Quantifying the Effect of Beta Blockers on Mortality: A Bayesian Hierarchical Meta-Analysis

Corrà Sara   Shaboian Goar
5206191   5217162

1. Introduction

Myocardial infarctions, defined as damage to the myocardium caused by lack of oxygen, result in a significant number of fatalities every year (Mechanic, Gavin, and Grossman (2017)). Among the available pharmacological agents used for infarction prevention and damage minimization are beta-blockers, which aim at slowing down the heart rate. However, there are certain risks of secondary effects associated with their application, hence it is essential to conduct comprehensive research to determine how effective this medication is in alleviating risks associated with heart attacks. The study conducted by Yusuf et al. (1985) aims at investigating the efficiency of beta blockers by aggregating results on 65 individual clinical trials, both on short-term and long-term time frames. Various endpoints, such as infarction size, arrhythmias, threatened infarctions, were considered, with mortality rates being the primary focus. Drawing meaningful conclusions based on data obtained from individual studies requires aggregating the results in an appropriate way, which implies implementing statistical methodology known as meta-analysis, the theoretical and methodological aspects of which will be the focus of this project.

2. Methodology

2.1.   Description of the dataset

For the original analysis, Yusuf et al. (1985) collected data from randomised clinical trials through computer search and personal contact in the research field. The studies were included irrespective of results that individual researchers obtained. For quantification of the findings, the Peto odds ratio was used, which is a common way of approaching meta-analysis in a frequentist perspective. The estimator is defined as follows:

$$log\left(\widehat{POR}\right) = \frac{\sum_{i=1}^{s}(O_i - E_i)}{\sum_{i=1}^{s} V_i}, \ where$$

O is the observed number of occurrences of the investigated event, E is the expected number in the treatment group, and V is the estimate of the variance (Brockhaus, Bender, and Skipka (2014)). The results of the estimation for the long-term (3 months to 1 year) framework provided evidence for a noticeable (20%) reduction in the total number of deaths for patients in the treatment groups was found. On the other hand, for short-term (at one week) beta-blocker application, Yusuf et al. (1985) discovered improvements in terms of infarction size and ventricular arrhythmia, however, no conclusive evidence of the effects of treatment on mortality was found: only 6% risk of death reduction was determined. Bayesian framework is used in this analysis to determine if efficiency can be conclusively determined.

The outcomes of 15 studies[1] that performed medical trials, with subjects in the treatment groups being administered beta-blockers intravenously upon admission to the coronary care units, were used. The observational data is binomial: total number of patients and number of deaths are recorded for both control (n_0) and treatment (n_1) groups. For statistical analysis, a common way to treat binomial data is by using the *empirical logit* estimator (Gelman et al. (2013)):

$$y_j = log\left(\frac{y_{1j}}{n_{1j} - y_{1j}}\right) - log\left(\frac{y_{0j}}{n_{0j} - y_{0j}}\right)$$

This method allows for easy interpretation, especially in studies concerning medical trials (Lipsitz, Laird, and Harrington (1991)). Additionally, the logarithmic transformation of the odds ratio allows to obtain approximately normally distributed sample even for small trials (Gelman et al. (2013)), which is used as a common approach in hierarchical Bayesian models, which will be described further (Schmid (2001)). To obtain a preliminary idea on the values of empirical logit for the studies included in the analysis, a forest plot was used.
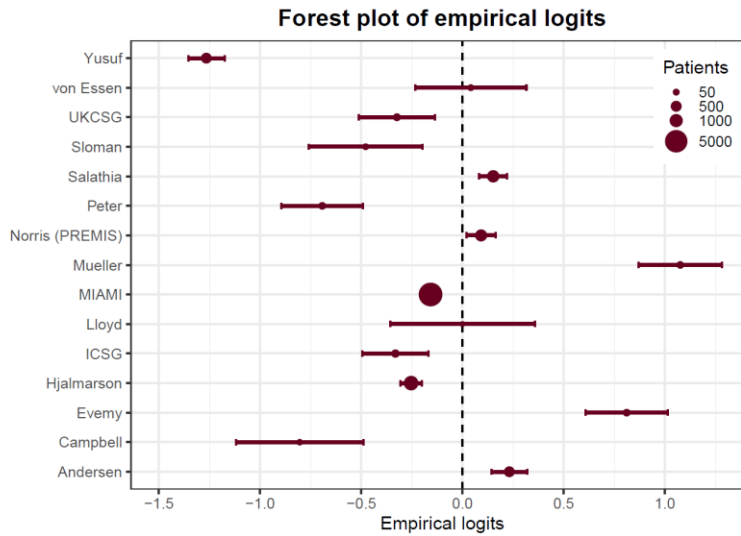

Figure 1: Forest plot for empirical logits for 15 studies and uncertainty associated to them

As evidenced by *Figure 1*, the studies differ in size and in the estimate they provide. It can be observed that bigger studies have lower uncertainty associated with them, the reasons for which will be discussed further. Additionally, it should be stressed that sample estimates provide that for some studies, the treatment reduces the risk of death significantly (for instance, the Yusuf study), whereas for other studies, such as Mueller or Evemy, the estimates demonstrate a harmful effect of administration of beta blockers. This observation warrants the need for further investigation of the observed effects.

2.2.   Meta-analysis in frequentist and Bayesian settings

Meta-analysis methodology implies aggregating results obtained in individual experiments, discerning magnitude and direction of the effects (Brockhaus, Bender, and Skipka (2014)). In holds that individual studies provide more moderate effects, and the differences between

---

[1] Studies that provided number of deaths in either treatment or control groups equal to 0 were not included in the analysis

them are likely to be quantitative rather than qualitative (Yusuf et al. (1985)), due to possibility of failing to obtain a statistically significant result, or even of producing contrary results because of non-availability of enough data. Examples of obtaining contrary results can be observed in *Figure 1*. Hence, combining studies with different sample size provides benefit. Furthermore, meta-analysis allows to minimize possible effects of individual selection and concealment bias (Yusuf et al. (1985)). However, for the benefits to hold, it is important to conclusively determine absence of bias in choosing the studies for meta-analysis. The process of choosing the studies for the original analysis comprehensively covered virtually the whole research field (with the possibility of several studies being overlooked), thus delivering a reliable dataset.

