# A Bayesian meta-analytic approach for safety signal detection in randomized clinical trials

Motoi Odani[1, 2*], Satoru Fukimbara[1], Tosiya Sato[2]

[1] *Data Science, Ono Pharmaceutical Co., Ltd., Osaka, Japan*

[2] *Department of Biostatistics, Kyoto University School of Public Health, Kyoto, Japan*

For submission as an original article to:

*Clinical Trials*: CT-16-0047

Revised, August 3, 2016

**Short title:** Bayesian safety meta-analysis in RCTs

**Total number of words (≤4000):** 3969

**Grant support:** This research received no grant support that requires acknowledgment.

*Corresponding author:

Motoi Odani

Ono Pharmaceutical Co., Ltd.

8-2, Kyutaromachi 1-chome, Chuo-ku

Osaka 541-8564, Japan

Tel: +81-6-6263-2966

Fax: +81-6-6263-2967

Email: m.odani@ono.co.jp

**Abstract (≤425 words)**

**Background:** Meta-analyses are frequently performed on adverse event (AE) data, and are primarily used for improving statistical power to detect safety signals. However, in the evaluation of drug safety for New Drug Applications, simple pooling of AE data from multiple clinical trials is still commonly used.

**Aim:** We sought to propose a new Bayesian hierarchical meta-analytic approach based on consideration of a hierarchical structure of reported individual AE data from multiple randomized clinical trials.

**Methods:** To develop our meta-analysis model, we extended an existing three-stage Bayesian hierarchical model by including an additional stage of the clinical trial level in the hierarchical model; this generated a four-stage Bayesian hierarchical model. We applied the proposed Bayesian meta-analysis models to published AE data from three premarketing randomized clinical trials of tadalafil and to a simulation study motivated by the case example to evaluate the characteristics of three alternative models.

**Results:** Comparison of the results from the Bayesian meta-analysis model with those from Fisher's exact test after simple pooling showed that 6 out of 10 AEs were the same within a top-10 ranking of individual AEs with regard to association with treatment. However, more individual AEs were detected in the Bayesian meta-analysis model than in

Fisher's exact test under the body system "Musculoskeletal and connective tissue disorders." Moreover, comparison of the overall trend of estimates between the Bayesian model and the standard approach (odds ratios after simple pooling methods) revealed that the posterior median odds ratios for the Bayesian model for most AEs shrank toward values for no association. Based on the simulation results, the Bayesian meta-analysis model could balance the false detection rate (FDR) and power to a better extent than Fisher's exact test. For example, when the threshold value of the posterior probability for signal detection was set to 0.8, the FDR was 41% and power was 88% in the Bayesian meta-analysis model, whereas the FDR was 56% and power was 86% in Fisher's exact test.

**Limitations:** AEs under the same body system were not necessarily positively related when we used "system organ class" and "preferred term" in the Medical Dictionary for Regulatory Activities as a hierarchical structure of AEs. For the Bayesian meta-analysis models to be effective, the validity of the hierarchical structure of AEs and the grouping of AEs are critical.

**Conclusions:** Our proposed meta-analysis models considered trial effects to avoid confounding by trial, and borrowed strength from both within and across body systems to obtain reasonable and stable estimates of an effect measure by considering a hierarchical structure of AEs.

**Keywords**

Bayesian hierarchical model, meta-analysis, randomized clinical trial, drug safety, adverse event data, signal detection

## Introduction

Statistical evaluation of drug-safety data is critical in both premarketing and postmarketing phases [1−5]. However, the following statistical challenges are associated with safety analyses performed on data from clinical trials, including adverse event (AE) data [3, 6]. First, statistical power: most clinical trials are designed to confirm primary efficacy endpoints, and the trials typically have insufficient power to detect between-treatment differences in safety endpoints, particularly for rare AEs. Second, multiplicity: the number of individual AEs, for example, can be extremely large, occasionally on the order of hundreds or even thousands in late-stage clinical trials; thus, an appropriate approach must be used that can effectively balance false-positive and false-negative rates. Third, medical classification: medically related AEs are grouped into categories, and these AEs in clinical trials are typically coded in terms by using a common dictionary (e.g., Medical Dictionary for Regulatory Activities (MedDRA)) with a specific hierarchical structure such as a body system (system organ class (SOC) in MedDRA) composed by individual AE terms (preferred term (PT) in MedDRA). The efficient use of such information is a statistical challenge. Here, we propose a Bayesian meta-analytic approach to overcome these three statistical challenges.

Bayesian approaches are useful for detecting safety signals [3, 7], and certain Bayesian methods have been proposed for analyzing the postmarketing spontaneous reports database (e.g., [8], [9]). From a regulatory perspective, Chi et al. noted the following: "Safety assessment is one area where frequentist strategies have been less applicable. Perhaps Bayesian approaches in this area have more promise." [10]. For analyzing AE data in clinical trials, Berry and Berry proposed a Bayesian hierarchical model [11]; they treated AE data as binary outcomes and modeled coded AE data with a hierarchical structure under the condition that AEs under the same body system are more similar and medically related than those under distinct body systems. Xia et al. extended the Bayesian hierarchical model to a Poisson model to account for differences in treatment durations between treatment groups [12]. By using this model, we can explicitly and concurrently model individual AEs with coding structures such as SOC and PT that are typically tabulated in clinical trial reports. The model also offers these other advantages: First, it suitably provides estimates that are partially corrected for multiplicity when most of the AEs are expected to be unassociated with treatment; this means that the model controls the detection of false positives and concomitantly adjusts multiplicity depending on the similarity of AEs within a hierarchical structure. Second, it efficiently analyzes the entire AE dataset and modulates the extremes. These features are particularly desirable for rare AEs.

Meta-analyses are frequently performed on AE data [13−15], and are primarily used as methods for improving statistical power to detect safety signals of rare AEs (e.g., [16], [17]). When performing meta-analyses of randomized clinical trials, the data analysis must preserve the structure of the trial design by using valid statistical methods that stratify by trial to preserve the randomization scheme [18, 19]. However, simple pooling of AE data from multiple clinical trials is still commonly adopted when evaluating drug safety for New Drug Applications. One reason for this is that the International Conference on Harmonization introduces simple pooling methods when data are presented in an Integrated Summary of Safety [20]. However, analysis results obtained from simple pooling methods can lead to inaccurate interpretations [21−24]. Furthermore, the potential for confounding by trial is strengthened because the imbalance in the distribution of participants' characteristics is occasionally not evaluable from aggregate trial data. Therefore, it is necessary to use a valid meta-analytic approach featuring appropriate adjustments and weights for distinct trials.

In this study, we aimed to develop a new Bayesian hierarchical meta-analytic approach by extending a Bayesian hierarchical model proposed by Berry and Berry [11] and Xia et al. [12] based on consideration of a hierarchical structure of reported individual AE data from multiple randomized clinical trials. We compared three alternative models by

using published AE data from three premarketing randomized clinical trials of tadalafil and through a small simulation study based on a case example. We focused on comparing the results obtained from simple pooling methods and the three models rather than model selection to assess the characteristics of the models and provide suggestions for practical use of the Bayesian meta-analysis models.

**Case example**

As a sample case, we analyzed a series of placebo-controlled trials for the treatment of lower urinary tract symptoms secondary to benign prostatic hyperplasia. Nishizawa et al. assessed the efficacy and safety of tadalafil (5 mg, administered once-daily) by pooling the data of 1199 patients administered placebo (n = 598) or tadalafil 5 mg (n = 601) from three premarketing randomized, double-blind, placebo-controlled, 12-week trials (a phase II trial in Japan (LVIA Study: placebo (n = 140) and tadalafil (n = 140)) and two phase III trials in Asia (LVHB Study: placebo (n = 154) and tadalafil (n = 155) and LVJF Study: placebo (n = 304) and tadalafil (n = 306))) [25]. These trials were similar with regard to design, participation criteria, and efficacy/safety endpoints.

Overall, tadalafil for up to 12 weeks of treatment was safe and well-tolerated in this pooled population. Because all of the individual AEs were not reported in the article, we extracted the AEs from the regulatory submission documents for the New Drug Application (available at: http://www.pmda.go.jp/drugs/2013/P201300159/index.html in Japanese); these are tabulated in the web appendix wTable 1. We found that 193 individual AEs (PTs) were present within 22 body systems (SOCs), which included 1–34 individual AEs. For example, 19 individual AEs such as myalgia and back pain were included under the body system "Musculoskeletal and connective tissue disorders," which constituted the

hierarchical structure of the AEs. In the following analyses, we focused on the placebo and tadalafil 5 mg once-daily groups as Nishizawa et al. did.

## Methods

### Bayesian meta-analysis models

To develop our meta-analysis model, we extended a three-stage Bayesian hierarchical model proposed by Berry and Berry [11] and Xia et al. [12] by including an additional stage of the clinical trial level in the hierarchical model; this resulted in the proposed four-stage Bayesian hierarchical model for AE data from multiple clinical trials. Here, we describe three Bayesian models: the first two are newly proposed four-stage Bayesian meta-analysis models, and the third is a non-hierarchical model for the structure of AE data and was used for comparison with the first two models.

### Model 1: Bayesian meta-analysis model with normal prior on log-odds ratio

We extended a Bayesian hierarchical model (Model 1a in [12]) to Model 1 as follows. $X_{ijk}$ and $Y_{ijk}$ were taken as the number of patients presenting an AE in the $j$th PT, for $j = 1, \ldots, p_i$, under the $i$th SOC, for $i = 1, \ldots, I$, in the $k$th clinical trial, for $k = 1, \ldots, K$, for placebo and treatment groups, respectively. $N_{ck}$ and $N_{tk}$ were taken as the number of patients in the $k$th clinical trial for placebo and treatment groups, respectively.

