SA80

Priors on between-study heterogeneity in Bayesian meta-analysis models: A simulation study using Hamiltonian Monte Carlo sampling

Dunnewind N[1], Ainsworth C[2], Ren S[3], Kroep S[1]

[1]OPEN Health HEOR & Market Access, Rotterdam, The Netherlands, [2]OPEN Health HEOR & Market Access, London, UK, [3]University of Sheffield, Sheffield, UK

OPEN HEALTH

INTRODUCTION & OBJECTIVE

- Meta-analysis (MA) plays a crucial role in health economic and outcomes research. Random effects (RE) models in MA allow for heterogeneity in treatment effects between studies.
- In a Bayesian framework, prior beliefs are assigned to model parameters, including the between-study heterogeneity, often using non-informative priors.[1]
- Previous research[2] explored the use of different priors for the heterogeneity parameter on MA results using WinBUGS[3] (Gibbs sampling).
- Hamiltonian Monte Carlo (HMC) sampling, implemented in Stan[4], is used increasingly more often in recent years and should be more efficient than Gibbs sampling.[5]

OBJECTIVES

- The aim of this study was to replicate Lambert et al. (2005)[2] instead using HMC sampling.
- Further, the study aimed to expand previous research by investigating performance of additional priors on the between-study heterogeneity parameter (i.e., an informative prior and improper flat prior), and the use of non-centred model parameterization.
- Finally, the study sought to explore model fit under RE models (for each prior investigated) and compare with that for fixed effect (FE) models.

METHODS

Model

- Three meta-analysis models were fitted:
  - **An RE model with centred parameterization**, as used for example in NICE technical support documents:[1]

$$y_i \sim N(\mu_i, s_i^2) \qquad \mu_i \sim N(\theta, \tau^2)$$

    where y_i and s_i are the observed log-odds ratio and its standard error for study i, μ_i is the modelled log-odds ratio, τ is the between-study standard deviation (SD) and θ is the pooled log-odds ratio.
  - **An RE model with non-centred parameterization**:[5]

$$y_i \sim N(\theta + \tau\mu_i, s_i^2) \qquad \mu_i \sim N(0,1)$$

  - **An FE model**:

$$y_i \sim N(\theta, s_i^2)$$

- Models were implemented in Stan.[4] Four chains with 2000 iterations (1000 warmup) each were sampled, with the target average acceptance probability set to 0.98.
- An uninformative prior was placed on the pooled log-odds ratio: $\theta \sim N(0,100)$.
- Fourteen priors for the between-study SD, varying in distribution and level of information, were tested (Table 1).
  - Prior 0 is an improper flat prior.
  - Prior 13 is the empirical Bayes[6] prior.
  - Prior 14 is the informative prior suggested by Turner et al.[7] for a pharmacological vs placebo comparison and a subjective outcome.

Table 1. Between-study SD priors

0.	$\boldsymbol{\tau \sim \text{Uniform}(0, \infty)}$	8.	$\frac{1}{\tau^2} \sim \text{Pareto}(1, 0.25)$
1.	$\boldsymbol{\frac{1}{\tau^2} \sim \text{Gamma}(0.001, 0.001)}$	9.	$\boldsymbol{\tau \sim \text{Uniform}(0, 100)}$
2.	$\boldsymbol{\frac{1}{\tau^2} \sim \text{Gamma}(0.1, 0.1)}$	10	$\tau \sim \text{Uniform}(0, 2)$
3.	$\log \tau^2 \sim \text{Uniform}(-10, 10)$	11.	$\boldsymbol{\tau \sim N(0, 100); \quad \tau > 0}$
4.	$\log \tau^2 \sim \text{Uniform}(-10, 1.386)$	12.	$\boldsymbol{\tau \sim N(0, 1); \quad \tau > 0}$
5.	$\boldsymbol{\tau^2 \sim \text{Uniform}(1/1000, 1000)}$	13.	$\boldsymbol{\tau \sim \text{LogLogistic}(S_0); \quad S_0 = \sqrt{K/\sum s_i^{-2}}}$
6.	$\tau^2 \sim \text{Uniform}(1/1000, 4)$	14.	$\boldsymbol{\tau^2 \sim \text{LogNormal}(-2.13, 1.58); \quad \tau > 0}$
7.	$\boldsymbol{\frac{1}{\tau^2} \sim \text{Pareto}(1, 0.001)}$		

Note: Lambert et al. (2005) used priors 1 to 13.[2] K is the number of studies.

