Prior distributions from meta-analytic predictions

Christian Röver

UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG

Department of Medical Statistics,
University Medical Center Göttingen,
Göttingen, Germany

IBS-DR workshop
"*Hierarchical models in preclinical research*"
Göttingen,
December 6, 2024

MAP priors

Outline

- the Bayesian hierarchical model
- estimating different quantities
- upstream priors and downstream likelihoods: the MAP prior
- investigating (prior) informativeness: ESS etc.
- being sceptical/conservative: anticipating prior-data conflict / robustification
- application areas:
  - predicting parameters
  - predicting data
  - predicting heterogeneity

Meta-analysis modeling

Normal approximation for “effect measures”

- single study’s outcome, often: **estimate** $\pm$ **standard error**
- **normal approximation** (“Wald” CI) often appropriate (“large” sample size within study)
- (standard errors are assumed **known**, fixed!)
- sometimes **transformations** are used—
  - improved normal approximation
  - (later: implications for “between-study” modeling)
- examples:
  - means, mean differences, standardized mean differences
  - (log-) proportions, (log-) odds
  - (log-) risk ratios, (log-) odds ratios
  - (log-) rate ratios, (log-) hazard ratios
  - (Fisher-z transformed) correlation coefficients
  - . . .

Meta-analysis modeling

Normal random effects

- variability (**heterogeneity**) between studies commonly anticipated
  - to reflect differing study characteristics
  - to implement stratification by study
  - to avoid overoptimism / naïve pooling
- especially for few studies (small k), heterogeneity is hard to detect (tests have low power) [1]
- empirically: heterogeneity commonly present [2]
- some amount of between-study **heterogeneity** should be anticipated [3] ($\rightarrow$ random-effects model)

[1] R. J. Hardy, S. G. Thompson. Detecting and describing heterogeneity in meta-analysis. *Statistics in Medicine*, **17**(8): 841–856, 1998.

[2] E. Kontopantelis, D. A. Springate, D. Reeves. A re-analysis of the Cochrane Library data: The dangers of unobserved heterogeneity in meta-analyses. *PLoS ONE*, **8**(7):e69930, 2013.

[3] J. P. T. Higgins. Commentary: Heterogeneity in meta-analysis should be expected and appropriately quantified. *International Journal of Epidemiology*, **37**(5):1158–1160, 2008.

Meta-analysis modeling

The generic normal-normal hierarchical model (NNHM)

- the meta-analysis **data** set:
  - effect estimates $y_i \quad (i = 1, \ldots, k)$
  - (standard errors s_i)
- model (likelihood):

$$\begin{aligned} y_i|\theta_i &\sim \text{Normal}(\theta_i, s_i^2) \\ \theta_i|\mu, \tau &\sim \text{Normal}(\mu, \tau^2) \end{aligned}$$

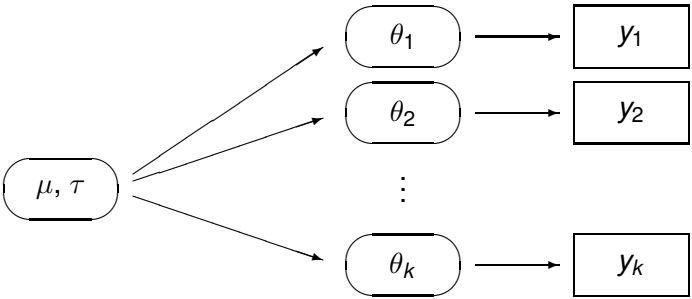
  or (marginally):

$$y_i|\mu, \tau \sim \text{Normal}(\mu, \tau^2 + s_i^2)$$

- parameters:
  - "**study-specific effects**" θ_i
  - "**overall mean effect**" μ
  - "**heterogeneity**" $\tau \geq 0$
- for $\tau = 0$, reduces to common-effect model $\quad (\tau = 0 \quad \Rightarrow \quad \theta_1 = \cdots = \theta_k = \mu)$

Meta-analysis modeling

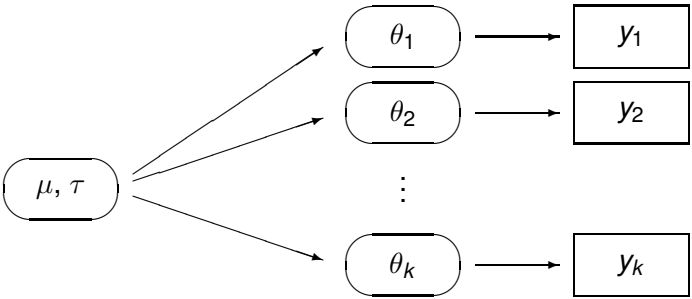
The generic normal-normal hierarchical model (NNHM)



- the NNHM as a *directed acyclic graph (DAG)*:
  overall parameters (μ, τ), study-specific effects (θ_i), data (y_i)

Meta-analysis modeling

The generic normal-normal hierarchical model (NNHM)



- the NNHM as a *directed acyclic graph (DAG)*:
  overall parameters (μ, τ), study-specific effects (θ_i), data (y_i), standard errors (s_i)

Example

11 historical trials (Neuenschwander *et al.*, 2010)

- example: *treatment failure* in transplantation;
  data from 11 “historical” **control groups** (930 patients) to support new trial [4]

