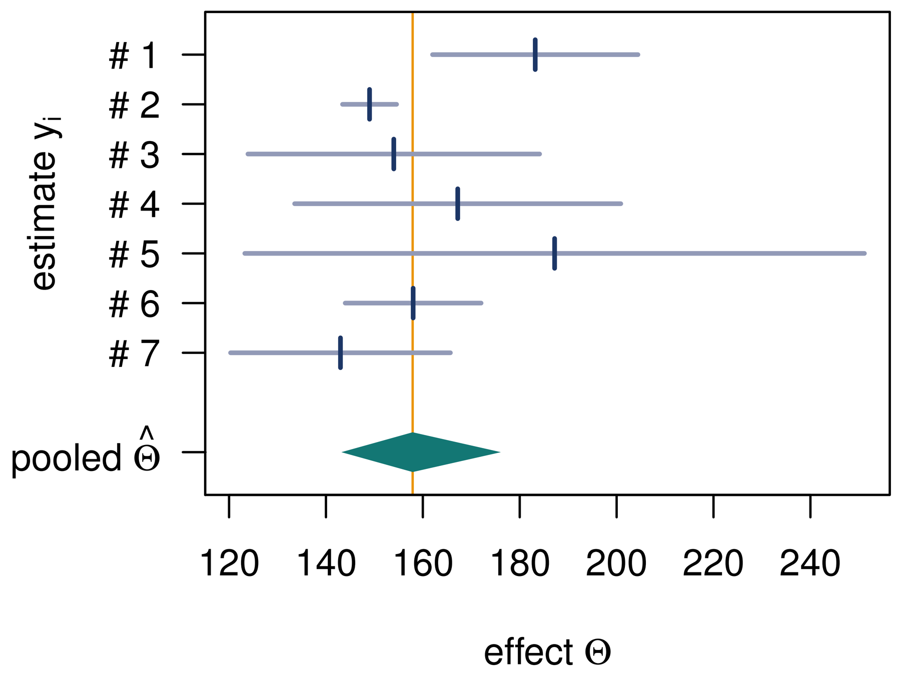
Bayesian random-effects meta-analysis made easy

Christian Röver[1] Simon Wandel[2] Beat Neuenschwander[2] Tim Friede[1]

[1]University Medical Center Göttingen, Germany [2]Novartis Pharma AG, Basel, Switzerland

The random-effects model

The common problem in meta analysis is to combine individual **parameter estimates** and **standard errors** into a pooled one.



It is usually reasonable and necessary to allow or account for **heterogeneity** between the estimates.

Parameters and likelihood

The inference problem essentially presents itself with the following key figures:

data:
- estimates y_i
- standard errors σ_i

parameters:
- true parameter value Θ
- heterogeneity τ

Most commonly, a simple **Normal model** is utilized, which may be stated as

$$y_i \;\sim\; \mathrm{Normal}(\Theta,\, \sigma_i^2 + \tau^2).$$

When we allow for heterogeneity ($\tau > 0$), this is a special case of a **random-effects** model. The likelihood function follows immediately as the sum-of-squares expression

$$p(\vec{y},\vec{\sigma}|\Theta,\tau) \;\propto\; -\tfrac{1}{2}\sum_i \left(\log(\tau^2{+}\sigma_i^2) + \frac{(y_i-\Theta)^2}{\tau^2+\sigma_i^2}\right).$$

