

Empirical Prior Distributions for Bayesian Meta-Analyses of Binary and Time to Event Outcomes

František Bartoš (f.bartos96@gmail.com; @BartosFra)

with Willem M. Otte, Quentin F. Gronau, Bram Timmers, Alexander Ly, Eric-Jan Wagenmakers

Outline

- Bayesian hypothesis testing
- Bayesian model-averaged meta-analysis
- Empirical prior distributions
- Example

Bayesian Hypothesis Testing

- Requires specification of both H₀ and H₁
 - vs. NHST requiring only H₀
- Bayes factors
 - Compares models via prior predictive performance



Bayesian Hypothesis Testing

Alternative hypothesis vs. null hypothesis

$$p(\text{data} \mid \mathcal{H}_0) = \int p(\text{data} \mid \theta_0, \mathcal{H}_0) p(\theta_0 \mid \mathcal{H}_0) \, \mathrm{d}\theta_0$$
$$p(\text{data} \mid \mathcal{H}_1) = \int p(\text{data} \mid \theta_1, \mathcal{H}_1) \, p(\theta_1 \mid \mathcal{H}_1) \, \mathrm{d}\theta_1$$



Bayesian Hypothesis Testing



$$BF_{10} = \frac{p(\text{data} \mid \mathcal{H}_1)}{p(\text{data} \mid \mathcal{H}_0)}$$

Bayesian Meta-Analyses (Normal-Normal Model)

Population mean effect μ Between-study heterogeneity τ
$$\begin{split} \mu &\sim f() \\ \tau &\sim g() \\ \mathbf{y}_i &\sim \mathrm{Normal}(\mu, \mathrm{se}_i^2 + \tau^2) \end{split}$$

Four possible models:

- 1. the fixed-effect null hypothesis \mathcal{H}_0^f : $\mu = 0$, $\tau = 0$
- 2. the fixed-effect alternative hypothesis \mathcal{H}_1^f : $\mu \sim f()$, $\tau = 0$
- 3. the random-effects null hypothesis \mathcal{H}_0^r : $\mu = 0, \tau \sim g()$
- 4. the random-effects alternative hypothesis \mathcal{H}_1^r : $\mu \sim f()$, $\tau \sim g()$





Hinne, M., Gronau, Q. F., van den Bergh, D., & Wagenmakers, E. J. (2020). A conceptual introduction to Bayesian model averaging. Advances in Methods and Practices in Psychological Science



Hinne, M., Gronau, Q. F., van den Bergh, D., & Wagenmakers, E. J. (2020). A conceptual introduction to Bayesian model averaging. Advances in Methods and Practices in Psychological Science



Hinne, M., Gronau, Q. F., van den Bergh, D., & Wagenmakers, E. J. (2020). A conceptual introduction to Bayesian model averaging. Advances in Methods and Practices in Psychological Science

Bayesian Model-Averaged Meta-Analysis

$$BF_{10} = \frac{p(\text{data} \mid \mathcal{H}_1)}{p(\text{data} \mid \mathcal{H}_0)}$$



Bayesian Model-Averaged Meta-Analysis



Gronau, Q. F., Heck, D. W., Berkhout, S. W., Haaf, J. M., & Wagenmakers, E. J. (2021). A primer on Bayesian model-averaged meta-analysis. *Advances in Methods and Practices in Psychological Science*, *4*(3) Bartoš, F., Gronau, Q. F., Timmers, B., Otte, W. M., Ly, A., & Wagenmakers, E. J. (2021). Bayesian model-averaged meta-analysis in medicine. *Statistics in Medicine*, *40*(30)



Bayesian Model-Averaged Meta-Analysis

Model-averaged posterior distributions:



Gronau, Q. F., Heck, D. W., Berkhout, S. W., Haaf, J. M., & Wagenmakers, E. J. (2021). A primer on Bayesian model-averaged meta-analysis. *Advances in Methods and Practices in Psychological Science*, *4*(3) Bartoš, F., Gronau, Q. F., Timmers, B., Otte, W. M., Ly, A., & Wagenmakers, E. J. (2021). Bayesian model-averaged meta-analysis in medicine. *Statistics in Medicine*, *40*(30)

