

Using the bayesmeta R package for Bayesian random-effects meta-regression

Christian Röver*

Department of Medical Statistics, University Medical Center Göttingen, Humboldtallee 32, 37073 Göttingen, Germany

Tim Friede

Department of Medical Statistics, University Medical Center Göttingen, Humboldtallee 32, 37073 Göttingen, Germany

Abstract

Background: Random-effects meta-analysis within a hierarchical normal modeling framework is commonly implemented in a wide range of evidence synthesis applications. More general problems may even be tackled when considering *meta-regression* approaches that in addition allow for the inclusion of study-level covariables. **Methods:** We describe the Bayesian meta-regression implementation provided in the `bayesmeta` R package including the choice of priors, and we illustrate its practical use. **Results:** A wide range of example applications are given, such as binary and continuous covariables, subgroup analysis, indirect comparisons, and model selection. Example R code is provided. **Conclusions:** The `bayesmeta` package provides a flexible implementation. Due to the avoidance of MCMC methods, computations are fast and reproducible, facilitating quick sensitivity checks or large-scale simulation studies.

Keywords: meta-analysis, subgroup analysis, covariables, moderators, heterogeneity

1. Introduction

In the course of scientific endeavour it is often necessary to assess the compiled evidence from several separate sources, e.g., from independent experiments. *Meta-analysis* methods have emerged as a popular class of tools to perform such evidence syntheses, which are nowadays commonplace in many scientific disciplines [8, 33].

A simple, versatile and common approach to meta-analysis is given by the *normal-normal hierarchical model (NNHM)*, where measurement uncertainty as well as variability between measurements are implemented using normal distributions [34, 36]. Inference within the NNHM framework may be tackled in different ways, and a Bayesian approach has proven particularly useful [66, 68, 72, 78]. The technical implementation is commonly facilitated using Markov chain Monte Carlo (MCMC) methods [27]. However, the relatively simple NNHM also lends itself to a semi-analytical solution using the DIRECT algorithm [61]. Meta-analysis within the generic NNHM is implemented in the `bayesmeta` R package [59, 60].

The simple NNHM is readily generalized to a *meta-regression* model that allows for the consideration of additional covariables (at the level of individual estimates, i.e. the level of the studies or experiments) in a meta-analysis [41, 48, 74, 75]. This common model extension may again also be analyzed via the DIRECT approach, technically by

extending from one-dimensional (conditional or marginal) posterior distributions of a single “overall mean” or “intercept” parameter to higher-dimensional posterior distributions of a set of regression coefficients. This approach was recently implemented and included in the `bayesmeta` R package; the present paper gives an overview of the new functionality and showcases its application in a range of different analysis scenarios.

Meta-regression methods aim to attribute differences apparent between individual empirical estimates to available covariables, and with that will reduce the between-study variance (heterogeneity) — just like consideration of additional covariables in an ordinary regression will generally improve the model fit and increase the *coefficient of determination* [50]. Meta-regression analyses are hence often seen in the context of the exploration of (potential) sources of heterogeneity [39, 73], with the intention to reduce or eliminate any unexplained variance and reduce bias [1, 19, 52, 65]. However, Thompson and Higgins [74] caution that associations derived from meta-regression are *observational* in nature and that data dredging may be an issue, while Hartung et al. [34] also point out the danger of overfitting due to the commonly small sample sizes (numbers of studies). With that, the *statistical power* to identify covariables will commonly also tend to be low [38]. Cooper [13] in this context distinguishes between the concepts of *study-generated* and *synthesis-generated evidence*, and cautions that in general only the former may allow to infer causal relationships. Such issues might to some extent be addressed by pre-specification of analyses [70],

*Corresponding author

while caution should be advised in general in order to avoid methodological problems (such as *ecological fallacy*) [24].

While the scope of meta-regression methods is very broad, in practice a large number of practical applications are concerned with the investigation of *subgroups* of estimates, which on the technical side means the consideration of binary “indicator” covariables. Such situations are often dealt with by simply analyzing groups of studies separately or jointly [21, 47]. Meta-regression provides an alternative approach by allowing for differences in subgroup means while assuming a common heterogeneity variance. The use of meta-regression methods here has the advantage that questions regarding *differences* between subgroup means are readily addressed, and that in particular in case of “small” subgroups, pathological behaviour is avoided through the borrowing of information on the heterogeneity (nuisance) parameter that is effectively taking place [19, 21, 40, 75].

The meta-regression implementation described here facilitates a range of applications, including parameter estimation, prediction, shrinkage estimation, indirect comparisons, sensitivity analyses, model selection or model averaging. In the following, familiarity with some basics of (Bayesian) random-effects meta-analysis is a bonus [60], but is not strictly necessary. Methods will be introduced with a focus on applications rather than on theoretical background. The remainder of the manuscript is structured as follows. In Section 2, the random-effects meta-regression model (i.e. NNHM with covariables) is introduced, including more guidance on model specification details, in particular the parametrization of covariable effects and prior distributions, while some technical details and aims are also covered. Example applications follow in Section 3; here we provide a wide range of potential applications with code snippets illustrated by real data. We close with a brief discussion in Section 4.

2. Methods

2.1. The data model

Meta-regression can be facilitated through a generalization of the normal-normal hierarchical model (NNHM) that is commonly used for random-effects meta-analysis [60, 78]. The model here is extended in order to consider linear effects of a set of study-level covariables.

Suppose that a set of estimates from k studies are to be modelled. We then have k estimates $y_i \in \mathbb{R}$ (where $i = 1, \dots, k$) with standard errors $\sigma_i \in \mathbb{R}_+$, which are assumed known. For each of the k estimates, we also have a set of corresponding covariables $x_i \in \mathbb{R}^d$ of dimension d . Such (study-level) covariables are sometimes also denoted as *moderators*.

It is then assumed that each estimate quantifies an underlying parameter θ_i with a normally distributed offset whose magnitude depends on the standard error σ_i :

$$y_i | \theta_i, \sigma_i \sim \text{Normal}(\theta_i, \sigma_i^2). \quad (1)$$

The study-specific mean (θ_i) then depends linearly on the covariables x_i via a d -dimensional coefficient vector β . However, even for an identical set of covariables, the mean may vary from study to study due to additional (normally distributed) variability:

$$\theta_i | x_i, \beta, \tau \sim \text{Normal}(\beta_1 x_{i1} + \dots + \beta_d x_{id}, \tau^2). \quad (2)$$

Besides measurement or sampling errors (σ_i), the between-study variation is hence determined both by effects of covariables x_i as well as by *heterogeneity* that is quantified through τ . The model may also be formulated via the marginal expression

$$y_i | x_i, \beta, \tau, \sigma_i \sim \text{Normal}(\beta_1 x_{i1} + \dots + \beta_d x_{id}, \sigma_i^2 + \tau^2). \quad (3)$$

It is often convenient to alternatively view the model in vector/matrix terminology; here we may re-write equations (1)–(3) as

$$y | \theta, \sigma \sim \text{Normal}(\theta, \Sigma) \quad \text{where } \Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_k^2), \quad (4)$$

$$\theta | X, \beta, \tau \sim \text{Normal}(X\beta, \tau^2 I), \quad \text{and} \quad (5)$$

$$y | X, \beta, \tau, \sigma \sim \text{Normal}(X\beta, \Sigma_\tau) \quad \text{where } \Sigma_\tau = \Sigma + \tau^2 I, \quad (6)$$

where the data are given in terms of the vectors of estimates $y \in \mathbb{R}^k$ and standard errors $\sigma \in \mathbb{R}_+^k$, and the set of covariables forms the regressor matrix $X \in \mathbb{R}^{k \times d}$, with rows corresponding to studies, and columns corresponding to different variables.

The unknowns in the model are the study-specific effects $\theta_i \in \mathbb{R}$, the heterogeneity $\tau \in \mathbb{R}_+$, and the vector of coefficients $\beta \in \mathbb{R}^d$ of dimension d . Prior distributions need to be specified for τ and β . In order to include an “intercept” (overall mean) term in the regression, one may specify one of the covariables (e.g., the first column of X) as $x_{i1} = 1$ for $i = 1, \dots, k$. Also, if *only* an intercept term is considered, the model again simplifies to the “plain” meta-analysis model [60].

2.2. Prior and data specification

2.2.1. Effect and heterogeneity priors

Prior specification works similarly to the simple random-effects meta-analysis model [60]; guidances provided for sensible specifications of heterogeneity (τ) priors largely apply here as well [60, 63].

Due to the implementation used in the `bayesmeta` package, priors for the β coefficients may be specified as (proper) multivariate normal or (improper) uniform only. These generic forms will however be appropriate to cover a majority of common applications.

2.2.2. Alternative model specifications

What becomes crucial in addition then is the *model specification*, i.e., the setup of covariables x_i eventually

Table 1: Illustration of different popular binary regressor matrix (X) codings for an example setting involving $k = 7$ studies in $d = 3$ subgroups. On the left-hand side, the three regression parameters β_1 to β_3 simply correspond to the three groups’ means. On the right-hand side, β_1^* again corresponds to the mean in group A, which also serves as a “reference”. Parameters β_2^* and β_3^* correspond to the differences (“contrasts”) between the means within groups A and B and groups A and C, respectively.

i	subgroup	“group mean” parametrisation			“intercept/offset” parametrisation		
		x_{i1}	x_{i2}	x_{i3}	x_{i1}^*	x_{i2}^*	x_{i3}^*
1	A	1	0	0	1	0	0
2	A	1	0	0	1	0	0
3	B	0	1	0	1	1	0
4	B	0	1	0	1	1	0
5	B	0	1	0	1	1	0
6	C	0	0	1	1	0	1
7	C	0	0	1	1	0	1

constituting the regressor matrix (or *design matrix*) X . Different, and to some extent equivalent, conventions are conceivable in order to approach the same analysis problem. This holds in particular in the context of binary covariables, which then imply different interpretations of the associated parameters (coefficients) and with that sometimes differing prior settings. A simple example involving binary covariables is illustrated in Table 1. Suppose three groups of studies (labelled A, B and C) are given, which may be identified using indicator variables as regressors in the design matrix X . Two possible setups are shown here; on the left-hand side, the first study’s mean is modeled as (see equation (3)) $\beta_1 x_{11} + \beta_2 x_{12} + \beta_3 x_{13} = \beta_1$, whereas the third study’s mean is $\beta_1 x_{31} + \beta_2 x_{32} + \beta_3 x_{33} = \beta_2$. The three β parameters hence directly correspond to the three group means. On the right-hand side, the first study is modelled in the same way, however, the third study’s mean is $\beta_1 x_{31} + \beta_2 x_{32} + \beta_3 x_{33} = \beta_1 + \beta_2$. The second coefficient (β_2) hence corresponds to the *difference* (or *contrast*) between groups A and B, while group A serves as a “reference”. Either way of formulating models may have its merits, and switching from one to another corresponds to a transformation between differing parameter spaces [26, Sec. 1.8]. Prior settings for the regression coefficients may have differing implications in different model setups, and prior settings in one parametrisation may be translated to another by considering the transformation step.

More specifically, in the example from Table 1, the right-hand side (“intercept / offset”) parametrization results from the left-hand side (“group means”) via a linear transformation $\beta \rightarrow \beta^*$, where

$$\beta^* = \begin{pmatrix} \beta_1^* \\ \beta_2^* \\ \beta_3^* \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 - \beta_1 \\ \beta_3 - \beta_1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = A\beta. \quad (7)$$

Implications of prior assumptions in one parametrization may be derived by considering the transformation’s effect on the transformed random variable [26, Sec. 1.8]. In this case, if a normal prior with mean μ and variance Σ was assumed for β , then this implies a normal prior with mean $A\mu$ and variance $A\Sigma A'$ for the transformed parameter set β^* .

