Meta-analysis using Bayesian methods: Applications in systematic reviews

Sibylle Sturtz

IQWiG Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Institute for Quality and Efficiency in Health Care

Agenda

- Introduction: meta-analysis according to IQWiG' methods paper
- Methods for meta-analysis
  - Frequentist methods
  - Bayesian methods
- Prior distributions for τ
  - Half-Normal
  - Jeffreys
  - Indication specific/general health care setting from Cochrane data base
- Data examples
  - Milk teeth, severe lung emphysema, motor-driven continuous passive motion devices
- Specific topics: shifted hypotheses, p-value/Bayes factor
- Conclusions

General Methods - Version 5.0

In principle, Bayesian methods may be regarded as an alternative to statistical significance tests. Depending on the research question posed, the Institute will, where necessary, also apply Bayesian methods (e.g. for indirect comparisons).

However, the use of meta-analyses with random effects reaches its limits in the event of very few studies (fewer than 5). As heterogeneity then cannot be reliably estimated, the use of meta-analyses with random effects can lead to very broad confidence intervals that potentially no longer allow conclusions on the evidence base. Especially in the event of very few studies, a fixed-effect model or a qualitative summary should be considered. Depending on the context, alternative procedures could also be an option, such as Bayesian approaches or methods from the area of generalized linear models.

IQWiG, 2017

Procedure for evidence synthesis

- Definition of PICOS
- Given a pool of studies from systematic literature search:
  - Calculation of heterogeneity of the study pool according to the statistical test for heterogeneity (Sutton, 2000)
  - Meta-analysis if p ≥ 0,05, otherwise perform qualitative evidence synthesis

For ≥ 5 studies: Knapp-Hartung with heterogeneity according to Paule-Mandel

  - Number of studies is typically rather small, i.e. less than 5
  - Therefore, the estimation of between-study variance τ^2 can be inaccurate

For < 5 studies: fixed effect models, Bayesian methods, generalized linear models, or qualitative evidence synthesis

IQWiG, 2017

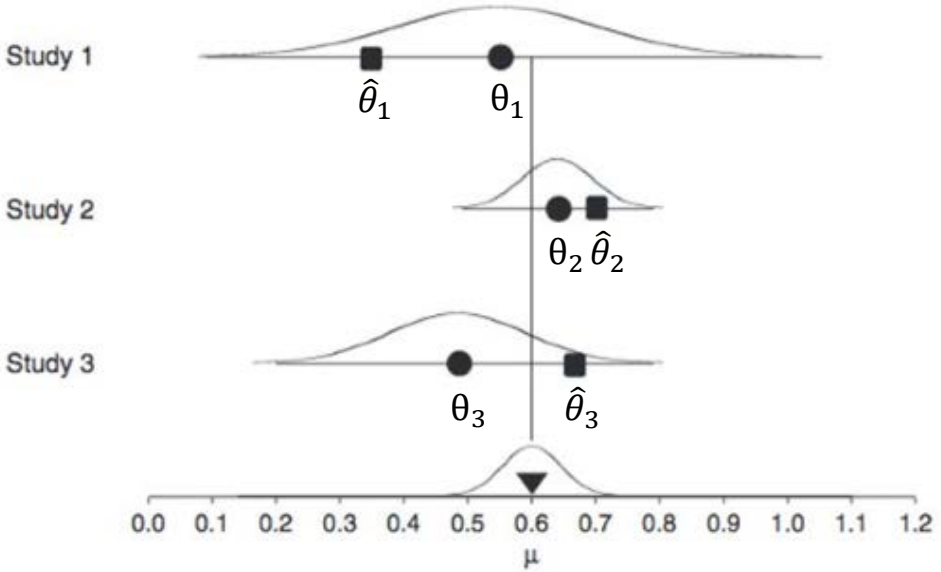
Meta-analysis

Meta-analysis of k studies with observations θ_i with standard error σ_i, $i = 1, \ldots, k$

$$\theta_i \sim N(\hat{\theta}_i, \sigma_i^2)$$
$$\hat{\theta}_i \sim N(\mu, \tau^2)$$

$$\Leftrightarrow$$

$$\theta_i \sim N(\mu, \sigma_i^2 + \tau^2)$$



known parameters, unknown parameters

Aus: Borenstein et al.(2009): Introduction to meta-analysis. Wiley.

Meta-analysis: frequentist methods

Meta-analysis of k studies with observations θ_i with standard error σ_i, $i = 1, \ldots, k$

$$\theta_i \sim N(\hat{\theta}_i, \sigma_i^2)$$
$$\hat{\theta}_i \sim N(\mu, 0)$$

Fixed effect/common effect model

assumes a common effect in all studies => between study variation $\tau^2 = 0$

continuous data: inverse variance approach

$$\hat{\mu} = \frac{\sum \theta_i w_i}{\sum w_i} \text{ with } w_i = \frac{1}{\widehat{\sigma_i}}$$

95 %-KI: $\hat{\theta} \pm z_{1-\alpha/2}\sqrt{1/\sum w_i}$.

binary data: Maentel-Haenszel approach: $OR_{MH} = \frac{\sum \frac{a_{I,i}(n_{C,i} - a_{C,i})}{n_{I,i}+n_{C,i}}}{\sum \frac{a_{C,i}(n_{I,i} - a_{I,i})}{n_{I,i}+n_{C,i}}}$

95 %-KI: $\exp(\ln(OR_{MH}) \pm z_{1-\frac{\alpha}{2}}\sqrt{Var(\ln(OR_{MH}))})$

known parameters, unknown parameters

Meta-analysis: frequentist methods

Meta-analysis of k studies with observations θ_i with standard error σ_i, $i = 1, \ldots, k$

$$\theta_i \sim N(\hat{\theta}_i, \sigma_i^2)$$
$$\hat{\theta}_i \sim N(\mu, \tau^2)$$

Random effects model: DerSimonian-Laird (1986)

$$\hat{\mu} = \frac{\sum \theta_i w_i}{\sum w_i} \text{ with } w_i = \frac{1}{\widehat{\sigma_i} + \widehat{\tau^2}}$$

$$\tau_{DL}^2 = \max\left\{0, \frac{\sum \sigma_i^{-2}(\theta_i - \hat{\mu})^2 - (k-1)}{\sum \sigma_i^{-2} - \frac{\sum(\sigma_i^{-2})^2}{\sum(\sigma_i^{-2})}}\right\}$$

95 %-KI: $\hat{\theta} \pm z_{1-\alpha/2}\sqrt{1/\sum w_i}$.

criticized due to its unfavourable statistical properties

especially in the case of very few studies too narrow confidence intervals, inflated type-I errors, estimation of uncertainty of τ is ignored

known parameters, unknown parameters

Meta-analysis: frequentist methods

Meta-analysis of k studies with observations θ_i with standard error σ_i, $i = 1, \ldots, k$

$$\theta_i \sim N\left(\hat{\theta}_i, \sigma_i^2\right)$$
$$\hat{\theta}_i \sim N(\mu, \tau^2)$$

Random effects model

Knapp-Hartung with heterogeneity according to Paule-Mandel (Veroniki, 2015)

$$\hat{\mu} = \frac{\sum \theta w_i}{\sum w_i} \text{ with } w_i = \frac{1}{\widehat{\sigma_i} + \widehat{\tau^2}}$$

$$\tau^2 = \frac{\sum w_i (y_i - \hat{\mu})^2 - \left(\sum w_i \sigma_i^2 - \frac{\sum w_i^2 \sigma_i^2}{\sum w_i}\right)}{\sum w_i - (\sum w_i^2)/(\sum w_i)}$$

95 %-KI: $\hat{\theta} \pm t_{k-1,1-\alpha/2} \sqrt{\sum w_i \left(\theta_i - \hat{\theta}\right)^2 / (k-1) \sum w_i}$.

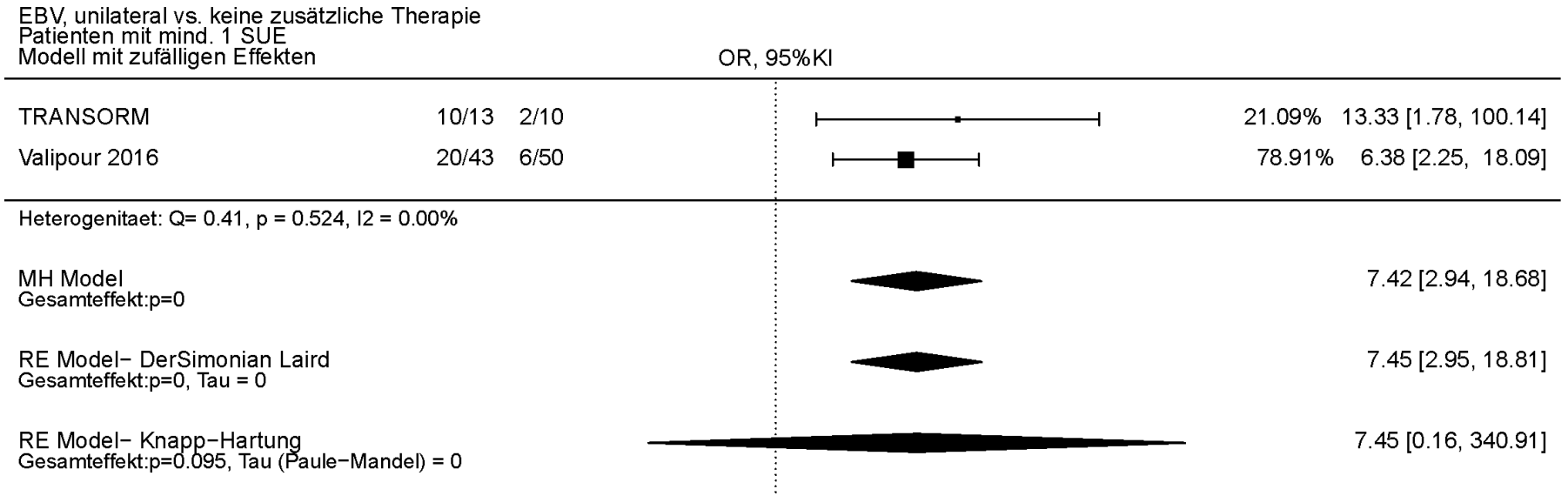
Tends to be over-conservative for $k = 2$ studies

May produce too narrow confidence intervals in very homogeneous situations
=> modified variance estimation available

known parameters, unknown parameters

Knapp-Hartung estimator with 2 studies

- low power, wide confidence intervals
- May lead to a non-significant pooled effect when two significant studies are pooled.