We assumed a binomial likelihood for the number of patients presenting an AE, i.e., $X_{ijk} \sim Binom(N_{ck}, \zeta_{ijk})$ and $Y_{ijk} \sim Binom(N_{tk}, \eta_{ijk})$, where $\zeta_{ijk}$ and $\eta_{ijk}$ are the probabilities of

incidence of an AE for the *j*th PT under the *i*th SOC in the *k*th clinical trial for placebo and treatment groups, respectively. As a mean structure, we considered a logistic regression: $\mathrm{log}it(\zeta_{ijk}) = \log(\zeta_{ijk}/(1 - \zeta_{ijk})) = \gamma_{ijk}$; $\mathrm{log}it(\eta_{ijk}) = \log(\eta_{ijk}/(1 - \eta_{ijk})) = \gamma_{ijk} + \theta_{ijk}$, where $\theta_{ijk}$ is the logarithm of odds ratio.

Listed below are the Stage-1 prior distributions; $\gamma_{ijk}$ and $\theta_{ijk}$ exhibited a normal prior distribution:

$$\gamma_{ijk} \sim N(\mu_{\gamma ij}, \sigma^2_{\gamma ij}) \qquad \theta_{ijk} \sim N(\mu_{\theta ij}, \sigma^2_{\theta ij})$$

Stage 1 corresponds to the clinical trial level, and we conducted a meta-analysis for individual AEs at this stage. Following Stage 1, we set the prior distributions to the hyperparameters through Stages 2–4 with the same principle used in the original Bayesian hierarchical model [11, 12]. Specifically, normal distributions and inverse gamma (IG) distributions were set to the means and the variances, respectively. The prior distributions of these stages were the following:

Stage 2, individual AE (PT) level:

$$\mu_{\gamma ij} \sim N(\mu_{\gamma i}, \tau^2_{\gamma i}) \qquad \sigma^2_{\gamma ij} \sim IG(\alpha_{\gamma i}, \beta_{\gamma i})$$

$$\mu_{\theta ij} \sim N(\mu_{\theta i}, \tau^2_{\theta i}) \qquad \sigma^2_{\theta ij} \sim IG(\alpha_{\theta i}, \beta_{\theta i})$$

Stage 3, body system (SOC) level:

$$\mu_{\gamma i} \sim N(\mu_{\gamma 0}, \tau^2_{\gamma 0}) \qquad \tau^2_{\gamma i} \sim IG(\alpha_{\gamma}, \beta_{\gamma})$$

$$\mu_{\theta i} \sim N(\mu_{\theta 0}, \tau^2_{\theta 0}) \qquad \tau^2_{\theta i} \sim IG(\alpha_\theta, \beta_\theta)$$

Stage 4, overall AE level:

$$\mu_{\gamma 0} \sim N(\mu_{\gamma 00}, \tau^2_{\gamma 00}) \qquad \tau^2_{\gamma 0} \sim IG(\alpha_{\gamma 00}, \beta_{\gamma 00})$$

$$\mu_{\theta 0} \sim N(\mu_{\theta 00}, \tau^2_{\theta 00}) \qquad \tau^2_{\theta 0} \sim IG(\alpha_{\theta 00}, \beta_{\theta 00})$$

The hyperparameters $\mu_{\gamma 00}$, $\tau^2_{\gamma 00}$, $\mu_{\theta 00}$, $\tau^2_{\theta 00}$, $\alpha_{\gamma 00}$, $\beta_{\gamma 00}$, $\alpha_{\theta 00}$, $\beta_{\theta 00}$, $\alpha_\gamma$, $\beta_\gamma$, $\alpha_\theta$, $\beta_\theta$, $\alpha_{\gamma i}$, $\beta_{\gamma i}$, $\alpha_{\theta i}$, and $\beta_{\theta i}$ were considered fixed constants. In our analysis, we used the same values used by Berry and Berry [11] and Xia et al. [12]: $\mu_{\gamma 00} = \mu_{\theta 00} = 0$, $\tau^2_{\gamma 00} = \tau^2_{\theta 00} = 10$, $\alpha_{\gamma 00} = \alpha_{\theta 00} = \alpha_\gamma = \alpha_\theta = \alpha_{\gamma i} = \alpha_{\theta i} = 3$; and $\beta_{\gamma 00} = \beta_{\theta 00} = \beta_\gamma = \beta_\theta = \beta_{\gamma i} = \beta_{\theta i} = 1$.

**Model 2: Bayesian meta-analysis model with mixture prior on log-odds ratio**

We extended Bayesian hierarchical models (the model proposed in [11] and Model 1b in [12]) to Model 2 as follows. We assumed the same likelihood and mean structure as in Model 1, but changed the prior distribution for the mean of the log-odds ratios to a mixture distribution:

$$\mu_{\theta ij} \sim \pi_i \delta(0) + (1 - \pi_i) N(\mu_{\theta i}, \tau^2_{\theta i})$$

where $\delta(0)$ is a distribution featuring a unit point mass at 0. Positive probability was assigned to the possibility of equality between the placebo and treatment proportions.

We used the same prior distributions for the common parameters as in Model 1. For the hyperparameters not included in Model 1, these were the prior distributions:

$\pi_i \sim Beta(\alpha_\pi, \beta_\pi)$

$\alpha_\pi \sim Exp(\lambda_\alpha)I[\alpha_\pi > 1]$ $\beta_\pi \sim Exp(\lambda_\beta)I[\beta_\pi > 1]$

The left-truncated exponential prior distributions for $\alpha_\pi$ and $\beta_\pi$ were chosen. Restricting the parameters to >1 prevents the posterior density of $\pi_i$ from becoming exceedingly concentrated at 0 and 1.

We used the same fixed values for the hyperparameters $\mu_{\gamma 00}$, $\tau^2_{\gamma 00}$, $\mu_{\theta 00}$, $\tau^2_{\theta 00}$, $\alpha_{\gamma 00}$, $\beta_{\gamma 00}$, $\alpha_{\theta 00}$, $\beta_{\theta 00}$, $\alpha_\gamma$, $\beta_\gamma$, $\alpha_\theta$, $\beta_\theta$, $\alpha_{\gamma i}$, $\beta_{\gamma i}$, $\alpha_{\theta i}$, and $\beta_{\theta i}$ as in Model 1; $\lambda_\alpha$ and $\lambda_\beta$ were considered fixed constants, and we used the same values of $\lambda_\alpha = \lambda_\beta = 0.1$ as previously defined [12].

**Model 3: Non-hierarchical model with mixture prior on log-odds ratio**

We extended a one-stage Bayesian model (Model 1c in [12]) to Model 3, a non-hierarchical model for the structure of AE data; in this model, no information is borrowed across different AEs within the same body system and all individual AEs are treated independently.

We assumed the same likelihood and mean structure as in Models 1 and 2, but changed the prior distributions at Stage 2 of Model 1 as in [12]:

$\mu_{\gamma ij} \sim N(0, 10^2)$ $\quad\quad$ $\sigma^2_{\gamma ij} \sim IG(3, 1)$

$\mu_{\theta ij} \sim 0.5\delta(0) + 0.5N(0, 10^2)$ $\quad$ $\sigma^2_{\theta ij} \sim IG(3, 1)$

**Evaluation of different models and a standard approach**

**Comparison of measures for the case example**

As a measure for safety signal detection, we calculated the following posterior probability to identify potential signals: Pr ($OR_{ij} > 1.0$ | Data) in an AE of the *j*th PT under the *i*th SOC, where $OR_{ij} = \exp(\mu_{\theta ij})$ and $\mu_{\theta ij}$ is the posterior mean of log-odds ratios in the Bayesian models for the *j*th PT under the *i*th SOC.

Fisher's exact test is a well-established test for investigating associations of AE data with treatment in clinical trials. Therefore, we employed the one-sided P values from Fisher's exact test obtained after simple pooling methods as a standard measure in current practice.

We compared the three models and the standard approach from two points of view: one, comparison of AEs that might exhibit strong associations with treatment; and two, overall comparison of the measures for safety signal detection.

**Simulation study**

We conducted a simulation study motivated by the data structure of the case example to evaluate the false detection rate (FDR) and power of our meta-analysis models. We generated 150 patients per group in the first and second trial and 300 patients per group in the third trial. Each trial had 193 PTs within 22 SOCs as in the case example. The incidence probability of an AE in the placebo group was 5%, 10%, 1%, 10%, 5%, and 15% for SOCs 1, 2, 3–6, 7–11, 12–16, and 17–22, respectively. We assumed that for a true signal the incidence probability in the treatment group was 5 percentage points higher than the corresponding probability in the placebo group. Five PTs in the 1st SOC and one PT in the 2nd SOC were found to be true signals, and all the other PTs were null. The 1st SOC and one PT in the 2nd SOC mimicked the SOC "Musculoskeletal and connective tissue disorders" and the PT "dyspepsia," respectively, in the case example. We estimated the FDR and power for our meta-analysis models using 1,000 simulated trials. In this simulation study, the FDR was the expected value of the proportion of falsely detected signals among all signals detected. The power was the expected value of the proportion of correctly detected signals among all true signals.