Transformations between alternative parametrisations, which may also be the result of transformations of the original *data*, are often useful to simplify interpretation, or in order to avoid numerical problems. These may also arise in the context of continuous covariables, for example, when re-expressing fractions as percentages, or when “centering” covariables by subtracting their mean levels.

2.3. Inference

2.3.1. Technical implementation

In the `bayesmeta` R package, the `DIRECT` algorithm is utilized to facilitate meta-analysis within the NNHM framework via the `bayesmeta()` function [61]. In contrast to the “simple” meta-analysis setup considered previously by Röver [60], instead of a single “overall mean” parameter μ , the *meta-regression* model now involves a d -dimensional coefficient vector β , which means that some analytic expressions need to be generalized to their multivariate analogues; the basic algorithm for deriving the posterior distributions however may still be applied analogously.

The meta-regression functionality is provided by the new `bmr()` function; its main input arguments are:

y: a vector of estimates (y_i) of length k

sigma: a vector of associated standard errors (σ_i) of length k

X: a regressor matrix (X) with k rows and d columns

tau.prior: a prior density function ($f(\tau)$) for the heterogeneity τ (or a character string denoting a specific form)

beta.prior.mean: a vector of prior means of dimension d

beta.prior.sd: a vector of prior standard deviations of dimension d

To a large extent, the behaviour is similar to the `bayesmeta()` function, especially with respect to the `y`, `sigma` and `tau.prior` arguments [60]. The major differences to the `bayesmeta()` function are that one may specify an additional “**X**” argument giving the regressor matrix (X), and that the posterior, instead of referring to only a single effect μ , now involves a d -dimensional parameter vector β . By default, if the `tau.prior`,

`beta.prior.mean` and `beta.prior.sd` arguments are left unspecified, (improper) uniform priors are assumed for τ and β . If an `X` argument is not supplied, a single-column matrix of ones is used, so that the analysis simplifies to fitting a single “intercept” parameter.

Inference is facilitated through a semi-analytical approach, which is based on noting that the problem essentially involves two parameters, namely, the heterogeneity τ and the coefficient vector β . For any fixed heterogeneity value, the *conditional* posterior distribution $p(\beta|\tau)$ results analytically as a (multivariate) normal distribution. The heterogeneity’s *marginal* posterior density function $p(\tau)$ on the other hand may also be expressed in analytical form. Noting that the joint posterior density may be written as a product, i.e., $p(\beta, \tau) = p(\beta|\tau)p(\tau)$, then implies that the coefficients’ marginal posterior results as a (continuous, normal) mixture distribution, i.e., $p(\beta) = \int p(\beta, \tau)d\tau = \int p(\beta|\tau)p(\tau)d\tau$. The DIRECT algorithm utilized here for posterior computations then approximates the continuous mixture distribution by a discrete mixture using a finite number of components; a strategic setup of the set of support points then allows to control the computational accuracy [61].

2.3.2. Aims

Inference within a meta-regression application may be aimed at a range of different aspects, e.g., joint or marginal distributions of regression coefficients (β_i), linear combinations of coefficients ($x'\beta$), investigation of heterogeneity (τ), shrinkage estimation (θ_i), or prediction ($\theta_{k+1}|x$). Posterior distributions of all these figures are available from the `bmr()` function’s output. It is possible to access these directly from the object returned by the `bmr()` function, however, many relevant figures are included in the default output, and it is often convenient to request certain figures to be included e.g. in a summary printout or a forest plot. Ways to retrieve such figures are illustrated alongside the example applications below.

3. Results

In this section a number of applications are presented to illustrate the versatile use of meta-regression with `bayesmeta`.

3.1. Binary covariable

3.1.1. Inferring two means

Crins et al. [14] reported on a meta-analysis of studies investigating the use of interleukin-2 receptor antagonists (IL-2RA) for immunosuppression in pediatric liver transplant recipients. Of primary interest was the occurrence of *acute rejection (AR)* reactions, a common adverse event that is supposed to be prevented by the medication. Two different types of IL-2RAs were used, namely, *basiliximab* and *daclizumab*. The rates of AR events in the studies’ treatment and control groups are summarized in terms of

Table 2: Data from the immunosuppression example. Each row here summarizes a 2x2 contingency table in terms of a derived log-OR (y_i) and its associated standard error (σ_i). Two different types of IL-2RA treatments were investigated (*basiliximab* and *daclizumab*).

<i>i</i>	study		log-OR	
	reference	IL-2RA	y_i	σ_i
1	Heffron (2003)	daclizumab	-2.31	0.60
2	Gibelli (2004)	basiliximab	-0.46	0.56
3	Schuller (2005)	daclizumab	-2.30	0.88
4	Ganschow (2005)	basiliximab	-1.76	0.46
5	Spada (2006)	basiliximab	-1.26	0.64
6	Gras (2008)	basiliximab	-2.42	1.53

odds ratios (see also [60]); the relevant data are shown in Table 2.

We will perform a meta-analysis aiming to investigate the mean effects for the two types of treatment; to that end, we specify the regressor matrix X reflecting the grouping of the data:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix}. \quad (8)$$

In analogy to the coding illustrated in Table 1, rows correspond to the six observations, and columns correspond to the two groups; the placement of zeroes and ones reflects the studies’ associations to one of the two medication types. The investigation of differences in treatment effects here technically implies the consideration of a possible *interaction* between treatment and IL-2RA type. For the heterogeneity parameter (τ), a half-normal prior distribution is appropriate in the context of a log-OR endpoint [22, 60, 63]. For the regression coefficients (β_1 and β_2), an (improper) uniform prior is used. First, the package and the example data set need to be loaded, and the effect measures (log-ORs) may be derived using the `metafor` package’s `escalc()` function:

```
R> library("bayesmeta")
R> data("CrinsEtAl2014")
R> crins.es <- escalc(measure="OR",
+ ai=exp.AR.events, n1i=exp.total,
+ ci=cont.AR.events, n2i=cont.total,
+ slab=publication, data=CrinsEtAl2014)
```

Then we may specify the regressor matrix X :

```
R> basil <- (crins.es$IL2RA=="basiliximab")
R> dacli <- (crins.es$IL2RA=="daclizumab")
R> X <- cbind("basiliximab"=as.numeric(basil),
+ "daclizumab" =as.numeric(dacli))
R> X
      basiliximab daclizumab
[1,]           0           1
[2,]           1           0
[3,]           0           1
[4,]           1           0
```

```
[5,]      1      0
[6,]      1      0
```

The `bmr()` function then works very similarly to the `bayesmeta()` function [60]; estimates and standard errors may be specified via the “`y`” and “`sigma`” arguments, or the data may simply be supplied in terms of the object returned from the `escalc()` function. The heterogeneity prior is specified in terms of its probability density function, and in addition the regressor matrix needs to be provided via the “`X`” argument. We may hence specify

```
R> bmr01 <- bmr(y=crins.es$yi,
+   sigma=sqrt(crins.es$vi), X=X,
+   tau.prior=function(t){dhalfnormal(t,s=0.5)})
```

or simply

```
R> bmr01 <- bmr(crins.es, X=X,
+   tau.prior=function(t){dhalfnormal(t,s=0.5)})
```

We may then have a closer look at the analysis output:

```
R> bmr01
' bmr ' object.

6 estimates:
Heffron (2003), Gibelli (2004), Schuller (2005),
Ganschow (2005), Spada (2006), Gras (2008)

2 regression parameters:
basiliximab, daclizumab

tau prior (proper):
function(t){dhalfnormal(t, scale=0.5)}
<bytecode: 0x555bbca28478 >

beta prior: (improper) uniform

MAP estimates:
              tau basiliximab daclizumab
joint      2.183638e-05    -1.287690    -2.307431
marginal   0.000000e+00    -1.283294    -2.307270
```

```
marginal posterior summary:
              tau basiliximab daclizumab
mode          0.0000000    -1.2832939    -2.3072701
median        0.2975462    -1.2832867    -2.3072414
mean          0.3420306    -1.2837755    -2.3072091
sd            0.2485756     0.3827227     0.5842337
95% lower    0.0000000    -2.0407937    -3.4589967
95% upper    0.8130176    -0.5252345    -1.1554177
```

(quoted intervals are shortest credible intervals.)

The function’s output again is very similar to the `bayesmeta()` function’s output (see also [60]); at the top we see details of the model specification, the number and labels for the included studies, the number of regression coefficients and the prior specification. The variable names (here: “`basiliximab`” and “`daclizumab`”) were taken from the column names of the regressor matrix. Then maximum-a-posteriori (MAP) estimates are shown, as well as summary statistics for the three parameters’ marginal posterior distributions. The median and 95% CIs for the *basiliximab* and *daclizumab* parameters (β_1 and β_2) are given by -1.38 $[-2.04, -0.53]$ and -2.31 $[-3.46, -1.16]$, respectively. The treatment hence

appears to be effective in reducing AR events in both study groups.

The (here: three) parameters’ marginal or joint posterior distributions may also be inspected using the `plot()` or `pairs()` functions. The posterior distributions may also be accessed e.g. via the functions contained in the returned object’s elements; for example, the `bmr01$qposterior()` function allows to compute posterior quantiles. A call of

```
R> bmr01$qposterior(tau.p=0.99)
[1] 1.072312
```

returns the heterogeneity posterior’s 99% quantile. Similarly, using

```
R> bmr01$qposterior(beta.p=0.99, which.beta=1)
[1] -0.3548454
```

the 99% quantile of the β_1 parameter may be determined. The “`which.beta`” argument here is used to specify the β parameter’s index. Analogously, the “`...$dposterior()`”, “`...$pposterior()`”, “`...$rposterior()`” and “`...$post.interval()`” functions may be used to determine posterior density, cumulative distribution function, random numbers or credible intervals (the naming of functions here follows the common R conventions, as e.g. known from the `dnorm()` or `pnorm()` functions).

Estimates of the θ_i parameters ($i = 1, \dots, k$), the *shrinkage estimates*, may also be derived. Some summary statistics are provided in the “`...$theta`” element. Access to the complete distributions (probability density, cumulative distribution function, quantile function, random number generation and credible intervals) is provided via the “`...$dshrink()`”, “`...$pshrink()`”, “`...$qshrink()`”, “`...$rshrink()`” and “`...$shrink.interval()`” functions.

To illustrate the results in a forest plot, one may call

```
R> forestplot(bmr01, xlab="log-OR")
```

The resulting plot is shown in Figure 1. The forest plot’s first six lines show the original data (estimates y_i and 95% CIs based on standard errors σ_i and the normal model), and the shrinkage estimates (θ_i). The table also includes the regressor matrix X in the columns that are labelled as

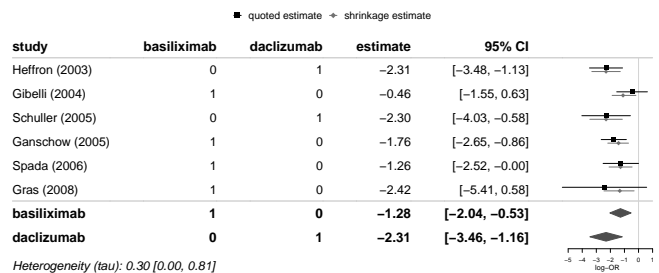


Figure 1: A forest plot illustrating the meta-regression results based on the transplantation data from Table 2. The two study groups are coded in terms of two binary indicator variables labelled “`basiliximab`” and “`daclizumab`”.