Meta-analysis: Bayesian methods

- Accounts for uncertainty in the estimation of the between-study variance τ^2

- Can incorporate external information as well as prior knowledge which in turn can reduce the length of interval estimates for the common mean.
- If no or only vague prior information is available, noninformative priors can be incorporated
- Produces a distribution for the quantities of interest, allows to calculate probability that e.g. the odds ratio is smaller/bigger than a prespecified threshold

- Choice of prior information can substantially influence final results (Weber, 2018)
- Carrying out sensitivity analyses is important to investigate how the results depend on any assumptions made (Cochrane Handbook 5.1)

- Some approaches depend on MCMC

Meta-analysis: Bayesian methods

Meta-analysis of k studies with observations θ_i with standard error σ_i, $i = 1, \ldots, k$

$$\theta_i \sim N(\hat{\theta}_i, \sigma_i^2)$$
$$\hat{\theta}_i \sim N(\mu, \tau^2)$$
$$\Leftrightarrow$$
$$\theta_i \sim N(\mu, \sigma_i^2 + \tau^2)$$

Bayesian approach:

Posterior	$\propto$	prior	$\times$	likelihood
$P((\mu, \tau^2)\mid \text{data})$	$\propto$	$P((\mu, \tau^2))$	$\times$	$P(\text{data} \mid (\mu, \tau^2))$

requires prior distributions for μ and τ^2

⇨ noninformative or (weakly) informative prior distributions

Prior choice for $\mu \sim N(0, 4^2)$

Prior choice for τ: choice is critical as the number of studies is small

known parameters, unknown parameters

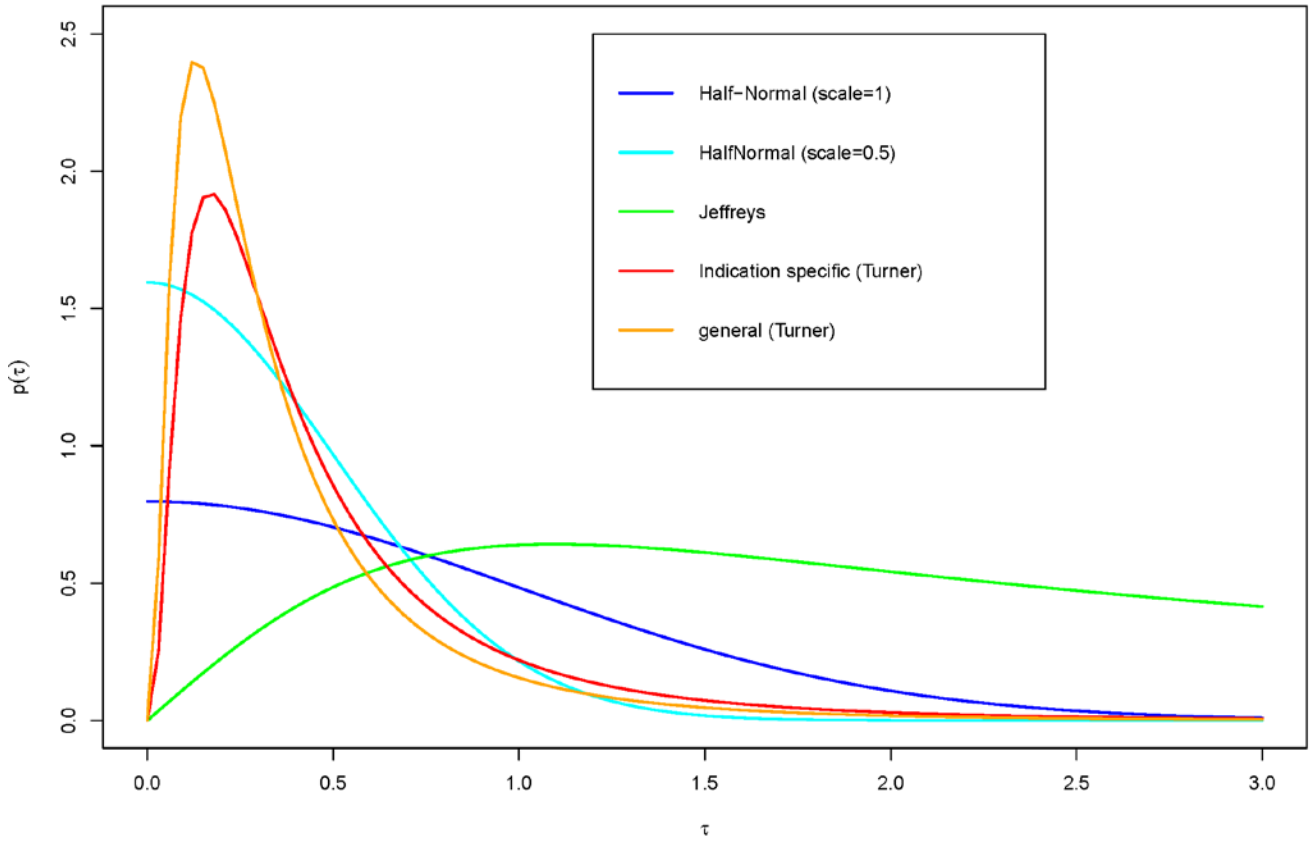
Chosen priors for between-trial heterogeneity τ

- Prior distributions for binary data
  1) Half-Normal prior with scale 0,5 (Friede et al., Biometrical Jornal, 2017)
  2) Half-Normal prior with scale 1 (Friede et al., Biometrical Jornal, 2017)
  3) Jeffreys prior (Bodnar et al., Stats in Med, 2016)
  4) Informed prior: indication and outcome specific (Turner et al., Stats in Med, 2015)
  5) Informed prior: general heath care setting (Turner et al., Stats in Med, 2015)

- Prior distributions for continuous data
  1) Half-Normal prior with scale 0,5
  2) Half-Normal prior with scale 1
  3) Jeffreys prior
  4) Informed prior: indication and outcome specific (Rhodes et al., Journal Clin Epi, 2018)
  5) Informed prior: general heath care setting (Rhodes et al., Journal Clin Epi, 2018)

- Calculations in R using `metafor` (Viechtbauer, 2010) and `bayesmeta` (Röver, 2017)

Chosen priors for binary data



Heterogeneity: small moderate substantial large very large

Friede et al., RSM, 2017

Half Normal distributions

Friede et al. (Biometrical Journal, 2017)

- Simulation study on log(OR) scale:
  - DerSimonian Laird
  - Knapp-Hartung with heterogeneity according to Paule-Mandel
  - Modified Knapp-Hartung
  - Bayesian random-effects meta-analysis
    - μ : noninformative (improper) uniform prior
    - τ : half Normal distribution with scale 1 or 0,5
  - $\tau \in \{0.0, 0.1, 0.2, 0.5, 1.0\}$; $k = 2; \mu = 0$, combinations $(n_1, n_2) \in \{25, 100, 400\}$
  - Fraction of τ estimated equal to 0, coverage probability, mean interval length

- Bayesian intervals: good coverage in the range of the prior

  (e.g. Half-Normal (scale =0.5) coverage is reasonable for τ up to 0.5, drops reasonably for larger values)
- Frequentist methods: too wide to allow reasonable conclusions (KH) or very narrow (DL); Bayesian intervals are much shorter, satisfying properties
- Bayesian RE meta-analysis = reasonable compromise between KH- and DL- methods

Half Normal distributions

Friede et al. (RSM, 2017)

- Simulation study on log(OR) scale:
  - DerSimonian Laird
  - REML
  - Knapp-Hartung with heterogeneity according to Paule-Mandel
  - Bayes modal estimate
  - Bayesian random-effects meta-analysis
    - τ : half-Normal distribution with scale 1 or 0,5; Uniform (0,4)
  - $\tau \in \{0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1.0, 2.0\};\ k = \{3, 5, 10\}$
  - Fraction of τ estimated equal to 0, bias, coverage probability, mean interval length
- Bayesian credible intervals with appropriate choice of prior performed well
- Bayesian credible intervals tend to be shorter compared to KH and coverage probability is either similar or closer to nominal level
- For very small k results are sensitive for prior specification
- Recommend half-Normal priors as values fit to typical applications of rare diseases and small populations

Jeffreys prior/Berger and Bernardo reference prior

Bodnar et al. (2016)

$$P(\theta) \propto \sqrt{\sum \frac{\tau}{\left(\sigma_i^2 + \tau^2\right)^2}}$$

- Least informative when compared with the posterior distribution in terms of Shannon‘s mutal information
- Most commonly employed in random-effects meta-analyses
- Improper prior, proper posterior for $k \geq 2$.
- For $k \geq 3$ first moment of the marginal posterior exists, for $k \geq 4$ second moment of the marginal posterior exists.
- Here: calculation of posterior median

Jeffreys prior/Berger and Bernardo reference prior

Bodnar et al. (2016)

- Simulation study on log(OR):
  - Knapp-Hartung with heterogeneity according to DerSimonian-Laird
  - Knapp-Hartung with heterogeneity according to Paule-Mandel
  - Profile likelihood estimation
  - Bayesian random-effects meta-analysis
    - τ: noninformative Jeffreys prior
  - $\tau \in \{0.0, 0.1, 0.2, 0.3, 0.5, 1.0, 2.0\}; \sigma_i \sim Unif(0.0, 0.5); \mu = 0; \mathrm{k} \in \{5, 10, 20, 30\}$
  - Coverage probability, average interval length, average bias, estimated heterogeneity

- Effective coverage probabilities of Bayesian nominal 95% credible intervals are less likely below 95% compared to the other 3 methods
- Effective coverage probabilities of the nominal 95% credible intervals were the closest to 95%
- Satisfying estimation of heterogeneity
- Bodnar et al. (2016) recommend this prior also for meta-analyses involving a small number of studies without relevant or credible prior information

Informed priors – binary data

Turner et al. (2015)

- Based on 14886 binary outcome meta-analyses from the Cochrane Database of Systematic Reviews
- Differentiates between intervention comparison types (pharmacological, placebo, non-pharmacological) and outcome categories (subjective, semi-objective, objective) in different therapeutic areas
- Candidates for distribution of τ^2_{new}:

  log-Normal-, inverse-Gamma-, Gamma-distributions
- Regression models for $\log(\tau^2_{new})$ identified Normal distribution to be best
- Normal distributions for $\log(\tau^2_{new})$ were fitted to each predictive distribution under a full Bayesian model using MCMC methods, using the posterior mean and standard deviation

  -> parametric distribution approximating the predictive distribution
- Overall average predictive distribution in general health care setting

  $$\tau^2_{new} \sim log\text{-}Normal\ (-2.56, 1.74^2)$$

  but also for specific outcome types and intervention comparison types.

Table IV. Predictive distributions* obtained for the between-study heterogeneity τ^2_{new} in a future meta-analysis, across 80 different settings.

Outcome type	Intervention comparison type				
	Pharmacological vs. Placebo/control	Pharmacological vs. Pharmacological	Non-pharmacological† vs. Placebo/control	Non-pharmacological† vs. Pharmacological	Non-pharma.† vs. Non-pharma.†
All-cause mortality	$LN(-3.95, 1.34^2)$	$LN(-4.18, 1.41^2)$	$LN(-4.17, 1.55^2)$	$LN(-2.92, 1.02^2)$	$LN(-3.50, 1.26^2)$
Obstetric outcomes	$LN(-3.52, 1.74^2)$	$LN(-3.75, 1.79^2)$	$LN(-3.74, 1.91^2)$	$LN(-2.49, 1.50^2)$	$LN(-3.08, 1.68^2)$
Cause-specific mortality/major morbidity event/composite (mortality or morbidity)	$LN(-3.71, 1.74^2)$	$LN(-3.95, 1.79^2)$	$LN(-3.93, 1.91^2)$	$LN(-2.68, 1.51^2)$	$LN(-3.27, 1.68^2)$
Resource use/hospital stay/process	$LN(-2.34, 1.74^2)$	$LN(-2.58, 1.79^2)$	$LN(-2.56, 1.91^2)$	$LN(-1.31, 1.50^2)$	$LN(-1.90, 1.68^2)$
Surgical/device related success/failure	$LN(-2.14, 1.74^2)$	$LN(-2.37, 1.79^2)$	$LN(-2.36, 1.91^2)$	$LN(-1.11, 1.50^2)$	$LN(-1.69, 1.68^2)$
Withdrawals/drop-outs	$LN(-2.99, 1.74^2)$	$LN(-3.23, 1.79^2)$	$LN(-3.21, 1.91^2)$	$LN(-1.96, 1.51^2)$	$LN(-2.55, 1.68^2)$
Internal/external structure-related outcomes	$LN(-2.71, 1.74^2)$	$LN(-2.94, 1.79^2)$	$LN(-2.93, 1.92^2)$	$LN(-1.67, 1.51^2)$	$LN(-2.26, 1.68^2)$
General physical health indicators	$LN(-2.29, 1.53^2)$	$LN(-2.53, 1.58^2)$	$LN(-2.51, 1.72^2)$	$LN(-1.26, 1.25^2)$	$LN(-1.85, 1.46^2)$
Adverse events	$LN(-1.87, 1.52^2)$	$LN(-2.10, 1.58^2)$	$LN(-2.10, 1.71^2)$	$LN(-0.84, 1.24^2)$	$LN(-1.43, 1.45^2)$
Infection/onset of new disease	$LN(-2.49, 1.52^2)$	$LN(-2.73, 1.58^2)$	$LN(-2.71, 1.71^2)$	$LN(-1.46, 1.24^2)$	$LN(-2.05, 1.45^2)$
Signs/symptoms reflecting continuation/end of condition	$LN(-2.06, 1.51^2)$	$LN(-2.29, 1.58^2)$	$LN(-2.28, 1.71^2)$	$LN(-1.03, 1.24^2)$	$LN(-1.61, 1.45^2)$
Pain	$LN(-1.83, 1.52^2)$	$LN(-2.06, 1.58^2)$	$LN(-2.05, 1.71^2)$	$LN(-0.80, 1.25^2)$	$LN(-1.38, 1.45^2)$
Quality of life/functioning (dichotomised)	$LN(-2.54, 1.54^2)$	$LN(-2.78, 1.60^2)$	$LN(-2.77, 1.73^2)$	$LN(-1.51, 1.27^2)$	$LN(-2.10, 1.47^2)$
Mental health indicators	$LN(-2.12, 1.53^2)$	$LN(-2.35, 1.60^2)$	$LN(-2.34, 1.72^2)$	$LN(-1.09, 1.27^2)$	$LN(-1.67, 1.47^2)$
Biological markers (dichotomised)	$LN(-1.77, 1.52^2)$	$LN(-2.00, 1.58^2)$	$LN(-1.99, 1.71^2)$	$LN(-0.74, 1.24^2)$	$LN(-1.33, 1.45^2)$
Subjective outcomes (various)†	$LN(-2.70, 1.52^2)$	$LN(-2.93, 1.58^2)$	$LN(-2.92, 1.71^2)$	$LN(-1.67, 1.25^2)$	$LN(-2.26, 1.45^2)$