Meta-analysis in Bayesian framework allows to account for all sources of variability of the data. Additionally, in such framework, it is possible to incorporate external information relevant to the scientific problem, while also guaranteeing model flexibility (Schmid (2001), M. L. Johnson (2010)). This implies including several parameters which may be useful in building a model, while avoiding the problem of overfitting (M. L. Johnson (2010)). The concept of meta-analysis in Bayesian setting intrinsically implies hierarchical modelling[2].

2.1. Hierarchical Bayesian modelling

The framework of a Bayesian hierarchical model is built on the assumption that the observations in each group are generated from a specific distribution, wherein the mean parameter of this distribution is itself drawn from a common distribution that is shared across all groups. This concept can be expressed in formula as follows:

$$\{y_{1,j},\ldots,y_{n_j,j}\} \stackrel{iid}{\sim} p(y|\phi_j) \tag{1}$$

where the index i represents the observation, whereas the index j references the group number.

$$\{\phi_1,\ldots,\phi_m|\psi\} \stackrel{iid}{\sim} p(\phi|\psi) \tag{2}$$

$$\psi \sim p(\psi)$$

Equation (1) denotes the inherent sampling variability present within each group, while equation (2) characterizes the sampling variability observed between groups. The final one reveals information about a specific, yet unknown, fixed quantity through its prior distribution.

The sample of observed trials can be interpreted as a sample from all the possible trials that may be obtained. Since the presence of a common population is assumed for the parameter θ_j, it follows that information contributed by one group can be used to estimate the mean in the other. Hence, a significant and attractive aspect of this approach lies in its capacity to concurrently learn from each study (group), while also gaining a deeper understanding of the broader population of studies, thus pooling information across groups.

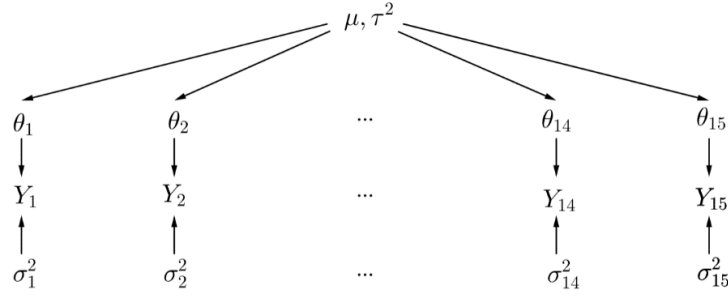[2] As discussed in Section Exchangeability.

Figure 2: Structure of the hierarchical model

From *Figure* 2, it appears evident that the hyperparameters only affect y through θ. Hence, the data y and the hyperparameters ψ are conditionally independent given the model parameters θ, so the likelihood satisfies (A. Johnson et al. (2020)):

$$f(y|\phi,\psi) = f(y|\phi)$$

2.1.1. Exchangeability

The term 'exchangeability' was introduced by de Finetti in the context of personal probability: it described a particular sense in which quantities treated as random variables in a probability specification are thought to be similar (Draper et al. (1993)).

In particular, $X_1, \ldots X_n$ are exchangeable random variables if they play a symmetrical role in relation to all problems of probability or, in other words, if the probability that $X_{i_1}, X_{i_2}, \ldots, X_{i_n}$, satisfy a given condition is always the same, however the distinct indices $i_1, \ldots, i_n$ are chosen (De Finetti (1937)). A probability distribution P for the processes is exchangeable if it's invariant under permutations:

$$PX_1 = e_1, \ldots, X_n = e_n = PX_{\pi(1)} = e_1, \ldots, X_{\pi(n)} = e_n,$$

Where $e_1, e_2, \ldots, e_n$ is any sequence of possible values, and π is any permutation (Diaconis (1988)). Since *de Finetti's Theorem* allows to extend the results to all *n*, the main ideas of this can be summarized as follows (Hoff (2009)):

$$\begin{cases} Y_1, \ldots, Y_n | \theta \ are \ i.i.d. \\ \theta \sim p(\theta) \end{cases} \Leftrightarrow \ Y_1, \ldots, Y_n \ are \ exchangeable \ for \ all \ n$$

Hence, the presence of exchangeable studies, as observed in meta-analysis, inherently lends itself to the formulation of a two-stage Bayesian hierarchical model.

2.1.2. Hierarchical Normal model

As discussed before, the choice to use empirical logits implies specifying a Normal hierarchical model. Additionally, it should be stressed that the analysis is conducted using point estimates, not single observations, as data points. Thus, the model is structured as follows:

$$y_j \sim N\big(\theta_j, \sigma_j^2\big),$$

where the parameter σ^2 is assumed to be a known quantity (Gelman et al. (2013)). It is calculated according to the following formula:

$$\sigma_j^2 = \frac{1}{y_{1j}} + \frac{1}{n_{1j} - y_{1j}} + \frac{1}{y_{0j}} + \frac{1}{n_{0j} - y_{0j}}$$

Thus, a prior distribution is not incorporated for it; this is a common assumption with underlying binomial sampling distributions. Moreover, attaching a prior distribution to the σ_j^2 independently of the θ_j or the underlying binomial parameters is not appropriate because of the functional relationship between the variance and mean of the binomial. In fact, where primary interest resides in the treatment effects, independent (beta) prior distributions for the two proportions involved in each study may not be appropriate (Carlin (1992)). It is worth noting that the assumption of normally distributed observed values with known variance is likely to be reasonable in most situations, as long as the studies are large and observed counts are not too small.

The sampling distribution of the parameter θ, under the assumption of exchangeability of the studies, is presented below (Gelman et al. (2013)):

$$\theta_1, \dots, \theta_m | \mu, \tau \sim \prod_{j=1}^{m} N\left(\theta_j | \mu, \tau^2\right)$$

Hence, once the parameters μ and τ are known, the parameters are independent of the other $\theta's$, as well as independent of the data from groups other than j.