“basiliximab” and “daclizumab”, as in the original specification of the “X” argument. The two lines at the bottom then show the estimates of the two associated regression parameters β_1 and β_2 . The heterogeneity (τ) distribution finally is also summarized at the bottom left.

Note that while in a simple meta-analysis shrinkage estimates are “shrunk” towards the common overall mean, in a meta-regression shrinkage acts in the direction of the corresponding predicted value; in this case this means that individual studies’ shrinkage estimates move towards the corresponding (basiliximab or daclizumab) group means.

3.1.2. Inferring means, contrasts, or predictions

Quite commonly, it is also of interest to evaluate the posterior distribution of *linear combinations* of the regression coefficients (β_i), or of *predictions* corresponding to such combinations. For such purposes, the “...\$dpredict()”, “...\$ppredict()”, “...\$qpredict()”, “...\$rpredict()” and “...\$pred.interval()” functions are available.

In the present example, it may be of interest to infer the *difference* of the two group means, $\beta_2 - \beta_1$. If its posterior includes zero, then the two medications may be equally effective; if zero is outside the plausible range, this indicates differing efficacies. The above difference is a *linear combination* of the two coefficients, i.e., a sum of $\beta_2 - \beta_1 = -1 \times \beta_1 + 1 \times \beta_2$ with coefficients -1 and $+1$ for β_1 and β_2 , respectively. We can request the linear combination’s distribution by specifying the two coefficients, e.g., in order to determine the median or a 95% CI:

```
R> bmr01$qpredict(p=0.5, x=c(-1,1))
[1] -1.023955
R> bmr01$pred.interval(level=0.95, x=c(-1,1))
      lower      upper
-2.402393  0.354489
attr(,"interval.type")
[1] "shortest"
```

Such coefficients were in fact already quoted along with the summary estimates in the forest plot in Figure 1, although only involving zeroes and ones as coefficients. Analogously, additional linear combinations can be specified for the plot; e.g., to include the group difference in the figure, we may specify

```
R> forestplot(bmr01, xlab="log-OR",
+ X.mean=rbind("basiliximab" = c(1, 0),
+ "daclizumab" = c(0, 1),
+ "group difference" = c(-1, 1)))
```

The resulting plot is shown in Figure 2 (top). It should be noted that the comparison of basiliximab vs. daclizumab constitutes an *indirect comparison* here, as it contrasts two treatments that have not been “directly” compared in a head-to-head comparison in any of the six trials considered [45]. For more on indirect comparisons, see also Section 3.2 below.

Besides the *mean* effects, *predictions* are often of interest, e.g., in order to assess plausible ranges for a “future” study’s mean parameter θ_{k+1} (which, in addition to the

β coefficients, also depends on the estimated amount of heterogeneity τ). In the present example, we might be interested in predicting the mean in a new study investigating basiliximab; we can check out some quantiles of the effect’s distribution via

```
R> bmr01$qpredict(p=c(0.025, 0.5, 0.975),
+ x=c(1,0), mean=FALSE)
[1] -2.4611428 -1.2835800 -0.1125025
```

We again need to specify the coefficients via the “x” argument, and the “mean” argument (which by default is TRUE) needs to be set to FALSE explicitly.

3.1.3. Alternative regressor matrix setups

The specification of the regressor matrix X (see equation (8) above) is not unique; a number of different approaches are conceivable and common, for example, one might as well specify

$$X = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix} \quad \text{or} \quad X = \begin{pmatrix} 1 & +0.5 \\ 1 & -0.5 \\ 1 & +0.5 \\ 1 & -0.5 \\ 1 & -0.5 \\ 1 & -0.5 \end{pmatrix} \quad (9)$$

to yield analogous results. Different setups will then imply different interpretations for the associated β parameters. Within R, the most common parametrization is also returned by the `model.matrix()` function (from the `stats` package); if we run

```
R> X.alt <- model.matrix(~ IL2RA, data=crins.es)
R> X.alt
      (Intercept) IL2RAdaclizumab
1                1                1
2                1                0
3                1                1
4                1                0
5                1                0
6                1                0
attr(,"assign")
[1] 0 1
attr(,"contrasts")
attr(,"contrasts")$IL2RA
[1] "contr.treatment"
```

we can see that we in fact yield the first of the above two versions, an “intercept/slope” parametrization, which would often be the default for many regression models within R. We may run the same analysis using this alternative regressor matrix:

```
R> bmr02 <- bmr(crins.es, X=X.alt,
+ tau.prior=function(t){dhalfnormal(t,s=0.5)})
```

and generate a corresponding forest plot:

```
R> forestplot(bmr02, xlab="log-OR",
+ X.mean=rbind("basiliximab" = c(1, 0),
+ "daclizumab" = c(1, 1),
+ "group difference" = c(0, 1)))
```

The resulting plot is shown in Figure 2 (bottom). We can see that we get essentially identical results here, and that we only need to specify the linear combinations differently

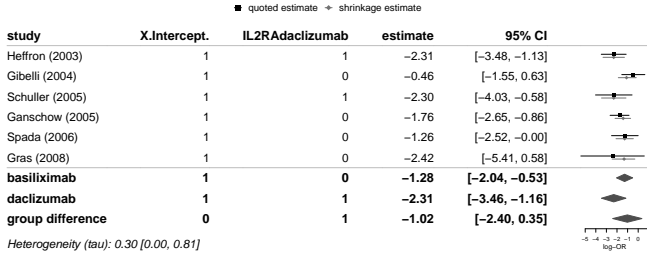
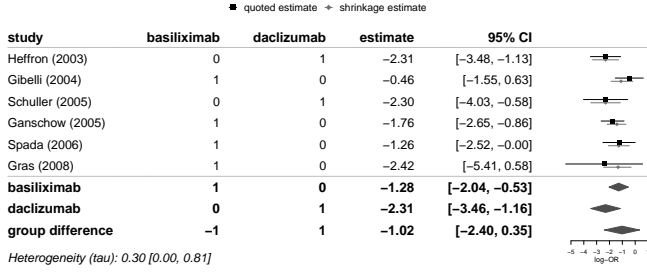


Figure 2: Two forest plots similar to the one shown in Figure 1 and illustrating analogous meta-regression results based on different parametrizations of the regressor matrix X (top: “group mean” parametrization, bottom: “intercept/slope” parametrization; note the differing setups in the 2nd and 3rd “regressor” columns).

in order to retrieve group means or group differences based on the differing parametrizations.

While in the above example the results are essentially equivalent (see also Section 2.2), either way of formulating the problem may have its advantages. Interpretation of parameters and prior specification may be easier or more obvious in one or another formulation. In case informative priors for the regression parameters β were to be used in the above example, this may either imply considerations of constraints on the two individual group means, or on their difference.

For example, consider the two alternative parametrizations in terms of

$$X = \begin{pmatrix} 1 & +0.5 \\ 1 & -0.5 \\ 1 & +0.5 \\ 1 & -0.5 \\ 1 & -0.5 \\ 1 & -0.5 \end{pmatrix} \quad \text{and} \quad X^* = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix}. \quad (10)$$

The implied parameters may be thought of as “overall mean / group difference” or “two group mean” parameters. The two associated parameter vectors β and β^* are related to one another as

$$\beta^* = \begin{pmatrix} \beta_1^* \\ \beta_2^* \end{pmatrix} = \begin{pmatrix} \beta_1 - 0.5\beta_2 \\ \beta_1 + 0.5\beta_2 \end{pmatrix} = A\beta, \quad \text{with} \quad A = \begin{pmatrix} 1 & -0.5 \\ 1 & +0.5 \end{pmatrix}. \quad (11)$$

Now suppose that in the former parametrization (β) we want to implement a vague prior for the overall mean, while the difference between groups is expected to be

rather small; for β we hence assume a normal prior with mean and covariance

$$\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} 100 & 0 \\ 0 & 1 \end{pmatrix}. \quad (12)$$

The prior specification in the “overall mean / group difference” parametrization has its counterpart in the “two group mean” parametrization; for the “transformed” parameter β^* , this implies a prior distribution with mean and covariance

$$\mu^* = A\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{and} \quad (13)$$

$$\Sigma^* = A\Sigma A' = \begin{pmatrix} 100.25 & 99.75 \\ 99.75 & 100.25 \end{pmatrix}. \quad (14)$$

The high correlation reflects the assumption implemented in the original parametrization that there is a constraint on the difference between the means while their common average level has greater uncertainty.

Performing the analysis using different parametrizations and matching proper, informative priors should then again yield identical sets of estimates as in the previous example (Figure 2). The two analyses may be performed via

```
R> bmr03 <- bmr(crins.es,
+ X=cbind("mean"=rep(1,6),
+ "difference"=c(0.5, -0.5, 0.5,
+ -0.5, -0.5, -0.5)),
+ beta.prior.mean=c(0,0),
+ beta.prior.sd=c(10,1),
+ tau.prior=function(t){dhalfnormal(t,s=0.5)})
R> bmr04 <- bmr(crins.es, X=X,
+ beta.prior.mean=c(0,0),
+ beta.prior.cov=cbind(c(100.25, 99.75),
+ c(99.75, 100.25)),
+ tau.prior=function(t){dhalfnormal(t,s=0.5)})
```

We may then check the corresponding estimates (of the two group means, the overall mean and the group difference) via the `summary()` function, which allows to specify an “X.mean” argument, similarly to the `forestplot()` function; for example:

```
R> smry03 <- summary(bmr03,
+ X.mean=rbind("basiliximab" = c(1, -0.5),
+ "daclizumab" = c(1, +0.5),
+ "average" = c(1, 0),
+ "difference" = c(0, 1)))
R> smry04 <- summary(bmr04,
+ X.mean=rbind("basiliximab" = c(1, 0),
+ "daclizumab" = c(0, 1),
+ "average" = c(0.5, 0.5),
+ "difference" = c(-1, 1)))
```

We may first double check the differing setups of the underlying regressor matrices:

```
R> smry03$X.mean
      mean difference
basiliximab 1 -0.5
daclizumab 1 0.5
average 1 0.0
difference 0 1.0
R> smry04$X.mean
```

	basiliximab	daclizumab
basiliximab	1.0	0.0
daclizumab	0.0	1.0
average	0.5	0.5
difference	-1.0	1.0

and then check the resulting estimates:

```
R> smry03$mean[,c(2,5,6)]
      median 95% lower 95% upper
basiliximab -1.3736003 -2.094845 -0.6670075
daclizumab  -2.0821175 -3.077314 -1.0710786
average      -1.7288302 -2.394382 -1.0591354
difference   -0.7062405 -1.811174  0.4232024
R> smry04$mean[,c(2,5,6)]
      median 95% lower 95% upper
basiliximab -1.3736003 -2.094845 -0.6670075
daclizumab  -2.0821175 -3.077314 -1.0710786
average      -1.7288302 -2.394382 -1.0591354
difference   -0.7062405 -1.811174  0.4232024
```

and indeed the corresponding estimates are identical under both model variations (and different from those shown in Figure 2, which were based on uniform priors).

3.1.4. Connection to meta-analysis without covariables

Meta-regression is a generalization of a “simple” meta-analysis, and the simple meta-analysis again constitutes the special case of a regression that does not consider additional covariables (besides an overall “intercept” term). We may also use the `bmr()` function without covariables by simply omitting the “X” argument:

```
R> bmr05 <- bmr(crins.es,
+ tau.prior=function(t){dhalfnormal(t, s=0.5)})
```

We can check the regressor matrix used internally (which here may be accessed as “`bmr05$X`”) to see that indeed this is by default a matrix consisting of a single column of ones, so that a single “intercept” parameter β_1 is fitted.

