# Implementing Bayesian random-effects meta-analysis

Christian Röver[1], Beat Neuenschwander[2], Tim Friede[1]

[1]Department of Medical Statistics
University Medical Center Göttingen

[2]Novartis Pharma AG,
Basel, Switzerland

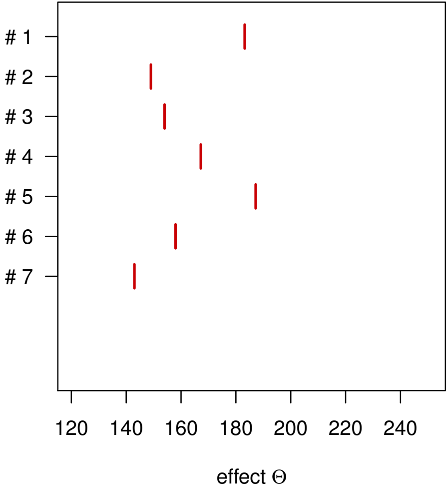
September 10, 2014

UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG

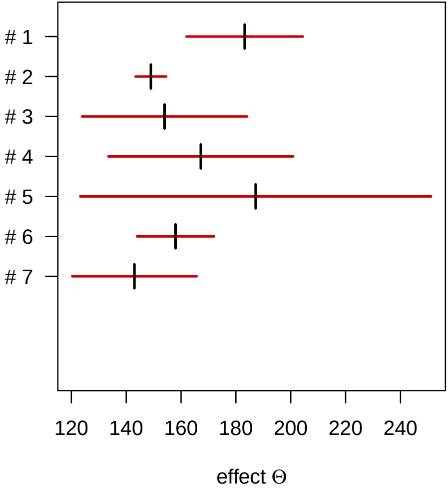
# Overview

- Meta analysis
  - the random-effects model
  - the common approach
- The Bayesian approach
  - prior, likelihood
  - marginal likelihood
  - posterior distribution
- Application
  - examples

# Meta analysis
## Context

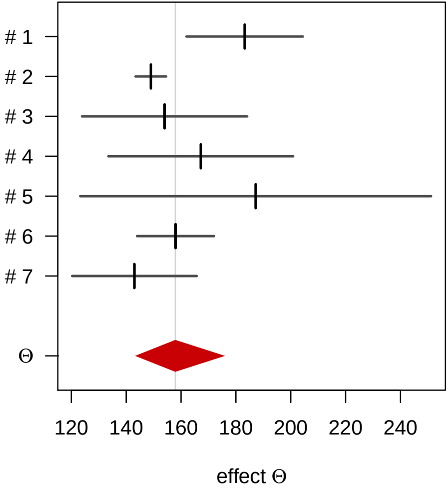



- have:
  - estimates $y_i$
  - standard errors $\sigma_i$
- want:
  - combined estimate $\hat{\Theta}$



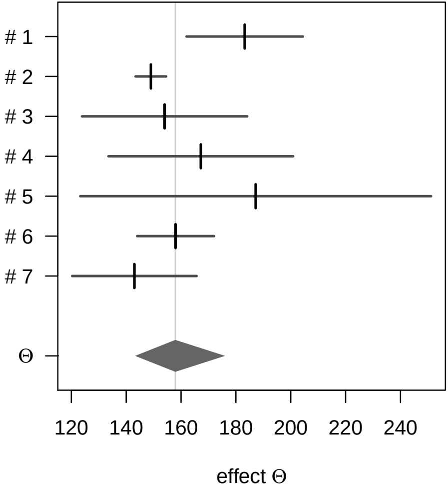

- have:
  - estimates $y_i$
  - standard errors $\sigma_i$
- want:
  - combined estimate $\hat{\Theta}$

# Meta analysis
## Context




- have:
  - estimates $y_i$
  - standard errors $\sigma_i$
- want:
  - combined estimate $\hat{\Theta}$

# Meta analysis
## Context




- have:
  - estimates $y_i$
  - standard errors $\sigma_i$
- want:
  - combined estimate $\hat{\Theta}$

# Meta analysis
## The random effects model

- assume[1,2]:

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

[1] L. V. Hedges, I. Olkin. *Statistical methods for meta-analysis*. Academic Press, 1985.

[2] J. Hartung, G. Knapp, B. K. Sinha. *Statistical meta-analysis with applications*. Wiley, 2008.

# Meta analysis
## The random effects model

- assume[1,2]:

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

- ingredients:

*Data*:

- estimates $y_i$
- standard errors $\sigma_i$

*Parameters*:

- true parameter value $\Theta$
- heterogeneity $\tau$

[1] L. V. Hedges, I. Olkin. *Statistical methods for meta-analysis*. Academic Press, 1985.
[2] J. Hartung, G. Knapp, B. K. Sinha. *Statistical meta-analysis with applications*. Wiley, 2008.

- assume[1,2]:

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

- ingredients:

*Data*:

- estimates $y_i$
- standard errors $\sigma_i$

*Parameters*:

- true parameter value $\Theta$
- heterogeneity $\tau$

[1] L. V. Hedges, I. Olkin. *Statistical methods for meta-analysis*. Academic Press, 1985.
[2] J. Hartung, G. Knapp, B. K. Sinha. *Statistical meta-analysis with applications*. Wiley, 2008.