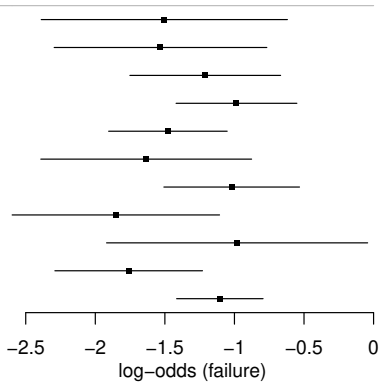
i	events	proportion (%)	odds	log-odds y_i (s_i)
1	6 / 33	18.2	0.222	-1.50 (0.45)
2	8 / 45	17.8	0.216	-1.53 (0.39)
3	17 / 74	23.0	0.298	-1.21 (0.28)
4	28 / 103	27.2	0.373	-0.99 (0.22)
5	26 / 140	18.6	0.228	-1.48 (0.22)
6	8 / 49	16.3	0.195	-1.63 (0.39)
7	22 / 83	26.5	0.361	-1.02 (0.25)
8	8 / 59	13.6	0.157	-1.85 (0.38)
9	6 / 22	27.3	0.375	-0.98 (0.48)
10	16 / 109	14.7	0.172	-1.76 (0.27)
11	53 / 213	24.9	0.331	-1.10 (0.16)

[4] B. Neuenschwander, G. Capkun-Niggli, A. Branson, D. J. Spiegelhalter. Summarizing historical information on controls in clinical trials. *Clinical Trials*, **7**(1):5–18, 2010.

Example

11 historical trials (Neuenschwander *et al.*, 2010)

study	log−odds	95% CI
study01	−1.50	[−2.39, −0.62]
study02	−1.53	[−2.30, −0.77]
study03	−1.21	[−1.75, −0.67]
study04	−0.99	[−1.42, −0.55]
study05	−1.48	[−1.90, −1.05]
study06	−1.63	[−2.39, −0.88]
study07	−1.02	[−1.51, −0.53]
study08	−1.85	[−2.60, −1.11]
study09	−0.98	[−1.92, −0.04]
study10	−1.76	[−2.29, −1.23]
study11	−1.10	[−1.42, −0.79]

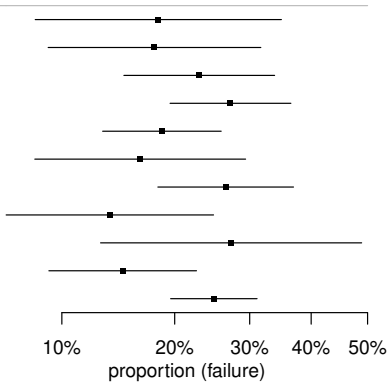


- analysis on log-odds scale

Example

11 historical trials (Neuenschwander *et al.*, 2010)

study	log−odds	95% CI
study01	−1.50	[−2.39, −0.62]
study02	−1.53	[−2.30, −0.77]
study03	−1.21	[−1.75, −0.67]
study04	−0.99	[−1.42, −0.55]
study05	−1.48	[−1.90, −1.05]
study06	−1.63	[−2.39, −0.88]
study07	−1.02	[−1.51, −0.53]
study08	−1.85	[−2.60, −1.11]
study09	−0.98	[−1.92, −0.04]
study10	−1.76	[−2.29, −1.23]
study11	−1.10	[−1.42, −0.79]



- analysis on log-odds scale (proportions $\approx$ 10–30%)

Example

Meta-analysis

- analyze using NNHM
- prior settings
  - (non-informative) uniform effect (μ) prior [5]
  - (weakly informative) half-Normal(1.0) heterogeneity (τ) prior [6]
- `bayesmeta` R package[7]

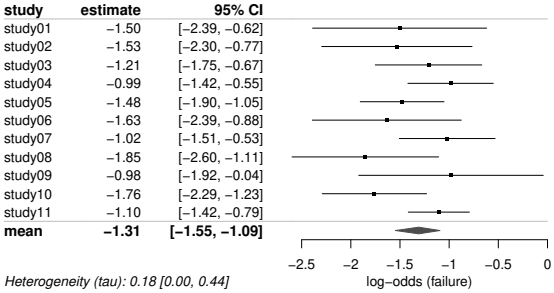
[5] C. Röver. Bayesian random-effects meta-analysis using the `bayesmeta` R package. *Journal of Statistical Software*, **93**(6), 2020.

[6] C. Röver, R. Bender, S. Dias, C. H. Schmid, H. Schmidli, S. Sturtz, S. Weber, T. Friede. On weakly informative prior distributions for the heterogeneity parameter in Bayesian random-effects meta-analysis. *Research Synthesis Methods*, **12**(4):448–474, 2021.

[7] `http://cran.r-project.org/package=bayesmeta`

Example

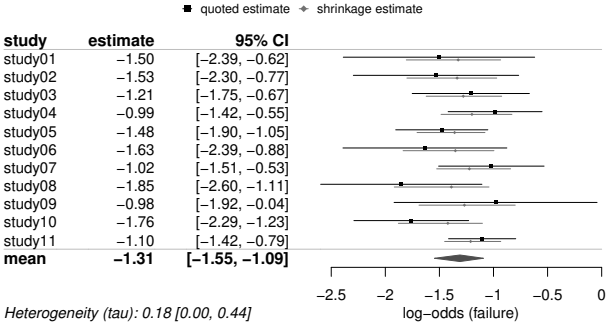
Meta-analysis: overall parameters



- **overall mean** (μ) estimate: -1.31 [-1.55, -1.09]
  heterogeneity (τ): 0.18 [0.00, 0.44]
- proportion (logit$^{-1}(\mu)$): 21% [18%, 25%]

Example

Meta-analysis: shrinkage estimates

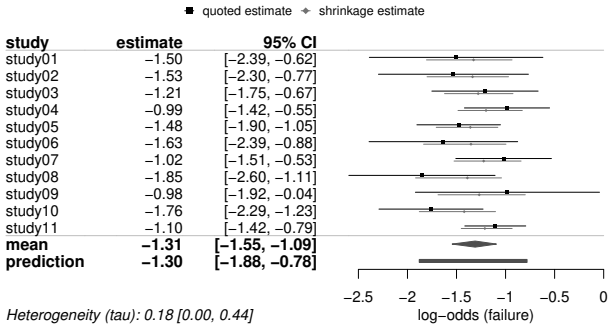


- **shrinkage estimates** (study-specific effects θ_i)
- joint analysis useful to support individual trials [8]

[8] S. Wandel, B. Neuenschwander, C. Röver, T. Friede. Using phase II data for the analysis of phase III studies: an application in rare diseases. *Clinical Trials*, **14**(3):277–285, 2017.