**Bayesian computations**

We implemented our meta-analysis models within a fully Bayesian framework by using Markov chain Monte Carlo methods with WinBUGS software 1.4.3 [26, 27]. The sample WinBUGS codes and a detailed method for Bayesian computations are presented in Supplementary appendix.

**Results**

Table 1 presents a top-10 ranking of individual AEs with regard to association with treatment. First, comparison of the results from Model 1 with those from Fisher's exact test after simple pooling revealed that 6/10 AEs were identical, and myalgia and dyspepsia were ranked first and second with both methods. The posterior median odds ratios and their 2.5% and 97.5% percentiles from Model 1 were 3.4 (1.1, 12.5) and 3.7 (1.0, 13.8), respectively. In the body system "Musculoskeletal and connective tissue disorders," only myalgia and musculoskeletal pain were detected with Fisher's exact test, whereas myalgia, back pain, musculoskeletal pain, pain in extremity, and periarthritis were detected with Model 1. Second, comparison of the results from Models 1 and 2 revealed that 8/10 AEs were identical, with the top 4 AEs detected being the same. The posterior median odds ratios and their 2.5% and 97.5% percentiles from Model 2 for myalgia and dyspepsia were 2.4 (1.0, 12.8) and 2.4 (1.0, 14.4), respectively. A notable difference between the two models was observed in the values of the posterior probability; with myalgia, for example, Pr ($OR_{ij}$ > 1.0 | Data) = 0.9846 for Model 1, and Pr ($OR_{ij}$ > 1.0 | Data) = 0.5942 for Model 2. Thus, the values obtained from Model 2 were considerably smaller than those from Model 1. Third, comparison of the results from Model 3 with those from Fisher's exact test after simple pooling revealed that 6/10 AEs were identical, with the top 5 AEs detected being the same.

However, the estimation of odds ratios from Model 3 was highly unstable: the standard deviations from the posterior samples were extremely large, particularly for AEs for which crude odds ratios cannot be defined (i.e., when the number of patients with AE = 0 in the placebo group). Therefore, Model 3 results are presented for illustration only.

Figure 1 shows the relationship between the odds ratios from simple pooling methods and Model 1. The posterior median odds ratios for most of the AEs shrank toward values for no association (i.e., the posterior odds ratios approached 1.0 in the Bayesian model). A typical example was pharyngitis, for which the crude odds ratio and its 95% confidence interval was 5.0 (0.6, 43.0), whereas the posterior median odds ratio and its 2.5% and 97.5% percentile was 1.3 (0.5, 4.5). This shrinkage was considered reasonable because most of the AEs showed either weak or no association with treatment in our case example. By contrast, the posterior median odds ratio and its 2.5% and 97.5% percentile (2.0 (0.8, 5.7)) were larger than the crude odds ratio and its 95% confidence interval (1.7 (0.6, 4.6)) for back pain, which is differently located above the diagonal line in the figure and is classified under "Musculoskeletal and connective tissue disorders" as stated above.

Table 3 presents the results of the simulation study by showing the FDR and power of Fisher's exact test and our meta-analysis models. Our Model 1 can balance the FDR and power to a better extent than Fisher's exact test. For example, when the threshold value of

the posterior probability for signal detection was set to 0.8, the FDR was 41% and power was 88% from Model 1, whereas the FDR was 56% and power was 86% in Fisher's exact test. When we restricted the inference on AEs under the 1st SOC, the power was higher compared to the overall result from Model 1. Both FDR and power were low from Models 2 and 3 in this scenario, and they were far lower than 20%.

## Discussion

We detected more individual AEs under the body system "Musculoskeletal and connective tissue disorders" in Model 1 than in Fisher's exact test. To elucidate this result, we examined an extracted tabulation of the incidence of individual AEs under this body system according to treatment group and trial (Table 2). More incidences were observed in the tadalafil group than in the placebo group for most of the individual AEs. In Model 1, as a property of a Bayesian hierarchical model, strength could be borrowed from both within and across body systems depending on the actual data.

Similar AEs were detected in Models 1 and 2, but a notable difference was observed in their posterior probability values. We consider this difference to be primarily caused by modeling assumptions: In Model 2, we incorporated the belief that a certain proportion of AEs is completely unassociated with treatment by assuming a mixture prior distribution for the mean of log-odds ratios. A similar result was obtained by Berry and Berry [11] and Xia et al. [12], and the posterior probabilities that the odds ratio exactly equals 1.0 were several dozen percent. Based on the simulation results, we recommend using Model 1 in an AE data structure similar to our case example. As indicated in our case example as a property of a Bayesian hierarchical model, Model 1 could suitably detect safety signals under the same SOC, although the model might detect more non-signals under the SOC in a

compensatory manner. Although Model 2 was not effective in our case example, one potential scenario in which we could use the model is that in which the number of AEs is extremely large, and it would be challenging to handle the large number of safety signals detected in Model 1. The threshold value of the posterior probability for signal detection might depend on the requirement for balancing false-positive and false-negative rates, but we could use 0.9 as a signal detection rule with Model 1 in an AE data structure similar to our case example, although the power might be slightly low (i.e., 63%).

In our case example, the estimation of odds ratios from Model 3 was highly unstable, but more stable estimates were obtained from Models 1 and 2. Model 3 did not work well also in the simulation study. Certain AEs were still rare at an individual event level for safety meta-analyses, and we expect to frequently encounter a similar type of situation in which the association of individual AEs with treatment is challenging to evaluate. Therefore, employing a Bayesian hierarchical model that considers a hierarchical structure of AEs can also be useful for obtaining stable estimates for an effect measure.

For our proposed Bayesian meta-analysis models to be effective, the validity of the hierarchical structure of AEs is also crucial. As discussed by Xia et al. [12], biological relationships among individual AEs are reflected in the medical coding structure, and AEs under the same SOC are more likely to be similar than other AEs. Thus, they can

reasonably borrow strength from each other; this means that ignoring the hierarchical structure will result in a loss of information regarding their relationships. In our case example, the AEs within a top-10 ranking from Model 3 were similar to those from Fisher's exact test, which is also a method that ignores the hierarchical structure. However, one limitation is that the AEs under the same SOC were not necessarily positively related when we used SOC and PT in MedDRA. For example, the positive relationship was questionable among certain AEs under "Investigations": in our case example, both "white blood cell count increased" and "white blood cell count decreased," which are clearly counter events, were included under this body system. Therefore, the critical point is the validity of the AE grouping. One possible method to identify events that reflect an identical or highly similar medical concept of interest is to use Standardized MedDRA Queries (SMQs); medically related AEs that might be under distinct SOCs are grouped into the same SMQ, and AEs can be modeled under the SMQ to enable them to borrow strength from each other [12, 18].

Our choice of priors was the same as previously described [11, 12]. Xia et al. [12] conducted some sensitivity analyses for the choice of priors, and they considered their results were robust to different prior distributions. In the case that informative prior distributions are selected, it is necessary to carefully evaluate how the estimates change

depending on the selection, although it might be difficult to put some information on specific SOCs or PTs when safety signal detection is performed.

Certain other multivariate meta-analyses for AE data have been proposed [28, 29], but the approaches adopted were distinct from ours; although AEs were modeled at a comparatively less granular classification of the body system level in one case [28], and AEs were modeled without explicitly considering a hierarchical structure of AEs in another case [29]. All analytic approaches must be selected depending on the objective of the safety meta-analysis and the given AE dataset. Our proposed meta-analysis models are particularly suitable for analyzing entire AE datasets to detect safety signals for further investigation as potential risks from among hundreds or thousands of AEs from multiple randomized clinical trials.

Another limitation of our analysis is that individual patient data from the randomized clinical trials were unavailable, and within-patient dependencies among the individual AEs could not be modeled. When individual patient data are available, such dependencies can be incorporated into Bayesian meta-analysis models, and our models can be extended to Poisson models by using an exposure time in person-years in each treatment group.

Recently, there has been considerable interest in comparing multiple treatments, and network meta-analysis is a well-adopted approach for integrating evidence of a complex network of multiple treatments [30–32]. For safety meta-analyses, a network meta-analytic approach has been proposed for AEs of interest with a pre-specified hypothesis [33], and general guidance for implementing a safety Bayesian network meta-analysis has been published [34]. Ohlssen et al. [34] discussed a network meta-analysis with a pre-specified event of interest, and then extended their model to a multivariate model with borrowing of strength across outcomes. Because their model does not account for a hierarchical structure of AEs, our proposed Bayesian meta-analysis models can be applied to extend their network meta-analytic framework for safety signal detection.

In summary, we have proposed and implemented a new Bayesian meta-analytic approach for individual AE data from multiple randomized clinical trials and for the simulation study motivated by the case example. Our meta-analysis models consider trial effects to avoid confounding by trial, and borrow strength from both within and across body systems to obtain reasonable and stable estimates of an effect measure by considering the hierarchical structure of AEs. Different safety signals could be detected among the proposed meta-analysis models and the standard approach depending on the underlying

assumptions of each approach. Such differences could lead to distinct interpretations regarding drug safety.

**Acknowledgments**

The authors would like to thank two anonymous reviewers for their critical evaluation of this manuscript and thoughtful suggestions.

**Funding**

This research received no funding from any public, for-profit, or not-for-profit sources, other than support of the authors' time spent on this manuscript preparation, which was provided by their employers.

**Conflict of interest**

Motoi Odani and Satoru Fukimbara are full-time employees of Ono Pharmaceutical Co., Ltd. Tosiya Sato received an honorarium from Eli Lilly Japan K.K.

