MAIN PAPER

Pharmaceutical Statistics

(wileyonlinelibrary.com) DOI: 10.1002/pst.1741

Published online 23 February 2016 in Wiley Online Library

Bayesian robustness in meta-analysis for studies with zero responses

F. J. Vázquez,[a]* E. Moreno,[b] M. A. Negrín,[a] and M. Martel[a]

Statistical meta-analysis is mostly carried out with the help of the random effect normal model, including the case of discrete random variables. We argue that the normal approximation is not always able to adequately capture the underlying uncertainty of the original discrete data. Furthermore, when we examine the influence of the prior distributions considered, in the presence of rare events, the results from this approximation can be very poor. In order to assess the robustness of the quantities of interest in meta-analysis with respect to the choice of priors, this paper proposes an alternative Bayesian model for binomial random variables with several zero responses. Particular attention is paid to the coherence between the prior distributions of the study model parameters and the meta-parameter. Thus, our method introduces a simple way to examine the sensitivity of these quantities to the structure dependence selected for study. For illustrative purposes, an example with real data is analysed, using the proposed Bayesian meta-analysis model for binomial sparse data.

Keywords: Bayesian inference; noninformative priors; Sarmanov and intrinsic link distribution; testing on meta-parameters

1. INTRODUCTION

Clinical trials for medical treatments conducted in different centres/studies provide heterogeneous sample information that needs to be combined. In the last few decades, the number of clinical trials, concerning an enormous variety of medical treatments, has risen exponentially, and methodologies have had to be developed to combine the sample information from disparate clinical trials. Meta-analysis is the statistical procedure to bring together heterogenous studies of a medical treatment for a single disease, to determine what can fairly be concluded about the fundamental question that each study seeks to address: the effectiveness of the treatment.

Among the most widely used methods in practice are those termed 'random effects'. The frequentist random effect model is based on the DerSimonian–Laird estimator [1], but the latter does present serious concerns, and alternative procedures have been proposed ([2–4], among many others). On the other hand, Bayesian random effects models, among many other advantages, are very suitable when data are sparse and the number of studies presented is small [5].

Following Sutton and Abrams [6], the Bayesian random effects model for meta-analyses can be accommodated to the following hierarchical model:

$$\begin{aligned} x_i &\sim \mathcal{N}\left(\theta_i, \sigma_i^2\right) \quad i = 1, \ldots, k \\ \theta_i &\sim \mathcal{N}\left(\theta, \sigma^2\right), \\ \theta &\sim [-,-]\ \sigma^2 \sim [-,-], \end{aligned} \tag{1}$$

where, for each study, $i \in \{1, ..., k\}$ and the observed effect x_i is assumed to be normally distributed with parameters θ_i and σ_i^2, where θ_i represents the treatment effectiveness conditional on study i and σ_i^2 its variance. Similarly, θ represents the unconditional treatment effectiveness, that is, the pooled effect size and σ^2 its variance, and $[-,-]$ indicates a prior density to be assigned.

Thus, the assignment of prior densities in meta-analyses constitutes a source of uncertainty, and it is important to develop straightforward methods to analyse the problem of expressing uncertainty regarding prior information. Different prior densities can be used, generating varying results; consequently, there is no definitive analysis, and a sensitivity analysis is always required [6]. The problem is aggravated in the case of rare events, that is, binary outcomes with a high frequency of occurrence of zero responses [7].

The Cochrane Collaboration recently considered the case of adverse events. Data on adverse effects tend to be sparse, and this is a clinical area where specific meta-analytical techniques need to be investigated in depth [8]. Sweeting *et al.* [9] pointed out that in the presence of rare events, the results obtained from the commonly used hierarchical normal model in (1) can be very poor, while the logit approach ([10–12], among many others) does not always adequately reflect the underlying uncertainty of the original discrete data when proportions are small. In this respect, too, Bradburn *et al.* [13] found that alternative procedures correcting the lack of continuity and relaxing the assumption of normality (DerSimonian and Laird, Mantel–Haenszel, inverse variance or Peto methods, among others) were also biassed when data were sparse.

[a] *Department of Quantitative Methods and TiDES Institute, University of Las Palmas de GC, Las Palmas de Gran Canaria, 35017, Spain*

[b] *Department of Statistics and O.R., University of Granada, Granada, 18071, Spain*

**Correspondence to: F. J. Vázquez, Department of Quantitative Methods and TiDES Institute, University of Las Palmas de GC, Las Palmas de Gran Canaria, 35017, Spain.*
E-mail: fjvpolo@dmc.ulpgc.es

Very few papers have specifically examined the question of Bayesian robustness with respect to priors in meta-analysis, and those who have considered this topic have mostly studied a catalogue of distributions containing a small finite number of proposed vague priors and then examined the behaviour of the quantities of interest over the finite family of priors considered [7]. Lambert *et al*. [5] conducted a simulation study based on several different vague priors and reported that in meta-analysis with a small number of studies, 'the choice of vague prior distribution can lead to a marked variation in results'. With respect to previous studies, we address the problem of sensitivity from a different standpoint. It is now known that robust Bayesian methodology can be used to study how a given model reacts to deviations of some inputs, usually related to the prior distribution. Thus, we can measure the model's reaction to doubts in the expert's opinion by considering a family Γ of plausible priors. It is then of interest to examine how the relativities for priors in such a class behave.

In the present study, we proceed step by step to develop an alternative Bayesian model with a real conceptual improvement in its assumptions. First, we present an objective Bayesian model for meta-analysis with sparse binomial random variables, in order to overcome the concerns detected in the usual logit transformed methods for meta-analysis. The binomial sampling models are complemented with well-established default or 'automatic' prior distributions for θ_i and θ. Second, to complement the Bayesian model for meta-analysis, we have to choose the link distribution between θ_i and θ. This is a key point in Bayesian analysis and one that has received little attention in the literature. Finally, we derive a simple relationship between the commonly used between-study variance and the variance of the proposed linking distribution, which allows us to introduce a class of plausible priors in order to study the Bayesian sensitivity of the model in a rigorous and simple way and also facilitates implementation.

