Empirical Priors for Between-Study Heterogeneity in Bayesian Meta-analysis

Eleanor Pullenayegum
Clin Epi & Biostats, McMaster University
Biostatistics Unit, St Joseph's

Outline

- Bayesian meta-analysis
- Vagueness when there is little data
- Being Informative
- Cochrane Review
- Empirical Priors
- Application to dataset
- Questions for discussion

Meta-analysis

- Meta-analysis: pooling results from similar studies in order to summarize evidence
- Typically, experimental treatment vs. control
- Often use a binary outcome
- Pool results on the OR or RR scale
- For the frequentist, methods include:
  - Mantel-Haenszel, Peto, inverse-variance
- Fixed or random effects

Example

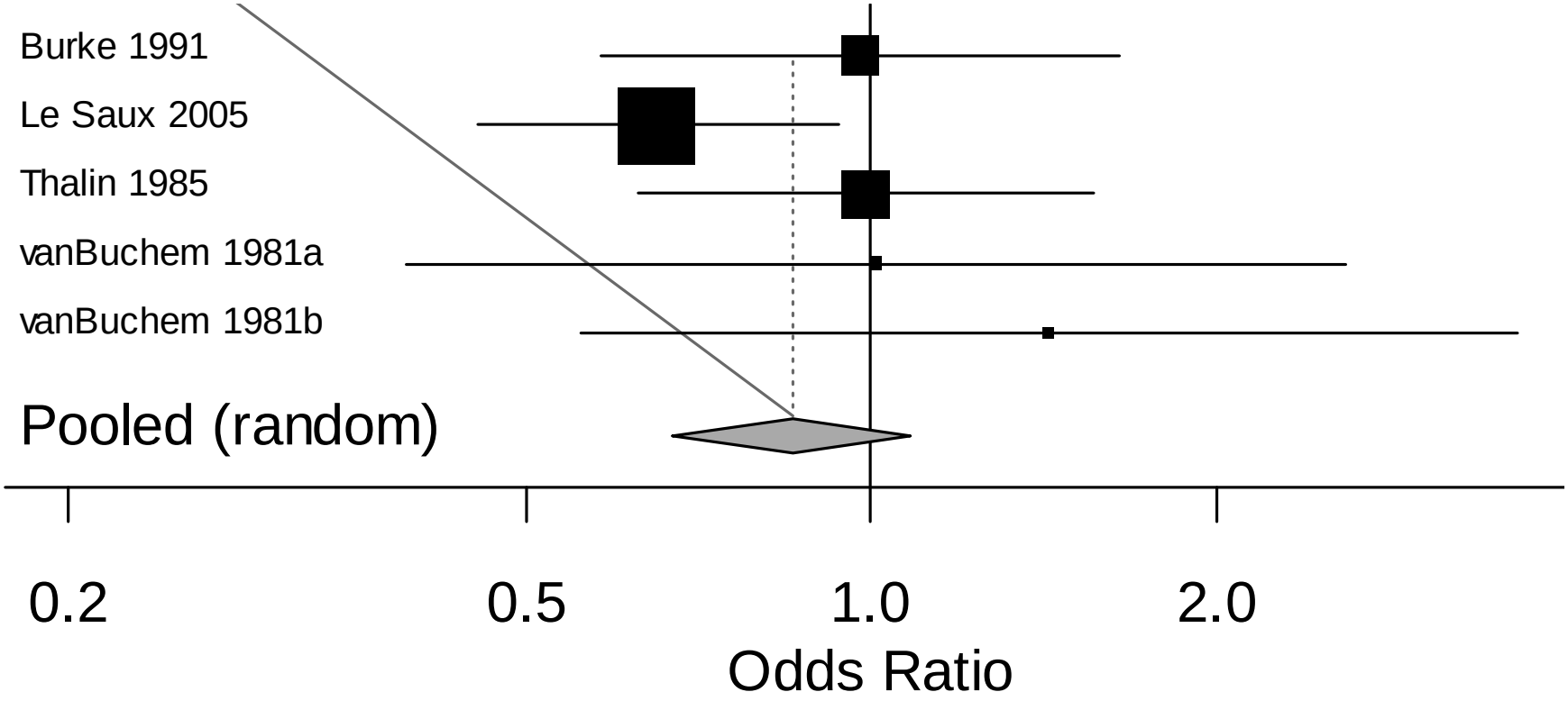
Antibiotics for acute otitis media in children (Review)