**References**


1. United States Food and Drug Administration. Guidance for industry: Premarketing risk assessment, http://www.fda.gov/downloads/RegulatoryInformation/Guidances/UCM126958.pdf (accessed 28 July 2016).

2. United States Food and Drug Administration. Guidance for industry: Good pharmacovigilance practices and pharmacoepidemiologic Assessment, http://www.fda.gov/downloads/RegulatoryInformation/Guidances/UCM126834.pdf (accessed 28 July 2016).

3. Xia HA, Jiang Q. Statistical evaluation of drug safety data. *Ther Innov Regul Sci* 2014; 48: 109–120.

4. Zhou Y, Ke C, Jiang Q, et al. Choosing appropriate metrics to evaluate adverse events in safety evaluation. *Ther Innov Regul Sci* 2015; 49: 398–404.

5. Ma H, Ke C, Jiang Q, et al. Statistical considerations on the evaluation of imbalances of adverse events in randomized clinical trials. *Ther Innov Regul Sci* 2015; 49: 957–965.

6. Council for International Organizations of Medical Sciences (CIOMS) Working Group VI. *Management of safety information from clinical trials*. CIOMS, Geneva, 2005.

7. Price KL, Xia HA, Lakshminarayanan M, et al. Bayesian methods for design and analysis of safety trials. *Pharm Stat* 2014; 13: 13–24.

8. Bate A, Lindquist M, Edwards IR, et al. A Bayesian neural network method for adverse drug reaction signal generation. *Eur J Clin Pharmacol* 1998; 54: 315–321.

9. DuMouchel W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. *Am Stat* 1999; 53: 177–190.

10. Chi G, Hung HMJ, O'Neill R. Some comments on "Adaptive Trials and Bayesian Statistics in Drug Development" by Don Berry. *Pharmaceutical Report* 2002; 9: 1–11.

11. Berry SM, Berry DA. Accounting for multiplicities in assessing drug safety: a three-level hierarchical mixture model. *Biometrics* 2004; 60: 418–426.

12. Xia HA, Ma H, Carlin BP. Bayesian hierarchical modeling for detecting safety signals in clinical trials. *J Biopharm Stat* 2011; 21: 1006–1029.

13. Cornelius VR, Perrio MJ, Shakir SA, et al. Systematic reviews of adverse effects of drug interventions: a survey of their conduct and reporting quality. *Pharmacoepidemiol Drug Saf* 2009; 18: 1223–1231.

14. Warren FC, Abrams KR, Golder S, et al. Systematic review of methods used in meta-analyses where a primary outcome is an adverse or unintended event. *BMC Med Res Methodol* 2012; 12: 64.

15. Berlin JA, Crowe BJ, Whalen E, et al. Meta-analysis of clinical trial safety data in a drug development program: answers to frequently asked questions. *Clin Trials* 2013; 10: 20–31.

16. Nissen SE, Wolski K. Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. *N Engl J Med* 2007; 356: 2457–2471.

17. Bridge JA, Iyengar S, Salary CB, et al. Clinical response and risk for reported suicidal ideation and suicide attempts in pediatric antidepressant treatment: a meta-analysis of randomized controlled trials. *JAMA* 2007; 297: 1683–1696.

18. Wittes J, Crowe B, Chuang-Stein C, et al. The FDA's final rule on expedited safety reporting: statistical considerations. *Stat Biopharm Res* 2015; 7: 174–190.

19. Chakravarty AG, Izem R, Keeton S, et al. The role of quantitative safety evaluation in regulatory decision making of drugs. *J Biopharm Stat* 2016; 26: 17–29.

20. International Conference on Harmonisation Expert Working Group. *The common technical document for the registration of pharmaceuticals for human use: efficacy M4e(R1)*, http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/CTD /M4__R1__Efficacy/M4E__R1_.pdf (accessed 28 July 2016).

21. Lièvre M, Cucherat M, Leizorovicz A. Pooling, meta-analysis, and the evaluation of drug safety. *Curr Control Trials Cardiovasc Med* 2002; 3: 6.

22. Chuang-Stein C, Beltangady M. Reporting cumulative proportion of subjects with an adverse event based on data from multiple studies. *Pharm Stat* 2011; 10: 3–7.

23. Egger M, Smith GD, Phillips AN. Meta-analysis: principles and procedures. *BMJ* 1997; 315: 1533–1537.

24. Hammad TA, Pinheiro SP, Neyarapally GA. Secondary use of randomized controlled trials to evaluate drug safety: a review of methodological considerations. *Clin Trials* 2011; 8: 559–570.

25. Nishizawa O, Yoshida M, Takeda M, et al. Tadalafil 5 mg once daily for the treatment of Asian men with lower urinary tract symptoms secondary to benign

prostatic hyperplasia: analyses of data pooled from three randomized, double-blind, placebo-controlled studies. *Int J Urol* 2015; 22: 378–384.

26. Lunn DJ, Thomas A, Best N, et al. WinBUGS - a Bayesian modelling framework: concepts, structure, and extensibility. *Stat Comput* 2000; 10: 325–337.

27. Lunn D, Spiegelhalter D, Thomas A, et al. The BUGS project: evolution, critique and future directions. *Stat Med* 2009; 28: 3049–3067.

28. Rosenkranz GK. An approach to integrated safety analyses from clinical studies. *Drug Inf J* 2010; 44: 649–657.

29. DuMouchel W. Multivariate Bayesian logistic regression for analysis of clinical study safety issues. *Stat Sci* 2012; 27: 319–339.

30. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004; 23: 3105–3124.

31. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005; 331: 897–900.

32. Cooper NJ, Sutton AJ, Lu G, et al. Mixed comparison of stroke prevention treatments in individuals with nonrheumatic atrial fibrillation. *Arch Intern Med* 2006; 166: 1269–1275.

33. Warren FC, Abrams KR, Sutton AJ. Hierarchical network meta-analysis models to address sparsity of events and differing treatment classifications with regard to adverse outcomes. *Stat Med* 2014; 33: 2449–2466.

34. Ohlssen D, Price KL, Xia HA, et al. Guidance on the implementation and reporting of a drug safety Bayesian network meta-analysis. *Pharm Stat* 2014; 13: 55–70.

**Figure Legends**

**Figure 1.** Comparison of the crude odds ratios and the posterior median odds ratios from the Bayesian meta-analysis model (Model 1).

One point in the plot corresponds to one AE, and the AEs for which crude odds ratios cannot be defined were omitted from this figure. Odds ratio = 1.0 if a point is on the dotted line in each axis.

**Table 1.** Top-10 ranking of individual AEs with regard to association with treatment

| Rank | Simple Pooling Method | | Model 1 | | Model 2 | | Model 3* | |
|---|---|---|---|---|---|---|---|---|
| | PT | P values | PT | PP | PT | PP | PT | PP |
| 1 | Dyspepsia | 0.0002 | Myalgia | 0.0154 | Myalgia | 0.4058 | Dyspepsia | 0.1447 |
| 2 | Myalgia | 0.0031 | Dyspepsia | 0.0232 | Dyspepsia | 0.4171 | Myalgia | 0.1982 |
| 3 | Nausea | 0.0628 | Back pain | 0.0809 | Musculoskeletal pain | 0.7184 | Rhinitis allergic | 0.2214 |
| 4 | Musculoskeletal pain | 0.0628 | Musculoskeletal pain | 0.0895 | Back pain | 0.7518 | Musculoskeletal pain | 0.2702 |
| 5 | Rhinitis allergic | 0.0628 | Hot flush | 0.1449 | Nausea | 0.8059 | Nausea | 0.3707 |
| 6 | Pharyngitis | 0.1100 | Pain in extremity | 0.1841 | Hot flush | 0.8085 | Cataract | 0.3750 |
| 7 | Cataract | 0.1256 | Headache | 0.2061 | Rhinitis allergic | 0.8213 | Muscle tightness | 0.5285 |
| 8 | Hot flush | 0.1454 | Rhinitis allergic | 0.2067 | Pain in extremity | 0.8305 | Periarthritis | 0.5316 |
| 9 | Creatinine renal clearance decreased | 0.1885 | Periarthritis | 0.2072 | Muscle tightness | 0.8432 | Fall | 0.5367 |
| 10 | Headache | 0.2134 | Diarrhoea | 0.2077 | Periarthritis | 0.8463 | Otitis media | 0.5440 |

PT: preferred term, PP: posterior probability.

P values were obtained from Fisher's exact test after simple pooling methods.

PTs were ranked in the order of the posterior probability Pr ($OR_{ij} > 1.0$ | Data) from the largest value for the Bayesian models and the one-sided P values from the smallest value for the simple pooling methods. In this table PPs were shown as Pr ($OR_{ij} \leq 1.0$ | Data) as they could be simply compared to the P values.

*The estimation of odds ratios from Model 3 was highly unstable, particularly for AEs for which crude odds ratios cannot be defined. Therefore, Model 3 results are presented for illustration only.