The rest of this paper is structured as follows. Section 2 presents the Bayesian model for binomial sparse data. This model requires us to derive the distributions linking the parameter θ_i of the binomial distribution in the *i*th study and the meta-parameter θ. The posterior distributions of θ and the meta-predictive distribution of the meta-variable x are then calculated. In Section 3, the Bayesian procedure for testing common hypotheses in meta-analysis is developed. Once the elements of this new model have been established, Section 4 then describes an improved way to study the sensitivity of the choice of priors in meta-analysis. In Section 5, the Bayesian meta-analysis procedure is applied to sparse real data sets from [14] and [15]. Finally, Section 6 contains some concluding remarks.

2. AN ALTERNATIVE OBJECTIVE BAYESIAN MODEL FOR BINOMIAL SPARSE DATA

In this section, we present an alternative model for the case of sparse count data, adopting the standpoint of the Bayesian model choice methodology [16]. Let us consider the Bayesian model for the observational study *i* in a collection of *k* studies, given by

$$M_i : \left\{\text{Bin}(x_i|n_i,\theta_i),\ \pi^J(\theta_i) = \text{Beta}(\theta_i|1/2,1/2)\right\}, \quad i = 1,\ldots,k, \tag{2}$$

where the outcome variable x_i is a binomial response variable taking a value between 0 and n_i, where n_i is the number of patients in study *i*. We assume that the prior information on the θ_i's is weak and, hence, the prior distribution $\pi^J(\theta_i)$ is the objective Jeffreys' prior beta distribution $\text{Beta}(\theta_i|1/2,1/2)$, this being one of the most commonly used forms as a non-informative prior distribution when non-elicited hyperparameters are required.

Then, a meta-model is directly induced by models for observational studies. The latent meta-variable x means the result we would obtain when a treatment is applied to a patient in a virtual centre that does not suffer of the random heterogeneity between studies effect. This meta-variable is a way of modelling those situations where we deal with heterogenous samples. That is, the Bayesian meta-model M_0 for the $0-1$ unobservable treatment effectiveness x is given by

$$M_0 : \left\{\text{Ber}(x|\theta),\ \pi^J(\theta)\right\}, \tag{3}$$

where the meta-parameter θ represents the unconditional probability of success of the treatment, $\text{Ber}(x|\theta) = \theta^x(1-\theta)^{1-x}$ is the Bernoulli distribution and $\pi^J(\theta)$ is the Jeffreys prior for θ. The consideration of the meta-model is important because the predictive distribution of x gives us the consequences for a new patient under the treatment whatever is the centre [17].

To carry out any posterior inference on the meta-parameter θ, its posterior distribution must be computed, and following the natural structure of random effects models in (1), we need a linking distribution between θ_i and θ.

2.1. Linking the experimental parameter θ_i with the meta-parameter θ

We argue that because the marginals of θ_i and θ are given, the bivariate distribution $\pi(\theta_i,\theta)$ must be chosen in the Fréchet class of bivariate distributions with given marginals. This is a large and well-studied class for which a considerable body of literature has been developed to reduce the class to subclasses, called copulas, that model particular dependence structures between θ_i and θ [18]. Properties of the copulas that are typically analysed are their lower and upper bounds and the range of correlation coefficients between θ_i and θ.

We follow the same strategy and consider classes of bidimensional distributions with given marginals. For obvious reasons, we only consider positive correlations between θ_i and θ. Mathematically, we are looking for a bivariate distribution $\pi(\theta_i,\theta)$ such that the integral equations

$$\int_0^1 \pi(\theta_i,\theta)d\theta_i = \pi(\theta),\ \int_0^1 \pi(\theta_i,\theta)d\theta = \pi(\theta_i) \tag{4}$$

hold for given $\pi(\theta)$ and $\pi(\theta_i)$.

2.2. The case of Jeffreys marginals

Let us now consider a meta-analysis with Jeffreys marginal prior distributions for θ_i and θ. If the Jeffreys priors $\pi^J(\theta_i)$ and $\pi^J(\theta)$ are used in (4), a family of distributions that is well adapted to our problem and easy to handle is the one described by Sarmanov [19],

$$\begin{aligned}\pi^{SJ}(\theta_i,\theta|\alpha) = {} & \frac{1}{\pi^2}(\theta(1-\theta))^{-1/2}(\theta_i(1-\theta_i))^{-1/2} \\ & \times\left(1+\frac{\alpha}{4}(2\theta_i-1)(2\theta-1)\right),\end{aligned} \tag{5}$$

for $0 \leqslant \alpha \leqslant 4$. To make explicit the dependency on the Jeffreys prior, we call this the Sarmanov–Jeffreys family. Interestingly, very

few papers in meta-analysis make use of the Sarmanov family, although in other scenarios, [20] and [21] are exceptions to this.

By construction, this family $\{\pi^{SJ}(\theta_i,\theta|\alpha), 0 \leqslant \alpha \leqslant 4\}$ is such that

$$\int_0^1 \pi^{SJ}(\theta_i,\theta|\alpha)d\theta_i = \pi^{J}(\theta), \int_0^1 \pi^{SJ}(\theta_i,\theta|\alpha)d\theta = \pi^{J}(\theta_i)$$

for any $0 \leqslant \alpha \leqslant 4$. Furthermore, the Pearson's correlation coefficient of θ_i and θ is given by $\rho^{SJ}(\theta_i,\theta|\alpha) = \frac{\alpha}{8}$, and hence,

$$0 \leqslant \rho^{SJ}(\theta_i,\theta|\alpha) \leqslant \frac{1}{2}.$$

Also, the conditional variance is obtained by

$$\mathbb{V}^{SJ}(\theta_i|\theta,\alpha) = \frac{32-\alpha^2(2\theta-1)^2}{256}. \quad (6)$$

Therefore, we observe that there is a relationship between the variance and the Pearson's correlation coefficient, which can help the practitioner model the heterogeneity.