Quelle: Turner et al. (2015)

Informed priors – continuous data

Rhodes et al. (2018)

- Based on 6492 continuous outcome meta-analyses from Cochrane Database of Systematic Reviews
- Differentiates between intervention comparison types (pharmacological, placebo, non-pharmacological) and outcome in different therapeutic areas
- Modelled on SMD scale
- Candidates for posterior distribution of τ^2_{new}:

  log-Normal-, inverse-Gamma-, log-*t*-distributions (df=5)
- Regression models for $\log(\tau^2_{new})$ identified t- distribution to be most appropriate
- t- distributions for $\log(\tau^2_{new})$ were fitted to each predictive distribution under a full Bayesian model using MCMC methods, using the posterior mean and standard deviation
- Overall average predictive distribution in general health care setting

  $$\log(\tau^2_{new}) \sim t(-3.44, 2.59^2, 5)$$

  but also for specific outcome types and intervention comparison types.

- If conjugate prior is preferred: authors report predictive inverse-gamma distributions

Data examples

- N17-03: *Assessment of the application of fluoride varnish on milk teeth to prevent the development and progression of initial caries or new carious lesions - rapid report*

  Children with milk teeth

  Application of fluoride varnish vs. common maintenance without fluoride varnish or with placebo

  Outcome: caries, side effects, tooth loss, toothache, dental abscesses or inflammation of the gums (gingivitis)

  RCTs

Data examples

- N14-04: *Procedure for lung volume reduction in severe lung emphysema*

  Patients with severe lung emphysema

  Surgical or bronchoscopic techniques vs. conservative approaches or other techniques for lung volume reduction

  Outcome: mortality, cardiovascular mortality and morbidity, COPD-symptoms, exacerbations, health-related quality of life, adverse events

  RCTs

Data Examples

- N16-03: *Motor-driven continuous passive motion (CPM) devices after interventions on the knee and shoulder joint*

  Patients who have had knee or shoulder surgery or who require conservative treatment

  CPM devices vs. treatment without CPM (=physiotherapy)

  Outcome: range of motion, pain, rupture, health-related quality of life, adverse events

  RCTs

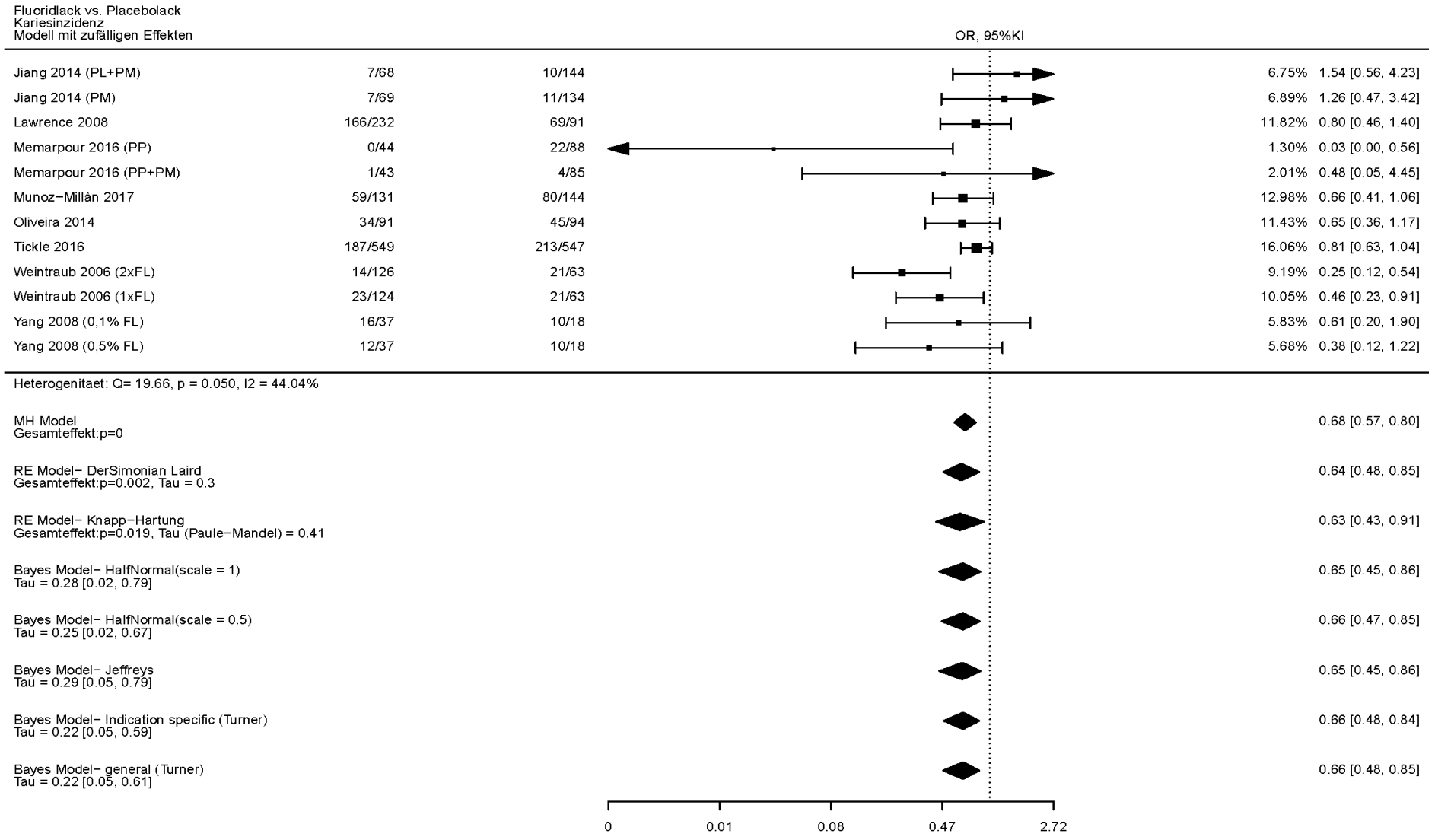
Chosen models: binary data

- Measure: Odds Ratio
- Frequentist models: MH, DSL, KH
- Bayesian models: Prior for μ: Normal$(0,4^2)$

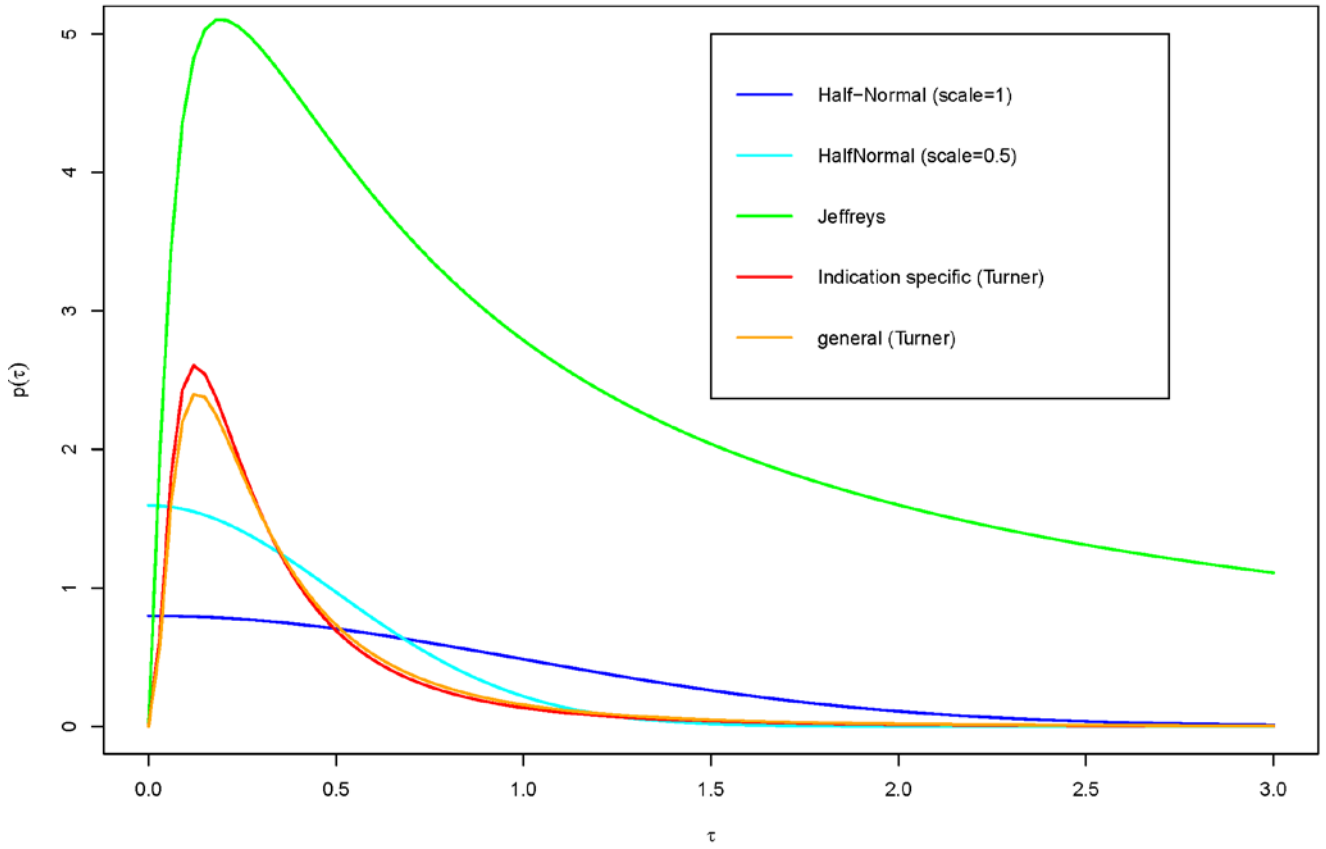
  Prior for τ :
  - half-Normal (scale=1)
  - half-Normal(scale=0.5)
  - Jeffreys prior
  - Turner, indication specific
  - Turner, general
- Estimation for Bayesian models: posterior median and 95% credible interval for the effect μ and between-study variation τ; central, equal-tailed credible intervals