**Table 2.** Extracted tabulation of individual AEs under the body system "Musculoskeletal and connective tissue disorders"

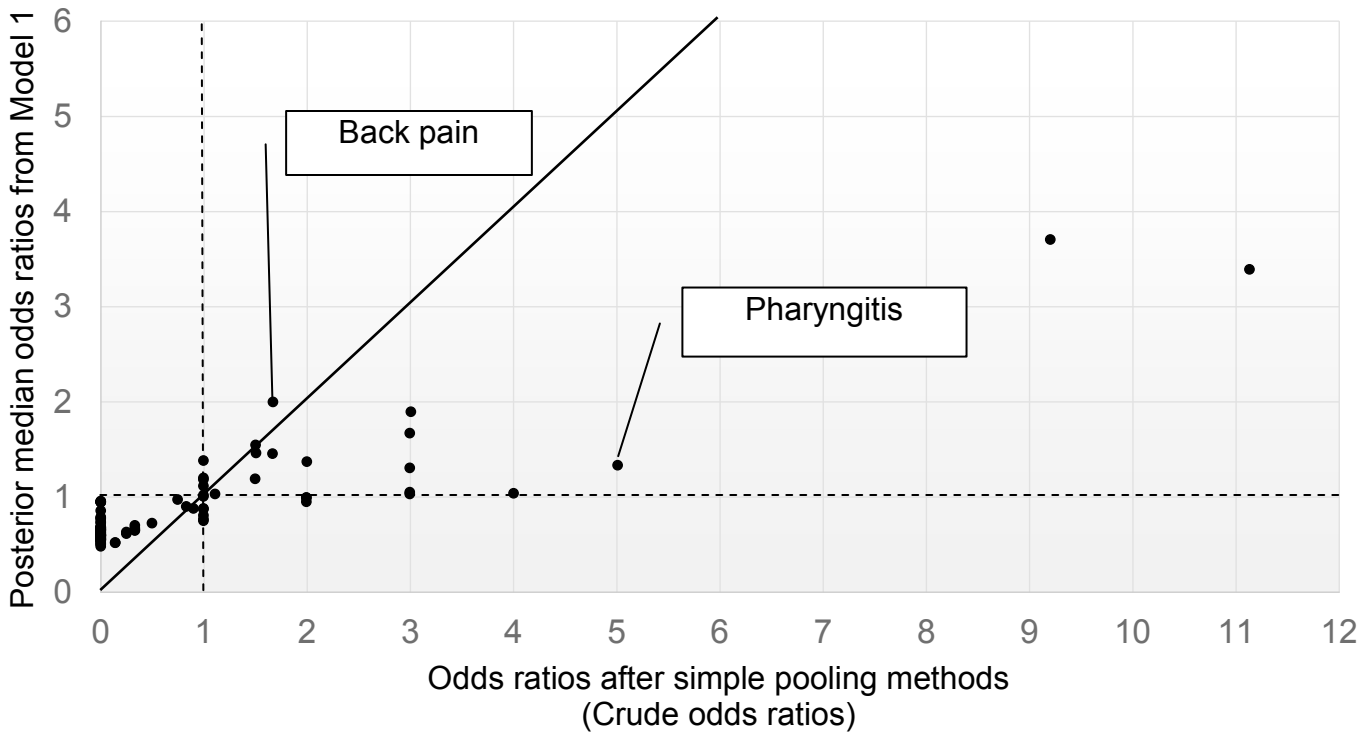
| Study | LVIA Study | | | | LVHB Study | | | | LVJF Study | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Treatment Group | Placebo | | Tadalafil 5 mg | | Placebo | | Tadalafil 5 mg | | Placebo | | Tadalafil 5 mg | |
| | n | (%) | n | (%) | n | (%) | n | (%) | n | (%) | n | (%) |
| N | 140 | | 140 | | 154 | | 155 | | 304 | | 306 | |
| Arthralgia | 2 | ( 1.4) | 1 | ( 0.7) | | | | | 1 | ( 0.3) | 2 | ( 0.7) |
| Arthritis | 1 | ( 0.7) | | | | | | | | | | |
| Back pain | 2 | ( 1.4) | 2 | ( 1.4) | 1 | ( 0.6) | 4 | ( 2.6) | 3 | ( 1.0) | 4 | ( 1.3) |
| Musculoskeletal pain | | | 2 | ( 1.4) | | | 2 | ( 1.3) | | | | |
| Myalgia | | | 2 | ( 1.4) | | | 6 | ( 3.9) | 1 | ( 0.3) | 3 | ( 1.0) |
| Pain in extremity | | | 1 | ( 0.7) | | | | | 1 | ( 0.3) | 2 | ( 0.7) |
| Periarthritis | | | 1 | ( 0.7) | | | | | | | 1 | ( 0.3) |
| Spinal column stenosis | | | 1 | ( 0.7) | | | | | | | | |
| Tenosynovitis | | | 1 | ( 0.7) | | | | | 1 | ( 0.3) | | |
| Muscle spasms | | | | | 1 | ( 0.6) | | | | | | |
| Myositis | | | | | | | 1 | ( 0.6) | | | | |
| Osteoarthritis | | | | | 1 | ( 0.6) | | | | | | |
| Joint range of motion decreased | | | | | | | 1 | ( 0.6) | | | | |
| Muscle tightness | | | | | | | 2 | ( 1.3) | | | | |
| Arthropathy | | | | | | | | | | | 1 | ( 0.3) |
| Gouty arthritis | | | | | | | | | 1 | ( 0.3) | | |
| Scoliosis | | | | | | | | | 1 | ( 0.3) | | |
| Musculoskeletal stiffness | | | | | | | | | 1 | ( 0.3) | 1 | ( 0.3) |
| Limb discomfort | | | | | | | | | | | 1 | ( 0.3) |

**Table 3.** False detection rate and power results for simulation study

| | Threshold value | Overall AEs | | AEs under the 1st SOC | |
|---|---|---|---|---|---|
| | | FDR (%) | Power (%) | FDR (%) | Power (%) |
| Fisher's exact test | Two-sided, P value < 0.05 | 56 | 86 | 9 | 89 |
| Model 1 | PP > 0.6 | 89 | 99 | 54 | 99 |
| | PP > 0.7 | 75 | 97 | 37 | 99 |
| | PP > 0.8 | 41 | 88 | 17 | 96 |
| | PP > 0.9 | 6 | 63 | 2 | 75 |
| Model 2 | PP > 0.6 | 0 | 0.12 | 0 | 0.14 |
| | PP > 0.7 | 0 | 0.04 | 0 | 0.05 |
| | PP > 0.8 | 0 | 0 | 0 | 0 |
| | PP > 0.9 | 0 | 0 | 0 | 0 |
| Model 3 | PP > 0.6 | 9 | 4 | 0 | 5 |
| | PP > 0.7 | 7 | 2 | 0 | 2 |
| | PP > 0.8 | 4 | 1 | 0 | 1 |
| | PP > 0.9 | 3 | 0.1 | 0 | 0.1 |

FDR: false detection rate, PP: posterior probability.
FDR was the expected value of the proportion of falsely detected signals among all signals detected. Power was the expected value of the proportion of correctly detected signals among all true signals.
For the Bayesian models, the inference was based on the posterior probability Pr ($OR_{ij}$ > 1.0 | Data) > (threshold value).
The 1st SOC mimicked the SOC "Musculoskeletal and connective tissue disorders" in the case example.

**Figure 1.**


Posterior median odds ratios from Model 1
0
1
2
3
4
5
6
Back pain
Pharyngitis
0
1
2
3
4
5
6
7
8
9
10
11
12
Odds ratios after simple pooling methods
(Crude odds ratios)

**Supplementary appendix**

Incidence of treatment-emergent AEs is given according to treatment group and trial in wTable 1. These individual AEs were extracted from the regulatory submission documents for the New Drug Application of tadalafil (5 mg, administered once-daily) for the treatment of lower urinary tract symptoms secondary to benign prostatic hyperplasia. The AE data are also provided in the CSV file format.

The sample WinBUGS codes for the proposed Bayesian meta-analysis models are presented in this appendix.

