# Bayesian hypothesis testing and estimation under the marginalized random-effects meta-analysis model

Robbie C.M. van Aert and Joris Mulder

September 30, 2021

European Research Council
Established by the European Commission

TILBURG
UNIVERSITY

# Bayesian meta-analysis

- Meta-analysis literature mainly focused on empirical Bayes and fully Bayesian estimation

- Bayes factors can be used for Bayesian hypothesis testing

- A Bayes factor quantifies the evidence for one model relative to a contrasting model

$$B_{12} = \frac{m_1(\mathbf{y})}{m_2(\mathbf{y})}$$

# Bayesian meta-analysis

- Meta-analysis literature mainly focused on empirical Bayes and fully Bayesian estimation
- Bayes factors can be used for Bayesian hypothesis testing
- A Bayes factor quantifies the evidence for one model relative to a contrasting model

$$B_{12} = \frac{m_1(\mathbf{y})}{m_2(\mathbf{y})}$$

- Existing meta-analytic Bayes factors either focus on a single parameter or are effect size measure dependent [1–4]
- **Goal:** Proposing a methodology for Bayesian estimation *and* hypothesis testing that can be used for any effect size measure

# MAREMA model

- We use the marginalized random-effects meta-analysis (MAREMA) model,

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

- The MAREMA model encompasses three meta-analysis models:
  - Equal-effect model → zero between-study variance
  - Random-effects model → positive between-study variance
  - Model with a negative between-study variance

- A negative between-study variance is not uncommon [5] and may be caused by chance or dependencies among the studies

# Estimation: Prior distributions

- A prior distribution is not placed on $\tau^2$ but on the $I^2$-statistic $\rightarrow I^2 = \tau^2/(\tau^2 + \tilde{\sigma}^2)$

- Reparameterizing the MAREMA model using the $I^2$-statistic and replacing it with $\rho$ yields

$$y_i \sim N\Big(\mu, \sigma_i^2 + \tilde{\sigma}^2 \rho/(1-\rho)\Big)$$

- The smallest possible value of $\rho$ is a function of the smallest sampling variance (i.e., $\sigma_{min}^2$)

$$\rho_{min} = \frac{-\sigma_{min}^2}{-\sigma_{min}^2 + \tilde{\sigma}^2}$$

# Estimation: Prior distributions

- Flat prior distributions are used:

$$\begin{aligned} \pi(\mu, \rho) &= \pi(\mu)\pi(\rho), \text{with} \\ \pi(\mu) &\propto 1 \\ \pi(\rho) &= U(\rho_{min}, 1) \end{aligned}$$

- Posterior distributions are obtained using a Gibbs sampler

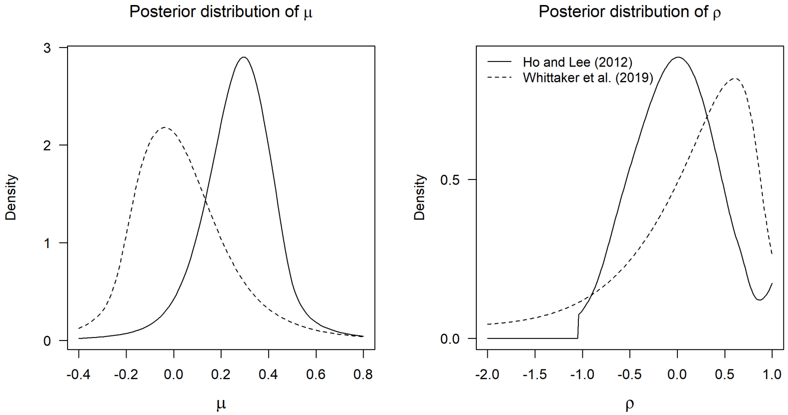
# Estimation: Prior distributions

- Flat prior distributions are used:

$$\begin{aligned} \pi(\mu,\rho) &= \pi(\mu)\pi(\rho), \text{with} \\ \pi(\mu) &\propto 1 \\ \pi(\rho) &= U(\rho_{min}, 1) \end{aligned}$$