Sanders S, Glasziou PP, Del Mar C, Rovers M

Example

- Objective: To assess the effects of antibiotics for children with acute otitis media
- Primary outcome measure: Pain (present/absent) at 24 hours

Study	Antibiotics (n/N)	Placebo (n/N)	OR	95% CI
Burke 1991	53/112	56/117	0.98	(0.58, 1.64)
Le Saux 2005	82/258	106/254	0.65	(0.45, 0.93)
Thalin 1985	58/159	58/158	0.99	(0.63, 1.56)
vanBuchem 1981a	13/47	11/40	1.01	(0.39, 2.59)
vanBuchem 1981b	17/48	10/36	1.43	(0.56, 3.65)


Burke 1991
Le Saux 2005
Thalin 1985
vanBuchem 1981a
vanBuchem 1981b
Pooled (random)
0.2
0.5
1.0
2.0
Odds Ratio

Fixed & Random effects

- Fixed effects model: each study is measuring the same odds ratio

$$\log \mathrm{OR}_i \sim \mathrm{N}(\theta, \mathrm{sd}_i^2)$$

- Random effects model: each study measures a slightly different odds ratio, however there is a common underlying odds ratio around which the individual ORs deviate

$$\log \mathrm{OR}_i \sim \mathrm{N}(\theta_i, \mathrm{sd}_i^2)$$

$$\theta_i \sim \mathrm{N}(\theta, \tau^2)$$

Uncertainty in τ

- Pooling can be accomplished by taking weighted means
  - For a fixed effects analysis, weights are the inverse of the variance of the within-study variance: $1/sd_i^2$
  - For a random effects analysis, weights are the inverse of the sum of the within and between study variances: $1/(sd_i^2+\tau^2)$
- When there are few studies in the analysis, τ is measured subject to considerable variability
- Frequentist analyses do not account for this.

Bayesian Methods

- Bayesian methods can account for all sources of variability

$$re_i \sim Bin(NE_i, pe_i)$$

$$rc_i \sim Bin(NC_i, pc_i)$$

$$logit(pc_i) = \alpha_i$$

$$logit(pe_i) = \alpha_i + \theta_i$$

$$\alpha_i \sim N(\alpha, \nu^2)$$

$$\theta_i \sim N(\theta, \tau^2)$$

Put priors on the parameters we're unsure about $(\alpha, \nu, \theta, \tau)$

Potential Advantages of Bayesian meta-analysis

- Accounts for all sources of variability
- Can handle zero events without continuity corrections
- Can make probability statements
  - E.g. Probability that the experimental treatment is better than the control is 0.8

Priors

- Need to put prior distributions on the unknown parameters (α,ν,θ,τ)
  - α is the overall log odds of success in the control groups
  - ν^2 is the between-study variance in the log odds of success in the control groups
  - θ is the overall log odds ratio
  - τ^2 is the between-study variance in the log odds ratios
- Most popular option is to use vague priors
- Fairly easy to do for α, θ
- For the variance parameters: How vague is vague?

Alternative formulation

To simplify discussion, shall use alternative formulation

-> Don't need to worry about distribution of underlying event rates

Standard formulation:

$ne_i \sim Bin(NE_i, pe_i)$

$nc_i \sim Bin(NC_i, pc_i)$

$logit(pc_i) = \alpha_i$

$logit(pe_i) = \alpha_i + \theta_i$

$\alpha_i \sim N(\alpha, \nu^2)$

$\theta_i \sim N(\theta, \tau^2)$

Alternative formulation:

$\log OR_i \sim N(\theta_i, sd_i^2)$

$\theta_i \sim N(\theta, \tau^2)$

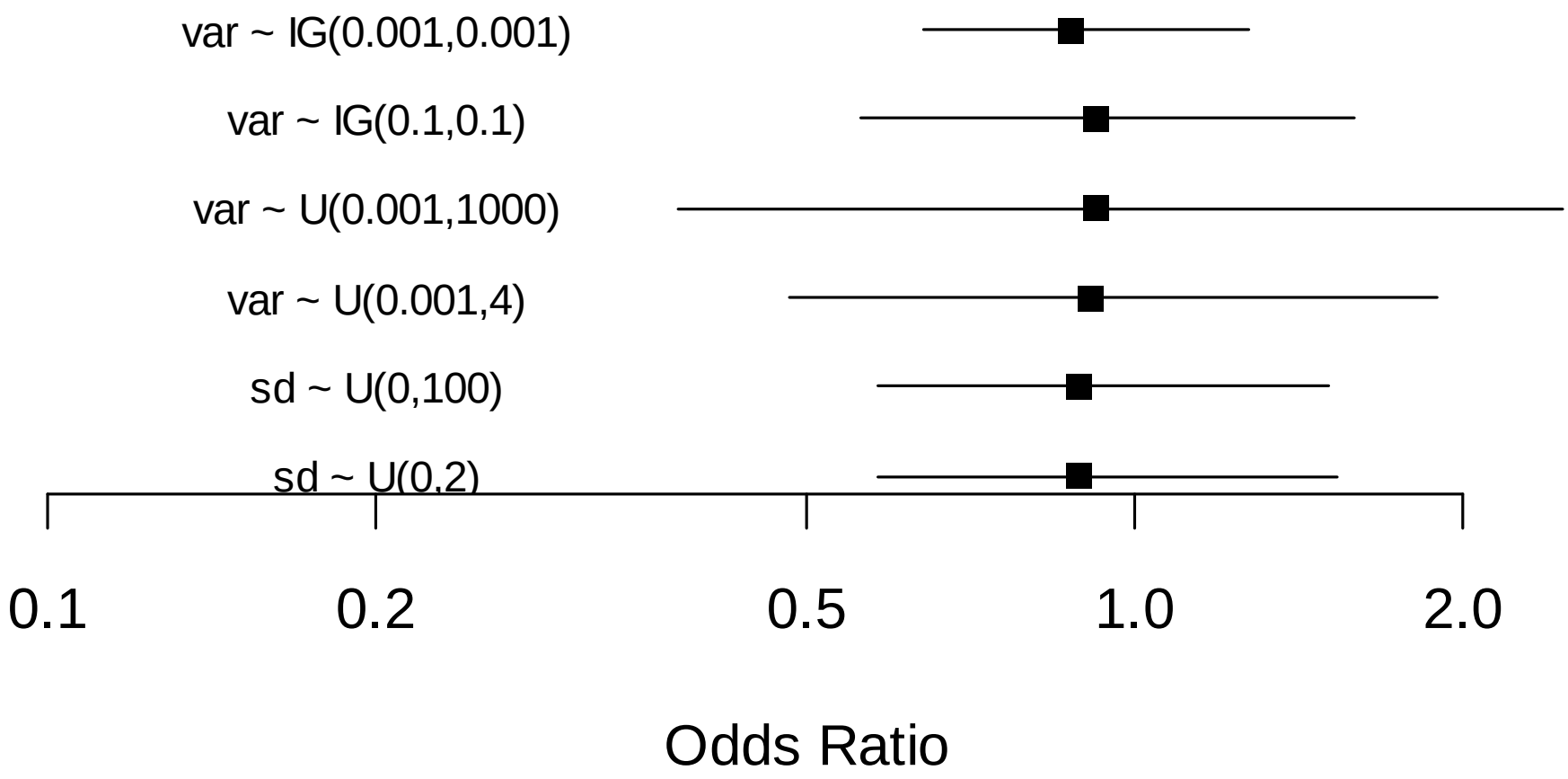
Alternative formulation:

- Cannot account for zero events
- Does not account for variability in estimates of sd_i