Following Moreno *et al*. [16], we need to enlarge this family in order to account for correlations of over $1/2$. For this purpose, the Intrinsic–Jeffreys (IJ) class is well-suited. It is obtained by applying the intrinsic methodology to the Jeffreys prior and is given by

$$\pi^{IJ}(\theta_i,\theta|t) = \frac{1}{\pi}\sum_{z=0}^{t}\binom{t}{z}\frac{(\theta_i\theta)^{z-1/2}[(1-\theta_i)(1-\theta)]^{t-z-1/2}}{B(z+1/2,t-z+1/2)}, \quad (7)$$

where $t = 1,2,\ldots$, and $B(a,b)$ is the beta function. Given t, each bidimensional intrinsic distribution in (7) is a beta-binomial mixture, and by integrating out with respect to either θ_i or θ, we can easily obtain the Jeffreys marginal distributions. Furthermore, the between-study heterogeneity captured by the class $\{\pi^{IJ}(\theta_i,\theta|t),\ t = 1,2,\ldots\}$ is considered by the conditional variance

$$\mathbb{V}^{IJ}(\theta_i|\theta,t) = \frac{(1+2t)(1-4t\theta(\theta-1))}{4(t+1)^2(t+2)} \quad (8)$$

and the correlation coefficient

$$\rho^{IJ} = \frac{t}{t+1}, \qquad t = 1,2,\ldots. \quad (9)$$

Observe that $1/2 \leqslant \rho^{IJ} \leqslant 1$ is an increasing function of t, as t ranges in $\{1,2,\ldots\}$. Equivalently, the variance in (8) is a decreasing function of ρ, for any θ.

2.3. Linking distributions

When the marginal densities of θ_i and θ are Jeffreys as in (5), we note that the value $\alpha = 0$, in the Sarmanov–Jeffreys case, models independence between θ_i and θ, a situation that is of no interest in meta-analysis. This suggests we should integrate out α in the family with respect to a hyperprior $\pi(\alpha)$ that penalises independence between θ_i and θ.

A natural hyperprior for $\alpha \in [0,4]$ is a transformed Beta$(\alpha|a,b)$, that is, $\pi(\alpha) \propto \alpha^{a-1}(4-\alpha)^{b-1}$. An appropriate assignment of the parameters a and b may help us 'penalise' unwanted scenarios of independence and favour more realistic situations. A particular choice of the hyperparameters a and b yields a beta J-shaped distribution with a left tail. The case $a > 2, b = 1$ is particularly well adjusted to the aforementioned comments; this case coincides with a convex power density function in which the mode is at 4. Furthermore, in order to model departures from the independence scenario more smoothly, we can consider beta distributions with $1 \leqslant a \leqslant 2$ and $b = 1$. Members of this selected beta family are the usual vague uniform prior, left skewed distributions and left skewed triangle distributions as a varies from 1 to 2, respectively.

The full Bayes Sarmanov–Jeffreys linking distribution derived by integrating out α with a transformed Beta$(\alpha|a,b = 1)$ is obtained as follows:

$$\pi^{SJ}(\theta_i|\theta) = \frac{1}{\pi}\theta_i^{-1/2}(1-\theta_i)^{-1/2}\left(1+\frac{a}{a+1}(2\theta_i-1)(2\theta-1)\right). \quad (10)$$

Analogously, a natural hyperprior for $\rho \in [1/2,1]$ is an inverted J-shaped beta distribution with a right tail by which unrealistic high-correlation scenarios are penalised and the usual uninformative uniform prior could be included. A beta distribution with parameters $a = 1$ and b is then selected, and the corresponding discrete distribution induced on t is given by

$$\pi(t) = \left(\frac{2}{t+1}\right)^b - \left(\frac{2}{t+2}\right)^b, \quad t = 1,2,\ldots. \quad (11)$$

The full Bayes Intrinsic–Jeffreys linking distribution derived by integrating out t in (7) with respect to prior (11) is obtained as follows:

$$\pi^{IJ}(\theta_i|\theta) = \sum_{t=1}^{\infty}\left(\left\{\left(\frac{2}{t+1}\right)^b - \left(\frac{2}{t+2}\right)^b\right\}\left(\sum_{z=0}^{t}\text{Beta}(\theta_i|\boldsymbol{\gamma})\cdot\text{Bin}(\theta|t,z)\right)\right), \quad (12)$$

where $\boldsymbol{\gamma} = (z+1/2, t-z+1/2)$ and Beta$(\cdot|r,s)$ and Bin$(\cdot|n,m)$ are the pdf and pmf of the beta and binomial distributions, respectively.

2.4. Likelihood and posterior distribution of the meta-parameter

Assuming that the samples $\{x_i, i = 1,\ldots,k\}$ are independent, conditional on θ_i, the probability distribution of the whole data sets $(\boldsymbol{x},\boldsymbol{n}) = \{(x_i,n_i),\ i = 1,2,\ldots,k\}$ for a specific link distribution $\pi(\theta_i|\theta)$ is obtained as

$$\Pr(\boldsymbol{x}|\boldsymbol{n},\theta) = \prod_{i=1}^{k}\int_0^1\binom{n_i}{x_i}\theta_i^{x_i}(1-\theta_i)^{n_i-x_i}\pi(\theta_i|\theta)d\theta_i.$$

This is the likelihood of θ for the data sets $(\boldsymbol{x},\boldsymbol{n})$. Thus, after some tedious algebra, for the Sarmanov–Jeffreys $\pi^{SJ}(\theta_i|\theta)$, the likelihood is

$$\Pr^{SJ}(\boldsymbol{x}|\boldsymbol{n},\theta)=\prod_{i=1}^{k}\left\{\frac{1}{\pi}\frac{\Gamma\left(x_i+\frac{1}{2}\right)\Gamma\left(n_i-x_i+\frac{1}{2}\right)}{\Gamma(x_i+1)\Gamma(n_i-x_i+1)}\left(1-\frac{a(2\theta-1)}{a+1}\right)+\frac{2a}{\pi(a+1)}\frac{2\theta-1}{n_i+1}\frac{\Gamma\left(x_i+\frac{3}{2}\right)\Gamma\left(n_i-x_i+\frac{1}{2}\right)}{\Gamma(x_i+1)\Gamma(n_i-x_i+1)}\right\}. \tag{13}$$