- Posterior distributions are obtained using a Gibbs sampler
- Illustrating estimation using two examples:
    - Ho et al. [6] contains 10 standardized mean differences on the efficacy of EMDR vs. CBT therapy to treat PTSD
    - Whittaker et al. [7] contains 3 log risk ratio on the difference between using a smartphone app and lower intensity support to quit smoking

# Application: Posterior distributions



# Application: Parameter estimates

**Ho et al.:** [6]

| | $\mu$ | | $\rho$ | |
|---|---|---|---|---|
| | Estimate | 95% CI/CrI | Estimate | 95% CI/CrI |
| MAREMA | 0.274 (0.327) | (-0.109;0.638) | -0.026 (-0.016) | (-0.837;0.812) |
| Frequentist | 0.249 | (-0.003;0.502) | 0.022 | (0;0.747) |

**Whittaker et al.:** [7]

| | $\mu$ | | $\rho$ | |
|---|---|---|---|---|
| | Estimate | 95% CI/CrI | Estimate | 95% CI/CrI |
| MAREMA | 0.033 (0.043) | (-0.413;0.625) | 0.089 (0.597) | (-1.752;0.922) |
| Frequentist | 0.114 | (-0.525;0.753) | 0.696 | (0;0.993) |

# Bayes factors: Prior distributions

- In the two examples, we test these hypotheses:

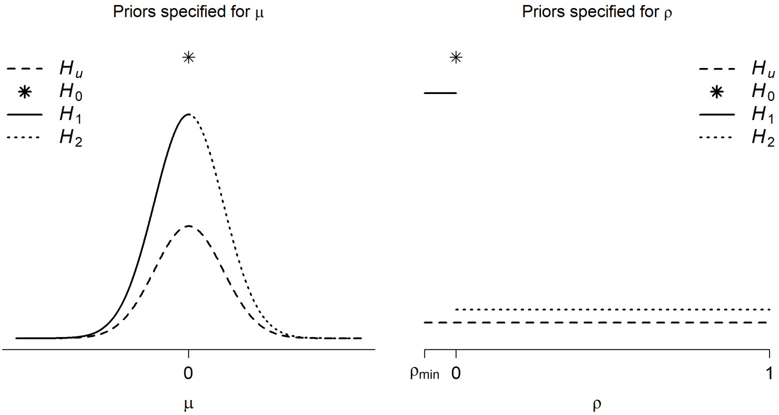
$$H_0 : \mu = 0 \qquad\qquad H_0 : \rho = 0$$
$$H_1 : \mu < 0 \qquad\qquad H_1 : \rho < 0$$
$$H_2 : \mu > 0 \qquad\qquad H_2 : \rho > 0$$

- A proper prior is needed for Bayes factors, so we cannot use the flat prior for $\mu$

- We propose a unit-information prior for $\mu$ and a uniform prior for $\rho$ under the unconstrained MAREMA model:

$$\pi_u(\mu, \rho) = \pi_u(\mu|\rho)\pi_u(\rho), \text{with}$$
$$\pi_u(\mu|\rho) = N(\mu, k(\mathbf{1}'{\textstyle\sum_\rho}^{-1}\mathbf{1})^{-1})$$
$$\pi(\rho) = U(\rho_{min}, 1)$$

# Bayes factors: Prior distributions



# Bayes factors: Computation

- Marginal likelihoods of the different hypotheses are needed to compute the Bayes factor

- For example, the marginal likelihood of $H_1 : \mu < 0$ is

$$m_1(\mathbf{y}) = \iint_{\mu<0} f(\mathbf{y}|\mu, \rho)\pi_1(\mu, \rho) d\mu d\rho$$

- Marginal likelihoods were approximated using importance sampling or a random walk procedure