We may then compare results against those from the `bayesmeta()` function; comparing the estimates for the overall mean, we get

```
R> bma <- bayesmeta(crins.es,
+ tau.prior=function(t){dhalfnormal(t,s=0.5)})
R> cbind("bmr" = bmr05$summary[, "intercept"],
+ "bayesmeta" = bma$summary[, "mu"])
      bmr bayesmeta
mode   -1.5771362 -1.5779534
median -1.5830833 -1.5834970
mean   -1.5871400 -1.5873257
sd      0.3310698  0.3314634
95% lower -2.2465774 -2.2464090
95% upper -0.9341840 -0.9338935
```

Differences between the two results are due to differences in numerical accuracy; if we check the number of support points used internally for the approximation of the posterior (via “`str(bmr05$support)`” or “`str(bma$support)`”), we can see that `bmr()` uses 9 support points, while `bayesmeta()` yields a grid of 17 points. The slight discrepancy arises since within the `bayesmeta()` function, the grid setup is determined based on the marginal distribution of the overall mean (μ) as well as the shrinkage estimates (θ_i) [60], while in

`bmr()` only the regression coefficients’ (multivariate) distribution is considered. If desired, accuracy may always be increased by adjusting the “delta” or “epsilon” parameters [61].

In the two-group comparison discussed above (see e.g. Figure 2), the two group means (basiliximab and daclizumab) are estimated “independently” in some sense, i.e., the estimates from one group of studies only help estimating the other group’s mean insofar as they provide information on the heterogeneity, but not on the actual *location*. The difference to performing two completely separate analyses of both group means is the assumption of a common heterogeneity parameter for both groups. This provides another connection to the “simple” meta-analysis model (without additional covariables): the two group mean estimates may also be recovered by performing two separate meta-analyses and propagating only the heterogeneity information. Consider the estimate of the *daclizumab* group, which was given by

```
R> bmr01$summary[, "daclizumab", drop=FALSE]
      daclizumab
mode          -2.3072701
median        -2.3072414
mean          -2.3072091
sd             0.5842337
95% lower    -3.4589967
95% upper    -1.1554177
```

The same estimate may be derived by first performing the analysis for the *basiliximab* group only, and then using the resulting heterogeneity posterior as the prior for the subsequent *daclizumab* analysis:

```
R> basil <- (crins.es$IL2RA=="basiliximab")
R> dacli <- (crins.es$IL2RA=="daclizumab")
R> bma.bas <- bayesmeta(crins.es[basil,],
+ tau.prior=function(t){dhalfnormal(t,s=0.5)})
R> bma.dac <- bayesmeta(crins.es[dacli,],
+ tau.prior=function(t){bma.bas$dposterior(
+ tau=t)})
R> bma.dac$summary[, "mu", drop=FALSE]
      mu
mode   -2.3072954
median -2.3072387
mean   -2.3072071
sd      0.5847048
95% lower -3.4596942
95% upper -1.1547187
```

One can see that the results are essentially identical, with slight discrepancies that may be attributed to numerical differences.

3.2. Indirect comparisons in a treatment network

It became already evident in the previous example (Section 3.1.2) that fitting individual coefficients for certain pairwise comparisons, along with the option to infer linear combinations of coefficients, allows to estimate certain *indirect comparisons* [45]. In fact, applicability of the meta-regression model to some degree extends into the domain of *network meta-analysis (NMA)* [40, Sec. 11.4.2]. However, the scope here is

somewhat limited, insofar as the model is *contrast-based*, only two-armed trials may be considered, and a single common heterogeneity parameter is assumed [18, 64, 79].

For illustration, we will consider the example data set due to Bucher et al. [5], which includes studies providing evidence for both the *direct* as well as the *indirect* comparison of two treatments. Bucher et al. considered the example of the comparison of *sulphamethoxazole-trimethoprim (TMP-SMX)* versus *dapsone/pyrimethamine (D/P)* for the prophylaxis of *Pneumocystis carinii* pneumonia (Pcp) in HIV patients. Eight studies had undertaken a head-to-head comparison of both medications, but an additional 14 studies were available investigating one of the medications with *aerosolized pentamidine (AP)* as a comparator. Nine studies compared TMP-SMX vs. AP, and five studies compared D/P vs. AP. Together these provide *indirect* evidence on the effect of TMP-SMX compared to D/P. The resulting triangular network of pairwise comparisons is illustrated in Figure 3.

We may load the data and compute effect sizes (log-ORs) for all 22 studies.

```
R> data("BucherEtA11997")
R> es <- escalc(measure="OR",
+             ai=events.A, n1i=total.A,
+             ci=events.B, n2i=total.B,
+             slab=study, data=BucherEtA11997)
```

We may then set up the regressor matrix to estimate the two relevant (non-redundant) treatment effects. Two coefficients (β_1 and β_2) are estimated; the first corresponds to the comparison of TMP-SMX against D/P, the second is for the comparison of AP against D/P, and the last remaining pairwise comparison (TMP-SMX vs. AP) then results as the difference of the former two.

```
R> X <- cbind("TMP.DP"=rep(c(1,0, 1), c(8,5,9)),
+           "AP.DP" =rep(c(0,1,-1), c(8,5,9)))
```

Again, there are alternative (equivalent) ways to set up the regressor matrix [64]. The actual analysis then is performed via a call of

```
R> bmr06 <- bmr(es, X=X)
```

The “`tau.prior`” argument is left unspecified, which means that the default of an (improper) uniform prior is

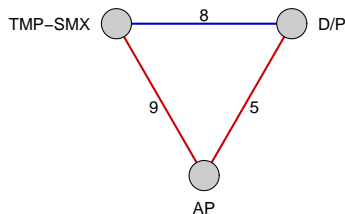


Figure 3: Illustration of the triangular network of comparisons within the Bucher et al. [5] example data set. 8 studies provide a *direct* head-to-head comparison of TMP-SMX vs. D/P (indicated by the blue edge); the remaining 14 studies (shown in red) provide *indirect* evidence on the effect via the comparison of either TMP-SMX versus AP (9 studies), or D/P versus AP (5 studies).

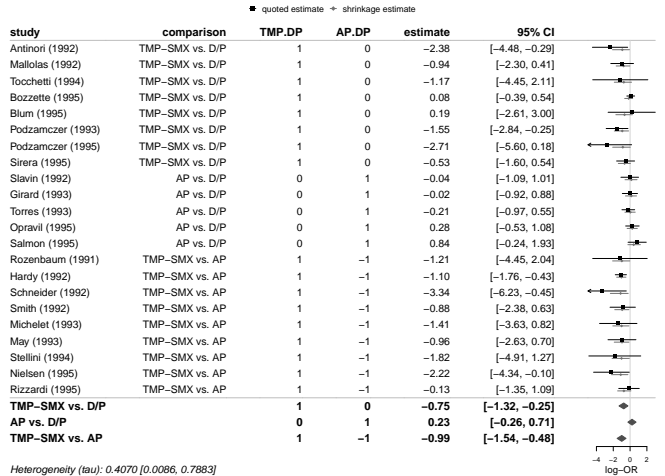


Figure 4: Forest plot for the network-MA data set [5]. The first 8 studies did a *direct* head-to-head comparison of TMP-SMX vs. D/P; the remaining studies provide *indirect* evidence on the effect via the comparison with AP. At the bottom, the estimates for all three pairwise comparisons are shown.

used for τ , which should be appropriate given the reasonably large number of studies included here ($k = 22$) [63]. Figure 4 illustrates the data along with the regressor matrix and the estimated coefficients in a corresponding forest plot.

In order to more closely investigate and contrast the “direct” and “indirect” contributions to specific estimates, we may have a closer look at the estimates resulting from considering subsets of the data. In addition to the analysis described above, we may run analyses based only on the subsets of studies providing direct or indirect evidence (studies 1–8 or studies 9–22):

```
R> bmr06.direct <- bmr(es[1:8,], X=X[1:8,1])
R> bmr06.indirect <- bmr(es[9:22,], X=X[9:22,1])
```

In the overall analysis, the log-OR for the effect of TMP-SMX vs. D/P is estimated at -0.75 with 95% CI $[-1.32, -0.25]$; the “direct” and “indirect” estimates are roughly similar and overlapping at -0.83 $[-1.91, 0.06]$ and -0.96 $[-1.73, -0.21]$. We may also inspect the corresponding posterior densities, which again are accessible via the returned “`...$dposterior()`” functions. In Figure 5, the three posteriors are contrasted side-by-side. All three estimates are consistent, and when combining direct and indirect evidence, the gain in precision becomes apparent.

3.3. Continuous covariable

Nicholas et al. [54] performed a systematic review and meta-analysis in order to examine how the characteristics of placebo groups of randomized controlled trials in multiple sclerosis may have evolved over time. A number of features were investigated, among these was the *proportion of patients experiencing disability progression* within

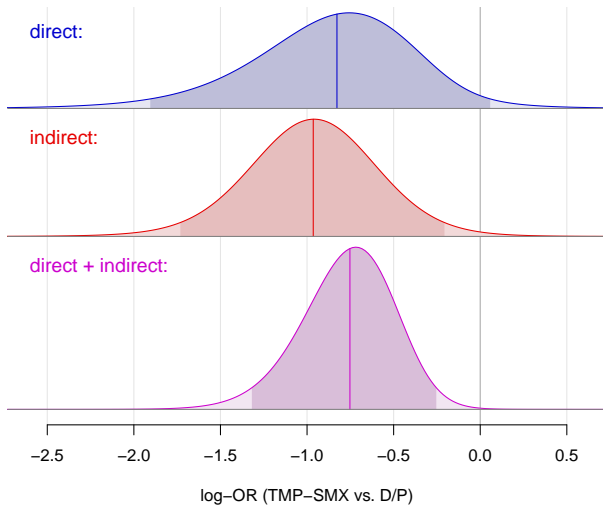


Figure 5: Posterior distributions for the β coefficient corresponding to the log-OR in the comparison of *TMP-SMX* vs. *D/P*, either considering only the studies providing direct or indirect evidence, or including all studies. Besides posterior densities, posterior medians and 95% CIs are indicated.

24 months. 28 studies with available information on disability progression were found, spanning the period from the year 1990 until 2018. A time trend in the progression rates would mean a tendency towards more severe or more benign cases being investigated over the years, and may have implications for the comparability of results from older or more recent studies, or also for the design of future studies. We can load the example data, and from the studies' placebo group sizes and the observed percentages of progressing patients, we can compute estimates of the logarithmic odds of disease progression using the `escalc()` function:

```
R> data("NicholasEtAl2019")
R> head(NicholasEtAl2019)
      study year patients prog.percent
1 Kastrukoff (1990) 1990    50      46.0
2 Wolinsky (1990) 1990   274      72.5
3 Bornstein (1991) 1991    55      29.5
4 Likosky (1991) 1991    20      80.0
5 Noseworthy (1991) 1991    56      22.0
6 Milanese (1992) 1992    21      55.0
R> es <- escalc(measure="PLO",
+ xi=patients*(prog.percent/100), ni=patients,
+ slab=study, data=NicholasEtAl2019)
R> head(es[,c(1,2,5,6)])
      study year   yi   vi
1 Kastrukoff (1990) 1990 -0.1603 0.0805
2 Wolinsky (1990) 1990  0.9694 0.0183
3 Bornstein (1991) 1991 -0.8712 0.0874
4 Likosky (1991) 1991  1.3863 0.3125
5 Noseworthy (1991) 1991 -1.2657 0.1041
6 Milanese (1992) 1992  0.2007 0.1924
```

The “*yi*” and “*vi*” columns here give the log-odds and their (squared) standard errors. The (continuous) covariable of interest is given in the “*year*” column. We may then specify the regressor matrix:

```
R> X <- cbind("intercept2000" = 1,
+            "year" = (es$year-2000))
```

```
R> head(X)
      intercept2000 year
[1,]              1 -10
[2,]              1 -10
[3,]              1  -9
[4,]              1  -9
[5,]              1  -9
[6,]              1  -8
```

Note that here we are using to a simple “intercept/slope” model, and that the “*year*” variable is re-coded so that the data are centered at the year 2000 (and the intercept parameter hence corresponds to the log-odds in 2000). We may then perform the analysis:

```
R> bmr07 <- bmr(es, X=X)
```

A prior for the heterogeneity (τ) again is not specified, implying that the default of an (improper) uniform prior is used [63].