Many studies:

different prior choices lead to comparable results

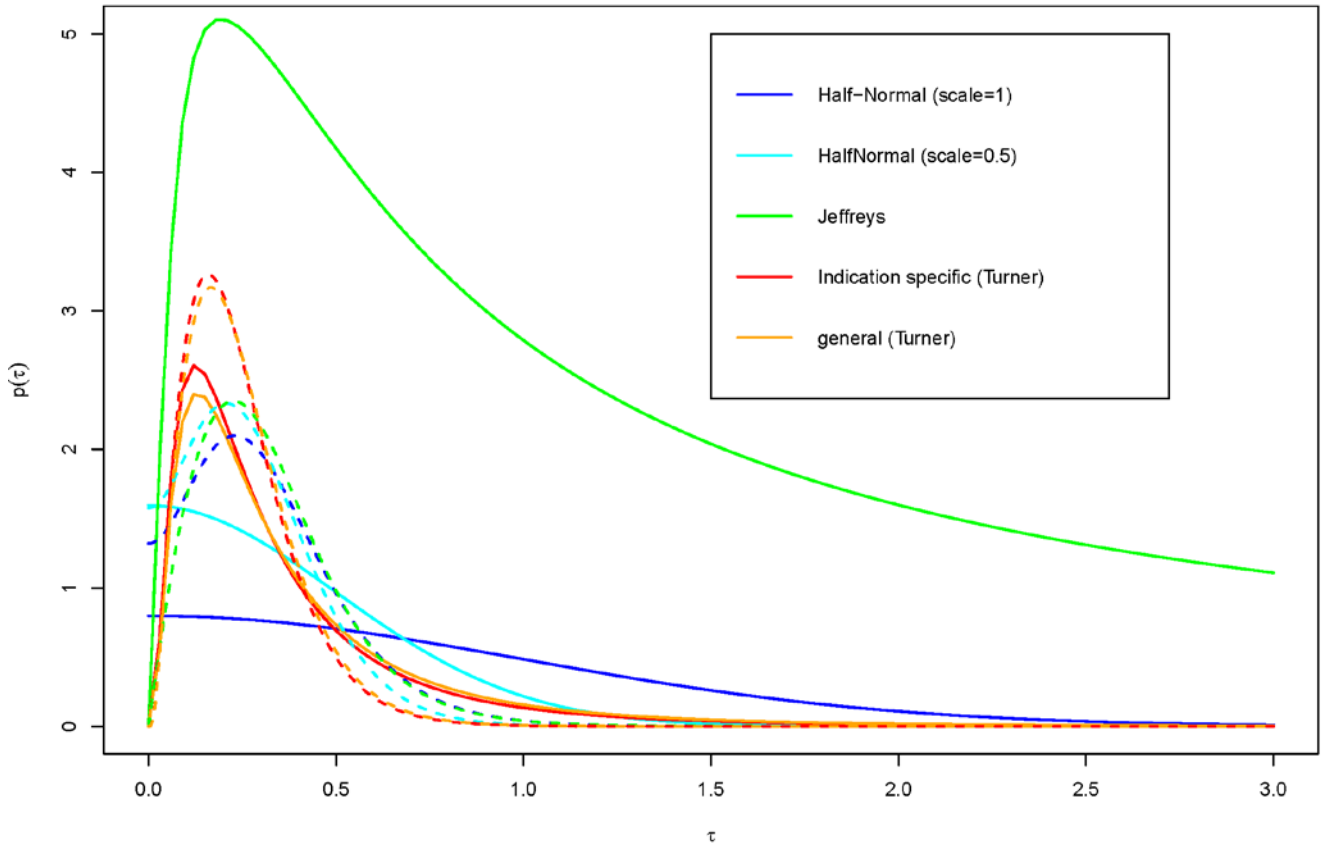


Prior and posterior distribution for τ



Indication specific (Turner): Endpunkt: infection / onset of new disease; Intervention: non−pharmacological; Vergleich: placebo / control

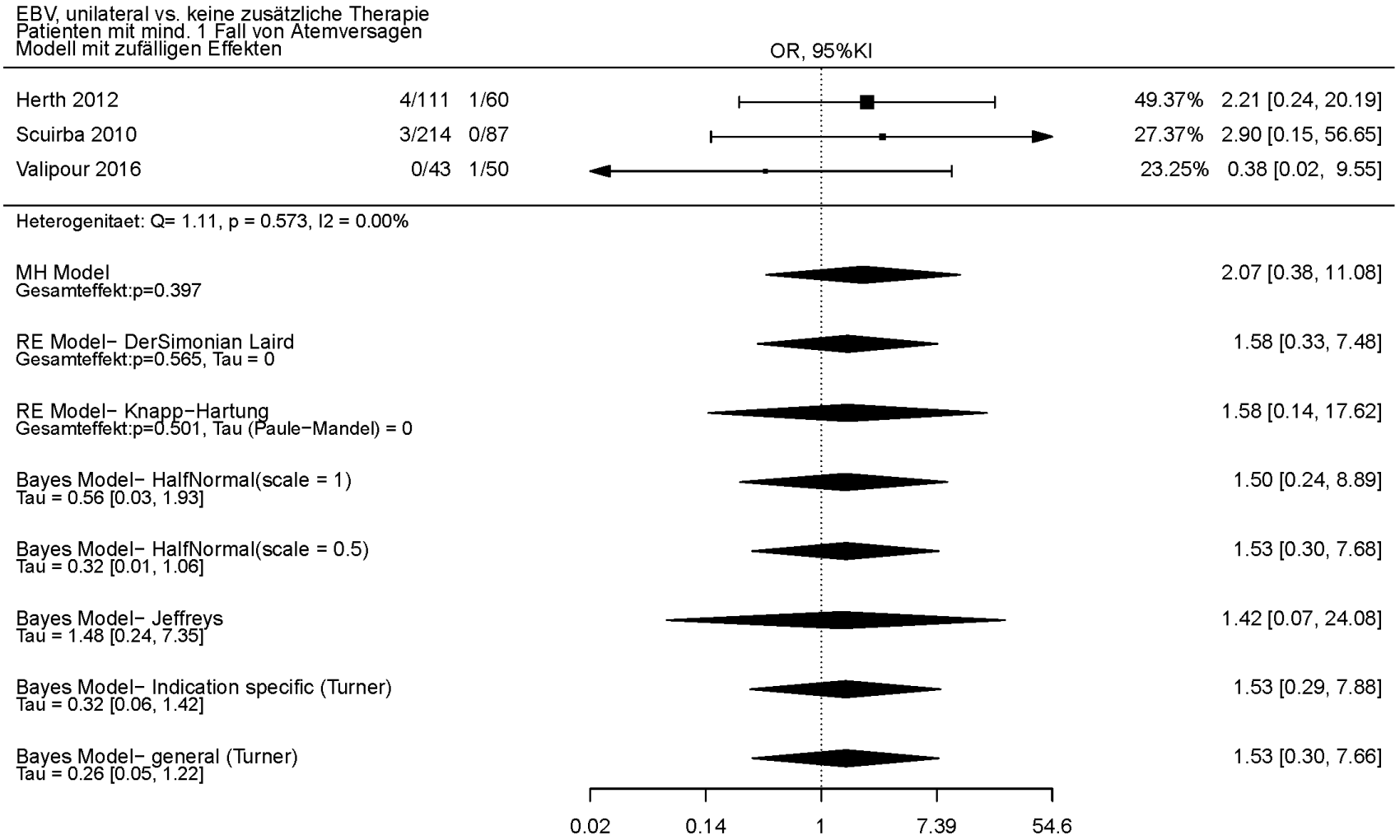
Prior and posterior distribution for τ



Indication specific (Turner): Endpunkt: infection / onset of new disease; Intervention: non−pharmacological; Vergleich: placebo / control

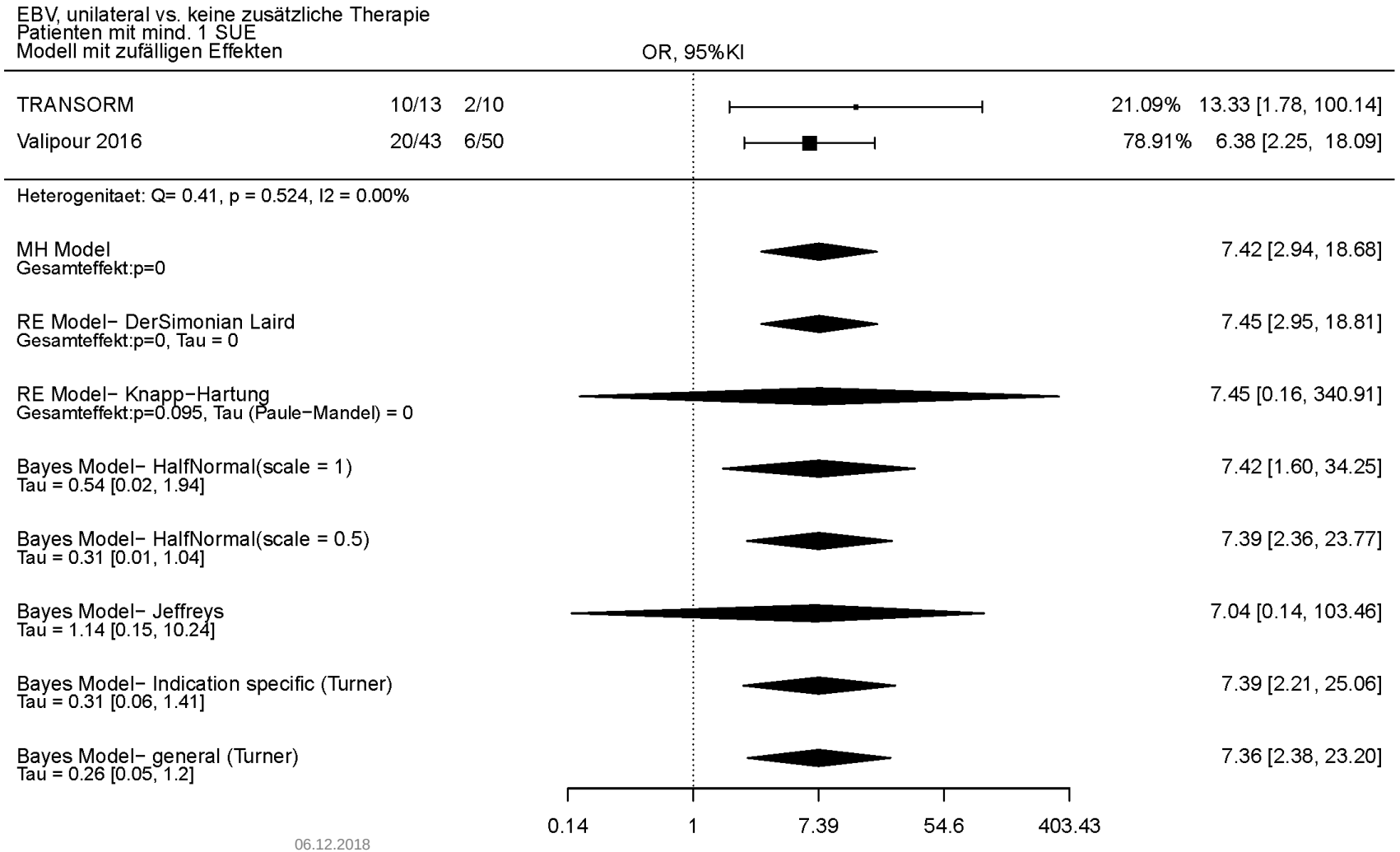
Few studies:

different prior choices may lead to comparable results



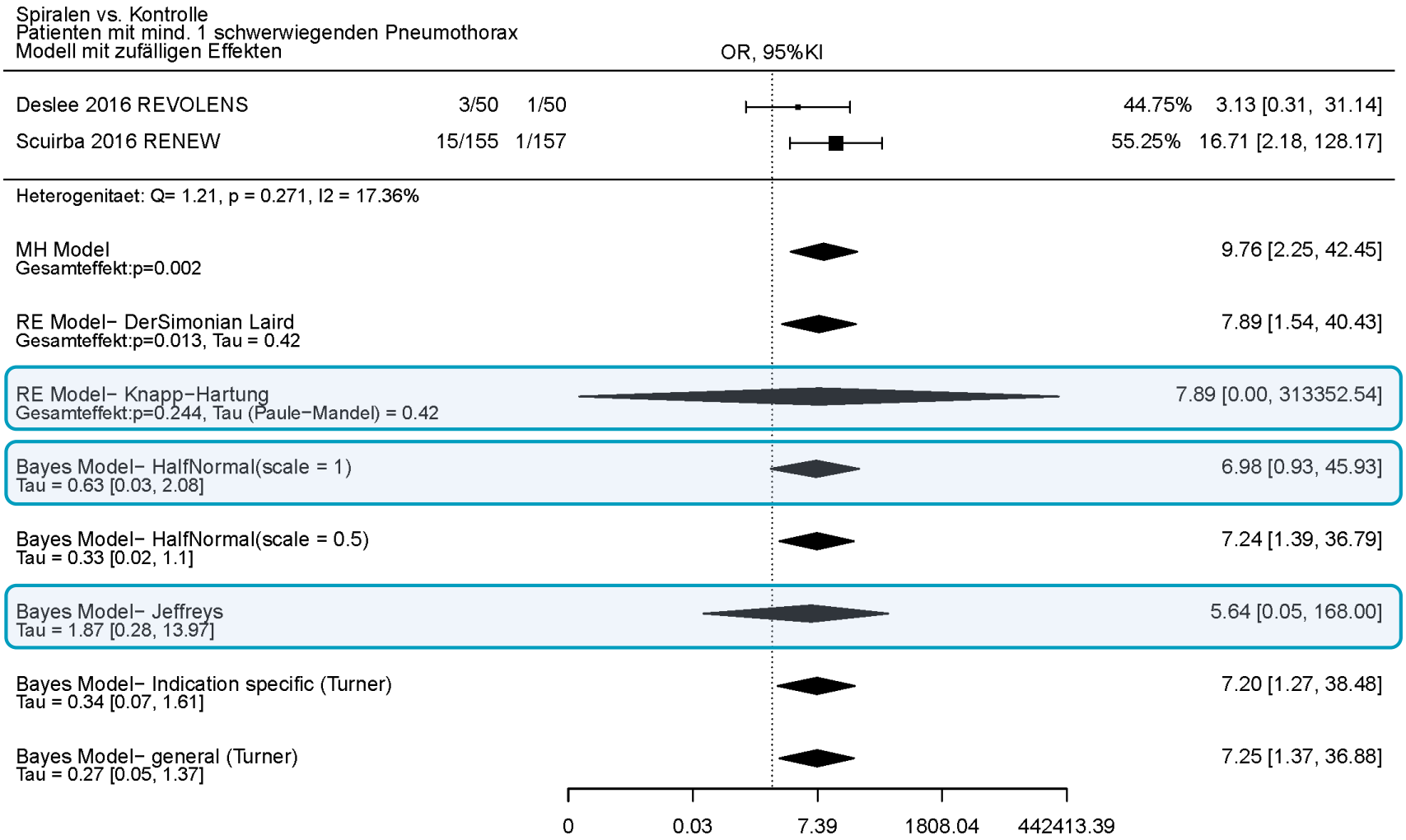
Few studies:

different prior choices may lead to comparable results… but also to divergent results



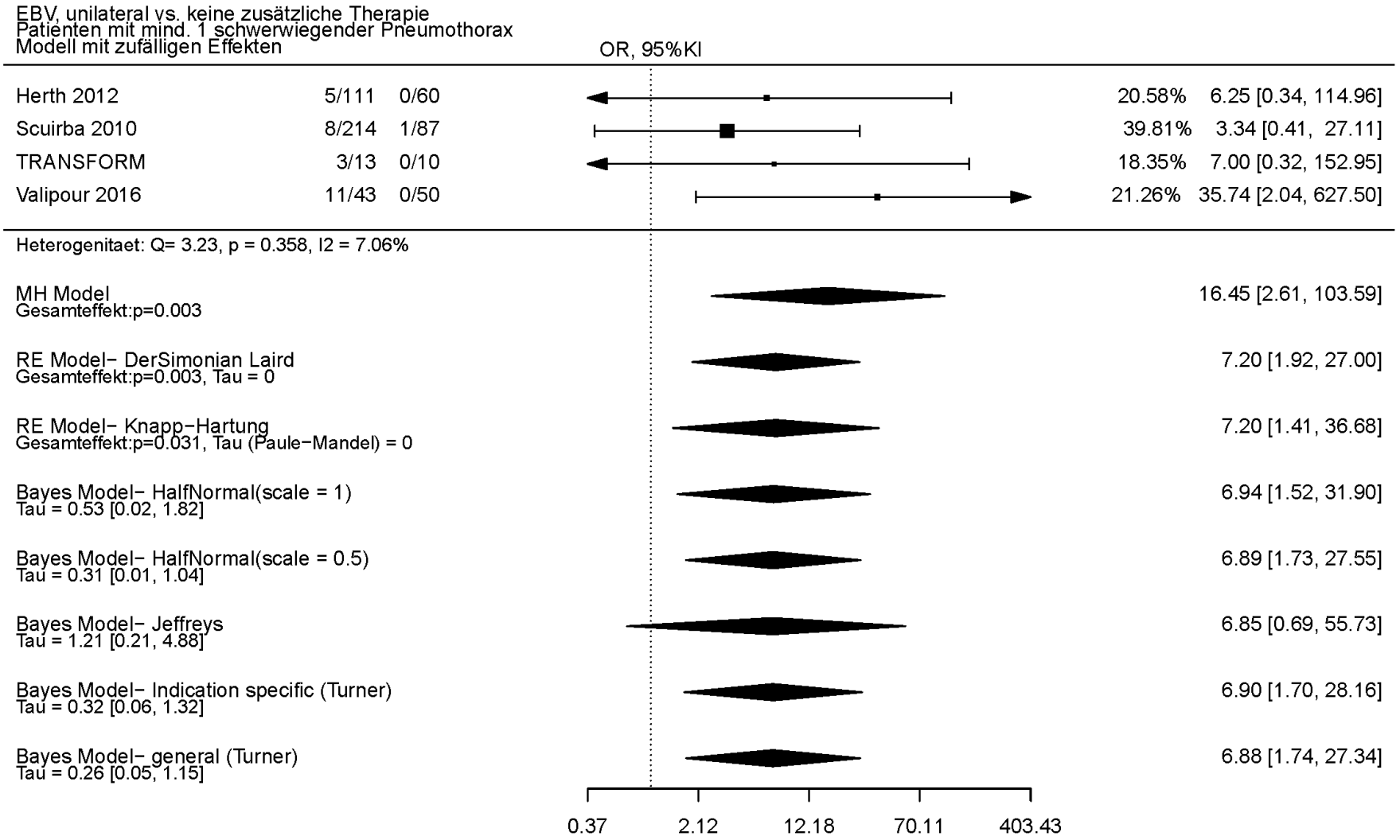
Few studies:

different prior choices may lead to comparable results… but also to divergent results



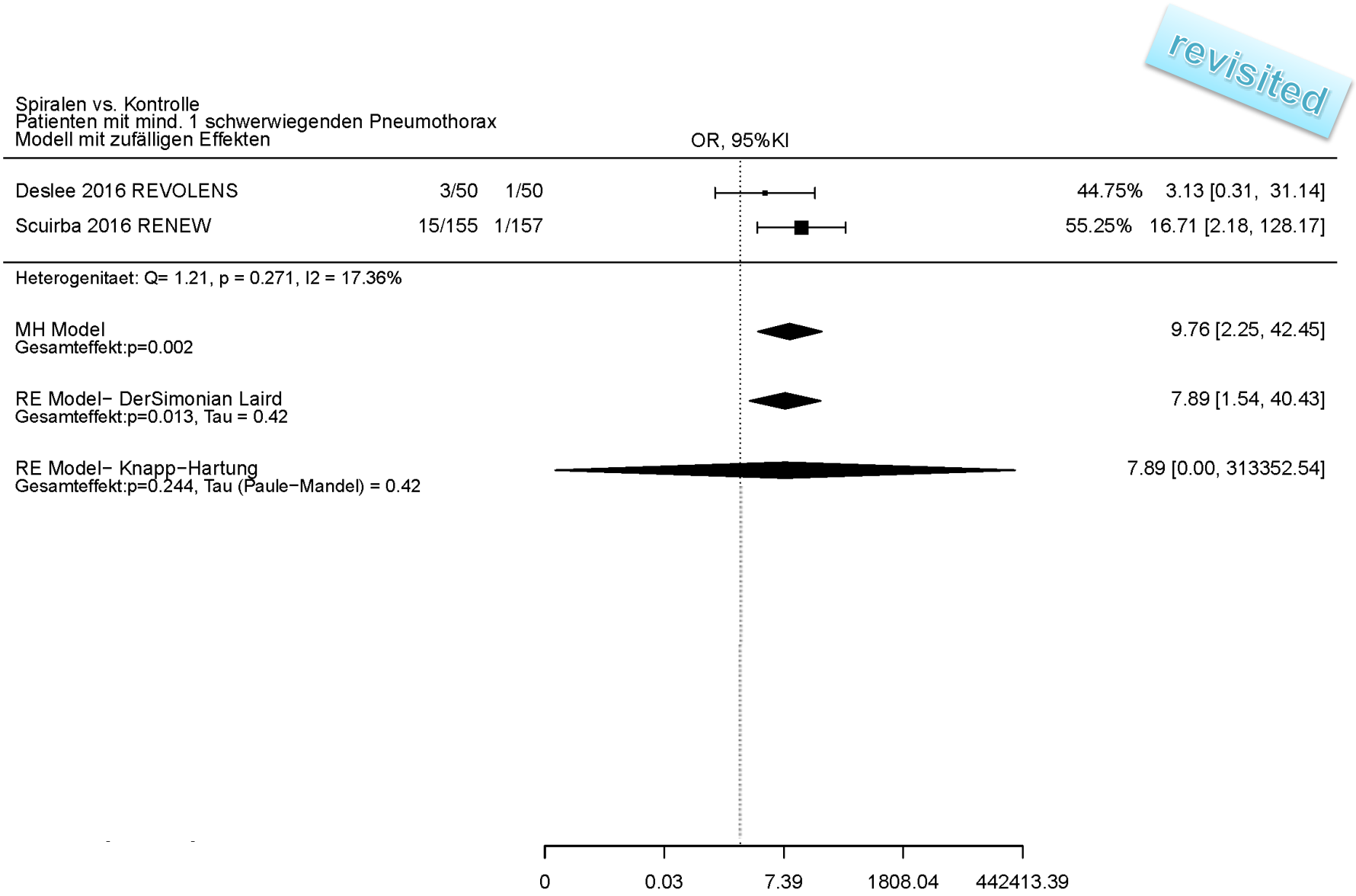
Few studies:

different prior choices may lead to comparable results… but also to divergent results may be even wider than Knapp-Hartung estimation



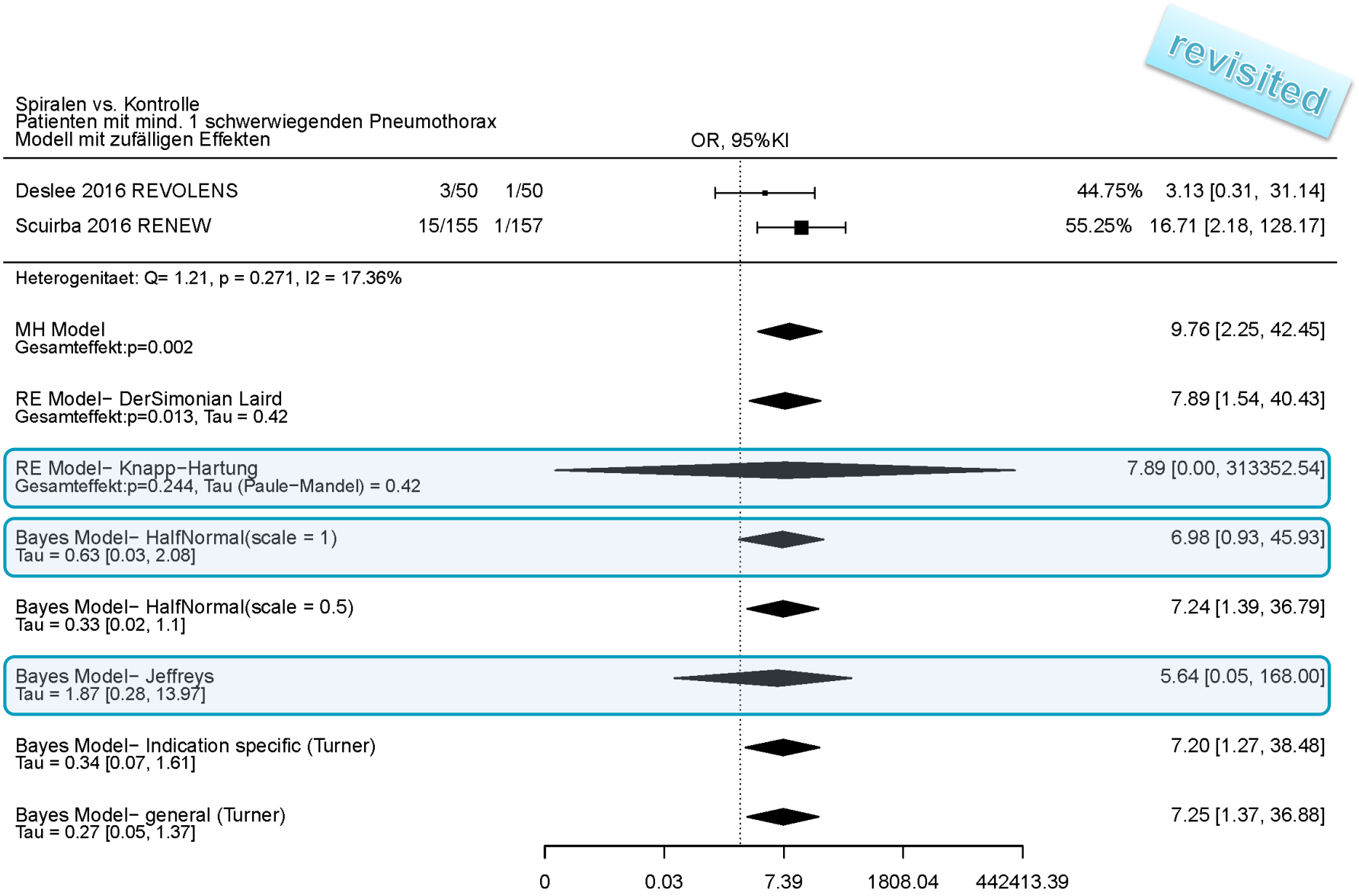
Few studies:

different prior choices may lead to comparable results… but also to divergent results