Example

Meta-analysis: prediction

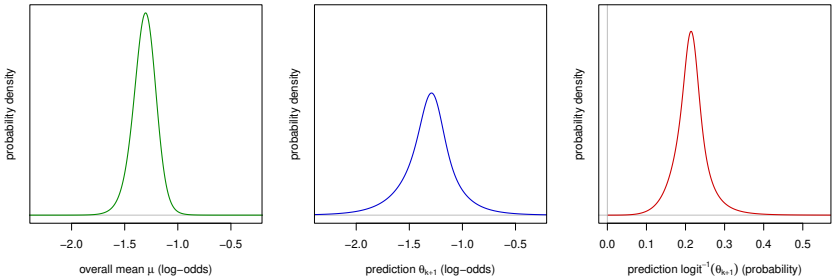


- **prediction**: effect in a new ("future") trial (θ_{k+1})
  ($\text{logit}^{-1}(\theta_{k+1})$: 21% [13%, 31%])
- useful e.g. for trial design [9]

[9] H. Schmidli, B. Neuenschwander, T. Friede. Meta-analytic-predictive use of historical variance data for the design and analysis of clinical trials. *Computational Statistics and Data Analysis*, **113**:100–110, 2017.

Example

Meta-analysis: predicting data

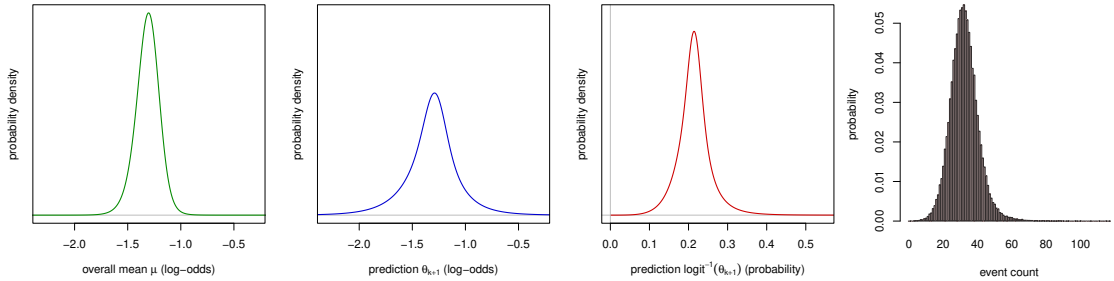


- predictions θ_{k+1} imply predicted probabilities $\text{logit}^{-1}(\theta_{k+1})$

10

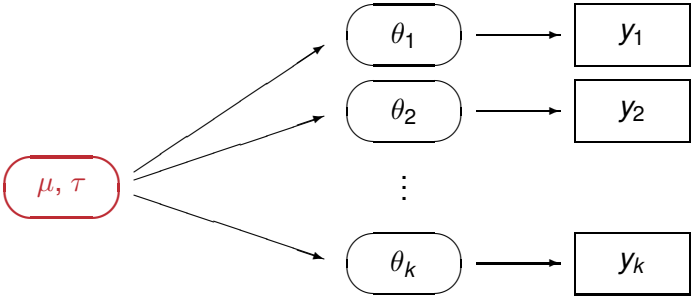
Example

Meta-analysis: predicting data

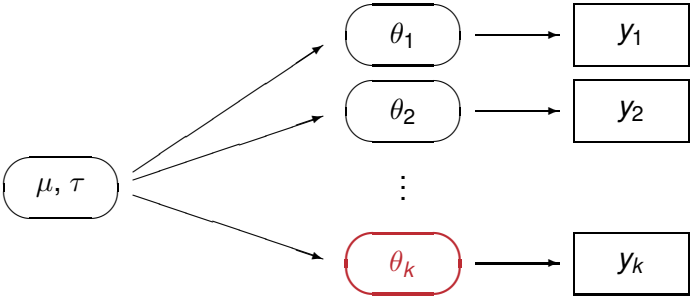


- predictions θ_{k+1} imply predicted probabilities $\text{logit}^{-1}(\theta_{k+1})$
- for a specific trial scenario (e.g. sample size $N_{k+1} = 150$), may derive predicted **data** (event counts)
- useful for checking consistency with historical data [10]

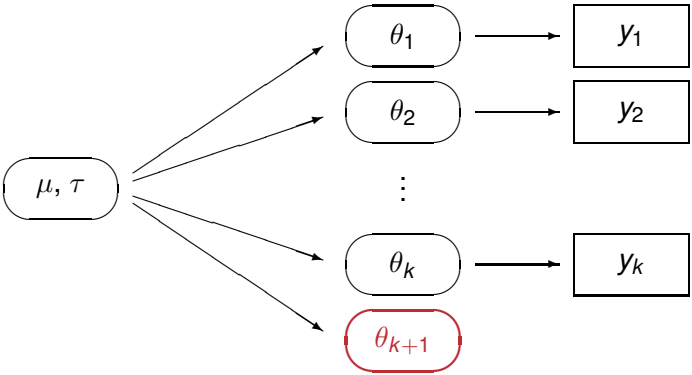
[10] F. M. Kluxen, K. Weber, C. Strupp, S. M. Jensen, L. A. Hothorn, J.-C. Garcin, T. Hofmann. Using historical control data in bioassays for regulatory toxicology. *Regulatory Toxicology and Pharmacology*, **126**:105024, 2021.