- assume[1,2]:

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

- ingredients:

*Data*:

- estimates $y_i$
- standard errors $\sigma_i$

*Parameters*:

- true parameter value $\Theta$
- heterogeneity $\tau$

[1] L. V. Hedges, I. Olkin. *Statistical methods for meta-analysis*. Academic Press, 1985.
[2] J. Hartung, G. Knapp, B. K. Sinha. *Statistical meta-analysis with applications*. Wiley, 2008.

# Meta analysis

## The random effects model

- assume[1,2]:

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

- ingredients:

*Data*:

- estimates $y_i$
- standard errors $\sigma_i$

*Parameters*:

- true parameter value $\Theta$
- heterogeneity $\tau$

[1] L. V. Hedges, I. Olkin. *Statistical methods for meta-analysis*. Academic Press, 1985.

[2] J. Hartung, G. Knapp, B. K. Sinha. *Statistical meta-analysis with applications*. Wiley, 2008.

# Meta analysis

## The random effects model

- assume[1,2]:

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

- ingredients:

*Data*:

- estimates $y_i$
- standard errors $\sigma_i$

*Parameters*:

- true parameter value $\Theta$
- heterogeneity $\tau$

- $\Theta \in \mathbb{R}$ of primary interest
- $\tau \in \mathbb{R}^+$ nuisance parameter: account for (potential) incompatibility

[1] L. V. Hedges, I. Olkin. *Statistical methods for meta-analysis*. Academic Press, 1985.

[2] J. Hartung, G. Knapp, B. K. Sinha. *Statistical meta-analysis with applications*. Wiley, 2008.

# Meta analysis

## Common approach to inference

- test for $\tau = 0$ vs. $\tau > 0$ (fixed vs. random effects)
- derive estimate $\hat{\tau}$
- derive estimate for $\Theta$ *conditional on* $\hat{\tau}$ *being actual heterogeneity* (plug-in estimate)

- test for $\tau = 0$ vs. $\tau > 0$ (fixed vs. random effects)
- derive estimate $\hat{\tau}$
- derive estimate for $\Theta$ *conditional on* $\hat{\tau}$ *being actual heterogeneity* (plug-in estimate)

- Problems:
  - significance tests have low power
  - $\tau = 0$ hypothesis questionable
  - how to estimate $\tau$?
    numerous approaches available,
    questionable properties, especially for (near-) zero $\tau$
  - conditioning on *fixed* $\tau$ value only makes sense in case of great accuracy
  - uncertainty in $\tau$ usually not accounted for

# Meta analysis

## The Bayesian approach

- Bayesian approach [3]
- consideration of prior information
- consideration of uncertainty
- straightforward interpretation
- computationally more expensive, usually done via stochastic integration (MCMC, BUGS)[4]

[3] A. J. Sutton, K. R. Abrams. *Bayesian methods in meta-analysis and evidence synthesis*. Statistical Methods in Medical Research, 10(4):277, 2001.

[4] T. C. Smith, D. J. Spiegelhalter, A. Thomas. *Bayesian approaches to random-effects meta-analysis: A comparative study*. Statistics in Medicine, 14(24):2685, 1995.

# The Bayesian approach

## Prior, likelihood

- likelihood follows from assumptions:

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

# The Bayesian approach

## Prior, likelihood

- likelihood follows from assumptions:

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

- assume a priori independence:

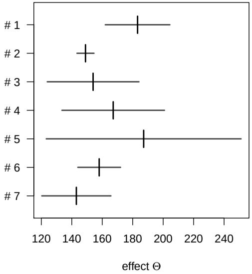
$$p(\Theta, \tau) \;=\; p(\Theta) \times p(\tau)$$

- $p(\Theta)$ uniform or normal
- $p(\tau)$ arbitrary (uniform or informative)[5]

[5] A. Gelman. *Prior distributions for variance parameters in hierarchical models*. Bayesian Analysis, 1(3):515, 2006.

# Example

## Cochran (1954) data[6]

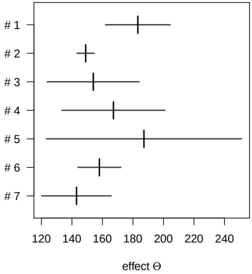



- data: 7 estimates
  and standard errors
- assume:
  uniform priors $p(\Theta)$, $p(\tau)$

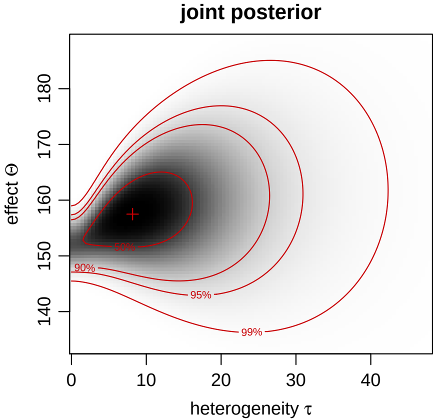
[6] W. G. Cochran. *The combination of estimates from different experiments*. Biometrics, 10(1):101, 1954.

# Example

## Cochran (1954) data[6]




- data: 7 estimates and standard errors
- assume: uniform priors $p(\Theta)$, $p(\tau)$




(uniform prior $\Rightarrow$ posterior $=$ likelihood)

[6] W. G. Cochran. *The combination of estimates from different experiments*. Biometrics, 10(1):101, 1954.