**wTable 1.** Incidence of treatment-emergent adverse events

Analysis Population: All Randomized Subjects

| Study | LVIA Study | | LVHB Study | | LVJF Study | |
|---|---|---|---|---|---|---|
| Treatment Group | Placebo | Tadalafil 5 mg | Placebo | Tadalafil 5 mg | Placebo | Tadalafil 5 mg |
| Treatment-emergent Adverse Events (Classification by MedDRA SOC/PT) | n (%) | n (%) | n (%) | n (%) | n (%) | n (%) |
| N | 140 | 140 | 154 | 155 | 304 | 306 |
| Overall | 53 ( 37.9) | 57 ( 40.7) | 30 ( 19.5) | 47 ( 30.3) | 76 ( 25.0) | 87 ( 28.4) |
| Blood and lymphatic system disorders | | | | | | 2 ( 0.7) |
| Anaemia | | | | | | 2 ( 0.7) |
| Cardiac disorders | 1 ( 0.7) | 1 ( 0.7) | 1 ( 0.6) | 1 ( 0.6) | 1 ( 0.3) | 1 ( 0.3) |
| Arrhythmia | 1 ( 0.7) | | | | | |
| Atrial fibrillation | | 1 ( 0.7) | | | | |
| Cardiac failure | | 1 ( 0.7) | | | | |
| Palpitations | | | 1 ( 0.6) | 1 ( 0.6) | | |
| Cardio-respiratory arrest | | | | | 1 ( 0.3) | |
| Supraventricular extrasystoles | | | | | | 1 ( 0.3) |
| Ear and labyrinth disorders | 2 ( 1.4) | 1 ( 0.7) | | 1 ( 0.6) | 2 ( 0.7) | |
| Meniere's disease | 1 ( 0.7) | | | | | |
| Tinnitus | 1 ( 0.7) | | | 1 ( 0.6) | | |
| Vertigo positional | | 1 ( 0.7) | | | | |
| Vertigo | | | | | 1 ( 0.3) | |
| Ear discomfort | | | | | 1 ( 0.3) | |
| Eye disorders | 2 ( 1.4) | 4 ( 2.9) | | 1 ( 0.6) | 1 ( 0.3) | 4 ( 1.3) |
| Cataract | | 3 ( 2.1) | | | | |
| Conjunctivitis allergic | 1 ( 0.7) | 1 ( 0.7) | | | | |
| Vision blurred | 1 ( 0.7) | | | | 1 ( 0.3) | |
| Ocular hyperaemia | | | | 1 ( 0.6) | | |
| Asthenopia | | | | | | 1 ( 0.3) |
| Astigmatism | | | | | | 1 ( 0.3) |
| Erythema of eyelid | | | | | | 1 ( 0.3) |
| Eyelid oedema | | | | | | 1 ( 0.3) |
| Vitreous detachment | | | | | | 1 ( 0.3) |
| Vitreous floaters | | | | | | 1 ( 0.3) |
| Gastrointestinal disorders | 18 ( 12.9) | 17 ( 12.1) | 4 ( 2.6) | 11 ( 7.1) | 15 ( 4.9) | 23 ( 7.5) |
| Abdominal discomfort | 1 ( 0.7) | | | 2 ( 1.3) | 1 ( 0.3) | 2 ( 0.7) |
| Abdominal distension | 1 ( 0.7) | 1 ( 0.7) | | | 1 ( 0.3) | |
| Abdominal pain | | 1 ( 0.7) | | | | |
| Abdominal pain upper | 1 ( 0.7) | 2 ( 1.4) | | | | |
| Constipation | 2 ( 1.4) | 1 ( 0.7) | 1 ( 0.6) | | 1 ( 0.3) | |
| Diarrhoea | 5 ( 3.6) | 2 ( 1.4) | | 2 ( 1.3) | 1 ( 0.3) | 5 ( 1.6) |
| Dyspepsia | | 4 ( 2.9) | | 2 ( 1.3) | 2 ( 0.7) | 12 ( 3.9) |
| Food poisoning | 1 ( 0.7) | | | | | |
| Gastritis | 2 ( 1.4) | 2 ( 1.4) | 1 ( 0.6) | 1 ( 0.6) | | 2 ( 0.7) |
| Gastrooesophageal reflux disease | | 2 ( 1.4) | | 1 ( 0.6) | 1 ( 0.3) | |

Reported adverse event terms were coded using MedDRA ver 14.1.

**wTable 1.** Incidence of treatment-emergent adverse events (continued)

Analysis Population: All Randomized Subjects

| Study | LVIA Study | | LVHB Study | | LVJF Study | |
|---|---|---|---|---|---|---|
| Treatment Group | Placebo | Tadalafil 5 mg | Placebo | Tadalafil 5 mg | Placebo | Tadalafil 5 mg |
| Treatment-emergent Adverse Events (Classification by MedDRA SOC/PT) | n (%) | n (%) | n (%) | n (%) | n (%) | n (%) |
| N | 140 | 140 | 154 | 155 | 304 | 306 |
| Gastrointestinal disorder | 1 ( 0.7) | | | | | |
| Gingival bleeding | | 1 ( 0.7) | | | | |
| Glossitis | 1 ( 0.7) | | | | | |
| Haemorrhoids | 1 ( 0.7) | | | | | |
| Hiatus hernia | | 1 ( 0.7) | | | | |
| Intestinal obstruction | 1 ( 0.7) | | | | | |
| Nausea | | 1 ( 0.7) | | 3 ( 1.9) | | |
| Periodontal disease | 1 ( 0.7) | | | | 2 ( 0.7) | |
| Stomatitis | 1 ( 0.7) | | | | 1 ( 0.3) | |
| Tooth loss | 1 ( 0.7) | | | | | |
| Toothache | 1 ( 0.7) | | | | | |
| Irritable bowel syndrome | | | | 1 ( 0.6) | | |
| Peptic ulcer | | | | 1 ( 0.6) | | |
| Vomiting | | | 1 ( 0.6) | | | 2 ( 0.7) |
| Epigastric discomfort | | | 1 ( 0.6) | | | |
| Abdominal pain lower | | | | | 1 ( 0.3) | |
| Dry mouth | | | | | | 1 ( 0.3) |
| Frequent bowel movements | | | | | | 1 ( 0.3) |
| Gastric ulcer | | | | | 2 ( 0.7) | |
| Gastritis atrophic | | | | | 1 ( 0.3) | |
| Gingivitis | | | | | | 1 ( 0.3) |
| Lip dry | | | | | 1 ( 0.3) | |
| Periodontitis | | | | | 2 ( 0.7) | |
| Gastrointestinal motility disorder | | | | | 1 ( 0.3) | |
| General disorders and administration site conditions | 3 ( 2.1) | | 1 ( 0.6) | 2 ( 1.3) | 1 ( 0.3) | 2 ( 0.7) |
| Chest pain | 1 ( 0.7) | | | | | |
| Pyrexia | 2 ( 1.4) | | | | | |
| Fatigue | | | 1 ( 0.6) | 1 ( 0.6) | | |
| Oedema peripheral | | | | 1 ( 0.6) | | 1 ( 0.3) |
| Thirst | | | 1 ( 0.6) | | | |
| Granuloma | | | | | 1 ( 0.3) | |
| Malaise | | | | | | 1 ( 0.3) |
| Therapeutic response unexpected | | | | | | 1 ( 0.3) |
| Hepatobiliary disorders | | | 4 ( 2.6) | 1 ( 0.6) | 3 ( 1.0) | 3 ( 1.0) |
| Hepatic function abnormal | | | 4 ( 2.6) | | 3 ( 1.0) | 1 ( 0.3) |
| Liver injury | | | | 1 ( 0.6) | | |
| Cholecystitis | | | | | | 1 ( 0.3) |
| Liver disorder | | | | | | 1 ( 0.3) |

Reported adverse event terms were coded using MedDRA ver 14.1.

**wTable 1.** Incidence of treatment-emergent adverse events (continued)

Analysis Population: All Randomized Subjects

| Study | LVIA Study | | LVHB Study | | LVJF Study | |
|---|---|---|---|---|---|---|
| Treatment Group | Placebo | Tadalafil 5 mg | Placebo | Tadalafil 5 mg | Placebo | Tadalafil 5 mg |
| Treatment-emergent Adverse Events (Classification by MedDRA SOC/PT) | n (%) | n (%) | n (%) | n (%) | n (%) | n (%) |
| N | 140 | 140 | 154 | 155 | 304 | 306 |
| Immune system disorders | 2 ( 1.4) | 1 ( 0.7) | 1 ( 0.6) | | | |
| Seasonal allergy | 2 ( 1.4) | 1 ( 0.7) | 1 ( 0.6) | | | |
| Infections and infestations | 22 ( 15.7) | 18 ( 12.9) | 8 ( 5.2) | 6 ( 3.9) | 22 ( 7.2) | 21 ( 6.9) |
| Acute sinusitis | 1 ( 0.7) | | | | | |
| Bronchitis | 1 ( 0.7) | | | 1 ( 0.6) | 1 ( 0.3) | 1 ( 0.3) |
| Chronic sinusitis | | 1 ( 0.7) | | | | |
| Empyema | 1 ( 0.7) | 1 ( 0.7) | | | | |
| Herpes zoster | 1 ( 0.7) | | | | 2 ( 0.7) | |
| Influenza | | 1 ( 0.7) | | | 1 ( 0.3) | 1 ( 0.3) |
| Nasopharyngitis | 18 ( 12.9) | 14 ( 10.0) | 4 ( 2.6) | 2 ( 1.3) | 10 ( 3.3) | 13 ( 4.2) |
| Otitis media | | 1 ( 0.7) | | | | 1 ( 0.3) |
| Pharyngitis | 1 ( 0.7) | 2 ( 1.4) | | | | 3 ( 1.0) |
| Upper respiratory tract infection | 1 ( 0.7) | | | | 5 ( 1.6) | 2 ( 0.7) |
| Oral herpes | | 1 ( 0.7) | | | | |
| Cellulitis | | | 1 ( 0.6) | | | |
| Herpes simplex | | | 1 ( 0.6) | | 1 ( 0.3) | |
| Infection | | | 1 ( 0.6) | | | |
| Otitis media chronic | | | | 1 ( 0.6) | | |
| Pneumonia | | | 1 ( 0.6) | | | |
| Tinea pedis | | | | 1 ( 0.6) | | |
| Adenoiditis | | | | 1 ( 0.6) | | |
| Fungal skin infection | | | | | | 1 ( 0.3) |
| Gastroenteritis | | | | | 1 ( 0.3) | |
| Tonsillitis | | | | | | 1 ( 0.3) |
| Tracheitis | | | | | | 1 ( 0.3) |
| Urinary tract infection | | | | | 1 ( 0.3) | |
| Injury, poisoning and procedural complications | 1 ( 0.7) | 3 ( 2.1) | 1 ( 0.6) | | 5 ( 1.6) | 3 ( 1.0) |
| Fall | | 2 ( 1.4) | | | | |
| Femur fracture | | 1 ( 0.7) | | | | |
| Fracture | | 1 ( 0.7) | | | | |
| Contusion | 1 ( 0.7) | | | | | |
| Arthropod sting | | | 1 ( 0.6) | | 1 ( 0.3) | |
| Hand fracture | | | | | 1 ( 0.3) | |
| Joint dislocation | | | | | | 1 ( 0.3) |
| Ligament sprain | | | | | 2 ( 0.7) | |
| Spinal cord injury cervical | | | | | 1 ( 0.3) | |
| Muscle strain | | | | | | 1 ( 0.3) |
| Upper limb fracture | | | | | | 1 ( 0.3) |

Reported adverse event terms were coded using MedDRA ver 14.1.