Binomial-Normal Model

Population mean effect μ Between-study heterogeneity τ Random-effects γ Baseline probability β

	Events	Non-Events	N
Group 1	a	b	n ₁
Group 2	С	d	n ₂

 $\begin{aligned} \mathbf{a}_i &\sim \mathrm{Binomial}(\pi_{1,i},\mathbf{n}_{1,i}), \\ \mathbf{c}_i &\sim \mathrm{Binomial}(\pi_{2,i},\mathbf{n}_{2,i}), \\ \mathrm{logit}(\pi_{1,i}) &= \beta_i + \gamma_i/2, \\ \mathrm{logit}(\pi_{2,i}) &= \beta_i - \gamma_i/2, \\ \gamma_i &\sim \mathrm{Normal}(\mu,\tau), \end{aligned}$

Prior Distributions



\mathcal{H}_{o}^{r} H; Inclusion Bayes factors: YES data) + $p(\mathscr{H}_1^r)$ data) $p(\mathscr{H}_1^f) + p(\mathscr{H}_1^r)$ BF_{10} data Inclusion Bl for effect inclusion odds Posterior inclusion odds for effect for effect Posterior Probabilities $p(\mathcal{H}_0^r \mid \text{data}) + p(\mathcal{H}_1^r)$ $p(\mathscr{H}_0^r) + p(\mathscr{H}_1^r)$ data) BFri $p(\mathscr{H}_0^f) + p(\mathscr{H}_1^f)$ $p(\mathcal{H}_0^f)$ data) + $p(\mathcal{H}_1)$ data Inclusion BF Prior inclusion odds for heterogeneity Posterior inclusion odds for heterogeneity for heterogeneity Gronau, Q. F., Heck, D. W., Berkhout, S. W., Haaf, J. M., & Wagenmakers, E. J. (2021). A primer on Bayesian nodel-averaged meta-analysis, Advances in Methods and Practices in Psychological Science, 4(3)

Bayesian Model-Averaged Meta-Analysis

Model-averaged posterior distributions:

$\underbrace{p(\mu \mid \text{data}, \mathcal{H}_1)}_{\text{Model-averaged}}$	$\propto \underline{p(\mu \mid \mathcal{H}_1^f, \text{data})}$	$p \times \underline{p(\mathcal{H}_1^f \mid \text{data})}_{\text{Posterior distr.}}$	+ $\underbrace{p(\mu \mid \mathcal{H}_1^r, \text{data})}_{\text{Posterior prob.}}$	$\times \underbrace{p(\mathcal{H}_1^r \mid \text{data})}_{\text{Posterior distr.}}$	ţ	8 - 6 -				—— Rand —— Avera	lom H, aged H, H,
posterior distribution for effect	of \mathcal{H}_1^f	of the effect under \mathcal{H}_1^f	of \mathcal{H}_1^r	of the effect under \mathcal{H}_1^r	Dens	4 - 2 -					
$\underbrace{p(\tau \mid \text{data}, \mathcal{H}^r)}_{\text{Model-averaged}}_{\text{posterior distribution}}_{\text{for heterogeneity}}$	$\propto \underbrace{p(\mu \mid \mathcal{H}_0^r, \text{data})}_{\text{Posterior prob.}}$	$\times \underbrace{p(\mathcal{H}_0^r \mid \text{data})}_{\substack{\text{Posterior distr.}\\ \text{of heterogeneity}\\ \text{under } \mathcal{H}_0^r}$	$+\underbrace{p(\mu \mid \mathcal{H}_{1}^{r}, \text{data})}_{\text{Posterior prob.}}_{\text{of } \mathcal{H}_{1}^{r}}$	$\times \underbrace{p(\mathcal{H}_1^r \mid \text{data})}_{\substack{\text{Posterior distr.}\\ \text{of heterogeneity}\\ \text{under } \mathcal{H}_1^r}$		 	-0.1	0.0	0.1	0.2	0.3
								Effect	Size µ		

10 ¬

Fixed H.

Bayesian Model-Averaged Meta-Analysis

Training Data Set

Initial training data set: 49,247 comparisons (243,257 estimates)

Removing comparisons with fewer than 10 estimates.

6,334 comparisons (127,584 estimates)

Estimating frequentist random-effects meta-analytic models.