Prior distributions to μ and τ are defined as follows:

$$\mu \sim N(\mu_0, \gamma_0^2)$$

$$\frac{1}{\tau^2} \sim Gamma\left(\frac{\eta_0}{2}, \frac{\eta_0 \tau_0^2}{2}\right)$$

The prior distribution applied to θ accounts for the uncertainty concerning an individual study. Conversely, the prior distributions assigned to μ and τ encompass the uncertainty surrounding the wider population of trials. The two densities are assumed to be independent. It is important to acknowledge that the full uncertainty regarding study-specific parameters is naturally propagated into the between-study model, and feedback is allowed from the between-study model to the estimation of study-specific parameters (Lunn et al. (2013)).

Consequently, the joint posterior distribution can be specified as:

$$p(\theta_1, \dots, \theta_m, \mu, \tau^2 | y_1, \dots, y_m)$$

$$\propto p(\mu, \tau^2) p(\theta_1, \dots, \theta_m | \mu, \tau^2) p(y_1, \dots, y_m | \theta_1, \dots, \theta_m, \mu, \tau^2)$$

$$= p(\mu) p(\tau^2) \left\{ \prod_{j=1}^{m} N\left(\theta_j | \mu, \tau^2\right) \right\} \left\{ \prod_{j=1}^{m} p\left(y_j | \theta_j, \sigma_j^2\right) \right\}$$

In particular, when we observe data from an individual study, we can directly derive conclusions about the parameter θ specific to that study. However, since θ and μ are interdependent, we can also utilize data from a single study to draw inferences about the

overall population of trials. This justifies employing the data from one study to inform the inference on other studies within the population.

2.2.   Gibbs sampling

The Gibbs Sampler algorithm is employed to iteratively draw values from the posterior distribution by sampling from the full conditionals. Fortunately, for this specific application, they can be readily recovered in a closed form[3]. The Gibbs sampler is a technique for generating random variables from the posterior distribution indirectly, without having to calculate the density. It generates a sample from $f(x)$ by sampling instead from the conditional distributions $f(x|y)$ and $f(y|x)$, which are often known in statistical models (Casella and George (1992)).

It can be confidently stated that, under fairly general conditions, which will almost always hold in meta-analysis, the simulation will eventually converge to the correct posterior distribution and enable random draws from it (Gilks, Richardson, and Spiegelhalter (1995)).

3. Implementation

The methodology described above was used to perform Bayesian meta-analysis. Applying hierarchical Bayesian techniques to meta-analysis allowed to synthesise the outcomes of 15 different studies, and utilise that information to obtain both study-specific effects, and the estimate for the average effect of the treatment (M. L. Johnson (2010)).

3.1.   Normality assumption

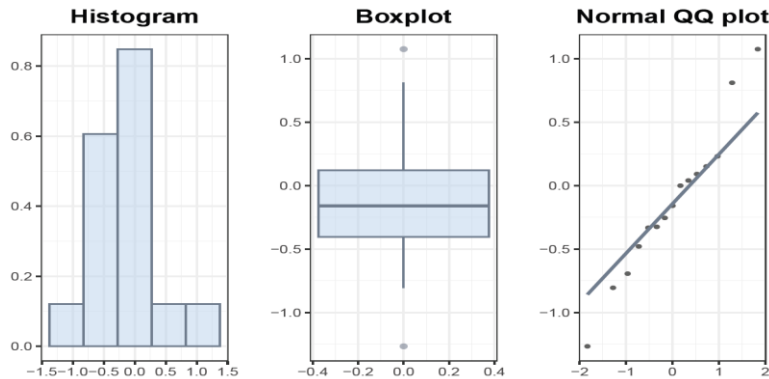First, the assumption on the normality of the distribution of the data was assessed.

Figure 3: In order: a histogram, a boxplot and a -plot are employed to assess the assumption of normality

In *Figure* 3, three techniques to evaluate the normality of the variable *log-odds ratio* were employed. Although the points do not exhibit a perfectly normal distribution, as evident in the *Q-Q plot*, we will still rely on this assumption for making inferences. With only 15 available point estimates, any departure from normality could be interpreted as a random mechanism rather than a systematic tendency towards values significantly deviating from the quantiles of a Normal distribution.

---

[3] The derivations of the full conditionals are lengthy, hence not provided in the summary, but are available upon request

3.2.   Prior distributions specification

Generally, in the medical research, information available on the μ and τ^2 hyperparameters is very scarce (Schmid (2001)), justifying assigning diffuse, non-informative priors, such as uniform, to represent the lack of knowledge on the subject. However, in this framework, a certain amount of information is available on the nature of the parameters. For instance, literature review has been conducted, with studies on the efficacy of beta blockage analysed (Yusuf et al. (1985), Schmid (2001), Al-Reesi et al. (2008)). Consequently, it is possible to estimate the prior expectation of the mean at a value of 0, with a small amount of variation, $\gamma_0^2 = 4$. The between-studies variation is also expected to assume relatively small values, hence the prior hyperparameters are set at $\tau_0^2 = 4$, $\eta_0 = 2$, where η_0 may be interpreted as the prior sample size. This choices for the hyperparameters allow to preserve the differences in the variation between the studies (Schmid (2001)).

At this stage, the Gibbs sampling was implemented, and samples from posterior distributions were obtained. In the following section, the resulting chained are assessed for convergence, and posterior inference is performed.