# The Bayesian approach
## Marginal likelihood

- interested in $\tau$,
  marginalize likelihood over $\Theta$:

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

# The Bayesian approach
## Marginal likelihood

- interested in $\tau$,
  marginalize likelihood over $\Theta$ (for uniform $p(\Theta)$):

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

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

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

# The Bayesian approach

## Marginal likelihood

- interested in $\tau$,
  marginalize likelihood over $\Theta$ (for uniform $p(\Theta)$):

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

$$\propto\; -\tfrac{1}{2} \sum_i \left( \log(\tau^2 + \sigma_i^2) + \frac{(y_i - \mu_{\Theta|\tau})^2}{\tau^2 + \sigma_i^2} \right) - \tfrac{1}{2} \log\left( \sum_i \frac{1}{\tau^2 + \sigma_i^2} \right)$$

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

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

# The Bayesian approach

## Marginal likelihood

- interested in $\tau$,
  marginalize likelihood over $\Theta$ (for uniform $p(\Theta)$):

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

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

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

- similar for normal prior $p(\Theta)$

# The Bayesian approach

## Inferring $\tau$

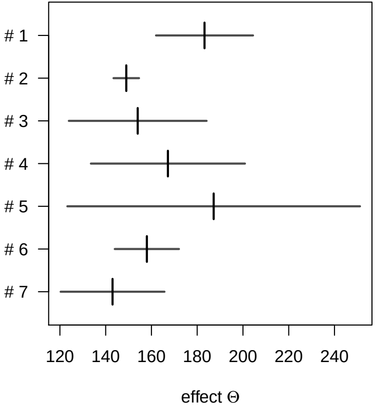
- posterior distribution of $\tau$ simply

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

- specify arbitrary prior $p(\tau)$
- use numerical integration for 1D posterior
- compute quantiles, moments, ...

# Example

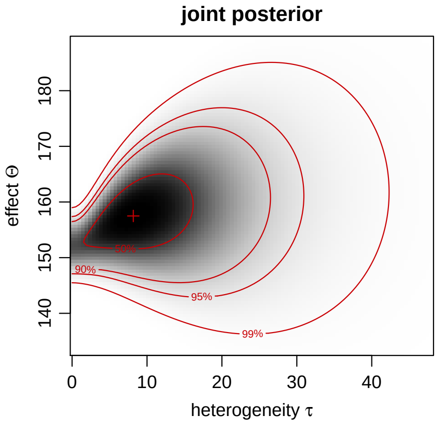
## Cochran (1954) data[7]




- data: 7 estimates and standard errors
- assume: random-effects model, uniform priors $p(\Theta)$, $p(\tau)$

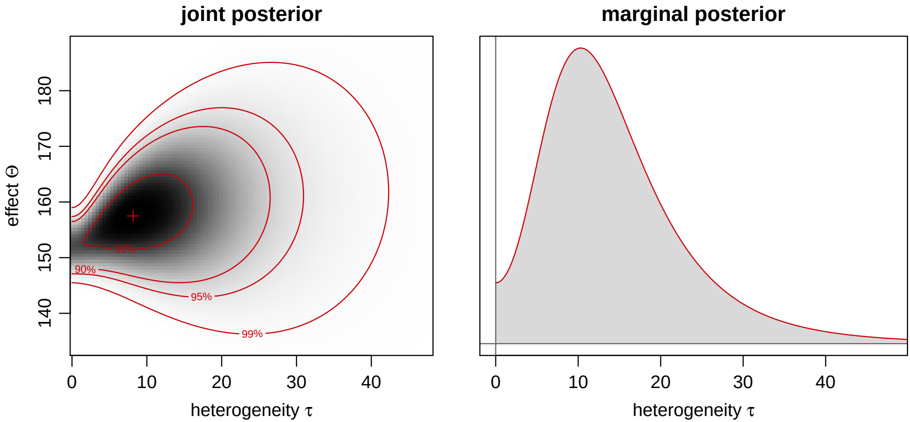
[7] W. G. Cochran. *The combination of estimates from different experiments*. Biometrics, 10(1):101, 1954.