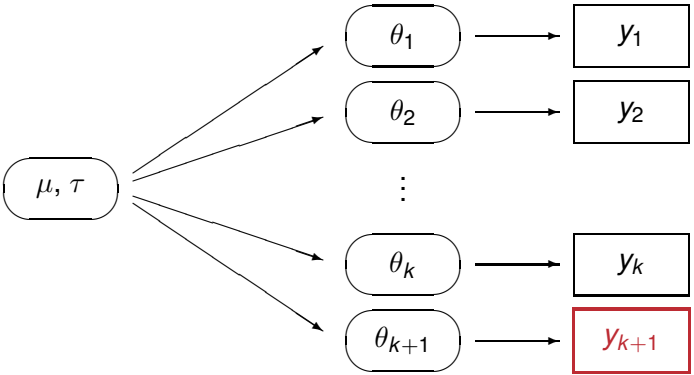
- overall mean (μ, τ)



- overall mean (μ, τ)
- shrinkage estimation (θ_i)



- overall mean (μ, τ)
- shrinkage estimation (θ_i)
- prediction (θ_{k+1})



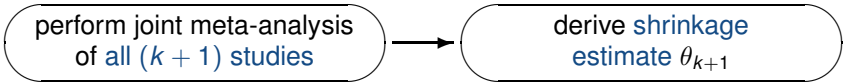
- overall mean (μ, τ)
- shrinkage estimation (θ_i)
- prediction (parameter θ_{k+1})
- prediction (data y_{k+1})

Meta-analysis of historical data

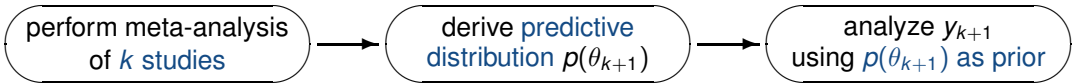
MAP vs. MAC

- suppose k "historical" + 1 current studies are given, interest is in new $(k+1)$th study (data y_{k+1}, effect θ_{k+1})
- two possibilities:

meta-analytic-combined (MAC) approach:



meta-analytic-predictive (MAP) approach:



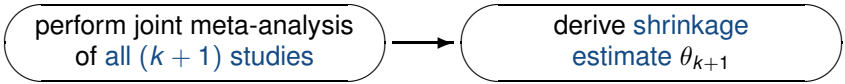
Meta-analysis of historical data

MAP vs. MAC

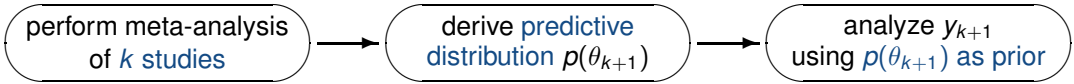
- suppose k "historical" + 1 current studies are given,
  interest is in new $(k+1)$th study (data y_{k+1}, effect θ_{k+1})

- two possibilities:

  meta-analytic-combined (MAC) approach:

  meta-analytic-predictive (MAP) approach:

  (which is preferable?)

- both (MAP and MAC) approaches are **equivalent**: [11]
  - analysis of $(k+1)$th study (estimation of θ_{k+1} using y_{k+1}) based on MAP prior may be interpreted as shrinkage estimation in a joint meta-analysis
  - information of remaining studies contributed to $(k+1)$th shrinkage estimate is expressed through MAP prior (predictive distribution) from k studies

- example of logical consistency of Bayesian methods — posterior may be factored:

$$\underbrace{p(\theta_{k+1} \mid y_1, \ldots, y_{k+1})}_{\text{MAC posterior}} \propto \underbrace{p(y_{k+1} \mid \theta_{k+1})}_{\text{likelihood}} \times \underbrace{p(\theta_{k+1} \mid y_1, \ldots y_k)}_{\text{MAP prior}}$$

[11] H. Schmidli *et al.* Robust meta-analytic-predictive priors in clinical trials with historical control information. *Biometrics*, **70**(4):1023–1032, 2014.

Example

Historical and current data

- *new* trial yields $x_c = 29$ events in a control group of size $N_c = 150$ ($\frac{x_c}{N_c} = 0.19$)
- new study consistent with historical studies?

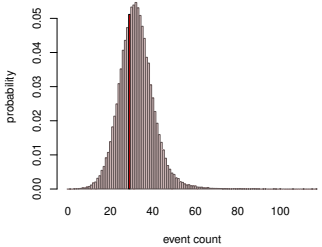
[12] A. Gelman, J. B. Carlin, H. Stern, D. B. Dunson, A. Vehtari, D. B. Rubin. *Bayesian data analysis*. Chapman & Hall / CRC, 2014.

[13] X. L. Meng. Posterior predictive *p*-values. *The Annals of Statistics*, **22**(3), 1142–1160, 1994.

Example

Historical and current data

- *new* trial yields $x_c = 29$ events in a control group of size $N_c = 150$ ($\frac{x_c}{N_c} = 0.19$)
- new study consistent with historical studies? $\rightarrow$ contrast with (MAP) predictions



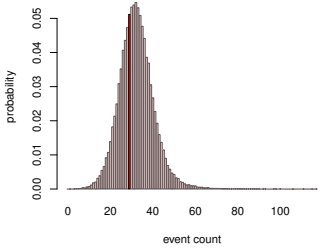
[12] A. Gelman, J. B. Carlin, H. Stern, D. B. Dunson, A. Vehtari, D. B. Rubin. *Bayesian data analysis*. Chapman & Hall / CRC, 2014.