**wTable 1.** Incidence of treatment-emergent adverse events (continued)

Analysis Population: All Randomized Subjects

| Study | LVIA Study | | LVHB Study | | LVJF Study | |
|---|---|---|---|---|---|---|
| Treatment Group | Placebo | Tadalafil 5 mg | Placebo | Tadalafil 5 mg | Placebo | Tadalafil 5 mg |
| Treatment-emergent Adverse Events (Classification by MedDRA SOC/PT) | n (%) | n (%) | n (%) | n (%) | n (%) | n (%) |
| N | 140 | 140 | 154 | 155 | 304 | 306 |
| Investigations | | 4 ( 2.9) | 9 ( 5.8) | 5 ( 3.2) | 15 ( 4.9) | 12 ( 3.9) |
| Alanine aminotransferase increased | | 1 ( 0.7) | | | 4 ( 1.3) | |
| Aspartate aminotransferase increased | | 1 ( 0.7) | | | 6 ( 2.0) | 1 ( 0.3) |
| Blood creatine phosphokinase increased | | 1 ( 0.7) | 2 ( 1.3) | 2 ( 1.3) | 7 ( 2.3) | 7 ( 2.3) |
| Prostatic specific antigen increased | | 1 ( 0.7) | 1 ( 0.6) | 2 ( 1.3) | | |
| Tumour marker increased | | 1 ( 0.7) | | | | |
| Urinary sediment present | | 1 ( 0.7) | | | | |
| Occult blood positive | | 1 ( 0.7) | | | | |
| Gamma-glutamyltransferase increased | | | 4 ( 2.6) | | 3 ( 1.0) | 1 ( 0.3) |
| Glucose urine present | | | 1 ( 0.6) | 1 ( 0.6) | | 2 ( 0.7) |
| White blood cell count increased | | | 1 ( 0.6) | | | |
| Blood bilirubin increased | | | | | 1 ( 0.3) | |
| Blood chloride decreased | | | | | | 1 ( 0.3) |
| Blood sodium decreased | | | | | | 1 ( 0.3) |
| Blood urine | | | | | 1 ( 0.3) | |
| Creatinine renal clearance decreased | | | | | 1 ( 0.3) | 4 ( 1.3) |
| White blood cell count decreased | | | | | 2 ( 0.7) | |
| Metabolism and nutrition disorders | 1 ( 0.7) | | | 1 ( 0.6) | | 2 ( 0.7) |
| Diabetes mellitus | 1 ( 0.7) | | | | | |
| Hypokalaemia | | | | 1 ( 0.6) | | |
| Glucose tolerance impaired | | | | | | 1 ( 0.3) |
| Hyperuricaemia | | | | | | 1 ( 0.3) |
| Musculoskeletal and connective tissue disorders | 5 ( 3.6) | 8 ( 5.7) | 3 ( 1.9) | 15 ( 9.7) | 10 ( 3.3) | 15 ( 4.9) |
| Arthralgia | 2 ( 1.4) | 1 ( 0.7) | | | 1 ( 0.3) | 2 ( 0.7) |
| Arthritis | 1 ( 0.7) | | | | | |
| Back pain | 2 ( 1.4) | 2 ( 1.4) | 1 ( 0.6) | 4 ( 2.6) | 3 ( 1.0) | 4 ( 1.3) |
| Musculoskeletal pain | | 2 ( 1.4) | | 2 ( 1.3) | | |
| Myalgia | | 2 ( 1.4) | | 6 ( 3.9) | 1 ( 0.3) | 3 ( 1.0) |
| Pain in extremity | | 1 ( 0.7) | | | 1 ( 0.3) | 2 ( 0.7) |
| Periarthritis | | 1 ( 0.7) | | | | 1 ( 0.3) |
| Spinal column stenosis | | 1 ( 0.7) | | | | |
| Tenosynovitis | | 1 ( 0.7) | | | 1 ( 0.3) | |
| Muscle spasms | | | 1 ( 0.6) | | | |

Reported adverse event terms were coded using MedDRA ver 14.1.

**wTable 1.** Incidence of treatment-emergent adverse events (continued)

Analysis Population: All Randomized Subjects

| Study | LVIA Study | | LVHB Study | | LVJF Study | |
|---|---|---|---|---|---|---|
| Treatment Group | Placebo | Tadalafil 5 mg | Placebo | Tadalafil 5 mg | Placebo | Tadalafil 5 mg |
| Treatment-emergent Adverse Events (Classification by MedDRA SOC/PT) | n (%) | n (%) | n (%) | n (%) | n (%) | n (%) |
| N | 140 | 140 | 154 | 155 | 304 | 306 |
| Myositis | | | | 1 ( 0.6) | | |
| Osteoarthritis | | | 1 ( 0.6) | | | |
| Joint range of motion decreased | | | | 1 ( 0.6) | | |
| Muscle tightness | | | | 2 ( 1.3) | | |
| Arthropathy | | | | | | 1 ( 0.3) |
| Gouty arthritis | | | | | 1 ( 0.3) | |
| Scoliosis | | | | | 1 ( 0.3) | |
| Musculoskeletal stiffness | | | | | 1 ( 0.3) | 1 ( 0.3) |
| Limb discomfort | | | | | | 1 ( 0.3) |
| Neoplasms benign, malignant and unspecified (incl cysts and polyps) | | | 1 ( 0.6) | | | |
| Lymphoma | | | 1 ( 0.6) | | | |
| Nervous system disorders | 6 ( 4.3) | 5 ( 3.6) | 2 ( 1.3) | 3 ( 1.9) | 6 ( 2.0) | 14 ( 4.6) |
| Dizziness | 2 ( 1.4) | 1 ( 0.7) | | | | 2 ( 0.7) |
| Headache | 3 ( 2.1) | 3 ( 2.1) | 1 ( 0.6) | 3 ( 1.9) | 6 ( 2.0) | 9 ( 2.9) |
| Hypoaesthesia | | 1 ( 0.7) | | | | 1 ( 0.3) |
| Sciatica | 1 ( 0.7) | | | | | |
| Cervicobrachial syndrome | | | 1 ( 0.6) | | | |
| Burning sensation | | | | | | 1 ( 0.3) |
| Carotid arteriosclerosis | | | | | | 1 ( 0.3) |
| Psychiatric disorders | | 1 ( 0.7) | | | 1 ( 0.3) | |
| Insomnia | | 1 ( 0.7) | | | | |
| Anxiety disorder | | | | | 1 ( 0.3) | |
| Renal and urinary disorders | 1 ( 0.7) | 2 ( 1.4) | 1 ( 0.6) | 1 ( 0.6) | 3 ( 1.0) | |
| Dysuria | | 1 ( 0.7) | | | | |
| Urinary retention | 1 ( 0.7) | 1 ( 0.7) | 1 ( 0.6) | | 1 ( 0.3) | |
| Calculus ureteric | | | | 1 ( 0.6) | | |
| Nephrolithiasis | | | | 1 ( 0.6) | | |
| Haematuria | | | | | 1 ( 0.3) | |
| Proteinuria | | | | | 1 ( 0.3) | |
| Renal impairment | | | | | 1 ( 0.3) | |
| Reproductive system and breast disorders | 1 ( 0.7) | 1 ( 0.7) | | 1 ( 0.6) | | 2 ( 0.7) |
| Nipple disorder | | 1 ( 0.7) | | | | |
| Prostatitis | 1 ( 0.7) | | | | | |
| Spontaneous penile erection | | | | 1 ( 0.6) | | 1 ( 0.3) |
| Erection increased | | | | | | 1 ( 0.3) |

Reported adverse event terms were coded using MedDRA ver 14.1.