# Example
## Cochran (1954) data



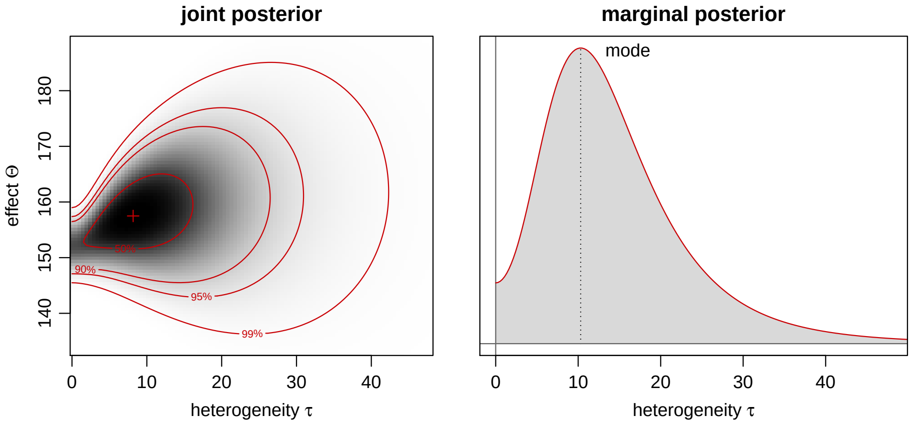
# Example

## Cochran (1954) data



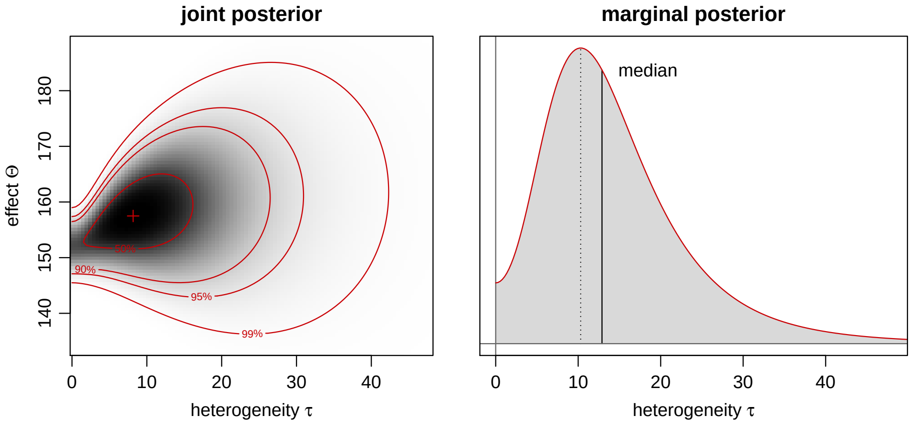
# Example

## Cochran (1954) data

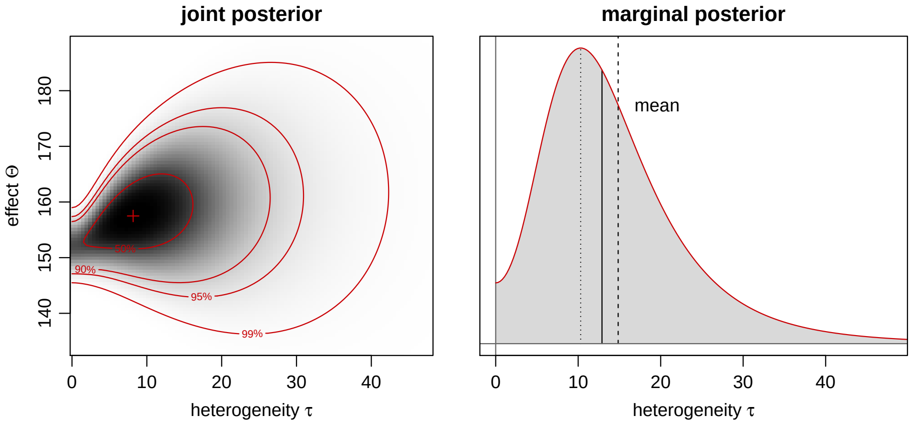


# Example

## Cochran (1954) data

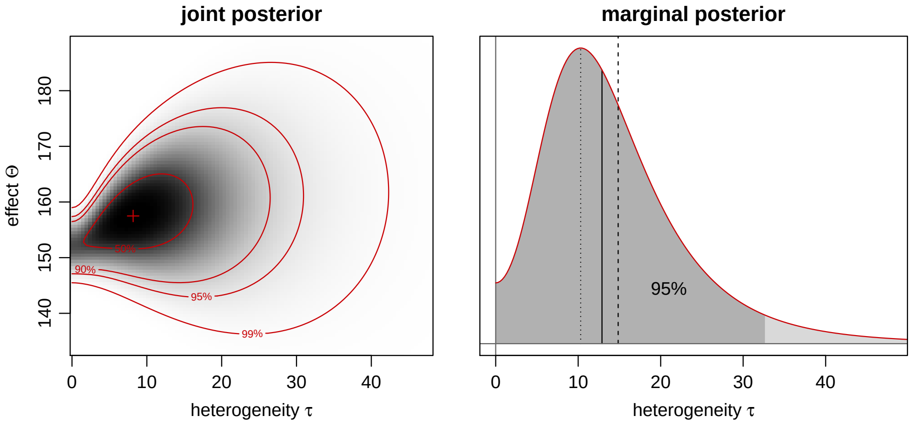


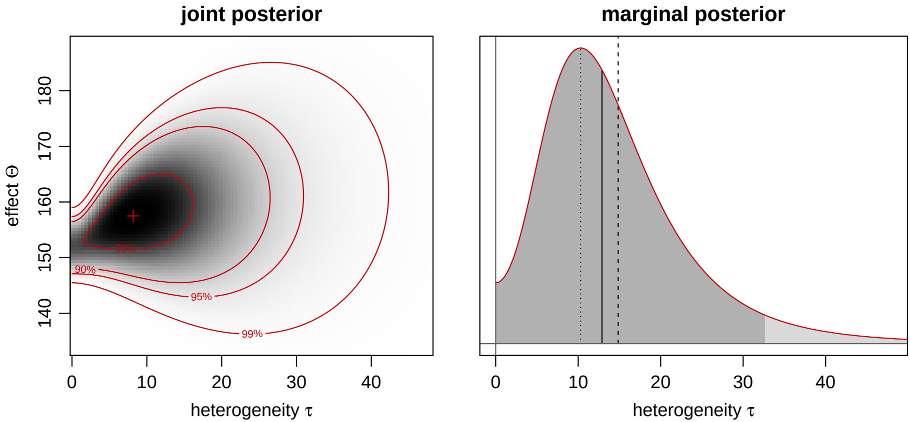
# Example
## Cochran (1954) data



# Example

## Cochran (1954) data

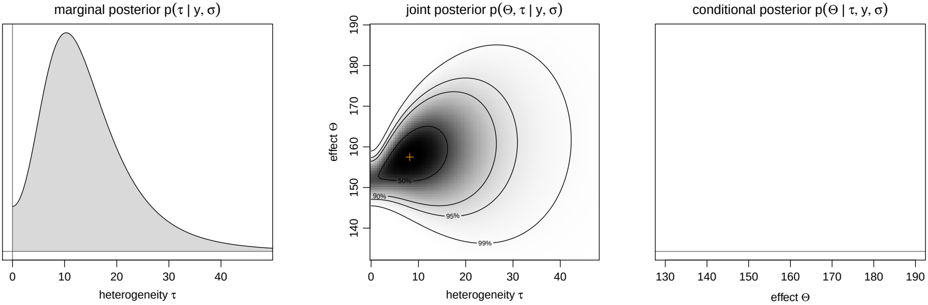




joint posterior
effect Θ
180
170
160
150
140
50%
90%
95%
99%
0
10
20
30
40
heterogeneity τ
marginal posterior
0
10
20
30
40
heterogeneity τ

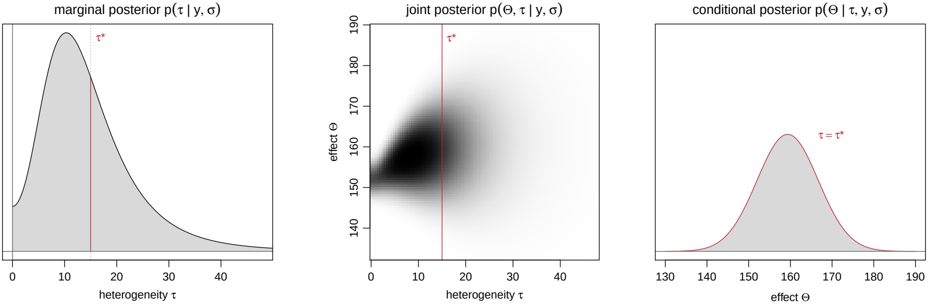
# Inferring Θ

- Note: fixing $\tau$ yields a *normal* conditional posterior $p\left(\Theta \,\middle|\, \tau, \vec{y}, \vec{\sigma}\right)$



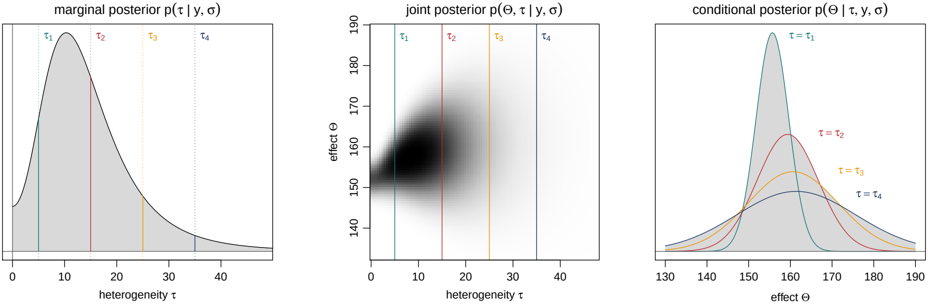