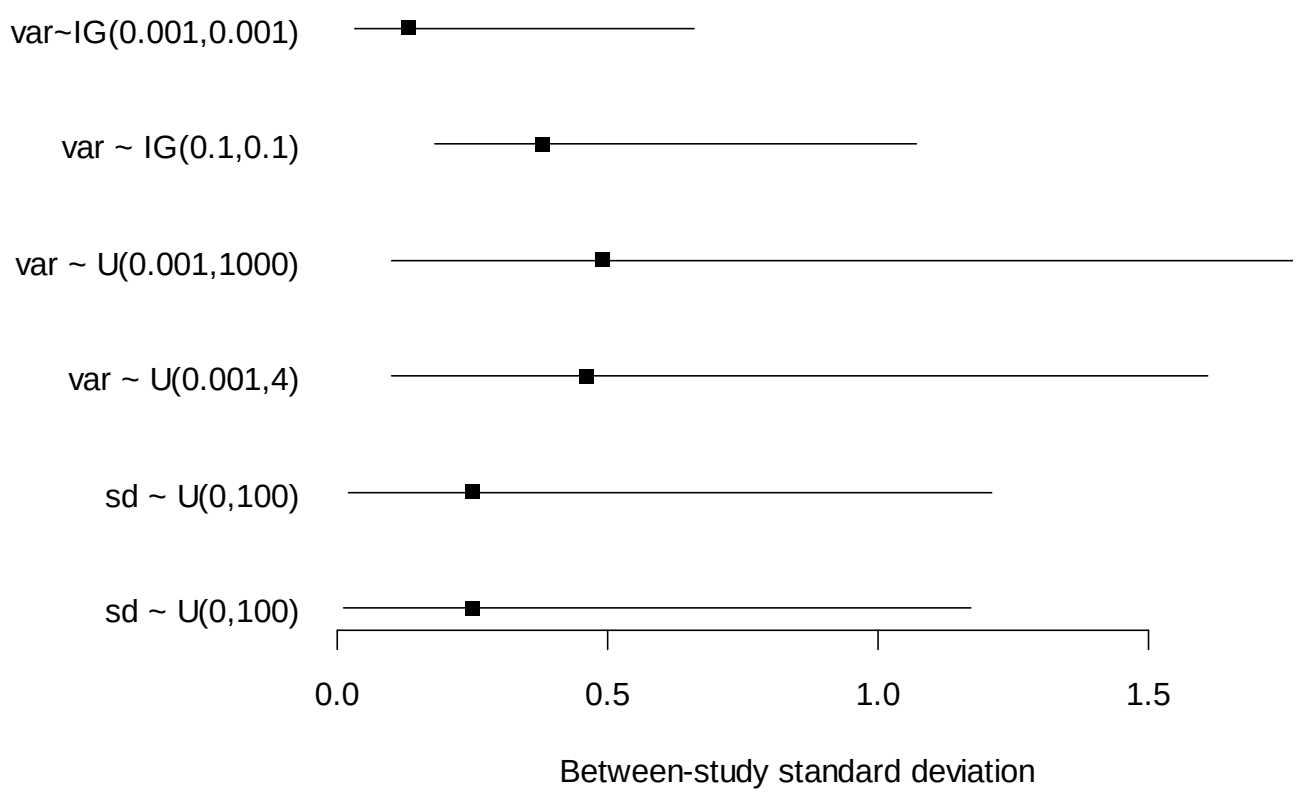
“How vague is vague?”

- Lambert et al, 2005
- Used 13 different vague priors
- These gave different results when the number of studies was small
  - Point estimates of pooled ORs were similar
  - 95% CrIs were very variable across priors

Pooled OR for a selection of vague priors for τ



What's going on?



Why be vague?

- Allow the data to dominate the prior
  - Less subjective
- Not much prior knowledge
- Requires less thought

BUT

- Not much data -> prior still matters
- There **is** prior knowledge
- Requires a lot of thought to be vague in this case!

What values of τ are reasonable?

$$\theta_i \sim N(\theta, \tau^2)$$

$$P(\text{Pooled OR} \times e^{-1.96\tau} < OR_i < \text{Pooled OR} \times e^{1.96\tau} \mid \theta, \tau) = 0.95$$

$$\text{Ratio of } 97.5\%\ \text{OR}_i \text{ to } 2.5\%\ \text{OR}_i = e^{3.92\tau}$$

- Usually the range of study ORs would be within an order of magnitude
  - i.e. $e^{3.92\tau} < 10$
  - i.e. $\tau < 0.59$
- Cf. Smith, 1995

What's reasonable?