**wTable 1.** Incidence of treatment-emergent adverse events (continued)

Analysis Population: All Randomized Subjects

| Study | LVIA Study | | LVHB Study | | LVJF Study | |
|---|---|---|---|---|---|---|
| Treatment Group | Placebo | Tadalafil 5 mg | Placebo | Tadalafil 5 mg | Placebo | Tadalafil 5 mg |
| Treatment-emergent Adverse Events (Classification by MedDRA SOC/PT) | n (%) | n (%) | n (%) | n (%) | n (%) | n (%) |
| N | 140 | 140 | 154 | 155 | 304 | 306 |
| Respiratory, thoracic and mediastinal disorders | 3 ( 2.1) | 2 ( 1.4) | 1 ( 0.6) | 4 ( 2.6) | 4 ( 1.3) | 2 ( 0.7) |
| Cough | | 1 ( 0.7) | | 1 ( 0.6) | 2 ( 0.7) | |
| Epistaxis | 1 ( 0.7) | | | | 1 ( 0.3) | |
| Rhinitis allergic | | 1 ( 0.7) | | 2 ( 1.3) | | 1 ( 0.3) |
| Rhinorrhoea | 2 ( 1.4) | | | 1 ( 0.6) | | 1 ( 0.3) |
| Asthma | | | | 1 ( 0.6) | | |
| Dysphonia | | | 1 ( 0.6) | | | |
| Nasal obstruction | | | 1 ( 0.6) | | | |
| Upper respiratory tract inflammation | | | | | 1 ( 0.3) | |
| Oropharyngeal discomfort | | | | | | 1 ( 0.3) |
| Skin and subcutaneous tissue disorders | 6 ( 4.3) | 4 ( 2.9) | 2 ( 1.3) | 1 ( 0.6) | 5 ( 1.6) | 7 ( 2.3) |
| Acne | | 1 ( 0.7) | | | | |
| Dermatitis contact | 2 ( 1.4) | | | | | |
| Drug eruption | | 1 ( 0.7) | | | | |
| Eczema | 3 ( 2.1) | | 1 ( 0.6) | 1 ( 0.6) | 2 ( 0.7) | 4 ( 1.3) |
| Erythema | | 1 ( 0.7) | | | | |
| Rash | 1 ( 0.7) | | 1 ( 0.6) | | | |
| Xeroderma | | 1 ( 0.7) | | | | |
| Dermatitis | | | | | | 1 ( 0.3) |
| Dermatitis allergic | | | | | 1 ( 0.3) | |
| Photosensitivity reaction | | | | | | 1 ( 0.3) |
| Pruritus | | | | | | 1 ( 0.3) |
| Urticaria | | | | | 2 ( 0.7) | |
| Surgical and medical procedures | 3 ( 2.1) | 5 ( 3.6) | | 2 ( 1.3) | 4 ( 1.3) | |
| Colon polypectomy | | 1 ( 0.7) | | 1 ( 0.6) | | |
| Internal fixation of fracture | | 1 ( 0.7) | | | | |
| Ureteral catheterisation | | 1 ( 0.7) | | | | |
| Gastrointestinal tube insertion | 1 ( 0.7) | | | | | |
| Rectal polypectomy | | | | | 1 ( 0.3) | |
| Tooth extraction | 2 ( 1.4) | 2 ( 1.4) | | 1 ( 0.6) | 2 ( 0.7) | |
| Electrocauterisation | | | | | 1 ( 0.3) | |
| Vascular disorders | 1 ( 0.7) | 2 ( 1.4) | 1 ( 0.6) | 2 ( 1.3) | 1 ( 0.3) | 2 ( 0.7) |
| Hypertension | 1 ( 0.7) | | | | | |
| Hot flush | | 2 ( 1.4) | 1 ( 0.6) | 2 ( 1.3) | 1 ( 0.3) | 2 ( 0.7) |

Reported adverse event terms were coded using MedDRA ver 14.1.

**The sample WinBUGS codes for the proposed Bayesian meta-analysis models and a detailed method for Bayesian computations**

We present here the sample WinBUGS codes for Models 1–3. Posterior samples were drawn with convergence assessed using trace plots, sample autocorrelations, and other standard convergence diagnostics. For each model, a burn-in period of 10,000 iterations was used, with 20,000 subsequent iterations retained for posterior estimations.

<Model 1>

```
#Model
model{
        for (i in 1:Nae){                                       #Loop over the Nae=193 adverse events
                for (k in 1:Nstud){                             #Loop over the Nstud=3 studies
                        X[k, i] ~ dbin(c[k, b[i], j[i]], Nc[k])
                        Y[k, i] ~ dbin(t[k, b[i], j[i]], Nt[k])
                        logit(c[k, b[i], j[i]]) <- gamma[k, b[i], j[i]]
                        logit(t[k, b[i], j[i]]) <- gamma[k, b[i], j[i]] + theta[k, b[i], j[i]]

                        gamma[k, b[i], j[i]] ~ dnorm(mu.gamma[b[i], j[i]], tau.gamma[b[i], j[i]])
                        theta[k, b[i], j[i]] ~ dnorm(mu.theta[b[i], j[i]], tau.theta[b[i], j[i]])
                }

                        mu.gamma[b[i], j[i]] ~ dnorm(mu.gamma.0[b[i]], tau.gamma.0[b[i]])
                        tau.gamma[b[i], j[i]] ~ dgamma(3,1)

                        mu.theta[b[i], j[i]] ~ dnorm(mu.theta.0[b[i]], tau.theta.0[b[i]])
                        tau.theta[b[i], j[i]] ~ dgamma(3,1)

                        OR[b[i], j[i]] <- exp(mu.theta[b[i], j[i]])
                        prob1.OR[b[i], j[i]] <- 1−step(1−OR[b[i], j[i]])        #Posterior mean of p(OR >1)
        }

                        #SOC level parameters
```

```
                    for (l in 1:B){                    #Loop over the B=22 body systems
                    mu.gamma.0[l] ~ dnorm(mu.gamma.00, tau.gamma.00)
                    tau.gamma.0[l] ~ dgamma(3,1)
                    mu.theta.0[l] ~ dnorm(mu.theta.00, tau.theta.00)
                    tau.theta.0[l] ~ dgamma(3,1)
                    }

                    #hyperpriors for gamma's
                    mu.gamma.00 ~ dnorm(0, 0.1)
                    tau.gamma.00 ~ dgamma(3,1)
                    #hyperpriors for theta's
                    mu.theta.00 ~ dnorm(0, 0.1)
                    tau.theta.00 ~ dgamma(3,1)
}
```

<Model 2>

```
#Model
model{
        for (i in 1:Nae){                              #Loop over the Nae=193 adverse events
                for (k in 1:Nstud){                    #Loop over the Nstud=3 studies
                        X[k, i] ~ dbin(c[k, b[i], j[i]], Nc[k])
                        Y[k, i] ~ dbin(t[k, b[i], j[i]], Nt[k])
                        logit(c[k, b[i], j[i]]) <- gamma[k, b[i], j[i]]
                        logit(t[k, b[i], j[i]]) <- gamma[k, b[i], j[i]] + theta[k, b[i], j[i]]

                        gamma[k, b[i], j[i]] ~ dnorm(mu.gamma[b[i], j[i]], tau.gamma[b[i], j[i]])
                        theta[k, b[i], j[i]] ~ dnorm(mu.theta[b[i], j[i]], tau.theta[b[i], j[i]])
                }

                        mu.gamma[b[i], j[i]] ~ dnorm(mu.gamma.0[b[i]], tau.gamma.0[b[i]])
                        tau.gamma[b[i], j[i]] ~ dgamma(3,1)

                        p0[i] ~ dbern(pi[b[i]])        # probability of having a unit point mass at 0
                        mu.theta1[b[i], j[i]] ~ dnorm(mu.theta.0[b[i]], tau.theta.0[b[i]])
                        #The second term of mixture prior
                        mu.theta[b[i], j[i]] <- (1−p0[i]) * mu.theta1[b[i], j[i]]
                        # theta=0 with probability of p0[i] and theta=theta1 with probability of 1−p0[i]

                        tau.theta[b[i], j[i]] ~ dgamma(3,1)
```

```
                OR[b[i], j[i]] <- exp(mu.theta[b[i], j[i]])
                prob1.OR[b[i], j[i]] <- 1−step(1−OR[b[i], j[i]])        #Posterior mean of p(OR >1)
        }

                #SOC level parameters
                for (l in 1:B){                         #Loop over the B=22 body systems
                pi[l] ~ dbeta(alpha.pi, beta.pi)
                mu.gamma.0[l] ~ dnorm(mu.gamma.00, tau.gamma.00)
                tau.gamma.0[l] ~ dgamma(3,1)
                mu.theta.0[l] ~ dnorm(mu.theta.00, tau.theta.00)
                tau.theta.0[l] ~ dgamma(3,1)
                }

                #hyperpriors for gamma's
                mu.gamma.00 ~ dnorm(0, 0.1)
                tau.gamma.00 ~ dgamma(3,1)
                #hyperpriors for theta's
                mu.theta.00 ~ dnorm(0, 0.1)
                tau.theta.00 ~ dgamma(3,1)
                # hyperpriors for pi's
                alpha.pi ~ dexp(0.1) I(1, )
                beta.pi ~ dexp(0.1) I(1, )
}
```

<Model 3>

```
#Model
model{
        for (i in 1:Nae){                                       #Loop over the Nae=193 adverse events
                for (k in 1:Nstud){                             #Loop over the Nstud=3 studies
                        X[k, i] ~ dbin(c[k, b[i], j[i]], Nc[k])
                        Y[k, i] ~ dbin(t[k, b[i], j[i]], Nt[k])
                        logit(c[k, b[i], j[i]]) <- gamma[k, b[i], j[i]]
                        logit(t[k, b[i], j[i]]) <- gamma[k, b[i], j[i]] + theta[k, b[i], j[i]]

                        gamma[k, b[i], j[i]] ~ dnorm(mu.gamma[b[i], j[i]], tau.gamma[b[i], j[i]])
                        theta[k, b[i], j[i]] ~ dnorm(mu.theta[b[i], j[i]], tau.theta[b[i], j[i]])
                }

                        mu.gamma[b[i], j[i]] ~ dnorm(0, 0.01)
                        tau.gamma[b[i], j[i]] ~ dgamma(3,1)
```

```
                p0[i] ~ dbern(pi[i])                    # probability of having a unit point mass at 0
                pi1[i] ~ dbern(1.0)
                pi[i] <- 0.5 * pi1[i]
                mu.theta1[b[i], j[i]] ~ dnorm(0, 0.01)
                #The second term of mixture prior
                mu.theta[b[i], j[i]] <- (1−p0[i]) * mu.theta1[b[i], j[i]]
                # theta=0 with probability of 0.5 and theta=theta1 with probability of 0.5

                tau.theta[b[i], j[i]] ~ dgamma(3,1)

                OR[b[i], j[i]] <- exp(mu.theta[b[i], j[i]])
                prob1.OR[b[i], j[i]] <- 1−step(1−OR[b[i], j[i]])        #Posterior mean of p(OR >1)
        }

}
```