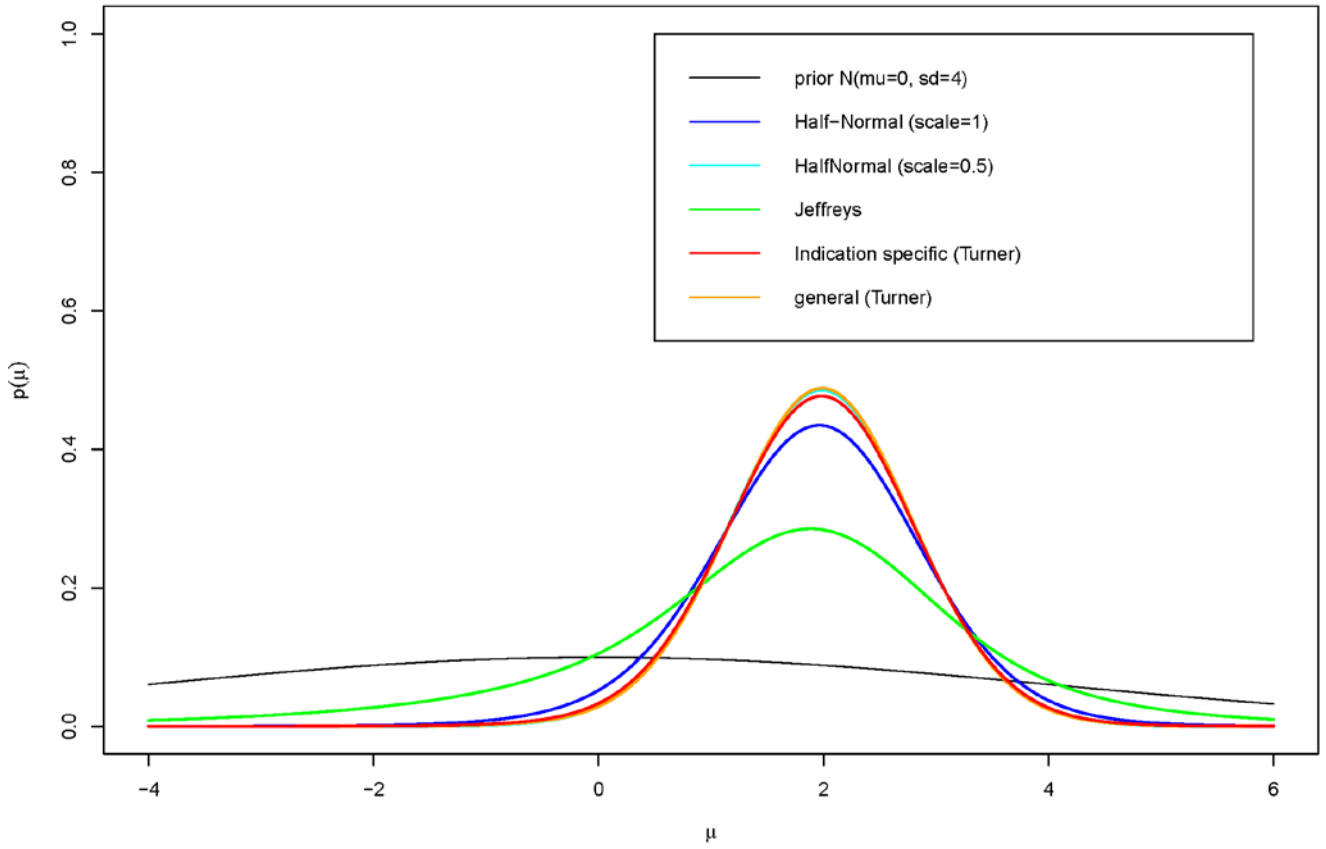


Few studies:

different prior choices may lead to comparable results… but also to divergent results



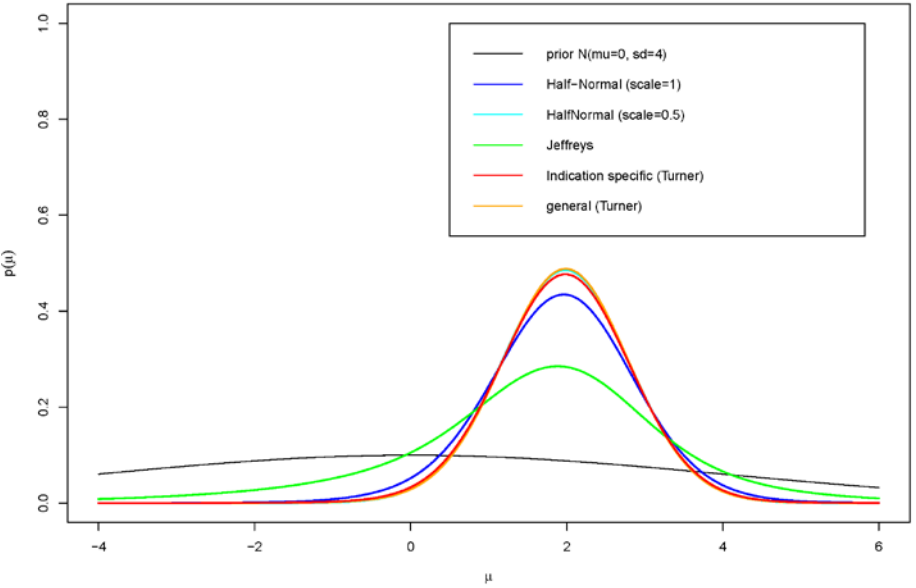
Prior distribution N(mu=0, sd=4) and posterior distribution for μ