# Inferring Θ

- Note: fixing $\tau$ yields a *normal* conditional posterior $p\left(\Theta \,\middle|\, \tau, \vec{y}, \vec{\sigma}\right)$



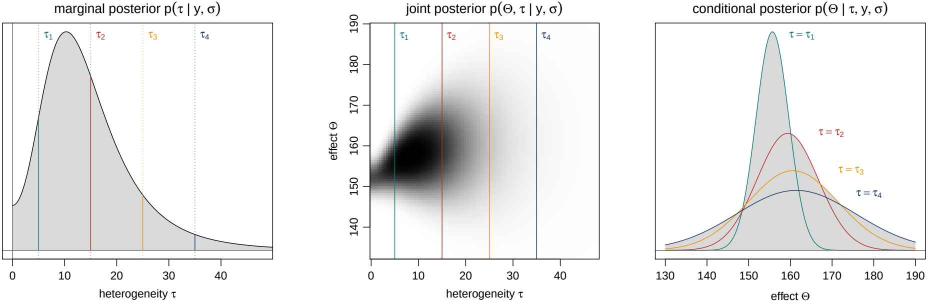
# Inferring Θ

- Note: fixing $\tau$ yields a *normal* conditional posterior $p\left(\Theta \,\middle|\, \tau, \vec{y}, \vec{\sigma}\right)$



# Inferring Θ

- Note: fixing $\tau$ yields a *normal* conditional posterior $p\left(\Theta \,\middle|\, \tau, \vec{y}, \vec{\sigma}\right)$




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

$$p\left(\Theta \,\middle|\, \vec{y}, \vec{\sigma}\right) \;=\; \int p\left(\Theta \,\middle|\, \tau, \vec{y}, \vec{\sigma}\right) \, p\left(\tau \,\middle|\, \vec{y}, \vec{\sigma}\right) \, \mathrm{d}\tau$$

- weights given by marginal posterior of $\tau$ ... $\rightarrow$ easy approximation