[13] X. L. Meng. Posterior predictive *p*-values. *The Annals of Statistics*, **22**(3), 1142–1160, 1994.

Example

Historical and current data

- *new* trial yields $x_c = 29$ events in a control group of size $N_c = 150$ ($\frac{x_c}{N_c} = 0.19$)
- new study consistent with historical studies? $\rightarrow$ contrast with (MAP) predictions



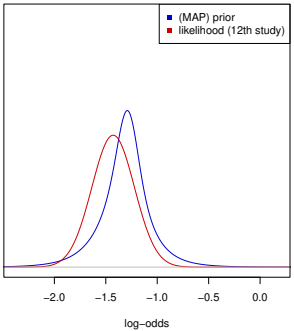
- in terms of MAP prior: a "**prior predictive check**" [12]
- may be turned into a (prior predictive) p-value [13] (here: $p = 0.78$)

[12] A. Gelman, J. B. Carlin, H. Stern, D. B. Dunson, A. Vehtari, D. B. Rubin. *Bayesian data analysis*. Chapman & Hall / CRC, 2014.

[13] X. L. Meng. Posterior predictive p-values. *The Annals of Statistics*, **22**(3), 1142–1160, 1994.

Example

Historical and current data

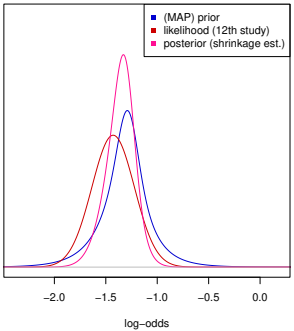


- MAP prior and estimate (y_{12}, s_{12}) combine to form shrinkage estimate.
  (probability: 0.19 [0.14, 0.26]

[14] B. Neuenschwander, S. Weber, H. Schmidli, A. O'Hagan. Predictively consistent prior effective sample sizes. *Biometrics*, **76**(2):578.587, 2020.

Example

Historical and current data

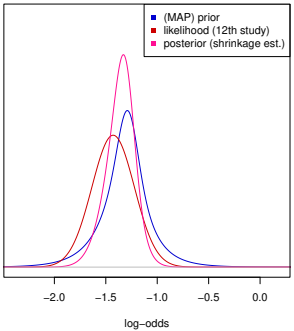


- MAP prior and estimate (y_{12}, s_{12}) combine to form shrinkage estimate.
  (probability: 0.19 [0.14, 0.26] $\rightarrow$ 0.21 [0.16, 0.25])

[14] B. Neuenschwander, S. Weber, H. Schmidli, A. O'Hagan. Predictively consistent prior effective sample sizes. *Biometrics*, **76**(2):578.587, 2020.

Example

Historical and current data

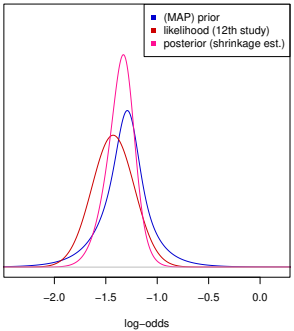


- MAP prior and estimate (y_{12}, s_{12}) combine to form shrinkage estimate.
  (probability: 0.19 [0.14, 0.26] $\rightarrow$ 0.21 [0.16, 0.25])
- substantial precision gain: posterior std.dev. only $0.7 \times s_{12}$
- would otherwise require 104% increase in sample size
  ("156 additional patients")

[14] B. Neuenschwander, S. Weber, H. Schmidli, A. O'Hagan. Predictively consistent prior effective sample sizes. *Biometrics*, **76**(2):578.587, 2020.

Example

Historical and current data



- MAP prior and estimate (y_{12}, s_{12}) combine to form shrinkage estimate.
  (probability: 0.19 [0.14, 0.26] $\rightarrow$ 0.21 [0.16, 0.25])
- substantial precision gain: posterior std.dev. only $0.7 \times s_{12}$
- would otherwise require 104% increase in sample size
  ("156 additional patients")

- more generally: **(MAP) prior** associated with "**effective sample size**" [14]
  (here: $\text{ESS}_{\text{ELIR}} = 153$)

[14] B. Neuenschwander, S. Weber, H. Schmidli, A. O'Hagan. Predictively consistent prior effective sample sizes. *Biometrics*, **76**(2):578.587, 2020.

Example

Historical and current data

- new trial yields $x_c = 29$ events in a **control** group of size $N_c = 150$ ($\frac{x_c}{N_c} = 0.19$)
- active arm: $x_t = 40$ events in a **treatment** group of size $N_c = 300$ ($\frac{x_t}{N_t} = 0.13$)

Example

Historical and current data

- new trial yields $x_c = 29$ events in a **control** group of size $N_c = 150$ ($\frac{x_c}{N_c} = 0.19$)
- active arm: $x_t = 40$ events in a **treatment** group of size $N_c = 300$ ($\frac{x_t}{N_t} = 0.13$)
- combine "plain" estimates / combine active with control + historical

	events		log-odds		log-OR	
	control	treatment	control	treatment	estimate	95% CI
RCT only	29/150	40/300	-1.43 (0.21)	-1.87 (0.17)	-0.44 (0.27)	[-0.97, 0.08]