3.3.   Convergence

Assessing the convergence of the chain is essential: as discussed above, the Gibbs sampling method introduces within-sequence correlation. Generally, the Markov Chain Monte Carlo chain results in independent and identically distributed draws from the posterior distributions; however, high degrees of correlation may result in decreased precision of the estimation (Gelman et al. (2013)).
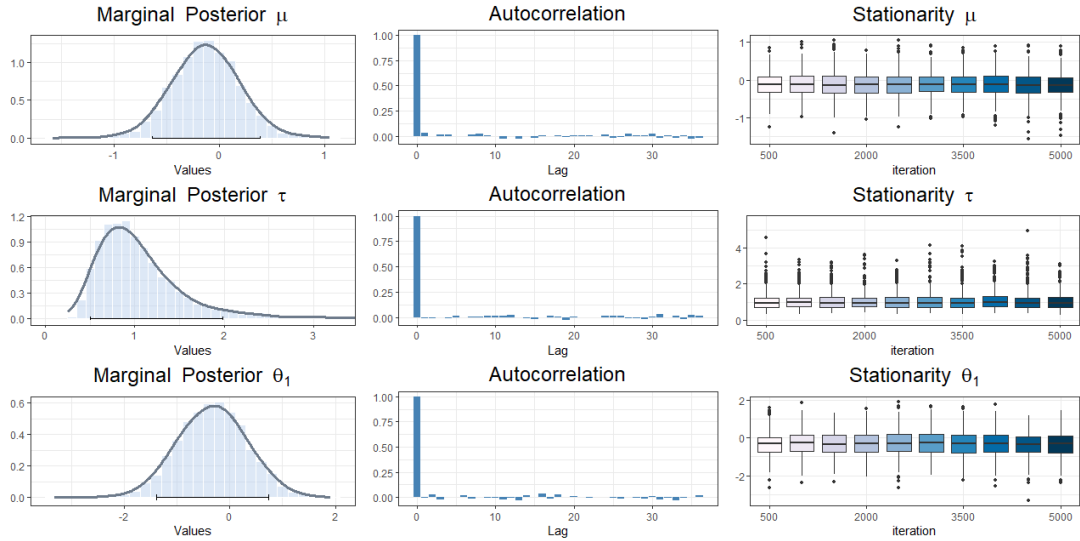

Figure 4: Diagnostic tools are provided for the parameters μ, τ and θ_1. For each parameter, three plots are shown: histograms, providing a visual representation of the parameter's posterior distribution together with credible intervals at 95% level; Autocorrelation plots indicating the chain's mixing and convergence properties; Boxplots used to check the stationarity of the chain

In *Figure* 4, it is evident that all three considered parameters have achieved convergence[4] . The boxplots illustrate the stationarity of the chains, while the autocorrelation plots assure

---

[4] Only one θ parameter is demonstrated for visual representation. However, all parameters have been investigated and convergence confirmed

that there are no concerns regarding dependence within the chain. These results were obtained by thinning the chain at intervals of four values. Additionally, we can examine the posterior distribution of the parameters through histograms. It is worth noting that both μ and θ_1 exhibit distributions resembling a normal distribution. Conversely, τ displays a right-skewed density line, indicating a high likelihood for the parameter to assume smaller values.

Effective sample size is a measure of the degree of autocorrelation in the chain:

$$ESS = \frac{G}{1 + 2 \sum_{g=1}^{G} acf_g}$$

If there is autocorrelation in the chain, the effective sample size will be reduced compared to G, the size of the chain after the burn-in period.

<p align="center">Table 1: Effective sample size</p>

μ	3787.206	θ_4	4737.798	θ_9	4490.897	θ_{14}	5148.474
τ	4408.406	θ_5	5000	θ_{10}	5000	θ_{15}	5000
θ_1	4776.792	θ_6	5000	θ_{11}	5000		
θ_2	5000	θ_7	5000	θ_{12}	5000		
θ_3	4780.391	θ_8	5000	θ_{13}	5000		

In *Table 1* we can observe the results of the effective sample size, which indicate favorable outcomes. This suggests that the algorithm is performing well, demonstrating its efficacy in capturing the underlying data patterns.

3.4.   Interpretation

To perform posterior inference, the samples drawn from the posterior distributions for the mean and variance parameters, as well as fifteen chains for the θ parameters. The obtained distributions allow to compute Monte Carlo approximations for estimates for the θ effects.

<p align="center">Table 2: Summary of the results of clinical trials</p>

Study	Treatment group[5]	Control group[5]	Log odds y_j	sd σ_j^2	Posterior mean of θ_j	Posterior variance of θ_j	Probability $\theta_j < 0$
Sloman	3/26	4/23	-0.48	0.68	-0.34	0.66	0.70
Evemy	6/46	3/48	0.81	0.55	0.50	0.63	0.22
Peter	1/47	2/48	-0.69	1.54	-0.36	0.84	0.67
Mueller	3/47	1/44	1.08	1.38	0.40	0.83	0.32
Lloyd	1/15	1/15	0.00	2.14	-0.09	0.90	0.54
Von Essen	1/25	1/26	0.04	2.08	-0.08	0.89	0.54
Yusuf	4/244	13/233	-1.27	0.34	-0.99	0.52	0.97
UKCSG	4/56	5/52	-0.32	0.49	-0.27	0.59	0.68
ICSG	3/73	4/71	-0.33	0.61	-0.26	0.64	0.66
Norris	15/364	14/371	0.09	0.14	0.07	0.36	0.42
Campbell	1/20	2/19	-0.80	1.61	-0.40	0.84	0.68
MIAMI	74/2877	87/2901	-0.16	0.03	-0.16	0.16	0.84
Andersen	29/238	24/242	0.23	0.09	0.21	0.28	0.23
Salathia	25/416	20/384	0.15	0.10	0.13	0.30	0.34
Hjalmarson	18/698	23/697	-0.25	0.10	-0.24	0.31	0.78

---

[5] Numbers recorded are: number of deaths / total number of patients in the group

Table 1 describes quantities related to the analysis. The estimates for the mean of empirical logits for each study, are discussed in detail below. The probabilities reported indicate how likely a Bayesian effect is to report reduction in odds of dying. For instance, the Yusuf study reports 97% posterior probability that the treatment reduces the chance of dying. However, for many studies, the probability is around 50%, providing indication of play by chance.

Furthermore, conclusions may be made on the values for posterior estimates of the prior parameters. The posterior approximation to the estimate of the mean is -0.13, which is the overall mean value for the estimator for all the studies included in the analysis (Gelman et al. (2013)). The 95% highest posterior density credible interval is [-0.71, 0.50] on the log scale, which also covers positive values of the empirical logit. With regard to the estimate of the between-study variance, the highest posterior density 95% credible interval is [0.38, 2.02]. The median value for τ is 0.97, which, compared to the values of variance in Table 2 provides some evidence of shrinkage in studies with low precision (Gelman et al. (2013), p. 127). This fact is further investigated below.