The same case study and simulation study setup as in Lambert et al. (2005) were used.[2]

Case study

- Models were first fitted on data from five studies comparing the effects of short course versus long course antibiotics for acute otitis media.[8]

Simulation study

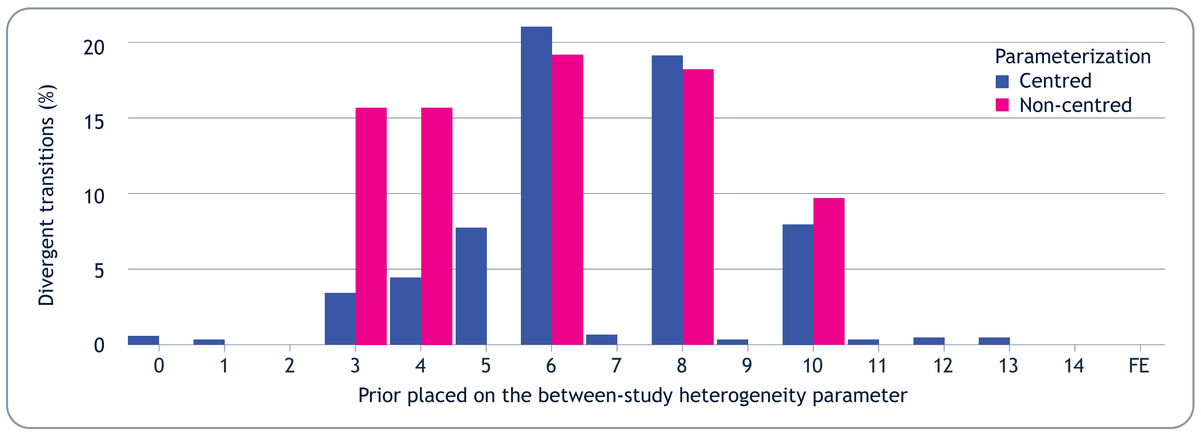
- One thousand simulated dichotomous outcome datasets, with differing true between-study SD and number of studies, were evaluated.
  - A random study effect was drawn from a normal distribution with mean zero, and a varied SD across three scenarios (0.001, 0.3, or 0.8).
  - The placebo log-odds, log-odds ratio and the random study effect were summed to obtain the odds ratio of the treatment arm, and the number of events is drawn from a binomial distribution. Scenarios with 5, 10 and 30 studies were evaluated, with 100 to 500 patients in each of the two study arms.
- Results were compared by assessing the bias, coverage, deviance information criterion (DIC), convergence (based on R-hats and Monte Carlo standard errors) and divergent transitions (an HMC diagnostic indicating Markov chain Monte Carlo estimators may be invalid and biased[8]).

RESULTS

Case study

- The case study demonstrated that priors with zero density at plausible values results in divergent transitions (Figure 1) and may therefore lead to biased results.[9]
- For most priors, the non-centred parametrization decreased the number of divergent transitions.

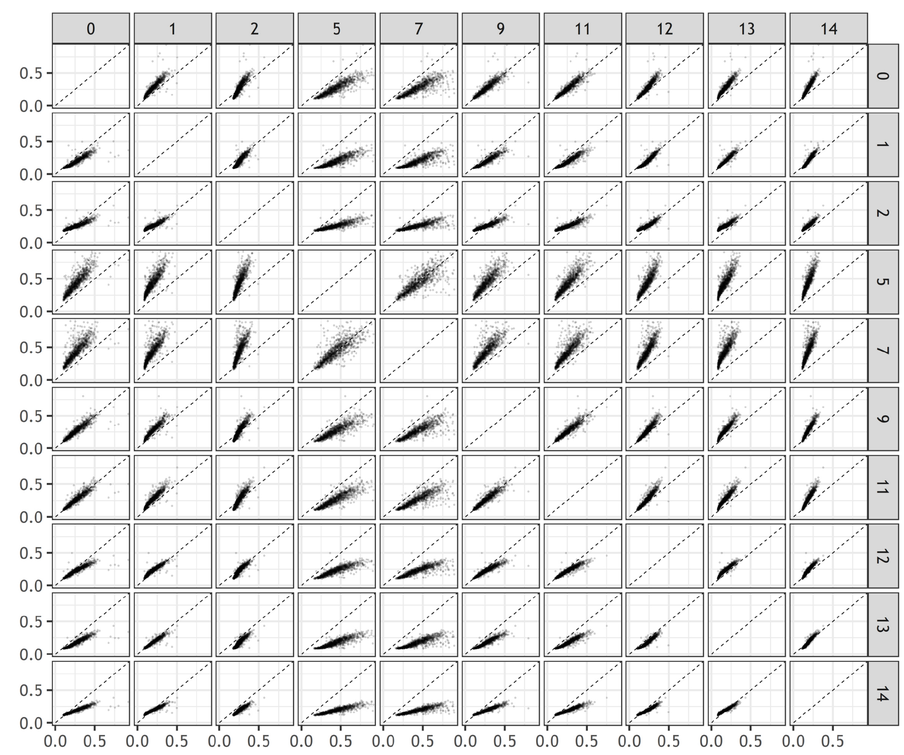
Figure 1. Divergent transitions as percentage of samples per prior and parameterization.