# Application: Bayes factors

**Ho et al.:** [6]

| | $\mu$ | | | $\rho$ | | |
|---|---|---|---|---|---|---|
| | $H_0$ | $H_1$ | $H_2$ | $H_0$ | $H_1$ | $H_2$ |
| $H_0$ | 1.000 | 4.183 | 0.265 | 1.000 | 3.977 | 4.979 |
| $H_1$ | 0.239 | 1.000 | 0.063 | 0.251 | 1.000 | 1.252 |
| $H_2$ | 3.779 | 15.810 | 1.000 | 0.201 | 0.799 | 1.000 |
| $P(H_q \mid \mathbf{y})$ | 0.199 | 0.048 | 0.753 | 0.689 | 0.173 | 0.138 |

*Note:* $H_0 : \mu = 0$; $H_1 : \mu < 0$; $H_2 : \mu > 0$

- $H_2 : \mu > 0$ is most likely compared to $H_0$ and $H_1$
- Frequentist test: $z = 1.936$, $p = 0.053$

# Application: Bayes factors

**Ho et al.:** [6]

|  | $\mu$ |  |  | $\rho$ |  |  |
|---|---|---|---|---|---|---|
|  | $H_0$ | $H_1$ | $H_2$ | $H_0$ | $H_1$ | $H_2$ |
| $H_0$ | 1.000 | 4.183 | 0.265 | 1.000 | 3.977 | 4.979 |
| $H_1$ | 0.239 | 1.000 | 0.063 | 0.251 | 1.000 | 1.252 |
| $H_2$ | 3.779 | 15.810 | 1.000 | 0.201 | 0.799 | 1.000 |
| $P(H_q\|\mathbf{y})$ | 0.199 | 0.048 | 0.753 | 0.689 | 0.173 | 0.138 |

*Note:* $H_0 : \rho = 0$; $H_1 : \rho < 0$; $H_2 : \rho > 0$

- $H_0 : \rho = 0$ is most likely compared to $H_1$ and $H_2$
- Frequentist test: $Q(9) = 9.417$, $p = 0.400$

# Application: Bayes factors

**Whittaker et al.:** [7]

| | $\mu$ | | | $\rho$ | | |
|---|---|---|---|---|---|---|
| | $H_0$ | $H_1$ | $H_2$ | $H_0$ | $H_1$ | $H_2$ |
| $H_0$ | 1.000 | 2.558 | 2.115 | 1.000 | 10.958 | 2.901 |
| $H_1$ | 0.391 | 1.000 | 0.827 | 0.091 | 1.000 | 0.265 |
| $H_2$ | 0.473 | 1.209 | 1.000 | 0.345 | 3.778 | 1.000 |
| $P(H_q\|\mathbf{y})$ | 0.537 | 0.210 | 0.254 | 0.696 | 0.064 | 0.240 |

*Note:* $H_0 : \mu = 0$; $H_1 : \mu < 0$; $H_2 : \mu > 0$

- $H_0 : \mu = 0$ is most likely compared to $H_1$ and $H_2$ but no strong evidence
- Frequentist test: $z = 0.349$, $p = 0.727$

# Application: Bayes factors

**Whittaker et al.:** [7]

| | $\mu$ | | | $\rho$ | | |
|---|---|---|---|---|---|---|
| | $H_0$ | $H_1$ | $H_2$ | $H_0$ | $H_1$ | $H_2$ |
| $H_0$ | 1.000 | 2.558 | 2.115 | 1.000 | 10.958 | 2.901 |
| $H_1$ | 0.391 | 1.000 | 0.827 | 0.091 | 1.000 | 0.265 |
| $H_2$ | 0.473 | 1.209 | 1.000 | 0.345 | 3.778 | 1.000 |
| $P(H_q \mid \mathbf{y})$ | 0.537 | 0.210 | 0.254 | 0.696 | 0.064 | 0.240 |

*Note:* $H_0 : \rho = 0$; $H_1 : \rho < 0$; $H_2 : \rho > 0$

- $H_0 : \rho = 0$ is most likely compared to $H_1$ and $H_2$
- Frequentist test: $Q(2) = 6.240$, $p = 0.044$