3.5.   Shrinkage

As discussed previously, one of the main benefits on hierarchical Bayesian modelling lies in the fact that the posterior estimates for each group borrow the strengths from the information of all the other groups. This implies that the studies that carry the most weight in the meta-analysis would influence the estimates for studies which weigh less more significantly. In this framework, the weights are defined by the precision of the sample estimates: the less variance is associated to the estimate of a group, the more reliable is the information it can convey to the estimates of the other groups, hence their values will tend to shrink to it.
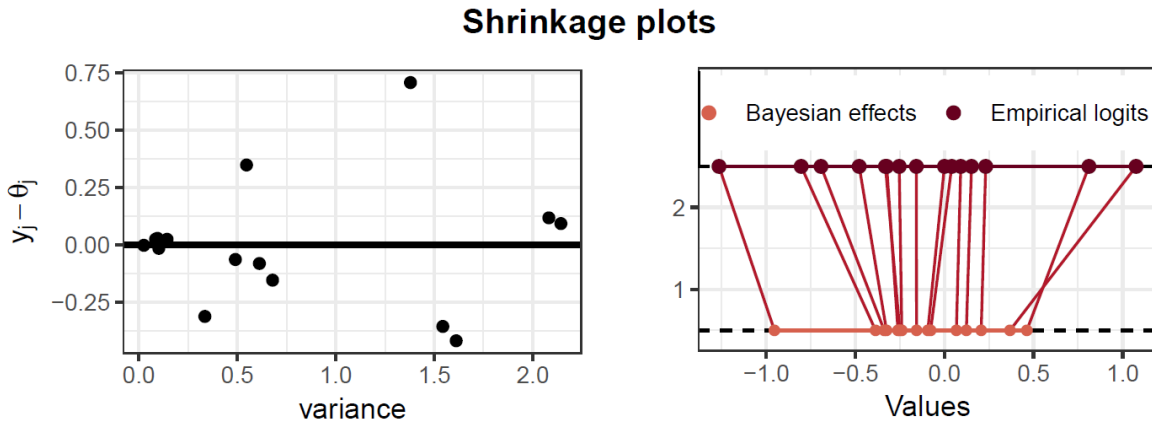

Figure 5: Shrinkage in the posterior estimates

Figure 5 provides that, for groups with higher variance, the sample estimates have a higher magnitude of difference from the posterior estimates. This demonstrates that for studies that do not have enough participants, their estimates are not sufficiently stable and reliable. Hence, the Bayesian framework allows less precise studies to "borrow" information from the other trials, improving the estimates. It is beneficial to investigate a particular example, namely two studies which demonstrated different degrees of shrinkage.
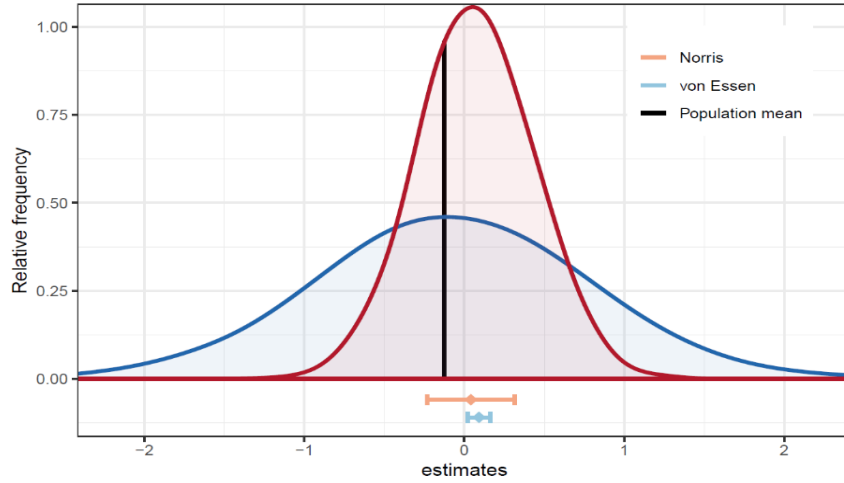
Figure 6: Comparison of posterior distributions and the sample estimates for studies of Norris and von Essen, with sample confidence intervals below the plot

Figure 6 compares two studies: von Essen and Norris. As demonstrated by the plot, the initial sample estimates were quite close to each other. However, after hierarchical Bayesian estimation, the posterior density for the less precise von Esssen study shrunk to the posterior overall mean, since it borrowed strength from more precise studies. On the other hand, the Bayesian estimate for the Norris study did not borrow much information from the population mean posterior estimate for it had a low degree of uncertainty associated to it.

3.6.   Sensitivity analysis

Although the diagnostics performed on the posterior Bayesian distributions above provided satisfactory results, there is a possibility that there exist specifications that would provide a good fit but lead to slightly different conclusions (Gelman et al. (2013)). One of those possibilities arises from the choice of the prior parameter specifications: inappropriate choices for priors may result in incorrect estimation of uncertainty (Roos et al. (2015)).
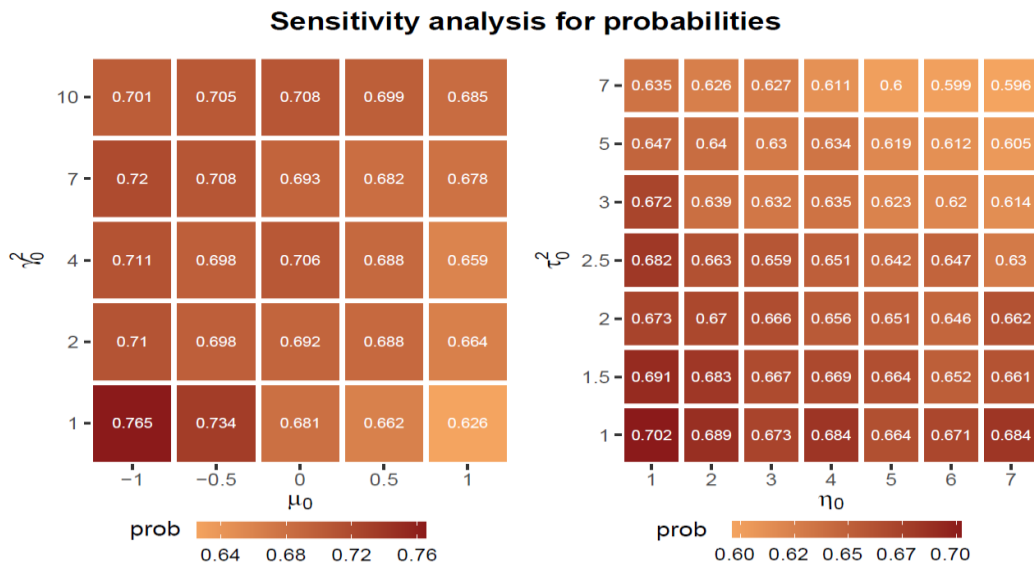