Simulation study

- Based on the observations in the case study, the non-centred parametrization and priors with no or very few divergent transitions were used in the simulation study.
- Figure 2 presents pairwise comparisons of between-study heterogeneity priors for the five-study scenario, illustrating the impact on the SD of the pooled log-odds ratio.
  - Within each pairwise comparison, dots (replications) which lie around the diagonal indicate similar results for the two priors being compared.
  - Conversely, dots which are consistently above (or below) the diagonal indicate a higher (or lower) SD of the pooled log-odds ratio for the row prior versus column prior in the comparison, i.e., more (or less) uncertainty in the treatment effect estimate under the row prior.
  - For example, in the pairwise comparison of informative prior 14 versus flat prior 0 (upper right subplot), there is decreased uncertainty in the treatment effect for prior 14 (dots are consistently above the diagonal).
- Among simulation scenarios, RE models with a gamma prior on precision (1), an empirical Bayes prior (13) and an informative prior on variance (14) were most likely to be favoured over FE models (lowest DIC).
- Convergence issues were more prominent when using the improper flat prior (0), the non-informative normal prior on SD (11), the non-informative priors on SD and variance (5, 9) and the non-informative Pareto prior on precision (7), especially in scenarios with fewer studies and higher true SD.
- The non-informative uniform prior on variance (5) still resulted in many divergent transitions, especially in scenarios with a true SD of 0.001.

Figure 2. Scatter plot matrix comparing all selected between-study SD prior distributions for the five-study scenario with a between-study SD of 0.3, displaying the posterior SD of the pooled log-odds ratio. Each point represents a simulation replication.



CONCLUSIONS

- Findings are very similar to results presented by Lambert et al. (2005)[2], despite the different parameterization and (more efficient[5]) sampling method.
- Selection of an appropriate between-study heterogeneity prior is crucial as it impacts the posterior treatment effect, particularly when dealing with a low number of studies.
- Caution must be exercised to avoid bias when using informative priors.
- In future studies, adopting non-centred model parameterization as standard practice is advisable over centred parameterization.
- HMC diagnostics, such as divergent transitions, are helpful and show that priors with zero density at plausible values should be avoided.

REFERENCES

1. Dias S, Welton NJ, Sutton AJ, Ades AE. NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials. London: National Institute for Health and Care Excellence (NICE); April 2014.
2. 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. *Stat Med*. 2005;24(15):2401-2428.
3. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS-a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and computing*. 2000;10:325-337.
4. Stan Development Team. RStan: The R interface to Stan. Published online 2022. https://mc-stan.org.
5. Betancourt M, Girolami M. Hamiltonian Monte Carlo for Hierarchical Models. *Current Trends in Bayesian Methodology with Applications*. 2015;79(30):2-4.
6. DuMouchel W, Normand S. Computer-Modeling and Graphical Strategies for Meta-Analysis. In: *Meta-Analysis in Medicine and Health Policy*. 2000:119-164.
7. Turner RM, Davey J, Clarke MJ, Thompson SG, Higgins JP. Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. *Int J Epidemiol*. 2012;41(3):818-827.8.
8. Venekamp RP, Sanders SL, Glasziou PP, Del Mar CB, Rovers MM. Antibiotics for acute otitis media in children. *Cochrane Database Syst Rev*. 2015;2015(6):CD0002199.
9. Betancourt M. A conceptual introduction to Hamiltonian Monte Carlo. Published online 2018. https://arxiv.org/abs/1701.02434

Presented at ISPOR Europe 2023