The likelihood of θ for the Intrinsic–Jeffreys link distribution is

$$\Pr^{IJ}(\boldsymbol{x}|\boldsymbol{n},\theta)=\prod_{i=1}^{k}\binom{n_i}{x_i}\sum_{t=1}^{\infty}\left(\left(\frac{2}{t+1}\right)^b-\left(\frac{2}{t+2}\right)^b\right)(1-\theta)^t\times\frac{\mathrm{B}(x_i+1/2,n_i+t-x_i+1/2)}{\mathrm{B}(t+1/2,1/2)}{}_3[F_2\left(\boldsymbol{\delta},\boldsymbol{\epsilon},\frac{\theta}{\theta-1}\right), \tag{14}$$

where ${}_3F_2\left(\boldsymbol{\delta},\boldsymbol{\epsilon},\frac{\theta}{\theta-1}\right)$ denotes the generalised hypergeometric function with $\boldsymbol{\delta}=(1/2-t,-t,1/2+x_i)$ and $\boldsymbol{\epsilon}=(1/2,1/2-t-n_i+x_i)$.

Finally, assuming that the links under the Jeffreys marginals are *a priori* equally likely, the likelihood of θ is given by

$$\Pr(\boldsymbol{x}|\boldsymbol{n},\theta)=\frac{1}{2}\Pr^{SJ}(\boldsymbol{x}|\boldsymbol{n},\theta)+\frac{1}{2}\Pr^{IJ}(\boldsymbol{x}|\boldsymbol{n},\theta). \tag{15}$$

Consequently, the posterior probability of θ for the Jeffreys prior for θ is given by

$$\pi^J(\theta|\boldsymbol{n},\boldsymbol{x})=\frac{\theta^{-1/2}(1-\theta)^{-1/2}\Pr(\boldsymbol{x}|\boldsymbol{n},\theta)}{\int_0^1\theta^{-1/2}(1-\theta)^{-1/2}\Pr(\boldsymbol{x}|\boldsymbol{n},\theta)d\theta}. \tag{16}$$

3. HYPOTHESIS TESTING: TESTING THE EQUALITY OF TREATMENT EFFECTIVENESS

Standard methodologies (classical and Bayesian) typically address this problem as an estimation problem addressing either the difference of the meta-effectiveness treatment parameters or the odds ratio (OR). The accept/reject criterion consists of whether the corresponding credible interval contains or does not contain the singular value 0 (similarly, the value 1 for the case of the OR). However, it is well established [22] that testing problems could be addressed using Bayes factors.

Then, assuming that we are interested in testing the equality of the meta-effectiveness of two treatments T_1 and T_2, given the samples $(\boldsymbol{x_1},\boldsymbol{n_1})$ and $(\boldsymbol{x_2},\boldsymbol{n_2})$ of the treatment effectiveness reported in k studies.

Let $\{\Pr(\boldsymbol{x_1}|\boldsymbol{n_1},\zeta),\ \pi(\zeta)\}$ and $\{\Pr(\boldsymbol{x_2}|\boldsymbol{n_2},\xi),\ \pi(\xi)\}$ be the likelihoods and priors of the meta-effectiveness ζ and ξ of treatments T_1 and T_2, respectively. Testing the null $H_0:\zeta=\xi$ versus the unrestricted alternative $H_1:(\zeta,\xi)$ is equivalent to the model selection problem between the Bayesian models M_0 and M_1 given by

$$M_0:\left\{\Pr(\boldsymbol{x_1}|\boldsymbol{n_1},\theta)\Pr(\boldsymbol{x_2}|\boldsymbol{n_2},\theta),\ \pi^J(\theta)\right\}$$

and

$$M_1:\left\{\Pr(\boldsymbol{x_1}|\boldsymbol{n_1},\zeta)\Pr(\boldsymbol{x_2}|\boldsymbol{n_2},\xi),\ \pi^J(\zeta)\pi^J(\xi)\right\}.$$

The model selector is usually taken as the posterior probability of the null model M_0. Assuming that *a priori* $\Pr(M_0)=\Pr(M_1)=1/2$, the posterior probability of M_0 is given by

$$\Pr(M_0|\boldsymbol{x_1},\boldsymbol{n_1},\boldsymbol{x_2},\boldsymbol{n_2})=\frac{1}{1+B_{10}}, \tag{17}$$

where B_{10}, the Bayes factor to compare M_1 with M_0, is given by

$$B_{10}(\boldsymbol{x_1},\boldsymbol{n_1},\boldsymbol{x_2},\boldsymbol{n_2})=\frac{m_1(\boldsymbol{x_1},\boldsymbol{n_1},\boldsymbol{x_2},\boldsymbol{n_2})}{m_0(\boldsymbol{x_1},\boldsymbol{n_1},\boldsymbol{x_2},\boldsymbol{n_2})}=\frac{\int_0^1\int_0^1\Pr(\boldsymbol{x_1}|\boldsymbol{n_1},\zeta)\Pr(\boldsymbol{x_2}|\boldsymbol{n_2},\xi)\pi^J(\zeta)\pi^J(\xi)d\zeta d\xi}{\int_0^1\Pr(\boldsymbol{x_1}|\boldsymbol{n_1},\theta)\Pr(\boldsymbol{x_2}|\boldsymbol{n_2},\theta)\pi^J(\theta)d\theta}.$$

The meaning of $B_{10}(\boldsymbol{x_1},\boldsymbol{n_1},\boldsymbol{x_2},\boldsymbol{n_2})$ is the ratio of the likelihood of model M_1 and model M_0 for the data $(\boldsymbol{x_1},\boldsymbol{n_1},\boldsymbol{x_2},\boldsymbol{n_2})$. The optimal decision of this model selection problem is, under the $0-1$ loss function, to reject the null H_0 iff $\Pr(M_0|\boldsymbol{x_1},\boldsymbol{n_1},\boldsymbol{x_2},\boldsymbol{n_2})<1/2$ or equivalently to reject the null iff $B_{10}(\boldsymbol{x_1},\boldsymbol{x_2})>1$.