Figure 7: Sensitivity analysis for the probability of the grand mean being less than zero for different prior specifications.

10

From *Figure* 7 it can be concluded that the probability of observing a net positive treatment effect is not very sensitive to changes in prior specifications, for either the parameter μ or the parameter τ^2.

3.6.1. Within and between variability

The ratio $\frac{\tau}{\tau+\sigma}$ is calculated to compare the two sources of variation. It can be interpreted as the proportion of total variability in the fifteen studies that can be attributed to differences between groups. If the value of the ratio exceeds 0.5, it indicates that the majority of the total variability is associated to between-group variability. This ratio provides valuable insights into the extent of shrinkage, which is determined by the estimate of between-study variability, τ, in relation to the within-study variances. Notably, greater shrinkage occurs when τ is smaller.

In our analysis the prior density assigns a higher probability (0.9942) to the ratio being greater than 0.5. This is because a dispersed prior hyperparameter is assigned to τ^2, reflecting the significant uncertainty associated with θ_j. However, the posterior probability drops significantly to 0.669, indicating that the updated belief, considering the observed data, is less inclined towards the ratio being greater than 0.5.

In *Figure* 8 we can notice that the probability associated with the posterior of the ratio $\frac{\tau}{\tau+\sigma}$ being greater than 0.5 highly depends on the prior assigned to τ_0 and η_0.
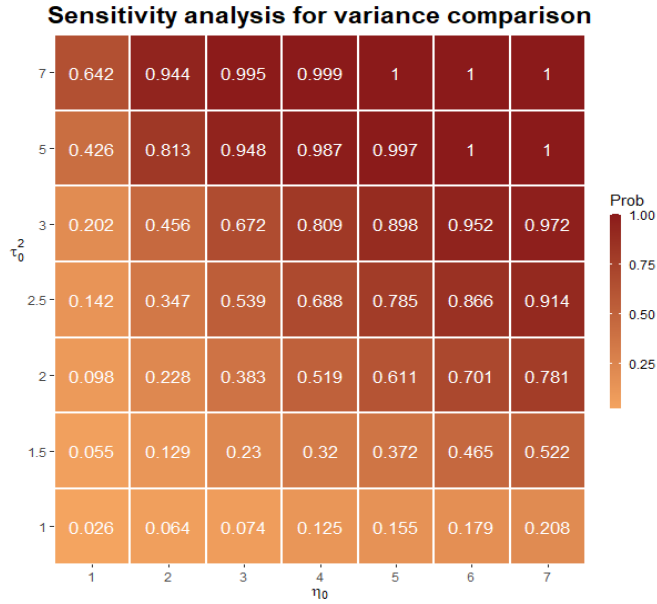

Sensitivity analysis for variance comparison

τ_0^2	$\eta_0=1$	2	3	4	5	6	7
7	0.642	0.944	0.995	0.999	1	1	1
5	0.426	0.813	0.948	0.987	0.997	1	1
3	0.202	0.456	0.672	0.809	0.898	0.952	0.972
2.5	0.142	0.347	0.539	0.688	0.785	0.866	0.914
2	0.098	0.228	0.383	0.519	0.611	0.701	0.781
1.5	0.055	0.129	0.23	0.32	0.372	0.465	0.522
1	0.026	0.064	0.074	0.125	0.155	0.179	0.208

Figure 8: Sensitivity analysis on the prior parameters η_0 and τ_0 illustrating the probability of the ratio exceeding 0.5 threshold.

3.7. Pooled and separate model

Instead of using a hierarchical model two different approaches can be followed. If the prior distributions of the individual parameters for the studies are assumed to be independent, fifteen separate Bayesian models should be fitted. As a consequence, the inference about one particular study will be independent of the inferences on the remaining studies (Albert and Hu (2019)).

On the other hand, if all the studies are assumed to come from the same population, it is possible to fit one single model, which ignores any differences between the fifteen trials. Hence, results from all studies are aggregated to perform one single analysis. In mathematical terms complete pooling makes the assumption that the parameter τ is exactly zero (Gelman et al. (2013)); thus, any variability between studies is entirely eliminated. In this kind of analysis θ_j parameters are assumed equal because each study estimates the same common treatment effect. This approach, however, ignores information that may be gained from between-study heterogeneity (Schmid (2001)).