Final training data set: log OR: 6,281 comparisons (126,794 estimates) log RR: 6,215 comparisons (125,002 estimates) RD: 6,318 comparisons (127,275 estimates)



Prior Distributions for log OR



Prior Distributions for log HR



Test Data Set





Binary outcomes as log OR







Comparisons

Comparisons

Topic	Comparisons μ	Comparisons τ	Prior μ	Prior τ
Acute Respiratory Infections	308 (5860)	197 (4061)	Student-t(0, 0.48, 3)	Inv-Gamma(1.67, 0.45)
Airways	450 (8702)	207 (3839)	Student-t(0, 0.37, 2)	Inv-Gamma(1.35, 0.27)
Anaesthesia	408 (9409)	261 (6781)	Student-t(0, 0.79, 6)	Inv-Gamma(2.12, 0.86)
Back and Neck	14 (295)	10 (238)	Student-t(0, 0.62, 4)	Inv-Gamma(1.84, 0.68)
Bone, Joint and Muscle Trauma	202 (3732)	103 (1937)	Student-t(0, 0.44, 2)	Inv-Gamma(1.01, 0.36)
Breast Cancer	137 (2646)	113 (2296)	Student-t(0, 0.39, 3)	Inv-Gamma(1.59, 0.48)
Childhood Cancer	1 (13)	_	Student-t(0, 0.47, 4)	
Colorectal	230 (4120)	141 (2685)	Student-t(0, 0.65, 4)	Inv-Gamma(1.71, 0.60)
Common Mental Disorders	649 (13322)	422 (9731)	Student-t(0, 0.54, 4)	Inv-Gamma(1.63, 0.45)
Consumers and Communication	44 (918)	41 (875)	Student-t(0, 0.48, 4)	Inv-Gamma(2.37, 0.86)
Cystic Fibrosis and Genetic Disorders	81 (1843)	40 (964)	Student-t(0, 0.40, 3)	Inv-Gamma(1.82, 0.58)
Dementia and Cognitive Improvement	127 (2097)	78 (1278)	Student-t(0, 0.49, 4)	Inv-Gamma(1.28, 0.25)
Developmental, Psychosocial and Learning Problems	106 (1457)	79 (1083)	Student-t(0, 0.83, 5)	Inv-Gamma(1.83, 0.82)
Drugs and Alcohol	103 (2027)	73 (1545)	Student-t(0, 0.44, 4)	Inv-Gamma(1.59, 0.42)
Effective Practice and Organisation of Care	59 (1106)	44 (878)	Student-t(0, 0.51, 4)	Inv-Gamma(1.90, 0.68)
Emergency and Critical Care	233 (4803)	152 (3466)	Student-t(0, 0.35, 3)	Inv-Gamma(1.46, 0.34)
ENT	17 (243)	10 (136)	Student-t(0, 0.81, 4)	Inv-Gamma(1.73, 0.71)
Epilepsy	114 (2052)	39 (839)	Student-t(0, 0.88, 6)	Inv-Gamma(1.71, 0.43)
Eyes and Vision	57 (1007)	41 (815)	Student-t(0, 0.77, 5)	Inv-Gamma(2.09, 0.94)
Fertility Regulation	62 (1072)	46 (818)	Student-t(0, 0,46, 5)	Inv-Gamma(2.21, 0.71)
Gut	314 (6056)	215 (4595)	Student-t(0, 0.63, 5)	Inv-Gamma(1.94, 0.62)
Gynaecological, Neuro-oncology and Orphan Cancer	245 (5811)	159 (4123)	Student-t(0, 0.53, 4)	Inv-Gamma(1.80, 0.56)
Gynaecology and Fertility	364 (6896)	185 (3714)	Student-t(0, 0.40, 2)	Inv-Gamma(1.24, 0.28)
Haematology	155 (4546)	101 (3080)	Student-t(0, 0.57, 4)	Inv-Gamma(2.91, 0.66)
Heart	979 (20760)	579 (14085)	Student-t(0, 0.20, 2)	Inv-Gamma(1.64, 0.29)