We note that the aforementioned decision is taken conditional on the data $\{(\boldsymbol{x_i},\boldsymbol{n_i}),i=1,2\}$. Furthermore, we note that $\Pr(M_0|\boldsymbol{x_1},\boldsymbol{n_1},\boldsymbol{x_2},\boldsymbol{n_2})$ depends explicitly on the prior $\pi(\cdot)$ and implicitly on the link between θ_i and θ utilised to construct the likelihoods $\Pr(\boldsymbol{x_1}|\boldsymbol{n_1},\zeta)$ and $\Pr(\boldsymbol{x_2}|\boldsymbol{n_2},\xi)$.

4. A WAY TO STUDY BAYESIAN ROBUSTNESS

As commented earlier, the few papers that have examined robustness with respect to priors in meta-analysis are mainly based on the study of a finite set of priors (frequently, no more than 10), and then examining the behaviour of some quantities of interest over this catalogue of distributions [5,7]. This empirical approach can be improved using a Bayesian robust philosophy.

Let us now observe that the use of the family of Sarmanov–Jeffreys priors introduces an easily handled relationship between the correlation coefficient ρ, the hyperparameter α (using (6)) and the between-study heterogeneity measured through the variance

$$\mathbb{V}^{SJ}(\theta_i|\theta,\rho)=\frac{1}{8}\left(1-2\rho^2(2\theta-1)^2\right)$$

or

$$\mathbb{V}^{SJ}(\theta_i|\theta,\alpha)=\frac{32-\alpha^2(2\theta-1)^2}{256}.$$

Analogously, we find in the Intrinsic–Jeffreys family the following relationship:

$$\mathbb{V}^{IJ}(\theta_i|\theta,\rho)=\frac{(\rho-1)^2(\rho+1)(\rho(1-2\theta)^2-1)}{4\rho(\rho-2)}$$

or

$$\mathbb{V}^{IJ}(\theta_i|\theta,t)=\frac{(1+2t)(1-4t\theta(\theta-1))}{4(t+1)^2(t+2)}.$$

According to Bayesian robustness analysis, the prior on the variance can be modelled by specifying a class Γ of priors instead of a single prior. An alternative way to do this consists in introducing a class of plausible priors on the hyperparameters α and t. Given that the Sarmanov/Intrinsic–Jeffreys class is well adapted, this is equivalent to choosing plausible ranges of values for the a and b parameters in the link distributions in (10) and (12).

Two further aspects characterise the problem described: on the one hand, the measure by which the sensitivity is evaluated is determined by the self-same structure of the problem, which in

this case is, in practice, point estimates, the predictive effectiveness distribution and/or the probability of equality of treatment effectiveness. Furthermore, the *a priori* class of possible distributions must embody a certain defensible, 'familiar' character for the practitioner. In this sense, the classes of transformed beta distributions given in Subsection 2.3 are presented as plausible alternatives. As commented earlier, this class penalises unrealistic independence situations and presents more realistic scenarios when a left tail is expected:

$$\Gamma_a = \{\,\text{Beta}(a,1) \mid a \geqslant 1\}\,. \tag{18}$$

Similarly, for larger values of correlation $\rho \geqslant 1/2$ where the Intrinsic–Jeffreys family works well for the class of beta distributions with parameters $a = 1$ and $b \geqslant 1$, it seems appropriate to conduct a Bayesian robustness study:

$$\Gamma_b = \{\text{Beta}(1,b) \mid b \geqslant 1\}\,. \tag{19}$$

The inference is assumed to be robust if the posterior quantities of interest do not change excessively over the class of priors selected. For instance, in testing the equality of effectiveness of two treatments, a robust Bayesian analysis of $\Pr(M_0|\text{data})$ over class $\Gamma_{a,b} = \Gamma_a + \Gamma_b$ can of course be performed by simply computing the interval

$$\mathcal{I} = \left(\inf_{a,b\geqslant 1} \Pr(M_0|\text{data},a,b),\ \sup_{a,b\geqslant 1} \Pr(M_0|\text{data},a,b) \right), \tag{20}$$

where the smaller the length of the interval $\mathcal{I}$, the larger the robustness of the posterior probability of equality with respect to $\Gamma_{a,b}$. We denote by $\Gamma_a + \Gamma_b$ to refer jointly to a prior from class Γ_a and one from Γ_b. Observe that the same is applied when we are interested in measuring the sensitivity of other measures of interest: effectiveness and OR, among many others.

As a measure of sensitivity (or its absence), we use the following useful adapted index [23]:

$$S_{a,b} = \frac{U-L}{2\cdot\psi_{a,b}} \times 100\%,$$

where U and L are the upper and lower bounds of the expected (posterior) effectiveness over the class of priors $\Gamma_{a,b}$ and $\psi_{a,b}$ is the expected (posterior) effect when base priors π_a and π_b in class $\Gamma_{a,b}$ are considered. Index S has a very simple interpretation and can be thought of as the percentage variation of effects around a particular selected prior, as the prior varies in $\Gamma_{a,b}$. A 'small' value of S reflects robustness with respect to a particular base prior selected, and then the practitioner will have few doubts concerning the effects obtained from the 'true' base prior. On the other hand, a 'large' value of S is an indication that additional refinement of the priors is needed. Similar comments could be applied when other quantities of interest are used, for instance, the *OR* and the probability of equality of treatment effects.

5. AN EXAMPLE WITH REAL DATA

As an illustration of the aforementioned considerations, we present the analysis of a real data set extracted from [14], corresponding to the use of a diuretic administered to women at risk of preeclampsia. This case study is frequently used to present classical procedures for estimating a common treatment effect in a heterogeneous sample. Cornell *et al*. [15] also presented the Bayesian fixed and random effects for estimating and comparing the results.

Example 1

We consider data from nine Randomised Controlled Trials (RCTs) provided in [14]. Two treatments were considered, termed 'treatment group' and 'control group', respectively.