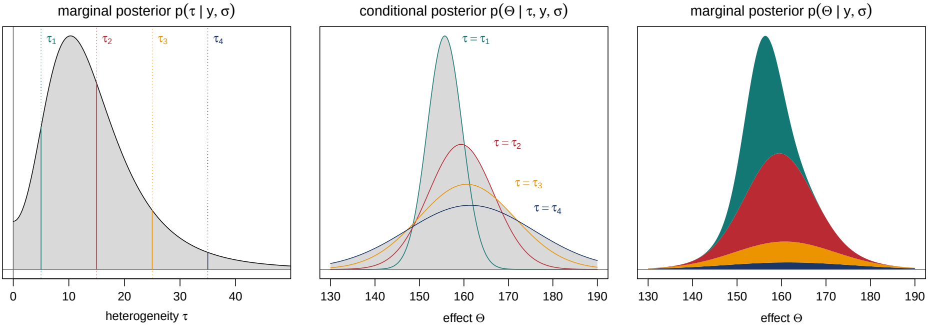
# Inferring Θ

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

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

$$\approx\; \sum_j p(\Theta \,|\, \tau_j, \vec{y}, \vec{\sigma})\; w_j$$

(*weights* $w_j$ via integration over marginal $p(\tau \,|\, \vec{y}, \vec{\sigma})$)



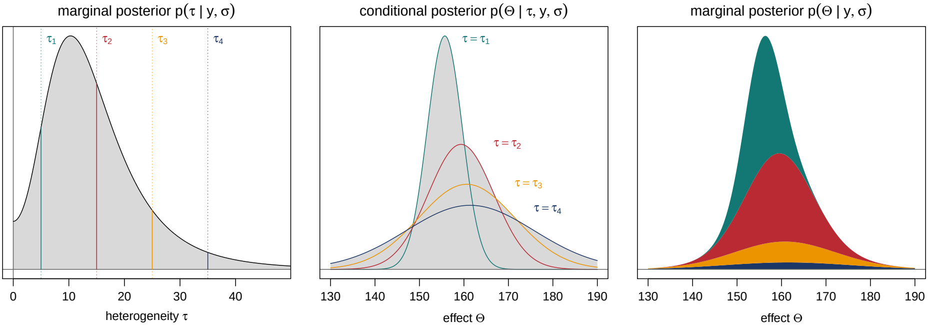
# Inferring Θ

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

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

$$\approx\; \sum_j p(\Theta \,|\, \tau_j, \vec{y}, \vec{\sigma})\; {\color{red} w_j}$$

(*weights* $w_j$ via integration over marginal $p(\tau \,|\, \vec{y}, \vec{\sigma})$)

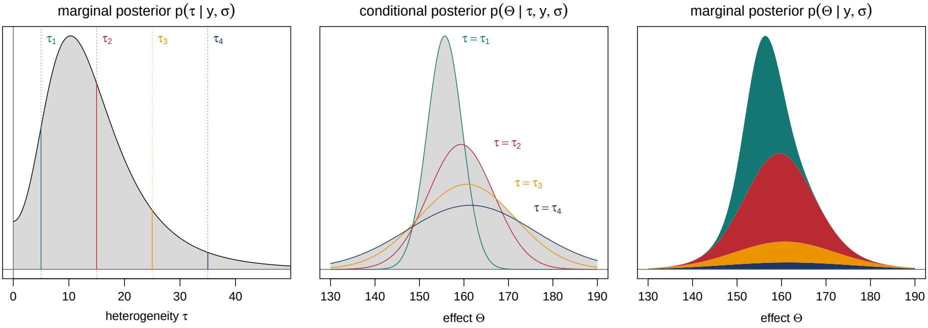


# Inferring Θ

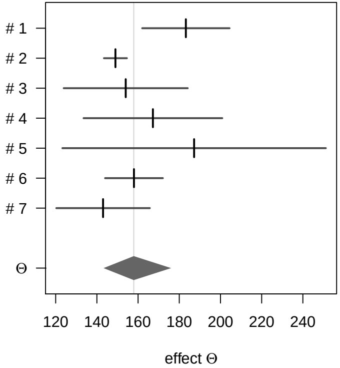
- marginal posterior of Θ is a *normal mixture*:

$$
\begin{aligned}
p(\Theta \,|\, \vec{y}, \vec{\sigma}) &= \int p(\Theta \,|\, \tau, \vec{y}, \vec{\sigma}) \; p(\tau \,|\, \vec{y}, \vec{\sigma}) \; \mathrm{d}\tau \\
&\approx \sum_j p(\Theta \,|\, \tau_j, \vec{y}, \vec{\sigma}) \; w_j
\end{aligned}
$$

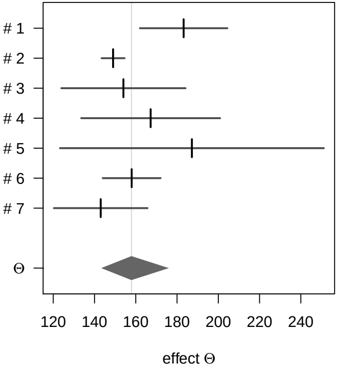
(*weights* $w_j$ via integration over marginal $p(\tau \,|\, \vec{y}, \vec{\sigma})$)