Heart: Vascular	8 (178)	7 (166)	Student-t(0, 0.95, 4)	Inv-Gamma(1.64, 0.83)
Hepato-Biliary	1042 (36665)	706 (25668)	Student-t(0, 0.43, 3)	Inv-Gamma(1.58, 0.40)
HIV/AIDS	32 (475)	22 (293)	Student-t(0, 0.32, 4)	Inv-Gamma(1.76, 0.36)
Hypertension	66 (1127)	28 (453)	Student-t(0, 0.28, 5)	Inv-Gamma(1.25, 0.10)
Incontinence	78 (1479)	53 (1023)	Student-t(0, 0.75, 3)	Inv-Gamma(2.07, 1.09)
Infectious Diseases	361 (6619)	251 (4787)	Student-t(0, 0.66, 3)	Inv-Gamma(2.08, 0.86)
Injuries	214 (5662)	143 (4182)	Student-t(0, 0.60, 4)	Inv-Gamma(1.52, 0.49)
Kidney and Transplant	479 (8828)	274 (5226)	Student-t(0, 0.53, 4)	Inv-Gamma(1.68, 0.44)
Lung Cancer	76 (1416)	68 (1294)	Student-t(0, 0.61, 5)	Inv-Gamma(2.04, 0.68)
Metabolic and Endocrine Disorders	130 (3092)	78 (2190)	Student-t(0, 0.29, 2)	Inv-Gamma(0.92, 0.11)
Methodology	74 (2098)	69 (1997)	Student-t(0, 0.60, 5)	Inv-Gamma(2.04, 0.50)
Movement Disorders	59 (1058)	35 (708)	Student-t(0, 0.73, 5)	Inv-Gamma(2.14, 0.64)
Multiple Sclerosis and Rare Diseases of the CNS	29 (564)	23 (487)	Student-t(0, 0, 76, 4)	Inv-Gamma(2.09, 0.71)
Musculoskeletal	139 (2403)	92 (1686)	Student-t(0, 0.59, 4)	Inv-Gamma(1.76, 0.59)
Neonatal	337 (6327)	153 (2738)	Student-t(0, 0.29, 3)	Inv-Gamma(1.80, 0.42)
Neuromuscular	43 (739)	20 (331)	Student-t(0, 0.70, 5)	Inv-Gamma(1.74, 0.49)
Oral Health	63 (990)	45 (724)	Student-t(0, 1.13, 4)	Inv-Gamma(1.85, 0.70)
Pain, Palliative and Supportive Care	315 (6062)	204 (4292)	Student-t(0, 1.00, 7)	Inv-Gamma(1.50, 0.40)
Pregnancy and Childbirth	1307 (24397)	832 (16588)	Student-t(0, 0.38, 2)	Inv-Gamma(1.73, 0.47)
Schizophrenia	609 (13515)	405 (9919)	Student-t(0, 0.58, 4)	Inv-Gamma(1.92, 0.69)
Sexually Transmitted Infections	25 (318)	19 (236)	Student-t(0, 0.61, 2)	Inv-Gamma(1.72, 0.47)
Skin	142 (2309)	76 (1294)	Student-t(0, 0.81, 2)	Inv-Gamma(1.64, 0.49)
Stroke	299 (5310)	131 (2357)	Student-t(0, 0.22, 2)	Inv-Gamma(1.36, 0.21)
Tobacco Addiction	213 (4990)	168 (4079)	Student-t(0, 0.49, 5)	Inv-Gamma(2.51, 0.63)
Urology	89 (1673)	60 (1271)	Student-t(0, 0.82, 5)	Inv-Gamma(1.72, 0.50)
Vascular	187 (2590)	116 (1711)	Student-t(0, 0.68, 5)	Inv-Gamma(1.70, 0.45)
Work	14 (208)	11 (178)	Student-t(0, 0.59, 4)	Inv-Gamma(1.79, 0.73)
Wounds	75 (1169)	53 (888)	Student-t(0, 0.60, 4)	Inv-Gamma(2.16, 0.86)
Pooled Estimate	11964 (253054)	7478 (170628)	Student-t(0, 0.58, 4)	Inv-Gamma(1.77, 0.55)

Example: Adverse effects of honey in treating acute cough in children

Two studies with the comparison honey and no treatment on the presence of nervousness, insomnia, hyperactivity

