicitation and construction discussed in this research will minimise the causes of biased assessments of experts which are identified in Bayesian analysis. This research suggests how people can improve their probability about the parameters used in survival analysis whilst reducing the burden on experts. The research gives a direct, intuitive and meaningful approach of constructing and computing a prior distribution which will allow us to tackle enormously more complex problems since the questions and the answers are computed exactly.

Acknowledgement. The authors thank the referee(s) for their individual and collective valuable suggestions that helped to produce this quality paper.

Competing interests: The manuscript was read and approved by all the authors. They therefore declare that there is no conflicts of interest.

Funding: The authors received no financial support for the research, authorship, and/or publication of this article.

References

[1] Abbas N. & Aslam M. (2013). Eliciting hyperparameters of prior distributions for the parameters of the paired comparison models. *Pakistan Journal of Statistics and Operations Research.* **9** (1), 71–78.

[2] Albert I., Donnet S., Guihenneuc-Jouyaux C., Low-Choy S., Mengersen K. & Rousseau J. (2012). Combining expert opinions in prior elicitation. *Bayesian Analysis.* **7** (3), 503–532.

[3] Alhussain Z. I. & Oakley J. (2019). Assurance for clinical trial design with normally distributed outcomes: eliciting uncertainty about variances. *arXiv.org ¿stat ¿ arXiv.* 1702.00978v2 1–21.

[4] Azzolina D., Berchialla P., Gregori D. & Baldi I. (2021). Prior elicitation for use in clinical trial design and analysis: A literature review. *International Journal of Environmental Research and Public Health.* **18** (1833), 1–22.

[5] Barrera-Causil C. J., Correa J. C. & Marmolejo-Ramos F. (2019). Experimental investigation on the elicitation of subjective distributions. *Fronteirs in Psychology.* **10** (862), 1–10.

[6] Clemen R. T., Fischer G. W. & Winkler R. L. (2000). Assessing dependence: some experimental results. *Management Science.* **46** (8), 1100–1115.

[7] Consul J. I. & Okrinya A. B. (2019). Bayesian inference of multinomial data using sequential updating: A case study of the nigerian presidential election. *International Journal of Applied Science and Research.* **2** (3), 1–9.

[8] Cox D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society.* **B 34**, 187–220.

[9] Cuevas J. R. T. (2012). Eliciting beta prior distributions for binomial sampling.. *Revista Brasileira de Biometria*. **30** (1), 159–172.

[10] Elfadaly F. G. & Garthwaite P. H. (2013). Eliciting Dirichlet and Connor-Mosimann prior distributions for multinomial models. *TEST*. **22** (4), 628–646.

[11] Garthwaite P. H. & Dickey J. M. (1992). Elicitation of Prior Distributions for Variable-Selection Problems in Regression. *The Annals of Statistics*. **20** (4), 1697–1719.

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

[13] Hemming V., Burgman M. A., Hanea A. M., McBride M. F. & Wintle B. C. (2018). A practical guide to structured expert elicitation using the idea protocol. *Methods of Ecology and Evaluation*. **9**, 169–180.

[14] James A., Choy S. L. & Mengersen K. (2010). Elicitator: An expert elicitation tool for regression in ecology. *Environmental Modelling & Software*. **25**, 129–145.

[15] John K. K. (2015). *Doing Bayesian Data Analysis, Second Edition: A Tutorial with R, JAGS, and Stan*. 2nd edition., Academic Press / Elsevier. ISBN: 9780124058880.

[16] Johnson S. R., Tomlinson G. A., Hawker G. A., Grantonf J. T. & Feldman B. M. (2010), 'Methods to elicit beliefs for bayesian priors: a systematic review. *Journal of Clinical Epidemiology*. **63**, 355–369.

[17] Kahle D., Stamey J., Natanegara F., Price K. & Han B. (2016). Facilitated prior elicitation with the wolfram cdf. *Biometrics & Biostatistics International Journal*. **3** (6), 1–21.

[18] Kiefer N. M. (2011). Default estimation, correlated defaults, and expert information. *Journal of Applied Econometrics*. **26** (2), 173–192.

[19] Lau H. S. & Lau A. H. L. (1998). An improved pert-type formula for standard deviation. *IIE Trans*. **30**, 273–275.

[20] Lee S. & Lim H. (2019). Review of statistical methods for survival analysis using genomic data. *Genomics and informatics*. **17** (4), – - –.

[21] Nogueira M. C. S. (2004). Orthogonal constrasts: Definitions and concepts. *Scientia Agricola*. **61**, 118–124.

[22] Oguntunde P., Adejumo A. & Okagbue H. (2017). Breast cancer patients in Nigeria: Data exploration approach. *Data in Brief*. **15**, 47–57.

[23] Raveendran N. & Sofronov G. (2021). A Markov Chain Monte Carlo Algorithm for Spatial Segmentation. *Information*. **12** (58), 1–14.

[24] Revie M., Bedford T. & Walls L. (2010). Evaluation of elicitation methods to quantify Bayes linear models. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*. **224** (4), 322–332.

[25] Sun C. Q., Prajna N. V., Krishnan T., Mascarenhas J., Rajaraman R., Srinivasan M., Raghavan A., O'Brien K. S., Ray K. J., McLeod S. D., Porco T. C., Acharya N. R. & Lietman T. M. (2013). Expert prior elicitation and Bayesian analysis of the mycotic ulcer treatment trial I. *Clinical and Epidemiologic Research*. **54** (6), 4167-4173.

[26] Thall P. F., Ursino M., Baudouin V., Alberti C. & Zohar S. (2019). Bayesian treatment comparison using parametric mixture priors computed from elicited histograms. *Statistical Methods in Medical Research*. **28** (2), 404–418.

[27] Vinaitheerthan R. (2016). Overview of frequentist and bayesian approach to survival analysis. *Applied Medical Informatics*. **38**, 25–38.

[28] Williams C. J., Wilson K. J. & Wilson N. (2021). A comparison of prior elicitation aggregation using the classical method and shelf. *Journal of the Royal Statistical Society:*. **Series A**, 1–21.

[29] WINKLER R. L. (1967). The assessment of prior distributions in Bayesian analysis. *Journal of the American Statistical Association*. **62**, 776–800.

[30] ZAPATA-VAZQUEZ R. E., O'HAGAN A. & SOARES BASTOS L. (2014). Eliciting expert judgements about a set of proportions. *Journal of Applied Statistics*. **41** (9), 1919–1933.