Conversely, when employing a hierarchical model, each study is assumed to have its own treatment effect θ_j differing from the effects of other studies. These effects are connected by a common distribution that represents how they stand in relation to the universe of all possible study effects. Furthermore, avoiding fully separate models helps mitigate the risk of overfitting. This is because information about each study is derived from multiple sources, leading to a more comprehensive and robust understanding of the data.
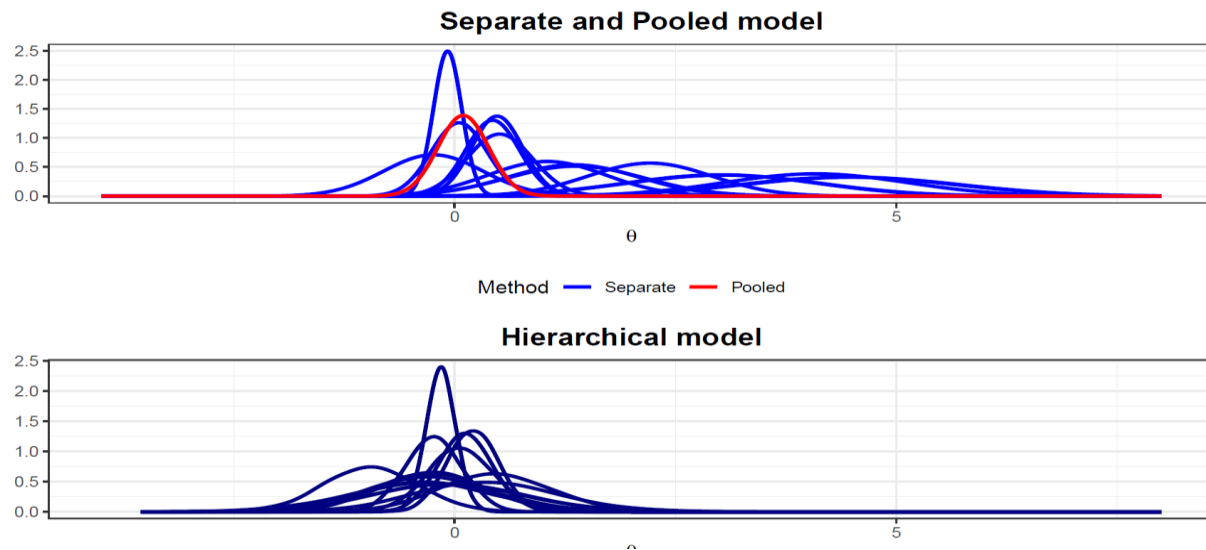

Figure 9: The first plot displays 15 lines representing the posterior density, each corresponding to a separate model. Additionally, the red line represents the posterior density of a pooled model. In the second plot, the posterior density is presented under a hierarchical setting.

In *Figure* 9 we can notice that employing a hierarchical model leads to a higher degree of similarity in terms of posterior density with respect to the complete separate model[6]. To understand the underlying reasons behind shrinkage, intuitively, if we can assume that certain studies are estimating the same thing or are exchangeable in statistical terminology, then we can gain more information about each one by using information from the others (Draper et al. (1993)).

3.7.1. Credible intervals

As depicted in *Figure 10*, the presence of shrinkage effect becomes evident as the credible intervals associated with the hierarchical model appear more tightly concentrated around

---

[6] Reference for the code https://github.com/avehtari/BDA_R_demos

the pooled estimate. Although not explicitly visible in the plot due to the utilization of different units of measurement on the y-axis, employing the hierarchical model results in a reduced posterior variance when estimating posterior parameters.

Through the utilization of the pooled estimate, we gain more certainty about the true effect in each study by leveraging information from other studies.

Moreover, the notable level of homogeneity among the studies is evident in the substantial reductions in posterior variance when transitioning from the study-specific estimates to the Bayesian estimates. This reduction is facilitated by the borrowing of strength and information from each study, resulting in more precise and consolidated estimates.
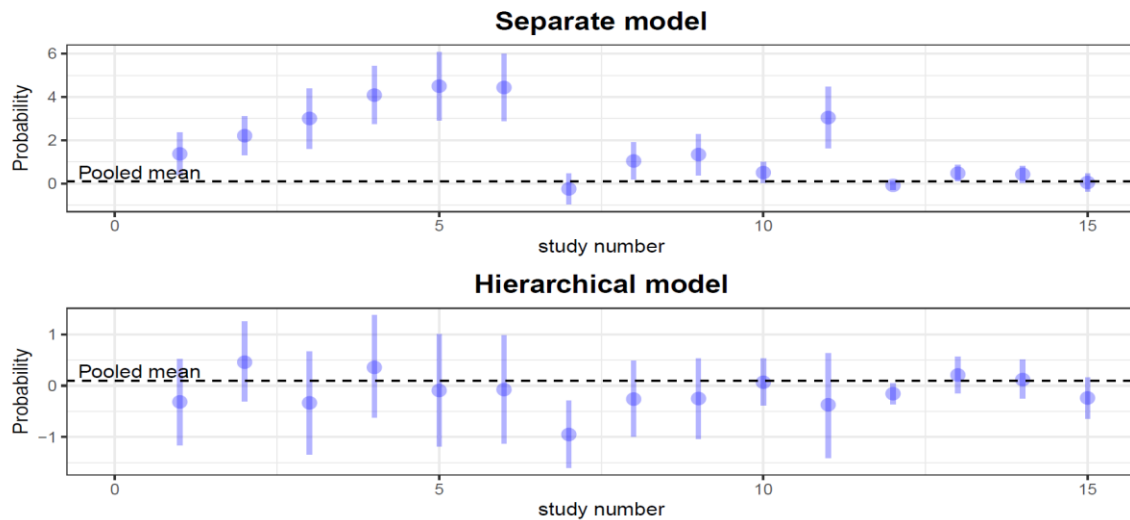