Commonly encountered issues

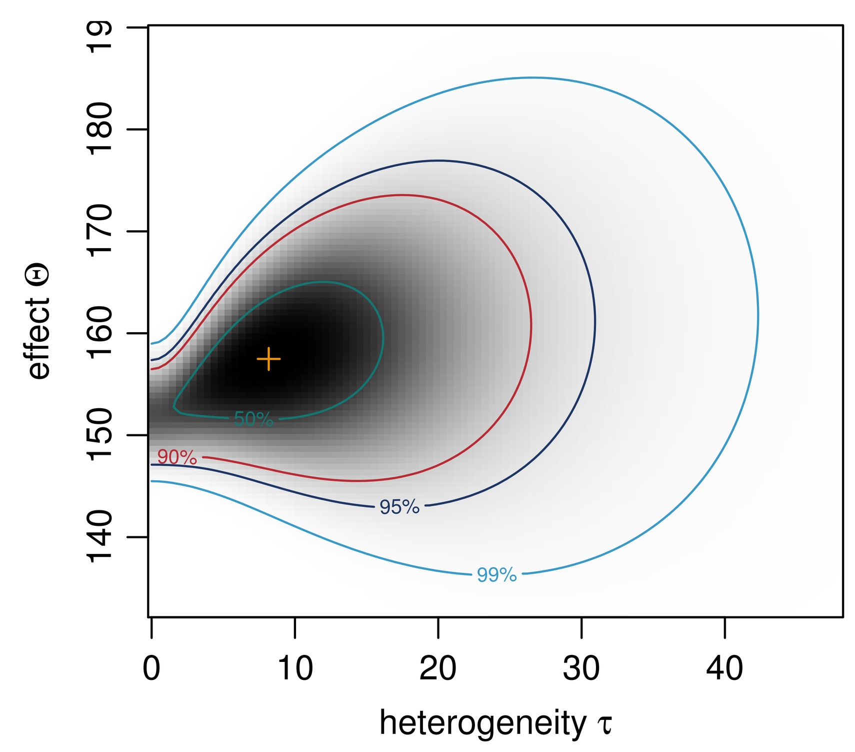
Although the problem is easily stated, a ("frequentist") solution is far from obvious. The nuisance parameter τ is commonly dealt with by deriving a **plug-in**-estimate, on which the following analysis is conditioned. One may **test** for zero heterogeneity, although such tests commonly have little power, and a plethora of **heterogeneity estimators** exists for τ, which sometimes may yield counter-intuitive results (e.g. zero estimates or confidence bounds).

The Bayesian approach

A Bayesian solution on the other hand is rather straightforward once the **prior** for the unknowns is specified. In the following we restrict ourselves to assuming a priori **independence**,

$$p(\Theta,\tau) \;=\; p(\Theta) \times p(\tau),$$

and to a **uniform or normal** prior $p(\Theta)$ for Θ, and an **arbitrary** prior $p(\tau)$ for the heterogeneity.



Posterior for example data (uniform priors).

Marginalization

It turns out that in this two-parameter problem we can **integrate out** the effect parameter Θ, leaving us with a one-dimensional **marginal likelihood** for the heterogeneity τ

$$p(\vec{y},\vec{\sigma}\,|\,\tau) = \int p(\vec{y},\vec{\sigma}\,|\,\Theta,\tau)\; p(\Theta)\; \mathrm{d}\Theta$$
$$\propto -\tfrac{1}{2}\sum_i \left(\log(\tau^2+\sigma_i^2) + \frac{(y_i-\mu_{\Theta|\tau})^2}{\tau^2+\sigma_i^2}\right)$$
$$-\tfrac{1}{2}\log\left(\sum_i \frac{1}{\tau^2+\sigma_i^2}\right)$$

where $\mu_{\Theta|\tau}$ is the *conditional posterior mean* of Θ for given τ:

$$\mu_{\Theta|\tau} \;=\; \mathrm{E}\big[\Theta\,\big|\,\tau,\vec{y},\vec{\sigma}\big] \;=\; \frac{\sum_i \frac{y_i}{\tau^2+\sigma_i^2}}{\sum_i \frac{1}{\tau^2+\sigma_i^2}}.$$

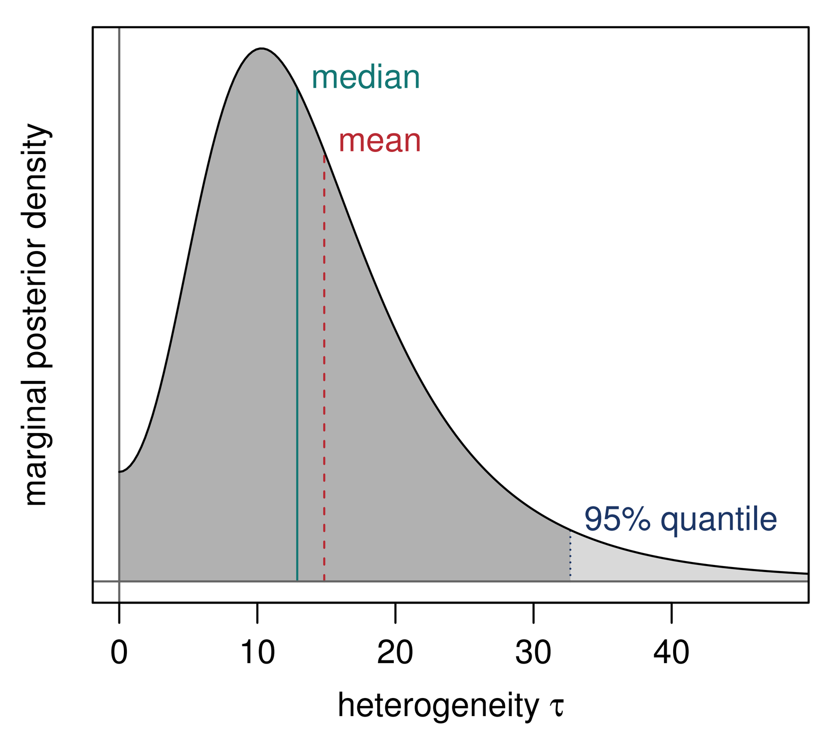
(Integration works similarly for a normal prior $p(\Theta)$).