The numbers of patients 'treated' in the nine studies were $\boldsymbol{n_1} = \{131, 335, 57, 34, 1011, 1370, 506, 108, 153\}$ and the numbers of observed perinatal deaths $\boldsymbol{x_1} = \{1, 6, 3, 1, 14, 24, 14, 0, 0\}$. Likewise, the numbers of patients in the control group were $\boldsymbol{n_2} = \{136, 110, 48, 40, 760, 1336, 524, 103, 102\}$, and the numbers of perinatal deaths observed were $\boldsymbol{x_2} = \{4, 3, 2, 3, 13, 19, 16, 0, 0\}$. The data set also presented stillbirths and neonatal deaths during the first month (in practice, in such trials, the perinatal mortality outcome is considered as 'stillbirths' + 'neonatal deaths'). The data for these two outcomes were $\boldsymbol{n_1^S} = \boldsymbol{n_1^N} = \{131, 335, 57, 34, 1011, 506, 108, 153\}$, and $\boldsymbol{x_1^S} = \{1, 3, 1, 0, 6, 6, 0, 0\}$, $\boldsymbol{x_1^N} = \{0, 3, 2, 1, 8, 8, 0, 0\}$; and $\boldsymbol{n_2^S} = \boldsymbol{n_2^N} = \{136, 110, 48, 40, 760, 524, 103, 102\}$, and $\boldsymbol{x_2^S} = \{2, 2, 1, 1, 5, 9, 0, 0\}$, and $\boldsymbol{x_2^N} = \{2, 1, 1, 2, 8, 7, 0, 0\}$, for stillbirths (superscript *S*) and neonatal deaths (superscript *N*), respectively. Observe that this last set of data only contains eight RCT studies because one of them only presented aggregated data for perinatal mortality.

A robustness Bayesian analysis with respect to the class of transformed beta priors for parameter α and t in (18) and (19) for the expected posterior effectiveness renders the graphics shown in Figure 1, referring to perinatal mortality data.

In practice, where unrealistically high correlations are obviously unattainable, the computational cost can be significantly reduced. For example, for the case study in this example, we can set an upper bound of $t_{\max} = 80$, for the parameter t, which represents an upper bound for the correlation of $80/81 \approx 0.99$, which we can obviously take as a case of maximum linear correlation. At the same time, the classes Γ_a and Γ_b must contain plausible distributions and/or those with which we want to make comparisons. They should be neither too small, suggesting a false degree of robustness, nor too large, in which case a high degree of sensitivity would be immediately apparent because of the presence of an inappropriate distribution. For example, in the case presented, with the aforementioned upper bound, it is not necessary to consider the entire possible range of values for b because with values $b \leqslant 4$, we are guaranteed full coverage of all possible values of t, because $\Pr(T \leqslant t_{\max}) \approx 1$. Therefore, we study the Bayesian robustness of the quantities of interest for this example, for the classes of priors:

$$\Gamma_a = \{\,\text{Beta}(a,1) |\, a \geqslant 1\} \text{ and } \Gamma_b = \{\text{Beta}(1,b)\,|1 \leqslant b \leqslant 4\}\,.$$

Obviously other, higher values can be considered for the bounds of t and b, but these would be much more costly in computation times, while the results obtained would be virtually the same.

All computations were conducted using `Mathematica` `package` (`v.9`). The code is available from the authors upon request.

Observe that Figure 1 shows the quasi-flat behaviour of the expected effectiveness for the 'treated group' and the control group as a increases, given b, obtaining ranges of 1.4–4.4% and 1.7–4.5%, respectively.

In this particular case of data from [14], the ranges obtained in each family of priors are shown in Table I. Table I also shows the OR estimates under standard methods extracted from [15].

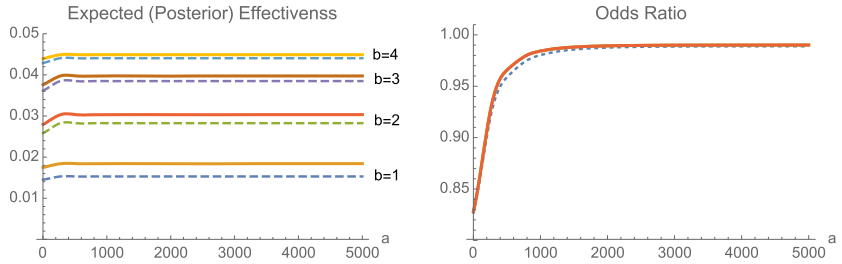

Figure 1. Left panel: Posterior expected effectiveness for treatment group (dashed line) and control (solid line) for values of the hyperparameter $a \geqslant 1$ and given discrete values of $b \in \{1, 2, 3, 4\}$ (bottom to top). Right panel: Odds ratio for the case in the left panel ($b = 1, 2, 3, 4$).

Table I. OR point estimation (extracted from [15]) and ranges for effectiveness, OR and index S in class $\Gamma_{a,b}$ in [14].

OR point estimation under common standard methods (95% CI)		
DerSimonian–Laird	0.60	(0.40, 0.89)
Fixed effects	0.67	(0.56, 0, 80)
Hierarchical Bayesian random effects	0.60	(0.34, 1.08)
Robust Bayesian estimation		
Group	Effectiveness	*S*
Treated	1.4–4.4%	34–106.1%
Control	1.7–4.5%	30.85–80.6%
OR	(0.82, 0.92)	8.3–9.9%

OR, odds ratio; CI, confidence interval.

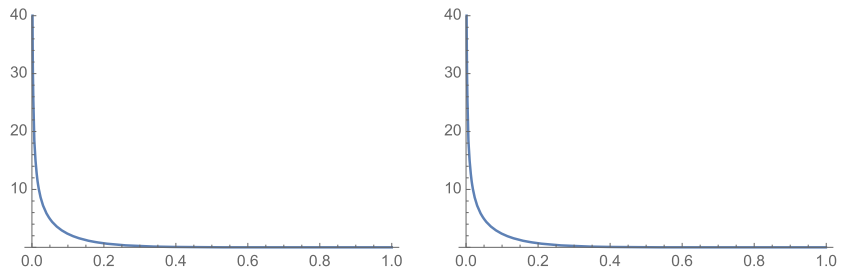

Figure 2. Posterior density of the meta-effectiveness for the treatment group as $a = 5$ and $b = 3$ for treatment group (left panel) and control group (right panel).