Figure 10: Comparison of credible intervals between pooled and separate model settings, revealing notable shrinkage effects.
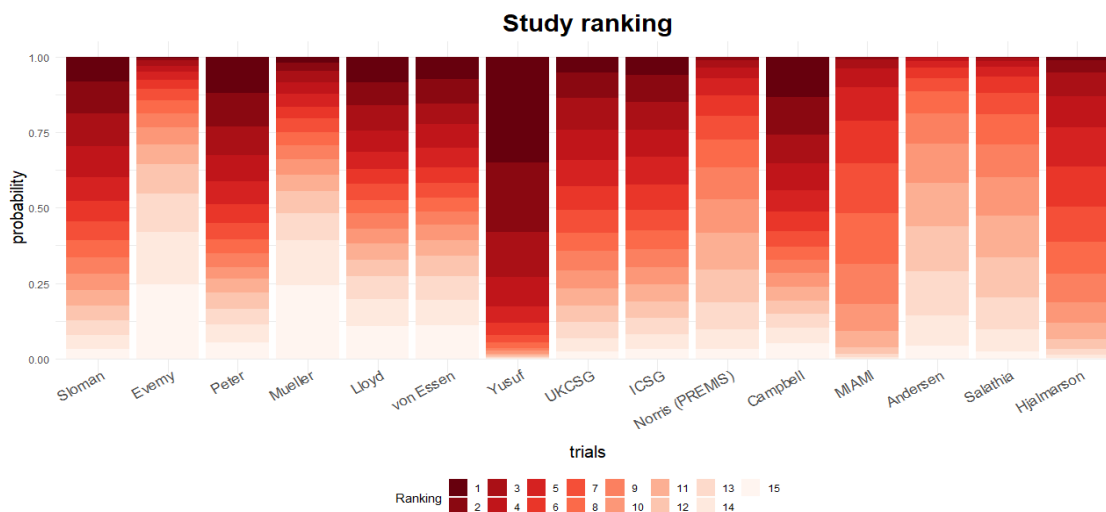
3.8. Ranking

Figure 11: The probability of each study ranking is depected. Darker colours represents a higher likelihood for the treatment to lead to a decrease in the number of deaths.

13

In *Figure 11* the objective was to rank the various studies based on their posterior analysis. The intention was to emphasize those studies that indicate a lower risk in terms of mortality associated with the administration of beta-blockers to patients. Within the figure, each bar is partitioned into fifteen smaller boxes, representing the probabilities assigned to the respective rankings of each study. These boxes illustrate the likelihood of a particular study being ranked first, second, or in subsequent positions.

The study conducted by *Yusuf et al*'s analysis stands out as having the lowest risk among the others. Notably, there is a substantial probability, precisely 73.06%, that this study is ranked within the top three positions.

4. Conclusions

To conclude, it is important to reiterate that adapting Bayesian framework for the meta-analysis of the mortality rates after beta-blocker treatment for patients who suffered from cardiac events did not yield conclusive results. Analysis on the set of 15 independent studies did not determine a positive net of treatment, with the possibility of attributing the results on play by chance.

An important problem in the meta-analysis framework is the "file drawer problem", also known as publication bias. This issue arises from the fact that studies that fail to find evidence of the presence of scientifically significant effect oftentimes are not published, hence creating a bias in the scientific field dedicated to researching a particular problem (DuMouchel (1994)). Hence, such studies cannot be included in the meta-analysis, which may cause for the overall estimates to be biased in favour of the presence of the effect.

A possible extension of the employed Bayesian hierarchical model could include incorporating the prior distribution on the variance component, thus no longer regarding it as fixed at the true value (Hoff (2009)). This approach would allow to incorporate information on the sampling distribution for each study and improve the study-specific estimates by accounting for shared information. This would be especially beneficial for trials with low numbers of participants.

References

1. Albert, Jim, and Jingchen Hu. 2019. *Probability and Bayesian Modeling*. CRC press.
2. Casella, George, and Edward I George. 1992. "Explaining the Gibbs Sampler." *The American Statistician* 46 (3): 167–74.
3. De Finetti, Bruno. 1937. "Foresight: Its Logical Laws, Its Subjective Sources." In *Breakthroughs in Statistics: Foundations and Basic Theory*, 134–74. Springer.
4. Diaconis, Persi. 1988. "Recent Progress on de Finetti's Notions of Exchangeability." Bayesian Statistics 3 (111-125): 13–14.
5. Draper, David, James S Hodges, Colin L Mallows, and Daryl Pregibon. 1993. "Exchangeability and Data Analysis." *Journal of the Royal Statistical Society Series A: Statistics in Society* 156 (1): 9–28.
6. DuMouchel, William. 1994. "Hierarchical Bayes Linear Models for Meta-Analysis." Technical Report 27, National Institute of Statistical Sciences.
7. Gelman, Andrew, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin. 2013. *Bayesian Data Analysis*. CRC press.
8. Gilks, Walter R, Sylvia Richardson, and David Spiegelhalter. 1995. *Markov Chain Monte Carlo in Practice*. CRC press.
9. Hoff, Peter D. 2009. A First Course in Bayesian Statistical Methods. Vol. 580. Springer.
10. Johnson, Alicia, Colin Rundel, Jingchen Hu, Kevin Ross, and Allan Rossman. 2020. "Teaching an Undergraduate Course in Bayesian Statistics: A Panel Discussion." Journal of Statistics Education 28 (3): 251–61.
11. Johnson, Alicia, Colin Rundel, Jingchen Hu, Kevin Ross, and Allan Rossman. 2020. "Teaching an Undergraduate Course in Bayesian Statistics: A Panel Discussion." *Journal of Statistics Education* 28 (3): 251–61.
12. Johnson, Michael L. 2010. *Essential Numerical Computer Methods*. Academic Press.
13. Lipsitz, Stuart R, Nan M Laird, and David P Harrington. 1991. "Generalized Estimating Equations for Correlated Binary Data: Using the Odds Ratio as a Measure of Association." *Biometrika* 78 (1): 153–60.
14. Lunn, David, Jessica Barrett, Michael Sweeting, and Simon Thompson. 2013. "Fully Bayesian Hierarchical Modelling in Two Stages, with Application to Meta-Analysis." *Journal of the Royal Statistical Society Series C: Applied Statistics* 62 (4): 551–72.
15. Mechanic, Oren J, Michael Gavin, and Shamai A Grossman. 2017. "Acute Myocardial Infarction."
16. Roos, Malgorzata, Thiago G Martins, Leonhard Held, and Haavard Rue. 2015. "Sensitivity Analysis for Bayesian Hierarchical Models."
17. Schmid, Christopher H. 2001. "Using Bayesian Inference to Perform Meta-Analysis." Evaluation & the Health Professions 24 (2): 165–89.
18. Yusuf, Salim, Richard Peto, John Lewis, Rory Collins, and Peter Sleight. 1985. "Beta Blockade During and After Myocardial Infarction: An Overview of the Randomized Trials." *Progress in Cardiovascular Diseases* 27 (5): 335–71.