We may inspect the regression results based on the returned parameter estimates, but it may in fact be more illustrative to present these in a forest plot. Besides the two “plain” parameter estimates (intercept β_1 and slope β_2) we may check estimates of certain linear combinations corresponding to the mean at certain time points, or also to predicted values. A call of

```
R> forestplot(bmr07, xlab="log-odds",
+ X.mean=rbind("intercept (2000)" =c(1, 0),
+             "annual change" =c(0, 1),
+             "change per decade" =c(0, 10),
+             "mean 1990" =c(1, -10),
+             "mean 2000" =c(1, 0),
+             "mean 2010" =c(1, 10),
+             "mean 2018" =c(1, 18)),
+ X.predict=rbind("prediction 2019"=c(1, 19)))
```

will generate a forest plot including the intercept and slope (annual change), the change per decade, the mean log-odds at several time points, as well as a prediction for the year 2019; the resulting plot is shown in Figure 6. The annual change is estimated to be negative, implying a reduction in the log-odds by 0.033 per year, or by 0.33 per decade. For the odds this means a reduction by 3.2% per year, or by 28% per decade.

Besides the forest plot, it is often useful to illustrate the data along with the model estimates graphically. To that end, we can compute predictions and credible intervals and combine these in a single plot:

```
R> # specify corresponding "regressor matrix":
R> newx <- cbind(1, (1989:2019)-2000)
R> # compute credible intervals for the mean:
R> pred <- cbind("median"=bmr07$qpred(0.5,x=newx),
+             bmr07$pred.interval(x=newx))
R> # compute prediction intervals:
R> map <- cbind("median"=bmr07$qpred(0.5,x=newx,
+             mean=FALSE),
+             bmr07$pred.interval(x=newx,
+             mean=FALSE))
R> # show the 26 studies' point estimates
R> # and 95 percent CIs:
R> plot(es$year-2000, es$yi,
+       xlim=range(newx[,2]), ylim=range(map),
+       xlab="publication year - 2000",
+       ylab="log-odds")
```

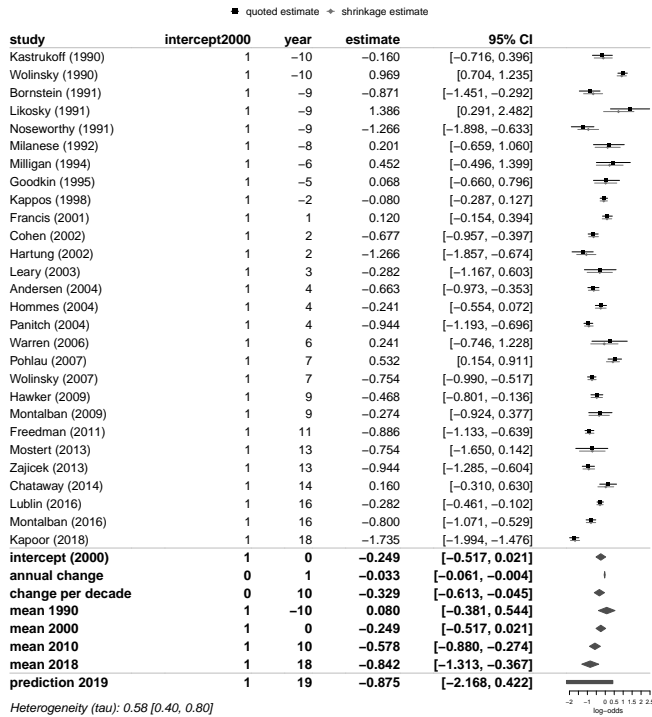


Figure 6: A forest plot illustrating the meta-regression results based on the multiple sclerosis data. The time trend is parameterized in terms of an *intercept* and a *year* variable that is centered at the year 2000. In addition to the “plain” intercept and (annual) slope, several linear combinations as well as a prediction for the year 2019 are also shown.

```
R> matlines(rbind(es$year, es$year)-2000,
+           rbind(es$yi-qnorm(0.975)*sqrt(es$vi),
+               es$yi+qnorm(0.975)*sqrt(es$vi)),
+           col=1, lty=1)
R> # show trend lines (and 95 percent intervals):
R> matlines(newx[,2], map, col="blue", lty=c(1,2,2))
R> matlines(newx[,2], pred, col="red", lty=c(1,2,2))
R> legend("topright", pch=15, col=c("red", "blue"),
+       c("mean", "prediction"))
```

The resulting trend plot is shown in Figure 7. All estimates (studies) along with their error bars are shown, and the estimated mean (along with credible and prediction intervals) is computed for a grid of values spanning the range from 1989 to 2019.

Nicholas et al. [54] pointed out the good agreement with a subsequently published study by Kappos et al. [43] in 2018, who (depending on the exact definition used) reported between 23% and 30% progressing patients, corresponding to log-odds of -1.18 or -0.83 , respectively.

The predicted log-odds were -0.87 $[-2.17, 0.42]$ for a “future” study in the year 2019 (see Figure 6), corresponding to probabilities of 0.29 $[0.10, 0.60]$. Such a prediction could be useful for study design [16, 23, 28] or sample size determination [15]; it might also be utilized to supplement a study’s sparse placebo data in terms of a meta-analytic-predictive (MAP) prior [67].

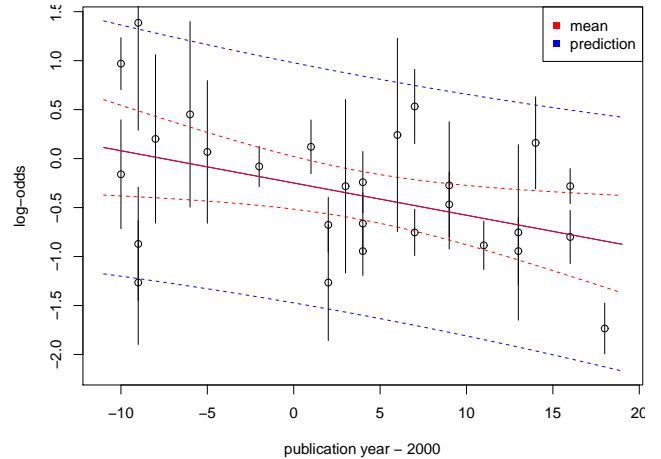


Figure 7: A plot illustrating the multiple sclerosis data along with credible intervals for the mean as well as prediction intervals.

3.4. Several covariables

Roberge et al. [58] performed a systematic literature review in order to summarize the evidence on effects of aspirin administered during pregnancy. Earlier research had already suggested that prophylactic administration of low-dose aspirin may reduce the prevalence of *fetal growth restriction (FGR)*, which is a common cause of perinatal morbidity and mortality [53]. While the exact mechanism by which aspirin works here is still unclear, it had become apparent that it is most effective when initiated early on, before 16 weeks of gestational age.

A total of 35 studies were included in the eventual analysis; in 17 studies, therapy was initiated early (≤ 16 weeks gestational age), and in 18 studies, onset was late (> 16 weeks). Doses differed between studies and ranged from 50 up to 150 mg daily. For each study, we have numbers of cases and FGR events in treatment and control groups.

We may load the example data and derive the log-ORs for all studies providing data on FGR events:

```
R> data("RobergeEtAl2017")
R> es.fgr <- escalc(measure="OR",
+ ai=asp.FGR.events, n1i=asp.FGR.total,
+ ci=cont.FGR.events, n2i=cont.FGR.total,
+ slab=study, data=RobergeEtAl2017,
+ subset=
+ complete.cases(RobergeEtAl2017[,11:14]))
```

At first we can check whether the aspirin dose appears to affect the chances of FGR; we can use the `model.matrix()` function to set up a corresponding regressor matrix and perform a simple analysis specifying an intercept and a linear effect for the dose. Again, due to the large number of studies included, we may utilize a non-informative (improper) uniform prior for the heterogeneity.

```
R> X01 <- model.matrix(~ dose, data=es.fgr)
R> colnames(X01) <- c("intercept", "dose")
R> head(X01)
      intercept dose
01           1    50
02           1    60
```

```

03      1    60
04      1    60
06      1    75
07      1    75
> bmr08 <- bmr(es.fgr, X=X01)
> bmr08$summary
      tau      intercept      dose
mode    0.27075394  0.08446175 -0.004604890
median  0.28349590  0.08799880 -0.004864835
mean    0.29005627  0.09099202 -0.005005075
sd      0.11591364  0.30196489  0.003685612
95% lower 0.06368288 -0.50870208 -0.012461472
95% upper 0.52626505  0.69503106  0.002157158

```

So far, this does not convincingly indicate an effect; while the “dose” effect is estimated to be negative, implying a reduction in FGR events with increasing dose, the 95% CI includes both negative as well as positive values.

Considering the earlier suggestion of the relevance of the timepoint of therapy initiation, we may then check whether the effect might differ between studies implementing an “early” or “late” onset. Such a model may be defined in different ways; here we will consider a setup including individual intercepts and slopes in both groups of studies:

```

R> X02 <- model.matrix(~ -1 + onset + onset:dose,
+                       data=es.fgr)
R> colnames(X02) <- c("intEarly", "intLate",
+                    "doseEarly", "doseLate")
R> head(X02)
  intEarly intLate doseEarly doseLate
01         1      0         50         0
02         1      0         60         0
03         1      0         60         0
04         1      0         60         0
06         1      0         75         0
07         1      0         75         0

```

We can see that the first studies within the data set all belong to the “early” group; in the second group, the regressor matrix entries corresponding to the “late” effects then are non-zero instead. We may then run the analysis based on the extended model:

```

R> bmr09 <- bmr(es.fgr, X=X02)
R> bmr09$summary[,1:3]
      tau      intEarly      intLate
mode    0.08758667  0.4290267  0.05426766
median  0.11720808  0.4263765  0.04654656
mean    0.13062301  0.4248359  0.04167659
sd      0.09016564  0.5201030  0.24672617
95% lower 0.00000000 -0.5970099 -0.45148496
95% upper 0.29774638  1.4449251  0.52534390
R> bmr09$summary[,4:5]
      doseEarly      doseLate
mode   -0.014014415 -0.001538704
median -0.013981904 -0.001491126
mean   -0.013970065 -0.001474899
sd     0.006281788  0.002990934
95% lower -0.026280559 -0.007383067
95% upper -0.001626998  0.004450449

```

From the analysis results, we now see a somewhat different picture; first of all, the heterogeneity (τ) is reduced, from a median of 0.28 down to 0.12. The “late” slope parameter still is small and centered near zero, while the “early” slope parameter along with its 95% CI is on the negative side.