Inferring the heterogeneity τ

The **marginal posterior density of τ** is simply

$$p(\tau\,|\,\vec{y},\vec{\sigma}) \;\propto\; p(\vec{y},\vec{\sigma}\,|\,\tau) \times p(\tau)$$

Now one may specify an arbitrary prior $p(\tau)$ and use **numerical integration** for the 1-dimensional posterior to compute quantiles, moments, . . .



Marginal posterior for the heterogeneity τ.

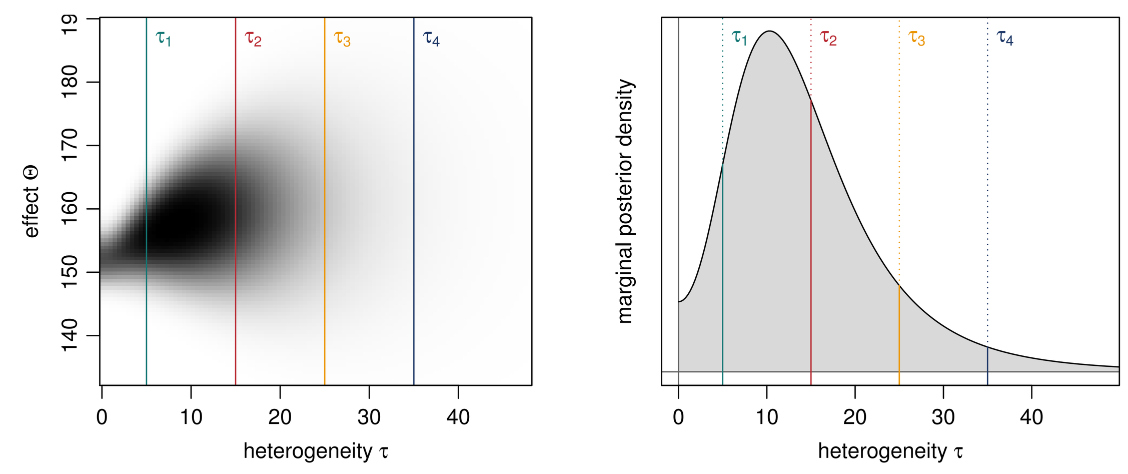
Inferring the effect Θ

The **marginal posterior of Θ** is a **normal mixture**:

$$p(\Theta\,|\,\vec{y},\vec{\sigma}) \;=\; \int \underbrace{p(\Theta\,|\,\tau,\vec{y},\vec{\sigma})}_{\text{Normal}}\; p(\tau\,|\,\vec{y},\vec{\sigma})\; \mathrm{d}\tau$$

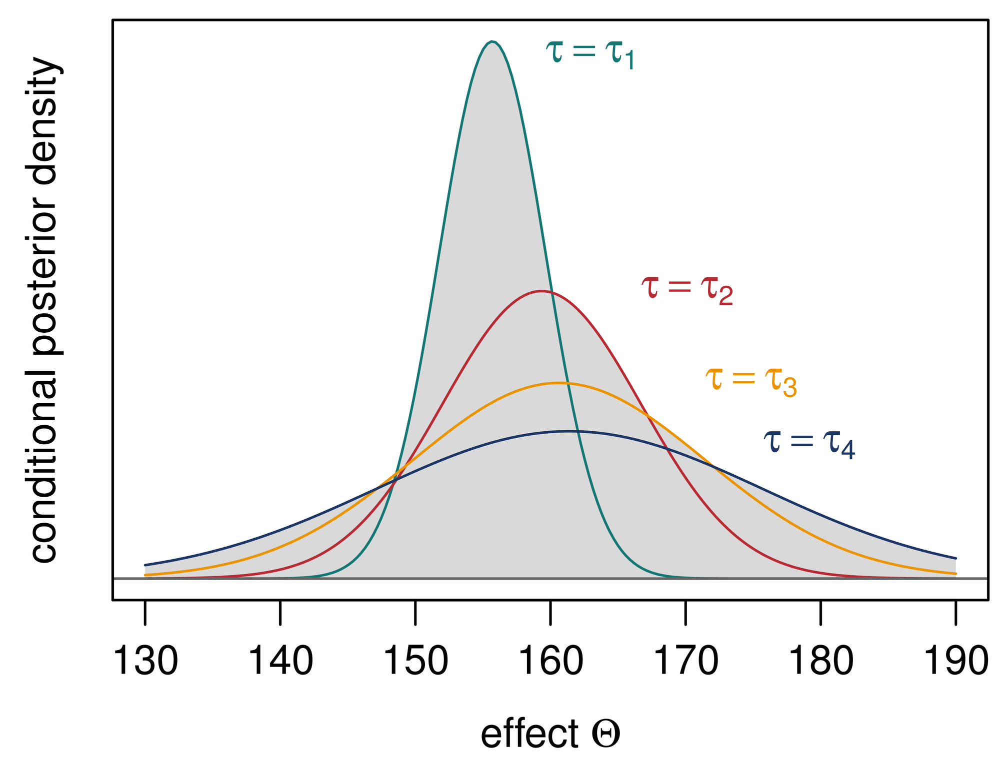
The mixture distribution may then easily be approximated via a discrete grid in τ:

$$p(\Theta\,|\,\vec{y},\vec{\sigma}) \;\approx\; \sum_{j=1}^{m} p(\Theta\,|\,\tau_j,\vec{y},\vec{\sigma})\; w_j.$$

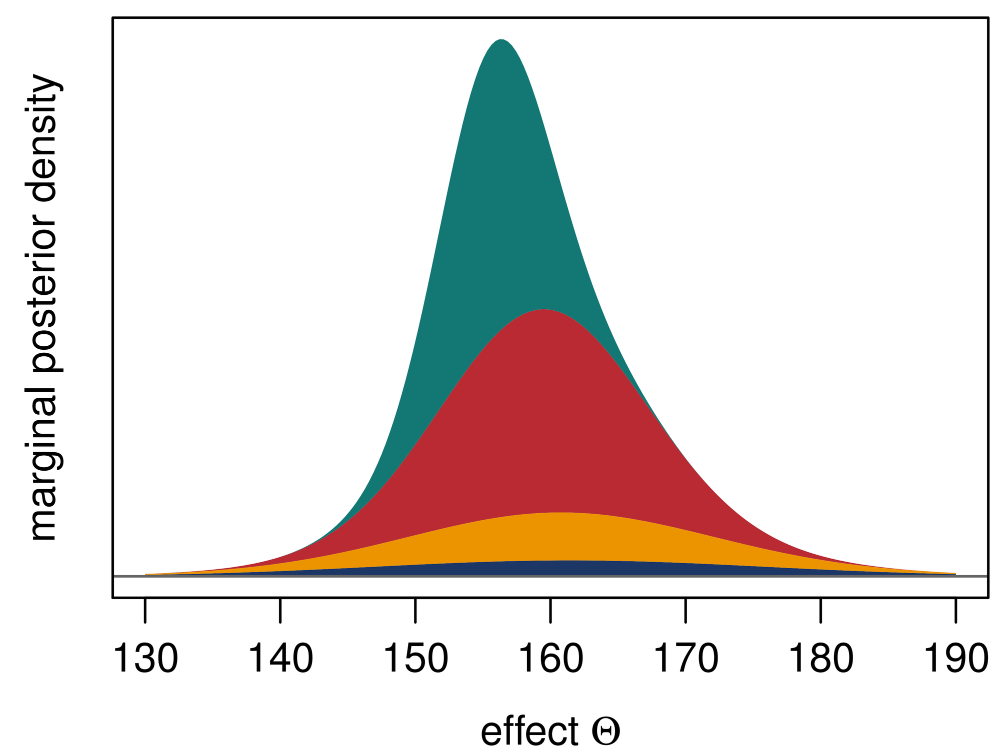


A discrete grid of τ_j values "slicing" the parameter space.

The mixture **weights w_j** are derived by integrating over the heterogeneity's marginal distribution. For a given set of τ_j values the conditional posterior distributions $p(\Theta\,|\,\tau_j,\vec{y},\vec{\sigma})$ again are **Normal**.



The effect's Normal conditional distributions for given τ values.



The effect's marginal posterior distribution as a weighted sum of the Normal conditionals.

Conclusions

For the common task of a random effects meta analysis, the Bayesian solution is **easily implemented**. Computations reduce to **seconds of CPU time**, the resulting estimates and credibility levels are **accurate**. The grid approximation may be set up so that a pre-specified accuracy is guaranteed. The use of (almost) **arbitrary priors** allows for **quick sensitivity checks**. Furthermore, calculation of the **prediction interval** for $\theta^\star$, the effect yet to be observed in a new study, is straightforward. The methods shown here are implemented in the `bmeta` **R package** which is to appear on CRAN soon.