- Cf. Spiegelhalter, 2004
- Imagine drawing θ_1, θ_2 at random from the distribution of θ_i
- Then, $\theta_1 - \theta_2 \sim N(0, 2\tau^2)$

  $|\theta_1 - \theta_2| \sim HN(2\tau^2)$
- Median ratio of smaller OR to larger is $e^{1.09\tau}$
- e.g. τ=0.64 leads to a median ratio of 2

What's reasonable?

τ	exp(3.92τ) (ratio of 97.5th to 2.5th percentile)	Median ratio of random pair
0	1	1
0.1	1.48	1.11
0.5	7.1	1.72
1	50.4	2.97
2	2540	8.84

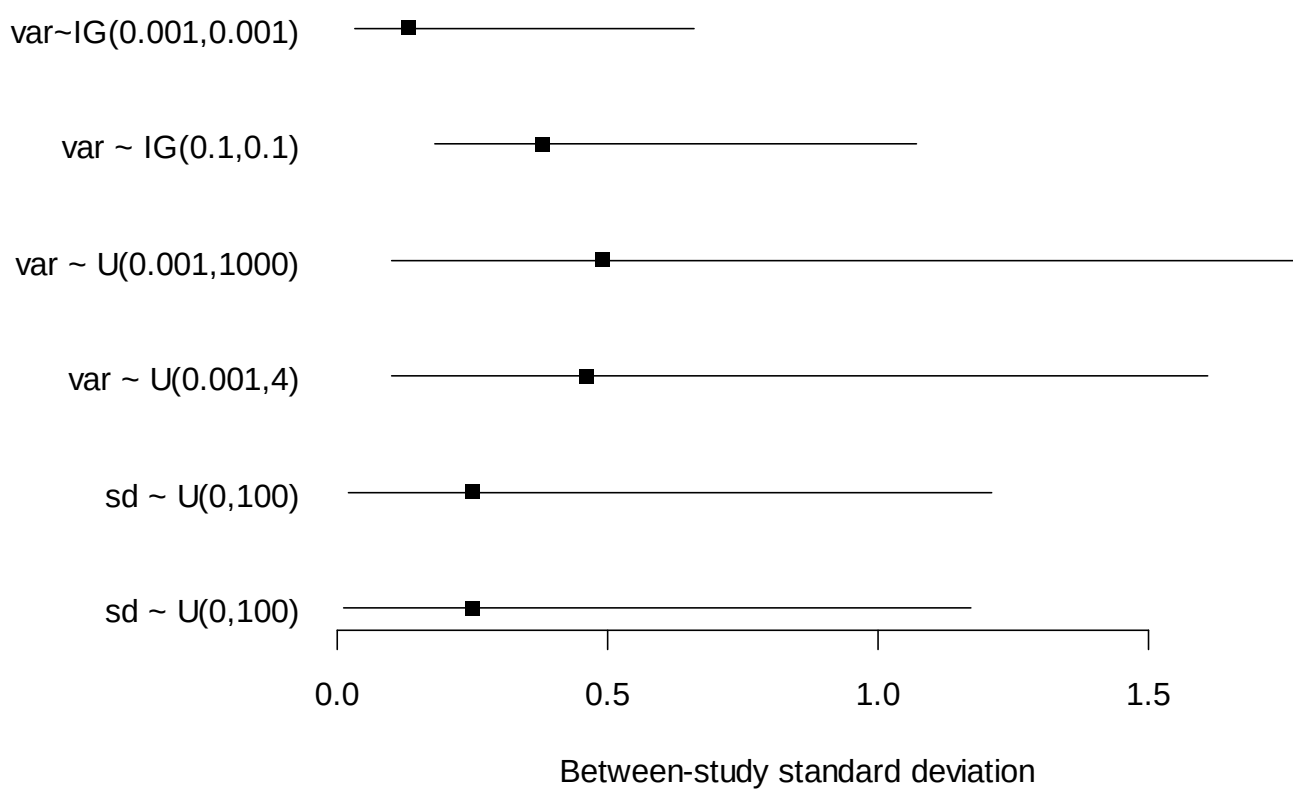
Extract from Spiegelhalter et al, Bayesian Approaches to Clinical Trials & Healthcare evalulation, p. 169

How reasonable were our vague priors?