In fact, the factor S is particularly high in all of the groups considered here with respect to the rate of mortality data set. However, as expected from Figure 1, the robustness is very high when we analyse the usual *OR* quantity. The ranges obtained for *OR* were $(0.82, 0.92)$ over class $\Gamma_{a,b}$. The corresponding values of the S index were roughly flat around 9%, indicating a robust behaviour of the *OR* parameter. Observe that, in Figure 1, it seems that only two values of b are depicted (corresponding to $b = 1$ and $b = 4$). The other values for $b = 2, 3$ are also plotted but, the difference on the OR scale is so small that the graphics are overlapped.

For the sake of illustration, let us examine a specific case. For $a = 5$ and $b = 3$, the estimates of the meta-effectiveness parameter for the 'treatment group' and the control group are 3.85% and 3.97%, respectively. Figure 2 shows the posterior density of the meta-parameter θ in both groups.

Observe that in contrast to the common meta-analysis procedures based on normal approximations, this posterior distribution is a skewed inverted-J distribution, such that one tail is longer than the other. Then, the objective Bayesian highest posterior density interval is also asymmetrical around the pooled estimate. In particular, the 95% highest posterior density for each treatment in this case is 0.003–18.85% and 0.002–19.4%, respectively. These intervals show that the uncertainty of the posterior expected effectiveness θ is very large, both for the treatment group and for the control one. For this case, the posterior distribution of the OR have an estimated mean value of 0.985 and a very small estimated standard variation of 0.0008, which indicates a very peaked posterior distribution around their mean value.

Next, we test the null hypothesis that the meta-efficacy of the treated group and the control group is equal (over the class of priors $\Gamma_{a,b}$). Using expression (17), we compute Bayes factors for every density on $\Gamma_{a,b}$. Then, a robust Bayesian analysis of the probability of the null as (a, b) varies can of course be carried out by computing the interval

$$\mathcal{I}(\boldsymbol{x_1}, \boldsymbol{n_1}, \boldsymbol{x_2}, \boldsymbol{n_2}) = \left(\inf_{a,b \geqslant 1} \Pr(M_0|\text{data}), \sup_{a,b \geqslant 1} \Pr(M_0|\text{data}) \right),$$

Table II. Ranges of expected (posterior) effectiveness and probabilities of equality of treatment effects and *S* index for class of priors $\Gamma_{a,b}$.

Stillbirths data				
Group	Exp. effects	*S* factor (%)	Post. prob. M_0	*S* factor (%)
Treated	(0.0116, 0.0481)	37.92–156.98	(0.7971, 0.8876)	5.1–5.7
Control	(0.0142, 0.0488)	35.43–121.54		
OR	(0.814, 0.984)	8.64–10.44		
Neonatal deaths data				
Group	Exp. effects	*S* factor (%)	Post. prob. M_0	*S* factor (%)
Treated	(0.0132, 0.0487)	36.49–134.01	(0.7963, 0.8839)	4.95–5.5
Control	(0.0145, 0.0491)	35.14–118.22		
OR	(0.906, 0.992)	4.34–4.75		

Exp., expected; Post. prob., posterior probabilities.

where data $= \{\boldsymbol{x_1}, \boldsymbol{n_1}, \boldsymbol{x_2}, \boldsymbol{n_2}\}$. Obviously, the smaller the length of the interval, the greater the robustness of the posterior probability of the equality being accepted.

The range for the posterior probabilities of the null model M_0 for the Sarmanov/Intrinsic–Jeffreys links obtained in $\Gamma_{a,b}$ is $(0.804, 0.874)$. The robustness is high, with the sensitivity index *S* ranging from 4.01% to 4.36%. Using the Jeffreys' scale of evidence for Bayes factors (really, we use a slight modification suggested by Wasserman [24]), it follows from these ranges that there is moderate empirical evidence for accepting the equality of effectiveness for the treatment group and the control studies. This conclusion contrasts with those obtained under the classical and hierarchical Bayesian analyses shown in Figure 1 in [15], based on the 'interval criterion' using a vague prior on precision.

An analogous data analysis was conducted for the data on 'stillbirths' and 'neonatal deaths'. Table II shows the posterior ranges obtained for expected effects and probabilities of the null in the equality test over the family of priors $\Gamma_{a,b}$. Once again, the ranges for posterior effects reflect a considerable lack of robustness, with very high values for the *S* factor. In contrast, the case of the posterior probabilities for testing equality of treatment effects is markedly more robust, with a sensitivity index *S* of around 5%, indicating considerable robustness with respect to the assertion that there is moderate empirical evidence for accepting the equality of effects for both groups of treatments.

Ranges for *OR* are also very robust for both data sets. In the stillbirths data set, the range obtained was $(0.814, 0.984)$. Similarly, when the neonatal deaths data set was analysed, the range obtained for *OR* was $(0.906, 0.992)$ with a quasi-flat *S* index of around 4.5%.

6. CONCLUDING REMARKS

In this paper, we propose a Bayesian model for meta-analyses with binomial sparse data. This model overcomes concerns presented by other methodologies where, among other *ad hoc* techniques, correcting data by continuity is used. Taking into account that link distribution cannot be arbitrarily chosen but must be compatible with the Bayesian models for the experiments and for the meta-model, we present a well-adapted Sarmanov/Intrinsic–Jeffreys family, where, in some sense, the higher the Pearson's correlation coefficient, the smaller the between-study heterogeneity. The particular dependence structure between θ_i and θ induced by the Sarmanov/Intrinsic family allows us to model a range of degrees of heterogeneity in a way that is very simple and easy to apply in practice. For instance, through the relationship obtained between linear correlation and variance in the Sarmanov family, we see that the proposed robust analysis is particularly appropriate in the commonly occurring situation in which heterogeneity variance is small. For example, it is common to find case study where treatment effectiveness does not differ appreciably between studies, and very similar clinical trials designs are applied.