We may illustrate data and regression lines jointly in a “bubble plot”; the estimated ORs as functions of the regressors may again be extracted from the `bmr()` function’s output:

```

R> # derive predictions from the model;
R> # specify corresponding "regressor matrices":
R> newx.early <- cbind(1, 0, seq(50,150,by=5), 0)
R> newx.late <- cbind(0, 1, 0, seq(50,150,by=5))
R> # compute predicted medians
R> # and 95 percent intervals:
R> pred.early <- cbind("median"=bmr09$qpred(0.5,
+                                           x=newx.early),
+
+ bmr09$pred.interval(x=newx.early))
R> pred.late <- cbind("median"=bmr09$qpred(0.5,
+                                           x=newx.late),
+
+                   bmr09$pred.interval(x=newx.late))
R> # draw "bubble plot":
R> plot(es.fgr$dose, es.fgr$yi,
+       cex=1/sqrt(es.fgr$vi),
+
+       col=c("blue", "red")[as.numeric(es.fgr$onset)],
+       xlab="dose (mg)", ylab="log-OR (FGR)")
R> legend("topright", col=c("blue", "red"),
+       c("early onset", "late onset"), pch=1)
R> # add predictions to bubble plot:
R> matlines(newx.early[,3], pred.early,
+          col="blue", lty=c(1,2,2))
R> matlines(newx.late[,4], pred.late,
+          col="red", lty=c(1,2,2))

```

The resulting trend plot is shown in Figure 8. “Larger” studies (those with smaller standard errors) are denoted by larger symbols.

Again, a range of alternative model or prior specifications may be sensible here; e.g., different parametrizations (common intercept or slope parameters, individual heterogeneity parameters, differing prior specifications, a zero intercept, ...), which may also suggest the use of a *model selection* approach (see following Section 3.5). In addition, it may also be interesting to investigate the sensitivity of

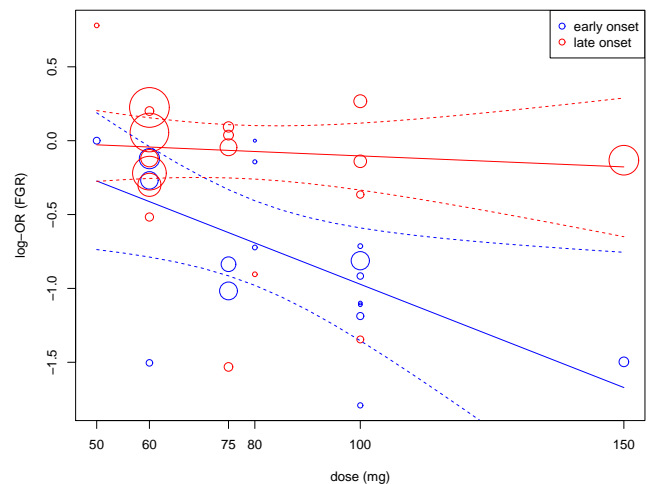


Figure 8: A “bubble plot” illustrating the data along with credible intervals for the mean effect as functions of the dose in “early” and “late” study groups. The point sizes are inversely proportional to the standard errors.

results to individual data points (studies).

3.5. Model selection

3.5.1. The example setup

Cinar et al. [10] discussed a meta-analysis problem involving a total of four potential covariables. Their example data set included 80 studies investigating biomass production of maize plants under different conditions; of interest were the effects of inoculation using symbiotic mycorrhizal fungi. Four dichotomous aspects varied between the studies, namely the type of fungus (FUN, *funneliformis* or *rhizophagus*), the use of phosphorus fertilizer (FP, yes/no), use of nitrogen fertilizer (FN, yes/no), and sterilization of the soil (STER, yes/no). The endpoint was expressed in terms of a logarithmic response ratio [37], and the problem was first of all to determine *which* of the variables affected the yield, making this a *variable selection* or *model selection* problem.

In the present case, combinations of the four (binary) variables allow to specify 16 different *models*, ranging from the model without covariables (besides an overall intercept) to the model with all four included. In a Bayesian context, after specification of prior probabilities (for all models themselves, as well as for parameters within models), one may then derive posterior probabilities for each model, or one may compare and rank models based on their associated *Bayes factors* [4, 26, 44, 69]. The uncertainty involved in the model selection may also be accounted for (or in fact, the *selection* of a single model is avoided) by using a *model averaging* approach [11, 12, 42, 57, 62]. Another approach may be to consider the *median probability model* based on all variables' marginal inclusion probabilities [2, 3]. Either way, computations hinge on the determination of *marginal likelihoods*, which first of all is often computationally challenging, and secondly, requires the specification of *proper* priors for all parameters within the 16 models. Unlike in many parameter estimation problems, the exact details of (non-informative or weakly informative) prior specifications are crucial and may affect results in sometimes unintuitive ways, as exemplified in *Lindley's paradox* [26, 49], so that particular caution is advised here. In some cases, it may be worth considering whether a model selection approach is in fact the method of choice [25].

3.5.2. Model specification

We assume all of the 16 possible models to be a priori equally likely; within each model we then assign a vague prior for the intercept (normal with mean zero and standard deviation 10), and weakly informative priors for the binary covariables' effects (normal with mean zero and standard deviation 2.82). The effect prior confines the likely effect magnitudes on their logarithmic scale so that back on the (exponentiated) scale of response ratios these roughly correspond to values within factors of 250 and $\frac{1}{250}$ [30]. For the heterogeneity parameter (τ), we assume a

weakly informative half-normal distribution with scale 0.5 [63].

Note that equal prior probabilities for all models imply that all variables a-priori have 50% inclusion probability, and that the prior expected number of parameters is $\frac{N}{2}$ (where N is the total number of variables). Alternatively, different specifications are also conceivable; for example, assigning a probability π for each variable to be included implies a probability $\pi^n(1-\pi)^{(N-n)}$ for each single model (where n is the number of variables included) and it implies a priori a binomially distributed total number of included parameters (with expectation πN).

3.5.3. Implementation

The example data are available in Cinar et al.'s online supplement [9]. We may download the data, read them into R, and then systematically apply the 16 possible meta-regression models. Computations may take a few minutes.

```
R> # load data:
R> CinarEtA12021 <- read.csv("CinarEtA12021.csv")
R> # convert effect sizes:
R> effsize <- escalc(measure="ROM", yi=yi, vi=vi,
+                   data=CinarEtA12021)
R> head(effsize)
      FUN FP FN STER      yi      vi
1      0  0  0   0  0.0950 0.4000
2      0  0  0   0 -1.8590 0.4000
3      0  0  0   0  0.3640 0.6667
4      0  0  0   0  0.2880 0.4000
5      0  0  0   0  0.1310 0.2500
6      0  0  0   0  0.0000 0.4000
R> # figure out possible models
R> # (factor combinations):
R> models <- expand.grid("FUN"=c(FALSE,TRUE),
+                      "FP"=c(FALSE,TRUE),
+                      "FN"=c(FALSE,TRUE),
+                      "STER"=c(FALSE,TRUE))
R> models <- as.matrix(models)
R> rownames(models) <- LETTERS[1:16]
R> # generate list of 16 regressor matrices:
R> Xlist <- vector(nrow(models), mode="list")
R> names(Xlist) <- rownames(models)
R> for (i in 1:nrow(models)) {
+   Xlist[[i]] <- matrix(
+     rep(1, nrow(CinarEtA12021)),
+     ncol=1, dimnames=list(NULL, "intercept"))
+   if (any(models[i,])) {
+     Xlist[[i]] <- cbind(Xlist[[i]],
+       as.matrix(CinarEtA12021[1:4][,models[i,]]),
+       colnames(Xlist[[i]]) <- c("intercept",
+         colnames(models)[models[i,]])
+   }
+ }
R> # perform 16 regression analyses:
R> bmrlist <- vector(nrow(models), mode="list")
R> names(bmrlist) <- rownames(models)
R> for (i in 1:nrow(models)) {
+   bmrlist[[i]] <- bmr(effsize, X=Xlist[[i]],
+     tau.prior=function(t){dhalfnormal(t,0.5)},
+     beta.prior.mean=rep(0, ncol(Xlist[[i]])),
+     beta.prior.sd=c(10,
+       rep(2.82,
+         ncol(Xlist[[i]])-1)))
+ }
```

The 16 regression outputs are now stored in the "bmrlist" object. The marginal likelihood for a `bmr()` output is

stored in the “...\$marginal.likelihood” element (presuming that proper priors have been used for the analysis). We may now assemble these numbers, and combine them with the models’ prior probabilities to derive the posterior probabilities.

```
R> # specify model prior probabilities:
R> priorprob <- rep(1/16, 16)
R> # assemble marginal likelihoods
R> margi <- sapply(bmrlist,
+   function(x){x$marginal.likelihood})
R> # determine posterior model probabilities:
R> prob <- (priorprob*margi)/sum(priorprob*margi)
```

Table 3 illustrates the 16 models along with their posterior probabilities. The a-posteriori most probable model at the top of the list is the one including (besides an intercept) the FP and FN variables, corresponding to influential effects for phosphorus and nitrogen fertilizers.

The *median probability model*, i.e., the model including all those variables that have a marginal inclusion probability ≥ 0.5 [2, 3], here also coincides with the most probable model. The three most probable models also match the top 3 models based on the *Akaike information criterion (AIC)* as quoted by Cinar et al. [10].

Table 4 shows the parameter estimates from the most probable model. Instead of singling out one “best” model for inference, one might now also utilize the results in a *model averaging* approach, effectively using all 16 models simultaneously and weighting predictions based to their

Table 3: The 16 models and their probabilities (in descending order). A dot (●) indicates that a variable is included in a model, an open circle (○) means that it is not included. The very last line shows the four variables’ marginal inclusion probabilities.

model	included variables				probability
	FUN	FP	FN	STER	
1	○	●	●	○	0.6293
2	○	●	●	●	0.1076
3	●	●	●	○	0.0907
4	●	●	○	○	0.0645
5	○	●	○	●	0.0380
6	●	●	○	●	0.0304
7	●	●	●	●	0.0148
8	○	●	○	○	0.0106
9	●	○	○	○	0.0077
10	●	○	○	●	0.0019
11	○	○	○	●	0.0014
12	○	○	○	○	0.0011
13	●	○	●	○	0.0009
14	○	○	●	○	0.0007
15	●	○	●	●	0.0002
16	○	○	●	●	0.0002
	0.2111	0.9859	0.8444	0.1946	

Table 4: Parameter estimates for the most probable model including the FP and FN variables (phosphorus and nitrogen fertilizers), which receives a posterior probability of 0.63.

parameter	median	95% CI
heterogeneity (τ)	0.510	[0.339, 0.690]
intercept (β_1)	0.227	[-0.097, 0.652]
FP (β_2)	-1.006	[-1.416, -0.582]
FN (β_3)	0.894	[0.429, 1.346]

associated probabilities [11, 12, 42, 57, 62].

It should be noted that, in a sense, the example shown here was particularly “simple” since all variables considered were in the same “units” (binary), so that it is relatively easy to specify a neutral prior without favouring any of the variables from the start. In practice, it might also be of interest to check the results’ sensitivity to any of the prior specifications, or to also investigate the possible relevance of interaction effects (as a simple additive effect of the two fertilizers may or may not be biologically plausible).

In the model selection context, the use of *penalized complexity priors* [46] may also play a more prominent role than in “simple” meta-analysis applications. Penalized complexity priors here correspond to exponential priors for the heterogeneity (τ) [63].