Posterior distribution for μ given different prior distributions

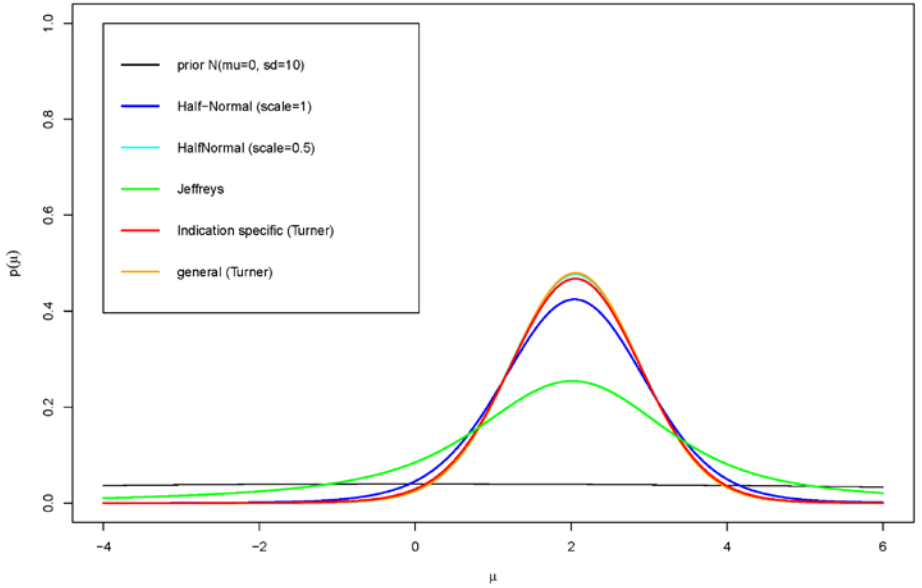
$\mu \sim N(0, sd = 4)$

Different settings for τ- prior



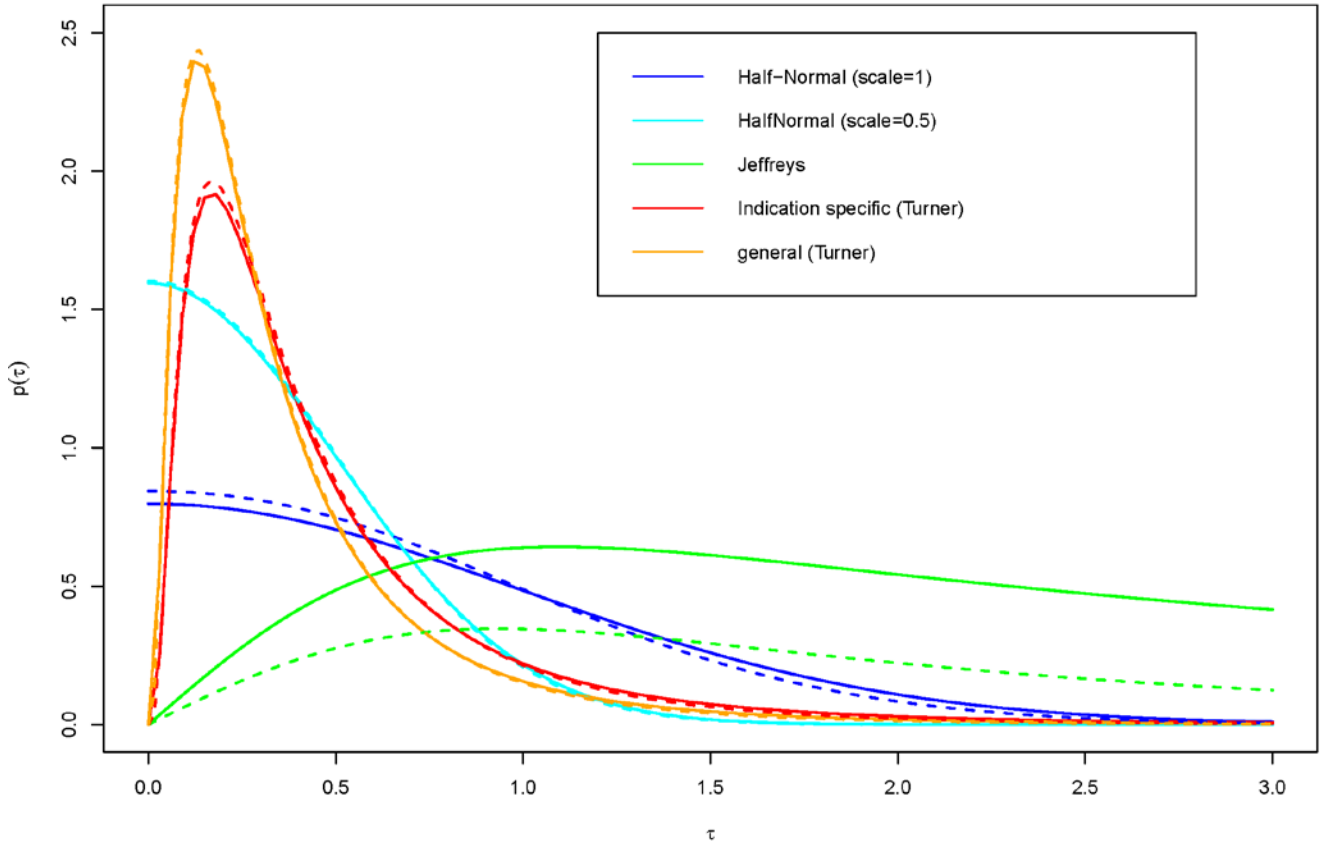
$\mu \sim N(0, sd = 10)$

Different settings for τ- prior



⇨ Almost identical results regardless the chosen prior variance for μ

Prior and posterior distribution for τ

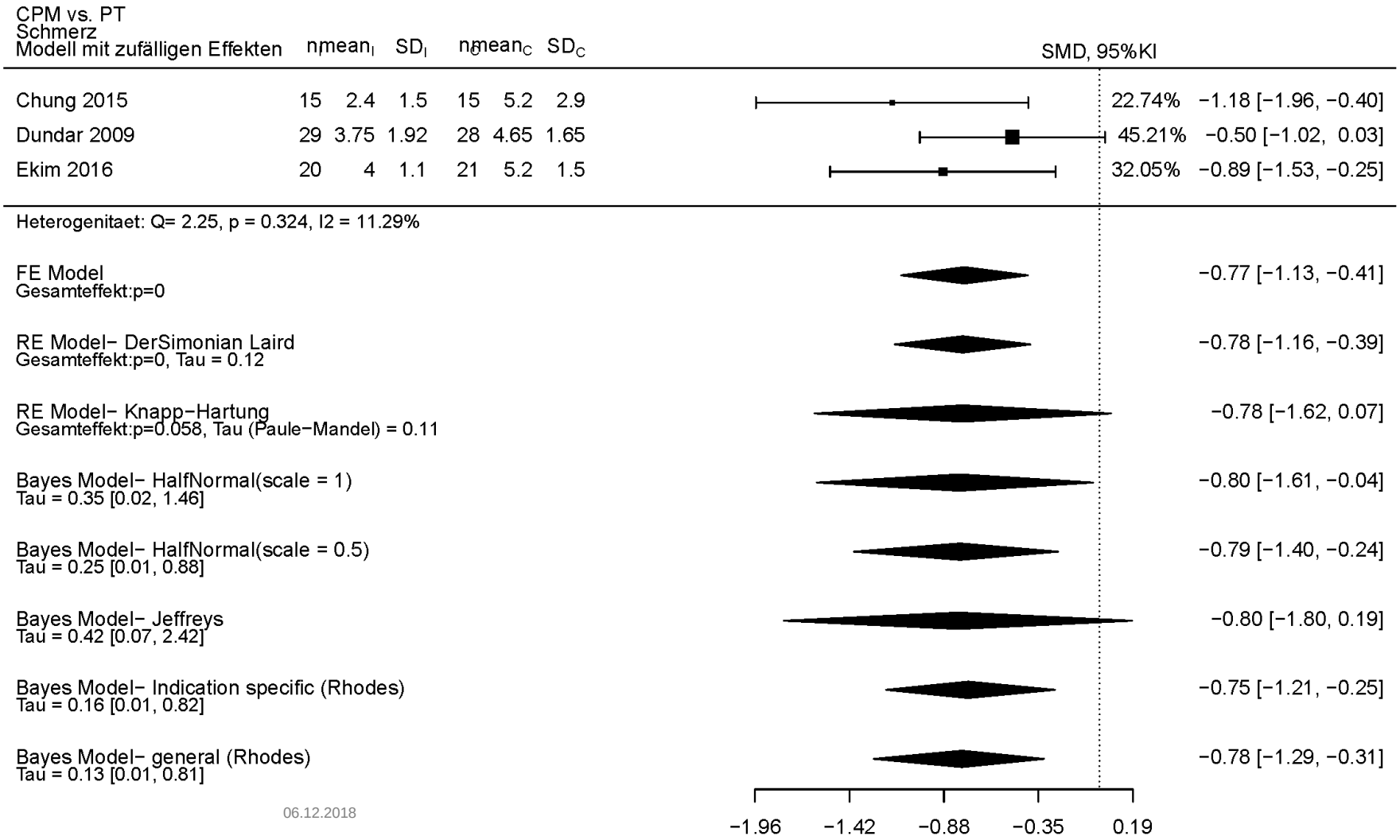


Indication specific (Turner): Endpunkt: adverse events; Intervention: non−pharmacological; Vergleich: placebo / control

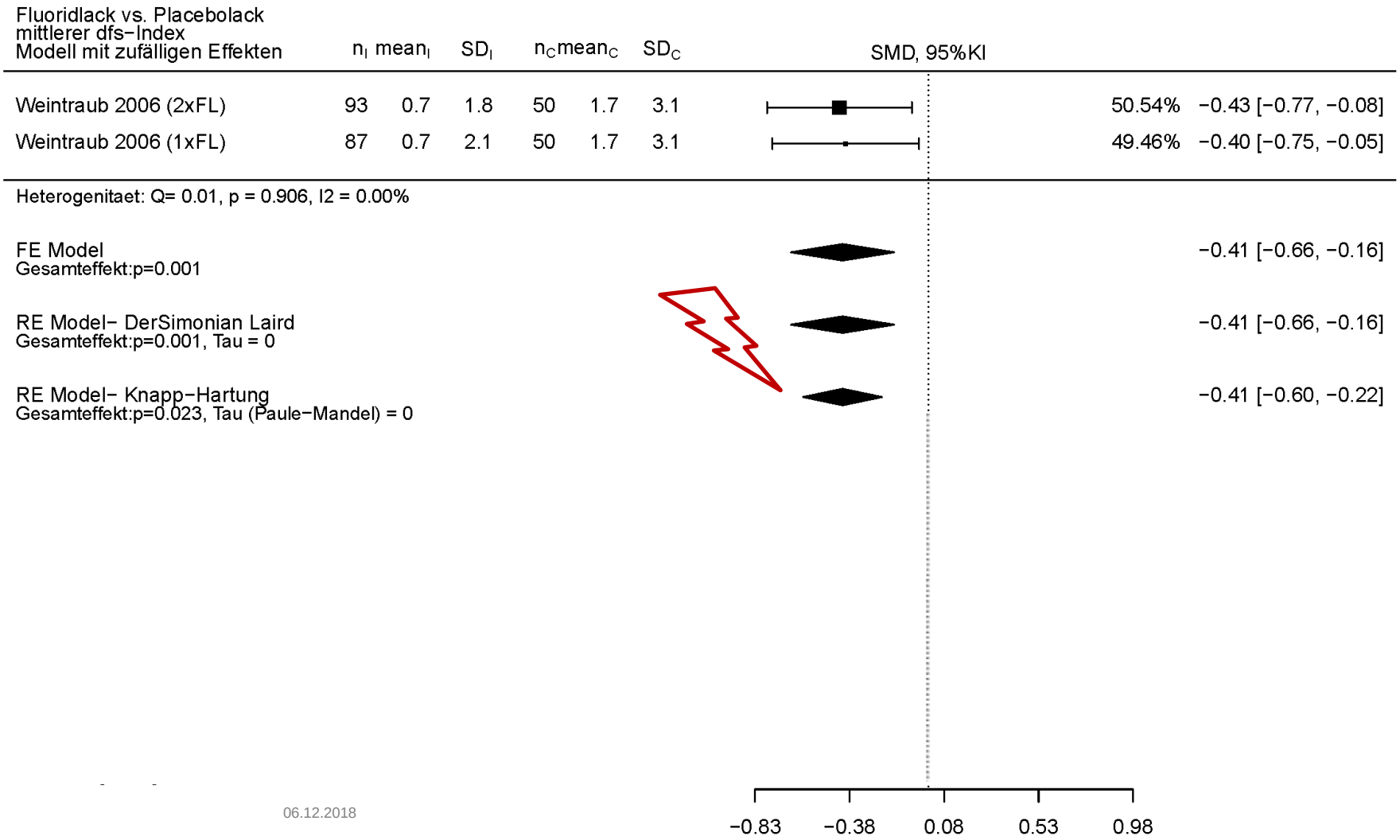
Continuous data

- Measure: standardised mean difference (SMD)
- Frequentist models: IV, DSL, KH
- Bayesian models: Prior for μ: Normal(0,4)

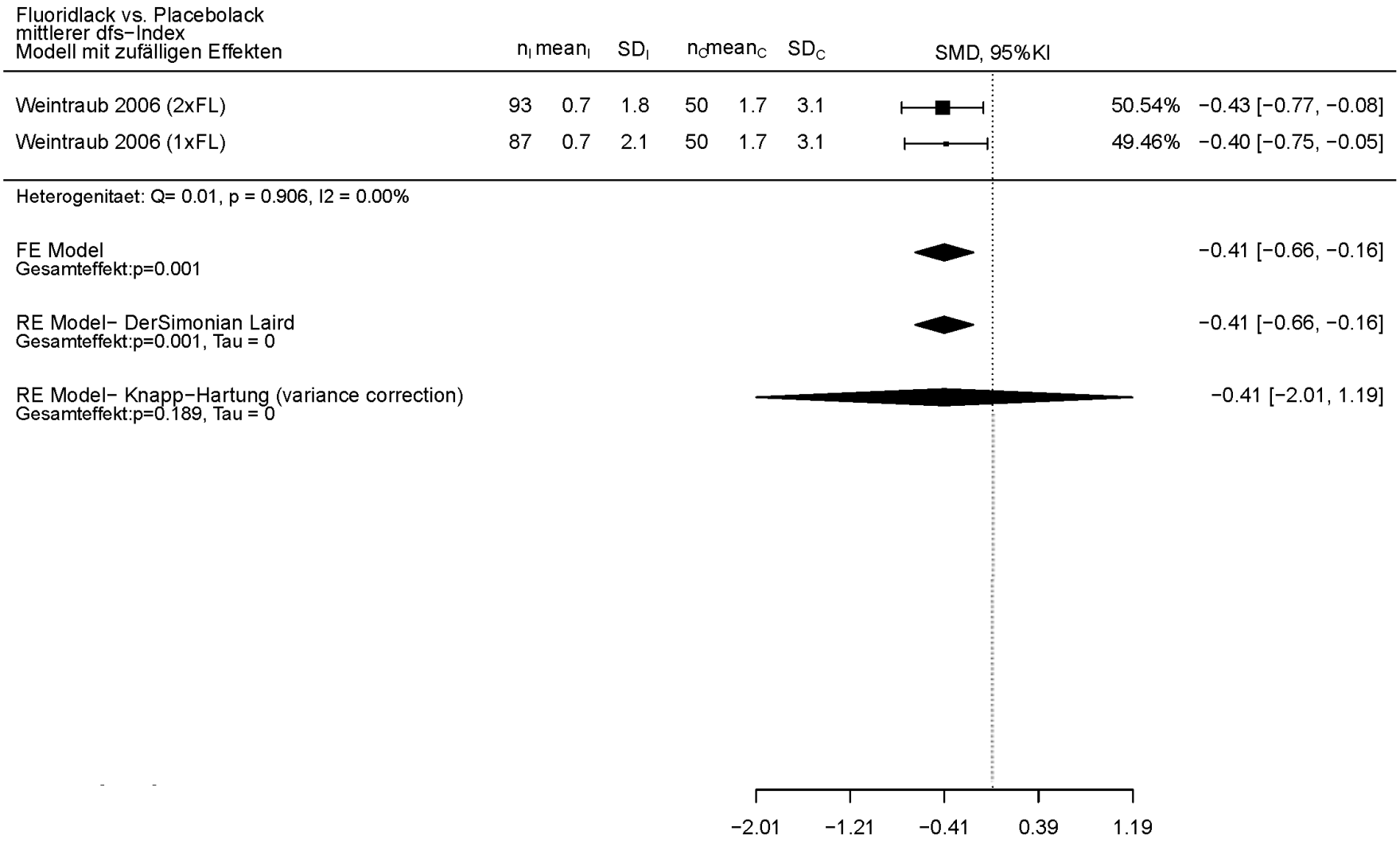
  Prior for τ :
  - half-Normal (scale=1)
  - half-Normal(scale=0.5)
  - Jeffreys prior
  - Rhodes, indication specific
  - Rodes, general
- Estimation for Bayesian models: posterior median and 95% credible interval for the effect μ and τ; central, equal-tailed credible intervals