The Sarmanov/Intrinsic–Jeffreys family used presents three significant advantages: (a) the problem can be approached mathematically, and solutions are readily achieved by intensive computation; (b) we can control the form or structure of the dependence arising between the parameters within the class; and (c) by introducing a prior density on the parameters α and *t* of the Sarmanov/Intrinsic family, we can analyse the sensitivity of the measures of interest in the meta-analysis. The analysis of other class of distributions and their properties can be an area for future research [25].

Bayesian robustness methodology can be used to study how a given model reacts to deviations of some inputs, usually related to the prior distribution. In this way, we can measure the model's reaction to different situations of prior information. In fact, the meta-analysis carried out of the real data shows a high degree of robustness of the results with respect to the priors (equivalently, hyperparameters *a* and *b*) when *OR* or the probability of equality of treatments is measured.

As Sutton and Abrams [6] point out, it is of practical interest to examine the sensitivity of the Bayesian meta-analysis models to the prior distributions and to explore different ways of doing so. The present paper may be considered a further step in this direction. Although in the example presented, the heterogeneity of treatments has been accepted, we observe that, when homogeneity is present, ignoring it can lead to misleading inferences being drawn. This fact has important consequences for treatment comparison, because the likelihoods of the meta-parameter under homogeneity and heterogeneity are different. More general scenarios need to be considered. This remains an open question, to be examined in a future study. Another potential extension of the methods presented here is in network meta-analysis where the data structure is sparse because one is interested in the comparison of several treatments, while only a few of them are compared in each trial, which depicts a matrix structure where most of the elements are zero.

Acknowledgements

Financial support for this study was provided in part by grants ECO2013–47092 (Ministerio de Economía y Competitividad, Spain) and MTM2011–28945 (Ministerio de Ciencia e Innovación, Spain). We thank two anonymous referees for their many helpful comments.

REFERENCES

[1] DerSimonian R, Larird N. Meta-analysis in clinical trials. *Controlled Clinical Trials* 1986; **7**:177–188.
[2] Hardy RJ, Thompson SG. A likelihood approach to meta-analysis with random effects. *Statistics in Medicine* 1996; **15**:619–629.
[3] Knapp G, Hartung J. Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine* 2003; **22**:2693–2710.
[4] DerSimonian R, Kacker R. Random-effects model for meta-analysis of clinical trials: an update. *Contemporary Clinical Trials* 2007; **28**:105–114.
[5] Lambert PC, Sutton AJ, Burton PR, Abrams KR, Jones DR. How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. *Statistics in Medicine* 2005; **24**:2401–2428.
[6] Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. *Statistical Methods in Medical Research* 2001; **10**:277–303.
[7] Gajic–Veljanoski O, Cheung AM, Bayoumi AM, Tomlinson G. The choice of a noninformative prior on between-study variance strongly affects predictions of future treatment effect. *Medical Decision Making* 2013; **33**:356–368.
[8] Warren FC, Abrams KR, Golder S, Sutton AJ. Systematic review of methods used in meta-analyses where a primary outcome is an adverse or unintended event. *BMC Medical Research Methodology* 2012; **12**:64.
[9] Sweeting MJ, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Statistics in Medicine* 2004; **23**:1351–1375.
[10] Higgins JPT, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society, Series A* 2009; **172**:137–159.
[11] Sutton AJ, Higgins JP. Recent developments in meta-analysis. *Statistics in Medicine* 2008; **27**:625–650.
[12] Stojanovski E, Mengersen KL. Bayesian methods in meta-analysis. *Encyclopedia of Biopharmaceutical Statistics* 2006; **1**:1–7.
[13] Bradburn MJ, Deeks JJ, Berlin JA, Localio AR. Much ado about nothing: a comparison of the performance of meta-analytical methods with rare errors. *Statistics in Medicine* 2007; **26**:53–77.
[14] Collins R, Yusuf S, Peto R. Overview of randomised trials of diuretics in pregnancy. *British Medical Journal* 1985; **290**:17–23.
[15] Cornell JE, Mulrow CD, Localio R, Stack CB, Meibohm AR, Guallar E, Goodman SN. Random-effects meta-analysis of inconsistent effects: a time for change. *Annals of Internal Medicine* 2014; **160**:267–270.
[16] Moreno E, Vázquez–Polo FJ, Negrín MA. Objective Bayesian meta-analysis for sparse discrete data. *Statistics in Medicine* 2014; **33**:3676–3692.
[17] Fienberg S. Comment to "Hierarchical models for combining information and for meta-analyses". In *Bayesian Statistics 4*. Clarendon Press: Oxford; 1992; pp. 321–344.
[18] Patton A. A review of copula models for economic time series. *Journal of Multivariate Analysis* 1992; **110**:4–18.
[19] Sarmanov O V. Generalized normal correlated and two-dimensional Fréchet classes. *Doklady (Soviet Mathematics)* 1966; **168**:596–599.
[20] Chen Y, Chu H, Luo S, Nie L, Chen S. Bayesian analysis on meta-analysis of case–control studies accounting for within-study correlation. *Statistical Methods in Medical Research* 2015; **24**: 836–855.
[21] Chu H, Nie L, Chen Y, Huang Y, Sun W. Bivariate random effects models for meta-analysis of comparative studies with binary outcomes: methods for the absolute risk difference and relative risk. *Statistical Methods in Medical Research* 2012; **21**:621–633.
[22] Jeffreys H. *Theory of probability* (3rd edition). Clarendon Press: Oxford, 1961.
[23] Sivaganesan S. Sensitivity of some posterior summaries when the prior is unimodal with specified quantiles. *The Canadian Journal of Statistics* 1991; **19**:57–65.
[24] Wasserman L. Bayesian model selection and model averaging. *Journal of Mathematical Psychology* 2000; **44**:92–107.
[25] Sarabia JM, Prieto F, Jorda V. Bivariate beta-generated distributions with applications to well-being data. *Journal of Statistical Distributions and Applications* 2014; **1**:15.