4. Discussion

In the present article, we demonstrated the use of Bayesian meta-regression as facilitated through the *bayesmeta* R package [59]. The implementation is conveniently based on the *DIRECT* algorithm [61] and constitutes a straightforward generalisation of “simple” meta-analysis within the *NNHM* framework [60]. This way, a wide range of extensions such as subgroup analysis, continuous covariables, indirect comparisons, or model selection are covered.

While Bayesian analyses sometimes tend to be technically demanding, through to its user-friendly interface, its generality and the quick and reproducible computation, the *bayesmeta* implementation provides a low-threshold entry point for a wider audience beyond computational experts. While a certain amount of preparation certainly is still required, it is relatively easy to extend an existing *bayesmeta* implementation to include covariables in addition, whereas the effort required to run, diagnose and possibly implement an MCMC approach would be substantially higher. In many applications, use of the *bmr()*, *forestplot()* and *summary()* functions may already be sufficient to address most relevant questions. More sophisticated investigations are possible using the comprehensive output available (such as custom plots (see Sections 3.2–3.4), model selection (see Section 3.5) or model averaging [62]).

The meta-regression approach presented here builds on the NNHM, which yields accurate inference in particular also in case of few studies [22]. A normal approximation at the study-level is often appropriate, for example when sample sizes are “large” and studies are sufficiently powered. However, the normal approximation may also deteriorate in certain circumstances, for example, for binary endpoints with zero (or near-zero) event counts. In such cases, use of an exact likelihood may be preferable [71, 76], as for example implemented in the `metaStan` R package [30]. This package also allows for *model-based meta-analysis (MBMA)*, where studies may contribute information on several (more than two) study arms that correspond to different exposures or dose levels [31, 51]. Similarly, the `bmeta` package provides meta-analysis and meta-regression functionalities based on MCMC sampling [20]. The `RBeST` package also supports meta-analysis based on a range of endpoint types, with a focus on deriving (“meta-analytic-predictive (MAP)”) prior distributions for use in a subsequent analysis [77]. In the example applications we included indirect comparisons as a very basic example of a treatment network; more complex models commonly applied for network meta-analysis (see e.g. [40, Sec. 11]) are currently not implemented in the `bayesmeta` package. Dedicated packages are available for network meta-analysis, for instance `nm-INLA`, which utilizes the *integrated nested Laplace approximation (INLA)* for posterior inference [29]. Recently, Williams et al. [80] proposed meta-analytic models with covariate effects on the heterogeneity variance besides the mean; these are implemented in the R package `blsmeta`. The `bspmma` and `metaBMA` R packages implement extensions of the “simple” NNHM, but currently without the option for meta-regression [7, 35]. Alternatively, many meta-analysis problems may also be formulated in terms of generalized linear mixed models (GLMMs), which can be fitted for instance using the R package `brms` [6]. Most of the above alternative Bayesian packages rely on MCMC methods for inference. In case a tailored solution is required, which is not covered by any of the mentioned packages, it is probably easiest to also resort to MCMC methods, which may be implemented e.g. using the JAGS or Stan engines, and the interfacing `rjags` or `rstan` packages [55, 56, 32]. For a comprehensive and up-to-date overview of available R packages, see also the corresponding *CRAN task view* [17].

Acknowledgment

Support from the *Deutsche Forschungsgemeinschaft (DFG)* is gratefully acknowledged (grant number FR 3070/3-1).

Conflict of interest statement

The authors declare no conflict of interest.

Ethics approval

Not applicable.

References

- [1] W. L. Baker, C. M. White, J. C. Cappelleri, J. Kluger, and C. I. Coleman. Understanding heterogeneity in meta-analysis: the role of meta-regression. *The International Journal of Clinical Practice*, 63(10):1426–1434, Oct. 2009. doi: 10.1111/j.1742-1241.2009.02168.x.
- [2] M. M. Barbieri and J. O. Berger. Optimal predictive model selection. *The Annals of Statistics*, 32(3):870–897, June 2004. doi: 10.1214/009053604000000238.
- [3] M. M. Barbieri, J. O. Berger, E. I. George, and V. Rockova. The median probability model and correlated variables. *Bayesian Analysis*, 0(0):0, Dec. 2021. doi: 10.1214/20-BA1249.
- [4] J. Berger and L. R. Pericchi. Objective Bayesian methods for model selection: Introduction and comparison. In P. Lahiri, editor, *Model Selection*, volume 38 of *IMS Lecture Notes*, pages 135–193. Institute of Mathematical Statistics, Beachwood, OH, 2001. doi: 10.1214/lnms/1215540968.
- [5] H. C. Bucher, G. H. Guyatt, L. E. Griffith, and S. D. Walter. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *Journal of Clinical Epidemiology*, 50(6):683–691, June 1997. doi: 10.1016/S0895-4356(97)00049-8.
- [6] P.-C. Bürkner. `brms`: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1), Aug. 2017. doi: 10.18637/jss.v080.i01.
- [7] D. Burr. `bspmma`: An R package for Bayesian semiparametric models for meta-analysis. *Journal of Statistical Software*, 50(4), July 2012. doi: 10.18637/jss.v050.i04.
- [8] I. Chalmers, L. V. Hedges, and H. Cooper. A brief history of research synthesis. *Evaluation & the Health Professions*, 25(1): 12–37, Mar. 2002. doi: 10.1177/0163278702025001003.
- [9] O. Cinar, J. Umbanhowar, J. D. Hoekesma, and W. Viechtbauer. Using information-theoretic approaches for model selection in meta-analysis; supplementary data. OSF, Nov. 2020. URL: <https://osf.io/3d8u5/>.
- [10] O. Cinar, J. Umbanhowar, J. D. Hoekesma, and W. Viechtbauer. Using information-theoretic approaches for model selection in meta-analysis. *Research Synthesis Methods*, 12(4): 537–556, July 2021. doi: 10.1002/jrsm.1489.
- [11] R. T. Clemen. Combining forecasts: a review and annotated bibliography. *International Journal of Forecasting*, 5(4):559–583, 1989. doi: 10.1016/0169-2070(89)90012-5.
- [12] M. Clyde and E. I. George. Model uncertainty. *Statistical Science*, 19(1):81–94, Feb. 2004. doi: 10.1214/088342304000000035.
- [13] H. Cooper. Hypotheses and problems in research synthesis. In H. Cooper, L. V. Hedges, and J. C. Valentine, editors, *Handbook of Research Synthesis and Meta-Analysis*, chapter 2, pages 19–35. Russell Sage Foundation, New York, NY, USA, 2nd edition, 2009.
- [14] N. D. Crins, C. Röver, A. D. Goralczyk, and T. Friede. Interleukin-2 receptor antagonists for pediatric liver transplant recipients: A systematic review and meta-analysis of controlled studies. *Pediatric Transplantation*, 18(8):839–850, Dec. 2014. doi: 10.1111/ptr.12362.
- [15] F. De Santis. Using historical data for Bayesian sample size determination. *Journal of the Royal Statistical Society A*, 170(1):95–113, Jan. 2007. doi: 10.1111/j.1467-985X.2006.00438.x.
- [16] R. DerSimonian. Meta-analysis in the design and monitoring of clinical trials. *Statistics in Medicine*, 15(12):1237–1248, June 1996.
- [17] M. Dewey. CRAN task view: Meta-analysis, 2022. URL: <https://cran.r-project.org/web/views/MetaAnalysis.html>.
- [18] S. Dias and A. E. Ades. Absolute or relative effects? Arm-based synthesis of trial data. *Research Synthesis Methods*, 7(1):23–28, Mar. 2016. doi: 10.1002/jrsm.1184.