# Discussion

- The proposed Bayesian estimation and hypothesis testing is novel, because
  - It is based on the MAREMA model
  - A prior is placed on $\rho$ (i.e., $I^2$-statistic) rather than on $\tau^2$
  - It does not depend on the effect size measure
- One-sided and point hypotheses were tested, but combined hypotheses can also be tested → $H : \mu > 0$ & $\rho > 0$
- Informative hypotheses can also be implemented
- Bayesian estimation and Bayes factors are included in the R package `BFpack` [8]

# Discussion

- Future research may focus on:
  - Extending the methodology to meta-regression models
  - Allowing for multiple outcomes per study and more complicated hierarchical structures
  - Taking uncertainty in the within-study variance into account
  - Studying to what extent the methodology gets distorted by publication bias

# Thank you for your attention

www.robbievanaert.com

www.metaresearch.nl

**Paper:**

Van Aert, R. C. M., & Mulder, J. (2021). Bayesian hypothesis testing and estimation under the marginalized random-effects meta-analysis model. Psychonomic Bulletin & Review. doi: 10.3758/s13423-021-01918-9

# References I

[1]
Berry SM. Understanding and testing for heterogeneity across 2×2 tables: Application to meta-analysis. Statistics in Medicine 1998;17:2353–69. https://doi.org/10.1002/(sici)1097-0258(19981030)17:20%3C2353::Aid-sim923%3E3.0.Co;2-y.

[2]
Rouder JN, Morey RD. A Bayes factor meta-analysis of Bems ESP claim. Psychonomic Bulletin & Review 2011;18:682–9.

[3]
Scheibehenne B, Gronau QF, Jamil T, Wagenmakers E-J. Fixed or random? A resolution through model averaging: Reply to Carlsson, Schimmack, Williams, and Bürkner (2017). Psychological Science 2017;28:1698–701. https://doi.org/10.1177/0956797617724426.

[4]
Gronau QF, Erp S van, Heck DW, Cesario J, Jonas KJ, Wagenmakers E-J. A Bayesian model-averaged meta-analysis of the power pose effect with informed and default priors: the case of felt power. Comprehensive Results in Social Psychology 2017;2:123–38. https://doi.org/10.1080/23743603.2017.1326760.

[5]
Ioannidis JP, Trikalinos TA, Zintzaras E. Extreme between-study homogeneity in meta-analyses could offer useful insights. Journal of Clinical Epidemiology 2006;59:1023–32. https://doi.org/10.1016/j.jclinepi.2006.02.013.

[6]
Ho MSK, Lee CW. Cognitive behaviour therapy versus eye movement desensitization and reprocessing for post-traumatic disorder - is it all in the homework then? Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology 2012;62:253–60. https://doi.org/10.1016/j.erap.2012.08.001.

[7]
Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y, Dobson R. Mobile phone text messaging and app-based interventions for smoking cessation. Cochrane Database of Systematic Reviews 2019. https://doi.org/10.1002/14651858.CD006611.pub5.

# References II

[8]
Mulder J, Williams DR, Gu X, Tomarken A, Boeing-Messing F, Olsson-Collentine A, et al. BFpack: Flexible Bayes factor testing of scientific theories in R. Journal of Statistical Software 2021.

[9]
Viechtbauer W. Conducting meta-analyses in R with the metafor package. Journal of Statistical Software 2010;36:1–48. https://doi.org/10.18637/jss.v036.i03.

# Software: BFpack

- Bayes factors and Bayesian estimation are included in the R package `BFpack` [8]

- `BF()` function only needs a fitted modeling object → object returned by a random-effects meta-analysis using `metafor` [9]:

```
res2 <- rma(yi = yi, vi = vi) # RE meta-analysis
BF(res2)
```

```
## Call:
## BF.rma.uni(x = res2)
##
## Bayesian hypothesis test
## Type: exploratory
## Object: rma.uni
## Parameter: between-study heterogeneity & effect size
## Method: Bayes factor using uniform prior for icc & unit information prior for effect
##
## Posterior probabilities:
##     Pr(=0) Pr(<0) Pr(>0)
## I^2  0.696  0.064  0.240
## mu   0.537  0.210  0.254
```