CPM vs. PT
Schmerz
Modell mit zufälligen Effekten
n_I mean_I SD_I
n_C mean_C SD_C
SMD, 95%KI
Chung 2015 15 2.4 1.5 15 5.2 2.9 22.74% −1.18 [−1.96, −0.40]
Dundar 2009 29 3.75 1.92 28 4.65 1.65 45.21% −0.50 [−1.02, 0.03]
Ekim 2016 20 4 1.1 21 5.2 1.5 32.05% −0.89 [−1.53, −0.25]
Heterogenitaet: Q= 2.25, p = 0.324, I2 = 11.29%
FE Model
Gesamteffekt:p=0
−0.77 [−1.13, −0.41]
RE Model− DerSimonian Laird
Gesamteffekt:p=0, Tau = 0.12
−0.78 [−1.16, −0.39]
RE Model− Knapp−Hartung
Gesamteffekt:p=0.058, Tau (Paule−Mandel) = 0.11
−0.78 [−1.62, 0.07]
Bayes Model− HalfNormal(scale = 1)
Tau = 0.35 [0.02, 1.46]
−0.80 [−1.61, −0.04]
Bayes Model− HalfNormal(scale = 0.5)
Tau = 0.25 [0.01, 0.88]
−0.79 [−1.40, −0.24]
Bayes Model− Jeffreys
Tau = 0.42 [0.07, 2.42]
−0.80 [−1.80, 0.19]
Bayes Model− Indication specific (Rhodes)
Tau = 0.16 [0.01, 0.82]
−0.75 [−1.21, −0.25]
Bayes Model− general (Rhodes)
Tau = 0.13 [0.01, 0.81]
−0.78 [−1.29, −0.31]
−1.96 −1.42 −0.88 −0.35 0.19

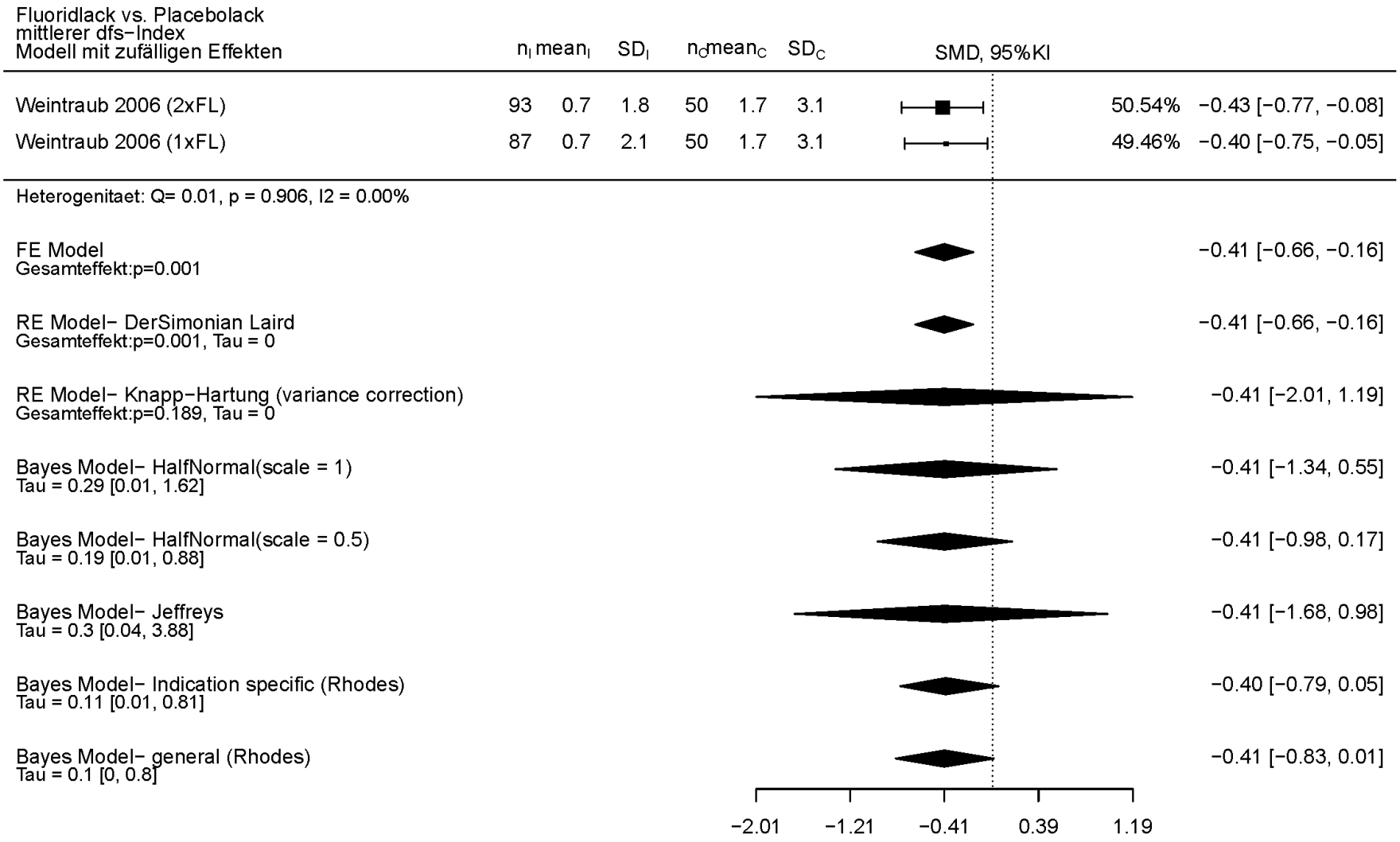
Meta-analysis of 2 significant studies



Meta-analysis of 2 significant studies



Meta-analysis of 2 significant studies



Alternative priors for τ

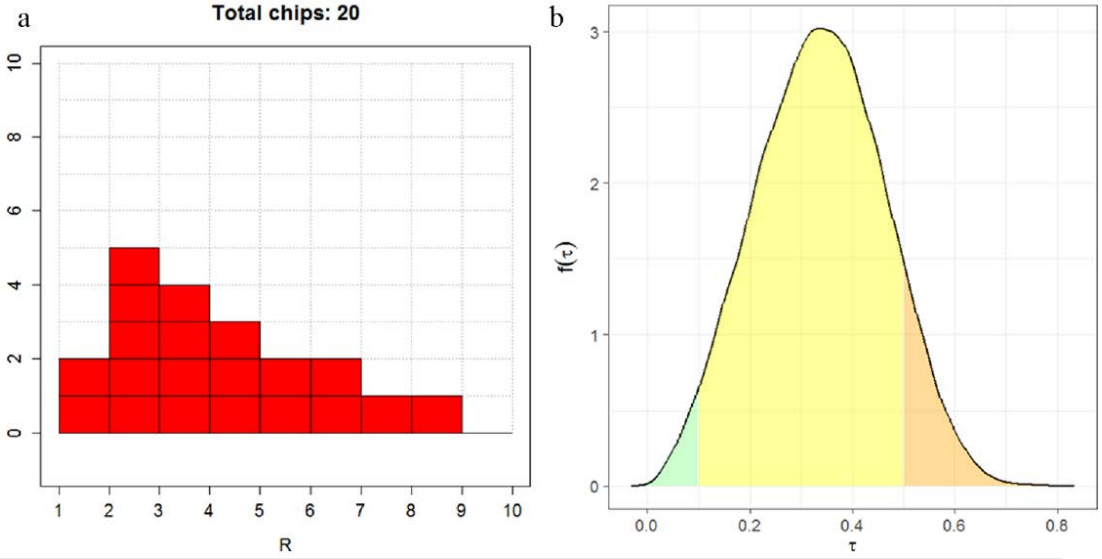
- Use informative priors for τ if solid information about trial heterogeneity is available

(Friede et al., RSM, 2017)

- Elicitate informative prior distribution for τ from expert‘s opinion (Ren et al., 2018)
  - Need for a random effect model?
  - Upper bound for $\mathrm{R} = OR_{97.5}/OR_{2.5}$
  - Consideration of a full distribution of R using a roulette elicitation method
  - Identify best fitting distribution
  - For log-Normal distribution:

$$\log(R-1) \sim Normal(m, v)$$

$$\tau = \log(R)\,/3.92$$



Shifted null hypotheses

- Confidence intervals are used for different decisions
  - Statistical significance to assess the benefit or harm
  - In dossier assessment also to determine the extend of the effect

⇨ confidence interval must lie completely below a certain threshold

  - For binary endpoints, RR

		Outcome category		
		All-cause mortality	Serious (or severe) symptoms (or late complications) and adverse events, as well as health-related quality of life[a]	Non-serious (or non-severe) symptoms (or late complications) and adverse events
Extent category	Major	0.85	0.75 and risk ≥ 5%[b]	Not applicable
	Considerable	0.95	0.90	0.80
	Minor	1.00	1.00	0.90

a: Precondition (as for all patient-reported outcomes): use of a validated or established instrument, as well as a validated or established response criterion.

b: Risk must be at least 5% for at least 1 of the 2 groups compared.

IQWiG, 2017

Summary

- For estimation of μ prior information has only little influence
- The less studies are included in meta analysis, the higher is the influence of the prior for τ
- For only two studies τ posterior equals almost the chosen prior

- This also holds when noninformative or vaguely informative priors are employed

⇨ Particularly the use of prior distribution should be handled with caution, sensitivity analyses for the choice of prior are required also when vague prior distributions are used.

⇨ Prespecification of methods as well as effect measures is important

References

- Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (2018, 10.07.2017). "General methods: version 5.0." Retrieved 06.06.2018, from https://www.iqwig.de/download/General-Methods_Version-5-0.pdf.
- Bodnar, O., et al. (2017). "Bayesian estimation in random effects meta-analysis using a non-informative prior." Statistics in Medicine **36**(2): 378-399.
- The Cochrane Collaboration (2011). “Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0” [updated March 2011]. J. P. T. Higgins and S. Green. http://www.cochrane-handbook.org/,.
- DerSimonian, R. and N. M. Laird (1986). "Meta-analysis in clinical trials." Controlled Clinical Trials **7**(3): 177-188.
- Friede, T., et al. (2017). "Meta-analysis of two studies in the presence of heterogeneity with applications in rare diseases." Biometrical Journal **59**(4): 658-671.
- Friede, T., et al. (2017). "Meta-analysis of few small studies in orphan diseases." Research Synthesis Methods 8(1): 79-91.
- Lunn, D., et al. (2013). The BUGS Book : A Practical Introduction to Bayesian Analysis. Boca Raton, FL, CRC Press.
- Ren, S., et al. (2018): “Incorporating genuine prior information about between-study heterogeneity in random effects pairwise and network meta-analyses.” Medical Decision Making **38**, 531-542.
- Rhodes, K. M., et al. (2018). "Between-trial heterogeneity in meta-analyses may be partially explained by reported design characteristics." Journal of Clinical Epidemiology **95**: 45-54.
- Röver, C (2017). Bayesian random-effects meta-analysis using the bayesmeta R package. arXiv preprint 1711.08683, URL: http://www.arxiv.org/abs/1711.08683
- Sutton, A. J., et al. (2000). Methods for Meta-Analysis in Medical Research. Chichester ; New York, J. Wiley.
- Turner, N. L., et al. (2015). "A Bayesian framework to account for uncertainty due to missing binary outcome data in pairwise meta-analysis." Statistics in Medicine **34**(12): 2062-2080.
- Veroniki, A. A., et al. (2015). "Recommendations for quantifying the uncertainty in the summary intervention effect and estimating the between-study heterogeneity variance in random-effectsmeta-analysis." Cochrane Database of Systematic Reviews(1): 1-72.
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, **36**(3), 1–48. URL: http://www.jstatsoft.org/v36/i03/.
- Weber, K., et al. (2018). "How to use prior knowledge and still give new data a chance?" Pharmaceutical Statistics 17(4): 329-341.

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)

Im Mediapark 8
50670 Köln

Telefon +49 221 35685-0
Telefax +49 221 35685-1

info@iqwig.de

www.iqwig.de
www.gesundheitsinformation.de
www.themencheck-medizin.de

Twitter: @iqwig und @iqwig_gi