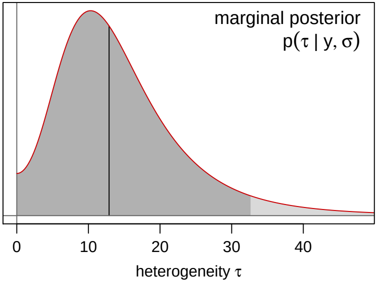


# Example
## Cochran (1954) data



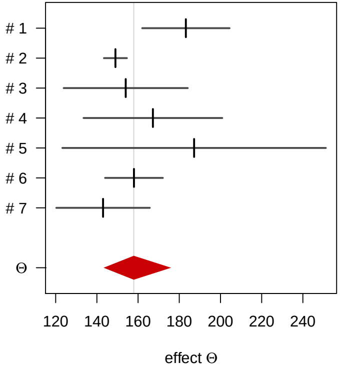
# Example
## Cochran (1954) data

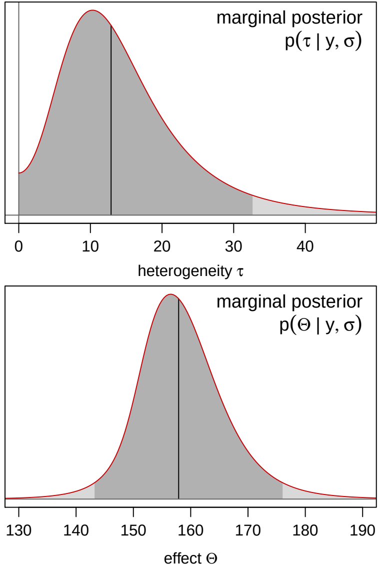


# Example

## Cochran (1954) data



# Implementation

## R package under development

```
> cochran01 <- bmeta(Cochran1954[,"mean"], sqrt(Cochran1954[,"se2"]))
> cochran02 <- bmeta(Cochran1954[,"mean"], sqrt(Cochran1954[,"se2"]),
+                    mu.prior.mean=150, mu.prior.sd=100,
+                    tau.prior=function(x){return(dexp(x, rate=0.05))})
>
> cochran01$summary
                  tau          mu   mu.pred
mode        10.303255 156.504954 154.16345
median      12.888735 157.896520 157.33321
mean        14.844457 158.547999 158.54800
sd           9.950631   8.358115  19.70028
95% lower    0.000000 143.180913 119.77459
95% upper   32.665117 176.106158 200.12309
>
> # compute posterior quantiles:
> cochran01$qposterior(mu.p=c(0.005, 0.995))
[1] 135.0429 187.3122
>
> # plot posterior density:
> x <- seq(from=130, to=190, length=100)
> plot(x, cochran02$dposterior(mu=x), type="l")
> lines(x, cochran01$dposterior(mu=x))
```

# Conclusions

- coherent inference, exact also for small number of estimates *k*
- applicable for wide range of effect measures
- flexible consideration of prior information (“default” options?)
- consideration of uncertainty
- straightforward interpretation
- simple implementation, fast computation
- grid approximation: accuracy under control
- no MCMCing necessary
  (implementation, tuning, diagnostics, post-processing,. . . )

- R package `bmeta` under construction (examples included)
- working on performance comparison (MSE, bias, prior choice,. . . )


- ACKNOWLEDGMENTS:
  partially funded by the EU through InSPiRe (FP HEALTH 2013 - 602144); thanks to Simon Wandel

+++ additional slides +++

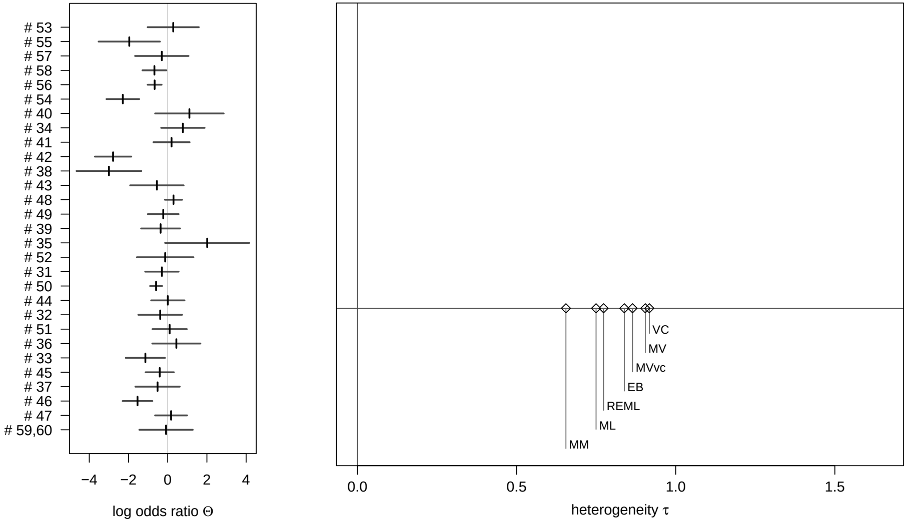
# Example

## Sidik / Jonkman data[8]

- Sidik / Jonkman investigated a range of heterogeneity estimators:
  - method of moments (MM)
  - variance component (VC)
  - maximum likelihood (ML)
  - restricted ML (REML)
  - empirical Bayes (EB)
  - model error variance (MV)
  - variation of MV (MVvc)
- used example for illustration: 29 log odds ratios

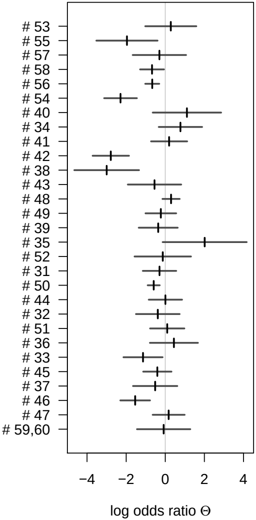
[8] K. Sidik and J. N. Jonkman. *A comparison of heterogeneity variance estimators in combining results of studies*. Statistics in Medicine, 26(9):1964, 2007.