Example

Historical and current data

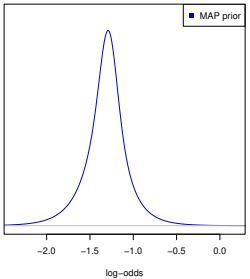
- new trial yields $x_c = 29$ events in a **control** group of size $N_c = 150$ ($\frac{x_c}{N_c} = 0.19$)
- active arm: $x_t = 40$ events in a **treatment** group of size $N_c = 300$ ($\frac{x_t}{N_t} = 0.13$)
- combine "plain" estimates / combine active with control + historical

	events		log-odds		log-OR	
	control	treatment	control	treatment	estimate	95% CI
RCT only	29/150	40/300	-1.43 (0.21)	-1.87 (0.17)	-0.44 (0.27)	[-0.97, 0.08]
RCT + MAP	$\frac{29+198}{150+930}$	40/300	-1.35 (0.14)	-1.87 (0.17)	-0.52 (0.22)	[-0.96, -0.09]

- historical controls:
  - more precise control group + effect estimates
  - fewer control patients

The MAP prior

Practical issues: simplification

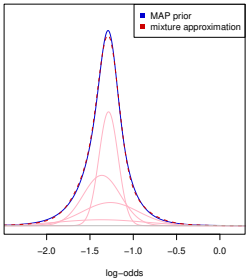


- for practical application (communication, pre-specification, . . .): “simple” summary of MAP prior required

[15] S. Weber, Y. Yi, J. W. Seaman, T. Kakizume, H. Schmidli. Applying meta-analytic-predictive priors with the R Bayesian evidence synthesis tools. *Journal of Statistical Software*, **100**(19):1–32, 2021.

The MAP prior

Practical issues: simplification



	weight	mean	std.dev.
1	0.37	-1.29	0.11
2	0.32	-1.36	0.22
3	0.22	-1.26	0.32
4	0.09	-1.37	0.50

- for practical application (communication, pre-specification, . . .): “simple” summary of MAP prior required
- idea: approximate by **mixture distribution** of few components [15] (implemented in `RBesT`; here: 4 normal components)

[15] S. Weber, Y. Yi, J. W. Seaman, T. Kakizume, H. Schmidli. Applying meta-analytic-predictive priors with the R Bayesian evidence synthesis tools. *Journal of Statistical Software*, **100**(19):1–32, 2021.

The MAP prior

Practical issues: robustification

- concern: analysis hinges on **exchangeability** of historical and current trials.
- pooling of control rates challenges randomization
- **consistency** check may be implemented (e.g., prior predictive p-value)
  — but probably with little "power"
- to safeguard against assumption violation: anticipate potential **prior/data conflict**
- "**robustification**" ideas:
  - "more conservative" priors (heavier tails, greater variance, ...)
  - include possibility of alternative models $\rightarrow$ **mixture prior**

$$p(\theta) = \begin{cases} \text{informative (MAP)} & \text{with probability } (1 - w_{\text{R}}) \\ \text{non-informative (vague)} & \text{with probability } w_{\text{R}} \end{cases}$$

- latter solution commonly preferred (easily motivated, elicited, ...)
- ESS considerations etc. may again be applied

Other MAP prior applications

Treatment effect estimation

- shrinkage estimation also useful for treatment **effects** (e.g., MAP prior from earlier-phase data) [16]

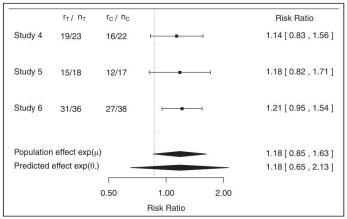


Figure 1. Data and results at end-of-phase II meeting.

[16] S. Wandel, B. Neuenschwander, C. Röver, T. Friede. Using phase II data for the analysis of phase III studies: an application in rare diseases. *Clinical Trials*, **14**(3):277–285, 2017.

[17] C. Röver, T. Friede. Dynamically borrowing strength from another study through shrinkage estimation. *Statistical Methods in Medical Research*, **29**(1):293–308, 2020.

Other MAP prior applications

Treatment effect estimation

- shrinkage estimation also useful for treatment **effects** (e.g., MAP prior from earlier-phase data) [16]

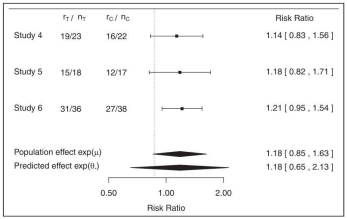


Figure 1. Data and results at end-of-phase II meeting.

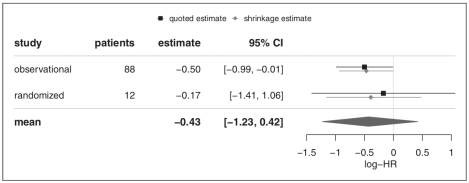


Figure 2. Forest plot for the CJD example (log-HR outcome). The shrinkage interval for the log-HR based on randomized evidence here is $[-1.16, 0.48]$, spanning only two-thirds of the original confidence interval width.

- borrowing of information also for a (heterogeneous) **pair** of estimates (i.e., $k = 2$) [17] (focus on shrinkage estimate θ_2, not overall mean μ)

[16] S. Wandel, B. Neuenschwander, C. Röver, T. Friede. Using phase II data for the analysis of phase III studies: an application in rare diseases. *Clinical Trials*, **14**(3):277–285, 2017.

[17] C. Röver, T. Friede. Dynamically borrowing strength from another study through shrinkage estimation. *Statistical Methods in Medical Research*, **29**(1):293–308, 2020.

Other MAP prior applications

Heterogeneity estimation

- important for meta-analysis: heterogeneity **prior** specification
  (especially for few studies)
- general guidance available for non-informative [18] or weakly informative priors [19]
- important aspect: empirical information — *what can we learn from past analyses?*

[18] C. Röver. Bayesian random-effects meta-analysis using the `bayesmeta` R package. *Journal of Statistical Software*, **93**(6), 2020.