- [19] S. Dias, A. J. Sutton, N. J. Welton, and A. E. Ades. Evidence synthesis for decision making 3: Heterogeneity—subgroups, meta-regression, bias, and bias-adjustment. *Medical Decision Making*, 33(5):618–640, July 2013. doi: 10.1177/0272989X13485157.
- [20] T. Ding and G. Baio. *bmata: Bayesian meta-analysis and meta-regression*, Feb. 2018. URL <https://github.com/giabaio/bmeta>. R package.
- [21] S. Donegan, L. Williams, S. Dias, C. Tudur-Smith, and N. Welton. Exploring treatment by covariate interactions using subgroup analysis and meta-regression in Cochrane reviews: A review of recent practice. *PLoS ONE*, 10(6):e0128804, June 2015. doi: 10.1371/journal.pone.0128804.
- [22] T. Friede, C. Röver, S. Wandel, and B. Neuenschwander. Meta-analysis of few small studies in orphan diseases. *Research Synthesis Methods*, 8(1):79–91, Mar. 2017. doi: 10.1002/jrsm.1217.
- [23] T. Friede, H. Pohlmann, and H. Schmidli. Blinded sample size reestimation in event-driven clinical trials: Methods and an application in multiple sclerosis. *Pharmaceutical Statistics*, 18(3):351–365, May 2019. doi: 10.1002/pst.1927.
- [24] M. Geissbühler, C. A. Hincapié, S. Aghlmandi, M. Zwahlen, P. Jüni, and B. R. da Costa. Most published meta-regression analyses based on aggregate data suffer from methodological pitfalls: a meta-epidemiological study. *BMC Medical Research Methodology*, 21:123, June 2021. doi: 10.1186/s12874-021-01310-0.
- [25] A. Gelman and D. B. Rubin. Avoiding model selection in Bayesian social research. *Sociological Methodology*, 25:165–173, 1995.
- [26] A. Gelman, J. B. Carlin, H. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. *Bayesian data analysis*. Chapman & Hall / CRC, Boca Raton, 3rd edition, 2014.
- [27] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter. *Markov chain Monte Carlo in practice*. Chapman & Hall / CRC, Boca Raton, 1996.
- [28] A. C. Goudie, A. J. Sutton, D. R. Jones, and A. Donald. Empirical assessment suggests that existing evidence could be used more fully in designing randomized controlled trials. *Journal of Clinical Epidemiology*, 63(9):983–991, Sept. 2010. doi: 10.1016/j.jclinepi.2010.01.022.
- [29] B. K. Günhan, T. Friede, and L. Held. A design-by-treatment interaction model for network meta-analysis and meta-regression with integrated nested Laplace approximations. *Research Synthesis Methods*, 9(2):179–194, June 2018. doi: 10.1002/jrsm.1285.
- [30] B. K. Günhan, C. Röver, and T. Friede. Random-effects meta-analysis of few studies involving rare events. *Research Synthesis Methods*, 11(1):74–90, Jan. 2020. doi: 10.1002/jrsm.1370.
- [31] B. K. Günhan, C. Röver, and T. Friede. MetaStan: An R package for Bayesian (model-based) meta-analysis using Stan. *arXiv preprint 2202.00502*, Feb. 2022. URL <http://www.arxiv.org/abs/2202.00502>.
- [32] J. Guo, J. Gabry, B. Goodrich, S. Weber, and D. a. o. Lee. *rstan: R interface to Stan*, July 2015. URL <http://cran.r-project.org/package=rstan>. R package.
- [33] J. Gurevitch, J. Koricheva, S. Nakagawa, and G. Stewart. Meta-analysis and the science of research synthesis. *Nature*, 555:175–182, Mar. 2018. doi: 10.1038/nature25753.
- [34] J. Hartung, G. Knapp, and B. K. Sinha. *Statistical meta-analysis with applications*. John Wiley & Sons, Hoboken, NJ, USA, 2008.
- [35] D. W. Heck, Q. F. Gronau, and E.-J. Wagenmakers. *metaBMA: Bayesian model averaging for random and fixed effects meta-analysis*, Mar. 2021. URL <https://cran.r-project.org/package=metaBMA>. R package.
- [36] L. V. Hedges and I. Olkin. *Statistical methods for meta-analysis*. Academic Press, San Diego, CA, USA, 1985.
- [37] L. V. Hedges, J. Gurevitch, and P. S. Curtis. The meta-analysis of response ratios in experimental ecology. *Ecology*, 80(4):1150–1156, June 1999. doi: 10.1890/0012-9658(1999)080[1150:TMAORR]2.0.CO;2.
- [38] S. Hempel, J. N. V. Miles, M. J. Booth, Z. Wang, S. C. Morton, and P. G. Shekelle. Risk of bias: a simulation study of power to detect study-level moderator effects in meta-analysis. *Systematic Reviews*, 2:107, Nov. 2013. doi: 10.1186/2046-4053-2-107.
- [39] J. P. T. Higgins. Commentary: Heterogeneity in meta-analysis should be expected and appropriately quantified. *International Journal of Epidemiology*, 37(5):1158–1160, Oct. 2008. doi: 10.1093/ije/dyn204.
- [40] J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, and V. A. Welch, editors. *Cochrane handbook for systematic reviews of interventions*. Wiley & Sons, Hoboken, NJ, USA, 2nd edition, 2019. doi: 10.1002/9781119536604. URL <http://handbook.cochrane.org/>.
- [41] J. P. T. Higgins, J. A. López-López, and A. M. Aloe. Meta-regression. In C. H. Schmid, I. White, and T. Stijnen, editors, *Handbook of meta-analysis*, chapter 7. Chapman and Hall / CRC, New York, 2021.
- [42] J. A. Hoeting, D. Madigan, A. Raftery, and C. T. Volinsky. Bayesian model averaging: a tutorial. *Statistical Science*, 14(4):382–401, Nov. 1999. doi: 10.1214/ss/1009212519.
- [43] L. Kappos, A. Bar-Or, B. A. Cree, R. J. Fox, G. Giovannoni, R. Gold, P. Vermersch, D. L. Arnold, S. Arnold, T. Scherz, C. Wolf, E. Wallström, and F. Dahlke. Siponimod versus placebo in secondary progressive multiple sclerosis (EXPAND): a double-blind, randomised, phase 3 study. *The Lancet*, 391(10127):1263–1273, Mar. 2018. doi: 10.1016/S0140-6736(18)30475-6.
- [44] A. E. Kass and R. E. Raftery. Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795, June 1995. doi: 10.2307/2291091.
- [45] C. Kiefer, S. Sturtz, and R. Bender. Indirect comparisons and network meta-analyses. *Deutsches Ärzteblatt International*, 112(47):803–808, Nov. 2015. doi: 10.3238/arztebl.2015.0803.
- [46] N. Klein and T. Kneib. Scale-dependent priors for variance parameters in structured additive distributional regression. *Bayesian Analysis*, 11(4):1071–1106, Dec. 2016. doi: 10.1214/15-BA983.
- [47] E. Kontopantelis, D. A. Springate, and D. Reeves. A re-analysis of the Cochrane Library data: The dangers of unobserved heterogeneity in meta-analyses. *PLoS ONE*, 8(7):e69930, July 2013. doi: 10.1371/journal.pone.0069930.
- [48] J. Lau, J. P. A. Ioannidis, and C. H. Schmid. Summing up evidence: one answer is not always enough. *The Lancet*, 351(9096):123–127, Jan. 1998. doi: 10.1016/S0140-6736(97)08468-7.
- [49] D. V. Lindley. A statistical paradox. *Biometrika*, 44(1/2):187–192, June 1957. doi: 10.1093/biomet/44.1-2.187.
- [50] C. Mahoney. Coefficient of determination. In M. S. Lewis-Beck, A. Bryman, and T. F. Liao, editors, *The Sage Encyclopedia of Social Science Research Methods*, pages 138–139. Sage Publications, Thousand Oaks, CA, USA, 2004. doi: 10.4135/9781412950589.n132.
- [51] D. Mawdsley, M. Bennetts, S. Dias, M. Boucher, and N. J. Welton. Model-based network meta-analysis: A framework for evidence synthesis of clinical trial data. *CPT: Pharmacometrics & Systems Pharmacology*, 5(8):393–401, Aug. 2016. doi: 10.1002/psp4.12091.
- [52] S. C. Morton, J. L. Adams, M. J. Suttorp, and P. G. Shekelle. Meta-regression approaches: What, why, when and how? AHRQ publication 04-0033, Agency for Healthcare Research and Quality (AHRQ), U.S. Department of Health and Human Services, Rockville, MD, USA, Mar. 2004.
- [53] L. M. M. Nardozza, A. C. R. Caetano, A. C. P. Zamarian, J. B. Mazzola, C. P. Silva, V. M. G. Marcal, T. F. Lobo, A. B. Peixoto, and E. A. Junior. Fetal growth restriction: current knowledge. *Archives of Gynecology and Obstetrics*, 295:1061–1077, May 2017. doi: 10.1007/s00404-017-4341-9.
- [54] R. S. Nicholas, E. Han, J. Raffel, J. Chataway, and T. Friede. Over three decades study populations in progressive multiple sclerosis have become older and more disabled, but have lower on-trial progression rates: A systematic review and meta-analysis of 43 randomised placebo-controlled trials. *Multiple*

- Sclerosis Journal*, 25(11):1462–1471, Oct. 2019. doi: 10.1177/1352458518794063.
- [55] M. Plummer. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In K. Hornik, F. Leisch, and A. Zeileis, editors, *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*. R Foundation for Statistical Computing, Austrian Association for Statistical Computing (AASC), Vienna, Austria, Mar. 2003. URL <http://www.r-project.org/conferences/DSC-2003/>.
- [56] M. Plummer. *rjags: Bayesian graphical models using MCMC*, May 2008. URL <http://cran.r-project.org/package=rjags>. R package.
- [57] A. E. Raftery, D. Madigan, and J. A. Hoeting. Bayesian model averaging for linear regression models. *Journal of the American Statistical Association*, 92(437):179–191, 1997. doi: 10.1080/01621459.1997.10473615.
- [58] S. Roberge, K. Nicolaides, S. Demers, J. Hyett, N. Chaillet, and E. Bujold. The role of aspirin dose on the prevention of preeclampsia and fetal growth restriction: systematic review and meta-analysis. *American Journal of Obstetrics & Gynecology*, 216(2):110–120, Feb. 2017. doi: 10.1016/j.ajog.2016.09.076.
- [59] C. Röver. bayesmeta: Bayesian random-effects meta analysis. R package. URL: <http://cran.r-project.org/package=bayesmeta>, Dec. 2015.
- [60] C. Röver. Bayesian random-effects meta-analysis using the bayesmeta R package. *Journal of Statistical Software*, 93(6):1–51, Apr. 2020. doi: 10.18637/jss.v093.i06.
- [61] C. Röver and T. Friede. Discrete approximation of a mixture distribution via restricted divergence. *Journal of Computational and Graphical Statistics*, 26(1):217–222, Feb. 2017. doi: 10.1080/10618600.2016.1276840.
- [62] C. Röver, S. Wandel, and T. Friede. Model averaging for robust extrapolation in evidence synthesis. *Statistics in Medicine*, 38(4):674–694, Oct. 2018. doi: 10.1002/sim.7991.
- [63] C. Röver, R. Bender, S. Dias, C. H. Schmid, H. Schmidli, S. Sturtz, S. Weber, and T. Friede. On weakly informative prior distributions for the heterogeneity parameter in Bayesian random-effects meta-analysis. *Research Synthesis Methods*, 12(4):448–474, July 2021. doi: 10.1002/jrsm.1475.
- [64] G. Salanti, J. P. T. Higgins, A. E. Ades, and P. A. Ioannidis. Evaluation of networks of randomized trials. *Statistical Methods in Medical Research*, 17(3):279–301, June 2008. doi: 10.1177/0962280207080643.
- [65] C. H. Schmid. Exploring heterogeneity in randomized trials via meta-analysis. *Drug Information Journal*, 33(1):211–224, Jan. 1999. doi: 10.1177/009286159903300124.
- [66] C. H. Schmid. Using Bayesian inference to perform meta-analysis. *Evaluation & the Health Professions*, 24(2):165–189, June 2001. doi: 10.1177/01632780122034867.
- [67] H. Schmidli, S. Gsteiger, S. Roychoudhury, A. O’Hagan, D. Spiegelhalter, and B. Neuenschwander. Robust meta-analytic-predictive priors in clinical trials with historical control information. *Biometrics*, 70(4):1023–1032, Dec. 2014. doi: 10.1111/biom.12242.
- [68] T. C. Smith, D. J. Spiegelhalter, and A. Thomas. Bayesian approaches to random-effects meta-analysis: A comparative study. *Statistics in Medicine*, 14(24):2685–2699, Dec. 1995. doi: 10.1002/sim.4780142408.
- [69] D. J. Spiegelhalter, K. R. Abrams, and J. P. Myles. *Bayesian approaches to clinical trials and health-care evaluation*. John Wiley & Sons, Chichester, UK, 2004.
- [70] L. Stewart, D. Moher, and P. Shekelle. Why prospective registration of systematic reviews makes sense. *Systematic Reviews*, 1:7, Feb. 2012. doi: 10.1186/2046-4053-1-7.
- [71] T. Stijnen, T. H. Hamza, and P. Özdemir. Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. *Statistics in Medicine*, 29(29):3046–3067, Dec. 2010. doi: 10.1002/sim.4040.
- [72] A. J. Sutton and K. R. Abrams. Bayesian methods in meta-analysis and evidence synthesis. *Statistical Methods in Medical Research*, 10(4):277–303, Aug. 2001. doi: 10.1177/096228020101000404.
- [73] S. G. Thompson. Why sources of heterogeneity in meta-analysis should be investigated. *BMJ*, 309(6965):1351–1355, Nov. 1994. doi: 10.1136/bmj.309.6965.1351.
- [74] S. G. Thompson and J. P. T. Higgins. How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine*, 21(11):1559–1573, June 2002. doi: 10.1002/sim.1187.
- [75] E. Tipton, J. E. Pustejovsky, and H. Ahmadi. A history of meta-regression: Technical, conceptual, and practical developments between 1974 and 2018. *Research Synthesis Methods*, 10(2):161–179, June 2019. doi: 10.1002/jrsm.1338.
- [76] Y.-K. Tu. Use of generalized linear mixed models for network meta-analysis. *Medical Decision Making*, 34(7):911–918, Sept. 2014. doi: 10.1177/0272989X14545789.
- [77] S. Weber, Y. Li, J. W. Seaman, T. Kakizume, and H. Schmidli. Applying meta-analytic-predictive priors with the R Bayesian evidence synthesis tools. *Journal of Statistical Software*, 100(19):1–32, Nov. 2021. doi: 10.18637/jss.v100.i19.
- [78] N. J. Welton, H. E. Jones, and S. Dias. Bayesian methods for meta-analysis. In E. Lesaffre, G. Baio, and B. Boulangier, editors, *Bayesian Methods in Pharmaceutical Research*, chapter 14, pages 273–300. Chapman & Hall / CRC, 2020. doi: 10.1201/9781315180212.
- [79] I. White, R. M. Turner, A. Karahalios, and G. Salanti. A comparison of arm-based and contrast-based models for network meta-analysis. *Statistics in Medicine*, 38(27):5197–5213, Nov. 2019. doi: 10.1002/sim.8360.
- [80] D. R. Williams, J. E. Rodriguez, and P.-C. Bürkner. Putting variation into variance: Modeling between-study heterogeneity in meta-analysis. *PsyArXiv*, May 2021. doi: 10.31234/osf.io/9vkqy.