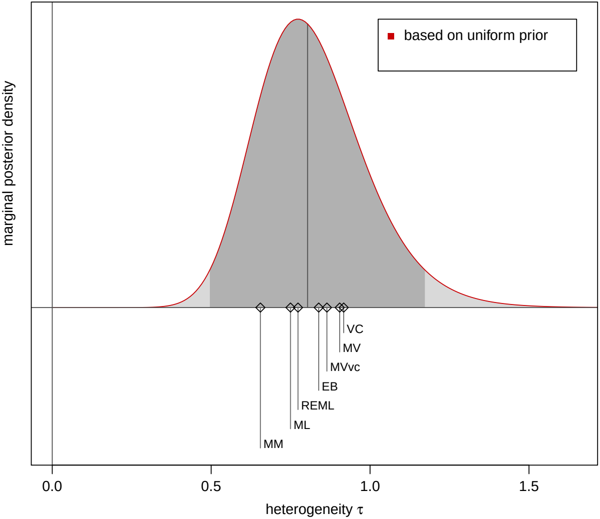
# Example
## Sidik / Jonkman data



# Example

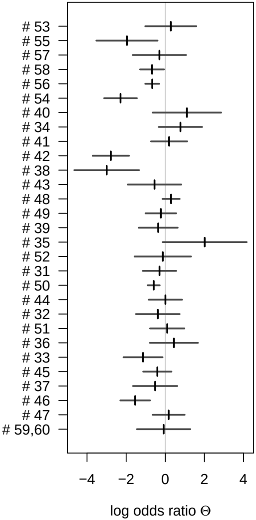
## Sidik / Jonkman data

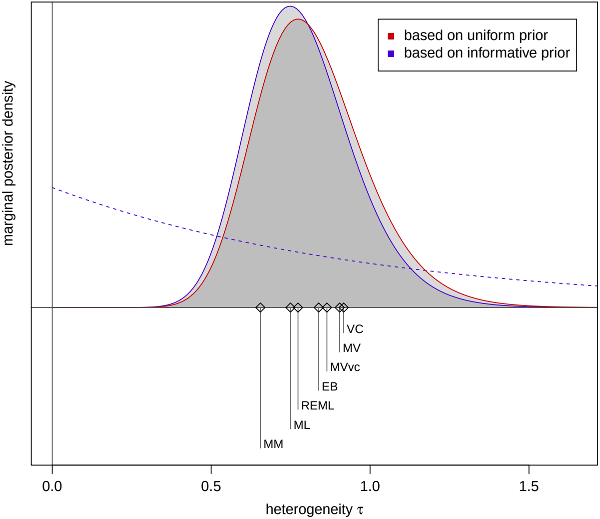


# Example

## Sidik / Jonkman data

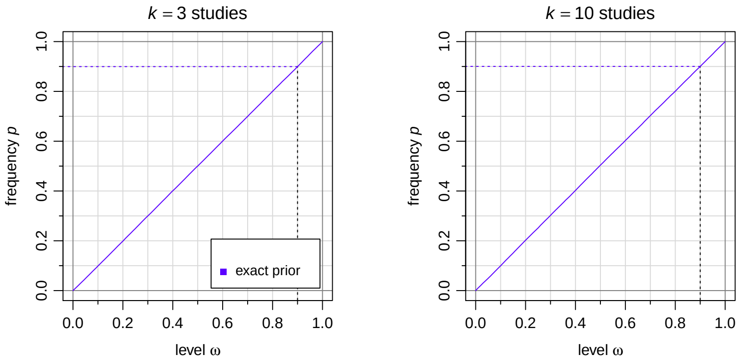


# Example

## Calibration: simulations

- posterior calibration: do upper limits on $\tau$ cover true values?
- simulate: generate parameter values, analyze data using matching prior or uniform prior

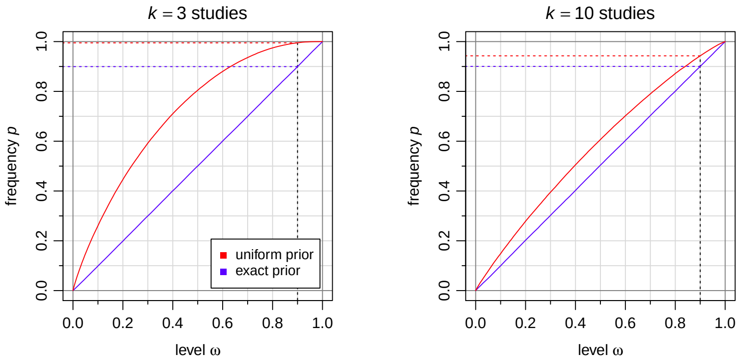



- calibration exact

# Example

## Calibration: simulations

- posterior calibration: do upper limits on $\tau$ cover true values?
- simulate: generate parameter values, analyze data using matching prior or uniform prior




- calibration exact, conservative for (improper) uniform prior