Honey condition: 5/35, 2/40 events No treatment: 0/39, 0/40 events

Frequentist analysis: OR = 9.40, 95% CI [1.16, 76.20], z = 2.10, p = 0.04

```
> library("RoBMA")
> fit <- BiBMA(</pre>
 x1 = events experimental,
 x^2 = events control,
  n1 = observations experimental,
  n2 = observations_control,
  priors effect
                      = prior(distribution="t", parameters=list(location=0, scale=0.48, df=3)),
  priors heterogeneity = prior(distribution="invgamma", parameters=list(shape=1.67, scale=0.45)),
  seed = 1)
> summary(fit, conditional = TRUE, output scale = "OR")
Bayesian model-averaged meta-analysis (binomial-normal model)
Components summary:
             Models Prior prob. Post. prob. Inclusion BF
                     0.500
Effect
                2/4
                                     0.725
                                                  2.630
Heterogeneity 2/4 0.500 0.564
                                                  1.296
Model-averaged estimates:
     Mean Median 0.025 0.975
mu 3.389 1.642 0.842 15.143
tau 0.420 0.158 0.000 2.594
Conditional estimates:
     Mean Median 0.025 0.975
mu 4.242 2.261 0.781 17.613
tau 0.747 0.426 0.097 3.233
The effect size estimates are summarized on the OR scale and heterogeneity is summarized on the logOR scale
(priors were specified on the log(OR) scale).
```

Bayesian Meta-Analysis (Binomial)

0 🖯 🚺 🔇



Plots

MCMC Diagnostics														
Models														
Effect		De						Trues				Drienw		
Distribution		Pa	ramet	ers				Irunca	ation			Prior w	eignts	
Student-t(μ , σ , v)	▼	μ	0	σ	0.48	ν	3	lower	-Inf	upper	Inf	Weight	1	×
							0							
Heterogeneity														
Distribution		Pa	ramet	ers	;			Trunca	ation			Prior we	eights	
Inverse-Gamma(α,β)	۷	α	1.67	β	0.45			lower	0	upper	Inf	Weight	1	×
							0							

Results

Bayesian Meta-Analysis (Binomial)

Summary

0

0

4

000

Model Summary

	Models	P(M)	P(M data)	Inclusion BF
Effect	2/4	0.500	0.725	2.640
Heterogeneity	2/4	0.500	0.565	1.300

Model Averaged Estimates

			959	% CI
	Mean	Median	Lower	Upper
Effect size (OR)	3.265	1.641	0.850	14.891
Heterogeneity (T)	0.415	0.156	0.000	2.464

Note. The effect size estimates are summarized on the OR scale and heterogeneity is summarized on the logOR scale (priors were specified on the log(OR) scale).

Conditional Estimates

			959	% CI	
	Mean	Median	Lower	Upper	
Effect size (OR)	4.248	2.299	0.795	18.194	
Heterogeneity (T)	0.725	0.420	0.096	3.127	

Note. The effect size estimates are summarized on the OR scale and heterogeneity is summarized on the logOR scale (priors were specified on the log(OR) scale).

Models Overview

	2	Prior Distribution	2				
#	Effect Size	Heterogeneity	Baseline	P(M)	P(M data)	log(MargLik)	Inclusion BF
1	Spike(0)	Spike(0)	independent contrast: Beta(1, 1)	0.250	0.096	-14.016	0.096
2	Spike(0)	InvGamma(1.67, 0.45)	independent contrast: Beta(1, 1)	0.250	0.179	-13.393	0.179
3	Student-t(0, 0.48, 3)	Spike(0)	independent contrast: Beta(1, 1)	0.250	0.339	-12.754	0.339
4	Student-t(0, 0.48, 3)	InvGamma(1.67, 0.45)	independent contrast: Beta(1, 1)	0.250	0.386	-12.623	0.386

Conclusions

- Differentiating between absence of evidence vs. evidence of absence
- Incorporating uncertainty about the specified model
- Incorporating historical knowledge
- Better convergence properties
- Sequential updating of evidence

Thank You for Your Attention

- Preprint: https://arxiv.org/abs/2306.11468
- R-package: <u>https://cran.r-project.org/package=RoBMA</u>
- JASP: <u>https://jasp-stats.org/</u>