[19] C. Röver, R. Bender, S. Dias, C. H. Schmid, H. Schmidli, S. Sturtz, S. Weber, T. Friede. On weakly informative prior distributions for the heterogeneity parameter in Bayesian random-effects meta-analysis. *Research Synthesis Methods*, **12**(4):448–474, 2021.

Other MAP prior applications

Heterogeneity estimation

- important for meta-analysis: heterogeneity **prior** specification
  (especially for few studies)
- general guidance available for non-informative [18] or weakly informative priors [19]
- important aspect: empirical information — *what can we learn from past analyses?*

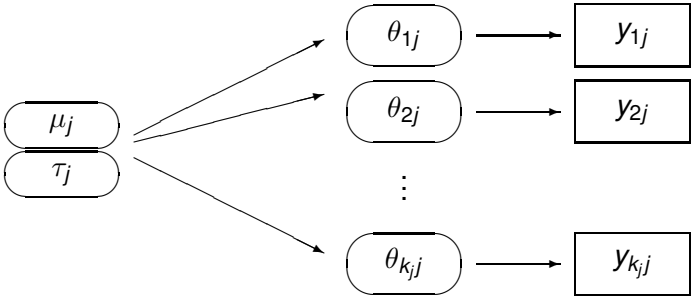
- pooling of heterogeneity *estimates* tricky:
  hard to summarize / model
- need: joint model for "historical" meta-analyses

[18] C. Röver. Bayesian random-effects meta-analysis using the `bayesmeta` R package. *Journal of Statistical Software*, **93**(6), 2020.

[19] C. Röver, R. Bender, S. Dias, C. H. Schmid, H. Schmidli, S. Sturtz, S. Weber, T. Friede. On weakly informative prior distributions for the heterogeneity parameter in Bayesian random-effects meta-analysis. *Research Synthesis Methods*, **12**(4):448–474, 2021.

Other MAP prior applications

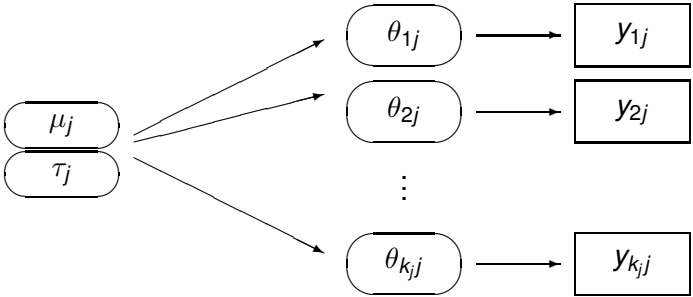
Heterogeneity prediction: the model



- model DAG for *j*th meta-analysis

Other MAP prior applications

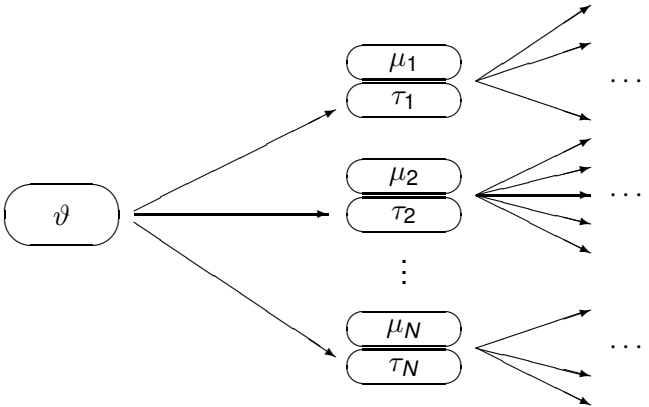
Heterogeneity prediction: the model



- model DAG for jth meta-analysis (out of several)
- idea: combine ($j = 1, \ldots, N$) meta-analyses, infer τ distribution

Other MAP prior applications

Heterogeneity prediction: the model



- additional overarching layer, combining N meta-analyses
- common heterogeneity distribution for $\tau_1, \ldots, \tau_N$ (e.g., half-Normal(ϑ))

Other MAP prior applications

Heterogeneity prediction: the model

- **data**: N meta-analyses, each involving k_j studies,
  effect estimates y_{ij}, standard errors s_{ij} ($i = 1, \ldots, k_j$, $j = 1, \ldots, N$),
- assume:

$$\begin{aligned} y_{ij}|\mu_j, \tau_j, s_{ij} &\sim \text{Normal}(\mu_j, s_{ij}^2 + \tau_j^2) \\ \mu_j|\mu_\text{p}, \sigma_\text{p} &\sim \text{Normal}(\mu_\text{p}, \sigma_\text{p}^2) \end{aligned}$$

  for fixed "neutral" μ_p and "large" σ_p ($\rightarrow$ stratification, no pooling)
- **heterogeneity stage**:

$$\tau_j|\vartheta \sim \text{P}(\vartheta)$$

  for some "heterogeneity distribution" $\text{P}(\vartheta)$
- **parameters**: N means μ_j and heterogeneities τ_j; "distribution" parameter(s) ϑ
- (hyper-) **prior** required for ϑ
- **aim**: prediction $\tau^\star$

Other MAP prior applications

Heterogeneity prediction: the model

- **data**: N meta-analyses, each involving k_j studies,
  effect estimates y_{ij}, standard errors s_{ij} ($i = 1, \ldots, k_j$, $j = 1, \ldots, N$),
- assume:

$$\begin{aligned} y_{ij}|\mu_j, \tau_j, s_{ij} &\sim \text{Normal}(\mu_j, s_{ij}^2 + \tau_j^2) \\ \mu_j|\mu_\text{p}, \sigma_\text{p} &\sim \text{Normal}(\mu_\text{p}, \sigma_\text{p}^2) \end{aligned}$$

  for fixed "neutral" μ_p and "large" σ_p ($\rightarrow$ stratification, no pooling)
- **heterogeneity stage**:

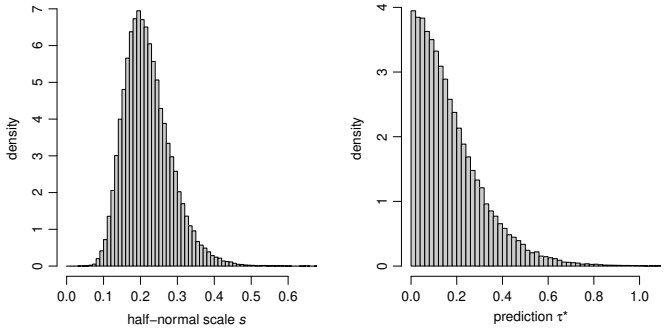
$$\tau_j|\vartheta \sim \text{P}(\vartheta)$$

  for some "heterogeneity distribution" $\text{P}(\vartheta)$ (e.g.: $\tau_j|\vartheta \sim$ half-Normal(ϑ))
- **parameters**: N means μ_j and heterogeneities τ_j; "distribution" parameter(s) ϑ
- (hyper-) **prior** required for ϑ (half-normal scale ϑ)
- **aim**: prediction $\tau^\star$

Other MAP prior applications

Heterogeneity prediction

- posterior: scale parameter s
- posterior predictive: $\tau^\star | s \sim$ half-Normal(s)



(figures from an example application)

- **posterior predictive** serves as **"MAP" prior** for new (($N+1$)th) analysis

Other MAP prior applications

Heterogeneity prediction: simplification, “robustification”

- simplification:
  - **posterior predictive** is a **mixture distribution**
    ($\tau^\star|\vartheta \sim$ half-Normal(ϑ), with uncertain ϑ)
  - half-normal example — obvious parametric approximation:
    by half-normal, or half-normal mixture
    (e.g., half-Student-*t* distribution)
- robustification:
  - rather “conservatization” (?)
  - generally: larger τ value yields “more conservative” meta-analysis
    (less shrinkage, wider intervals, . . .)
  - stochastically larger or heavier-tailed prior
    usually considered a *conservative* choice

Other MAP prior applications

Heterogeneity prediction: simplification, “robustification”

- original idea and first implementation: Rhodes *et al.* (2015) [20] and Turner *et al.* (2015) [21]
- general approach detailed [22]

[20] K.M. Rhodes *et al.* Predictive distributions were developed for the extent of heterogeneity in meta-analyses of continuous outcome data. *Journal of Clinical Epidemiology*, **68**(1):52–60, 2015.

[21] R.M. Turner *et al.* Predictive distributions for between-study heterogeneity and simple methods for their application in Bayesian meta-analysis. *Statistics in Medicine*, **34**(6):984–998, 2015.

[22] C. Röver, S. Sturtz, J. Lilienthal, R. Bender, T. Friede. Summarizing empirical information on between-study heterogeneity for Bayesian random-effects meta-analysis. *Statistics in Medicine*, **42**(14):2439–2454, 2023.

[23] J. Lilienthal, S. Sturtz, C. Schürmann, M. Maiworm, C. Röver, T. Friede, R. Bender. Bayesian random-effects meta-analysis with empirical heterogeneity priors for application in health technology assessment with very few studies. *Research Synthesis Methods*, **15**(2):275–287, 2024.

Other MAP prior applications

Heterogeneity prediction: simplification, “robustification”

- original idea and first implementation: Rhodes *et al.* (2015) [20] and Turner *et al.* (2015) [21]
- general approach detailed [22]
- applied to IQWiG data [23] to help pre-specifying analyses in regulatory context

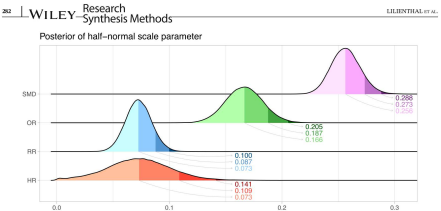


FIGURE 7 Posterior distributions of the half-normal scale parameter for different effect measures. Different color shades and the given numerical values indicate the 50%, 90%, and 99% quantiles, respectively.

[20] K.M. Rhodes *et al.* Predictive distributions were developed for the extent of heterogeneity in meta-analyses of continuous outcome data. *Journal of Clinical Epidemiology*, **68**(1):52–60, 2015.

[21] R.M. Turner *et al.* Predictive distributions for between-study heterogeneity and simple methods for their application in Bayesian meta-analysis. *Statistics in Medicine*, **34**(6):984–998, 2015.

[22] C. Röver, S. Sturtz, J. Lilienthal, R. Bender, T. Friede. Summarizing empirical information on between-study heterogeneity for Bayesian random-effects meta-analysis. *Statistics in Medicine*, **42**(14):2439–2454, 2023.

[23] J. Lilienthal, S. Sturtz, C. Schürmann, M. Maiworm, C. Röver, T. Friede, R. Bender. Bayesian random-effects meta-analysis with empirical heterogeneity priors for application in health technology assessment with very few studies. *Research Synthesis Methods*, **15**(2):275–287, 2024.

Prior distributions from meta-analytic predictions

Summary

- **hierarchical models** established for meta-analysis
- **Bayesian** models advantageous for sparse data and advanced applications
- besides main effect: **shrinkage**, **prediction**
- **MAP priors** as data-informed priors
- useful in many contexts (controls, effects, nuisance parameters, . . .)
- option to implement scepticism via **robustification**
- analogous “MAP” approach for empirically motivated **heterogeneity priors**