- τ=0.1 small amount of heterogeneity
- τ=0.5 (median ratio larger:smaller =1.72) reasonable, but quite large
- τ=1 (median ratio of larger:smaller ≈3) a lot of heterogeneity
- τ=2 (median ratio of larger:smaller >8) extreme

Prior	P(τ<2)	P(τ<1)
τ^2 ~ IG(0.001,0.001)	0.0077	0.0063
τ^2 ~ IG(0.1,0.1)	0.27	0.17
τ^2 ~ U(0.001,1000)	0.0040	0.0010
τ^2 ~ U(0.001,4)	1	0.25
τ ~ U(0,100)	0.02	0.01
τ ~ U(0,2)	1	0.5

Example: Posterior Distribution of τ



Summary so far

- Bayesian methods can incorporate uncertainty about τ
- Need a prior distribution for τ
- Vague priors supposedly let the data dominate
  - But we don't have much data
  - Different vague priors give different posterior inferences about τ
- As a Bayesian, I should believe my posterior distribution provided that I believe my prior
- I don't believe any of the vague priors!

Informative prior

- We can use mathematical reasoning to argue what is and is not theoretically plausible
- We can use past meta-analyses to establish what is and is not likely

Empirical Evidence

- Propose to describe the distribution of between-study heterogeneity in past meta-analyses
- Use this distribution as a prior distribution for τ in a new meta-analysis
- Should result in better calibration
  - Bayesians are well calibrated if their probability statements are borne out in practice
  - E.g. if, amongst the days for which the Bayesian forecaster says it will rain with probability 0.4, it rains 40% of the time

Cochrane Review

- From the Cochrane Database of Systematic Reviews, extracted all reviews published between Jan 2008 and Jul 31st 2009.
- Reviews were included in the sample if
  - They included data from two are more studies
  - The primary outcome was binary
  - The first forest plot included a pooled result

Cochrane Review

- Only the pooled result from the first forest plot was included.
  - If there was a total result pooled across all subgroups, this was used
  - If pooling was done within subgroups, the first subgroup with pooling was used

Summary of findings

- Search retrieved a total of 942 records
- 314 provided valid data
- Of those excluded:
  - 103 did not include any studies
  - 320 did not pool results for the primary comparison
  - 198 did not have a binary outcome
  - In 4 reviews, all the included studies had either 100% or 0% event rates
  - 3 studies were excluded for other reasons (meta-analysis of cross-over trials, numbers of events/patients not reported)

Fitting Distributions

- Candidate distributions
  - Inverse Gamma
  - Gamma
  - Log-Normal
- Option 1: Maximum Likelihood, ignoring uncertainty in parameters
- Option 2: Incorporate uncertainty in parameters through Bayesian model
- Option 3: Incorporate uncertainty in estimates of τ by including study-level data from each included review

Option 1

- Derive estimate of τ^2 from each individual review through method-of-moments

$$\hat{\tau}^2 = \left. Q-(k-1) \middle/ \left(\sum_{i=1}^{k} w_i - \frac{\sum_{i=1}^{k} w_i^2}{\sum_{i=1}^{k} w_i} \right) \right. \text{ if } Q \geq k\text{-}1$$

$$0 \text{ if } Q < k-1$$

- Distribution of τ^2:
  - Point mass at zero
  - Estimate distributional parameters by maximum likelihood

Option 2

- Use method-of-moments to estimate τ^2
- Point mass at zero
- Use model to express uncertainty about distributional parameters:

$$x = I(\tau^2_{review} = 0)$$

$$x \sim \text{Bernoulli}(p)$$

$$\text{for } \tau^2_{review} > 0:$$

$$y = \log(\tau^2_{review})$$

$$y \sim N(\mu, \sigma^2)$$

Option 3

- Use individual-level study results from each review in order to capture uncertainty in estimates of τ
- Useful if want to constrain $\tau>0$

New meta-analysis

$$\log \mathrm{OR}_i \sim \mathrm{N}(\theta_i, \mathrm{sd}_i^2)$$

$$\theta_i \sim \mathrm{N}(\theta, \tau^2)$$

$$\tau^2 = \exp(\log \mathrm{tau})$$

$$\log \mathrm{tau} \sim \mathrm{N}(\mu, \sigma^2)$$

Cochrane review data

$$\log \mathrm{OR}_{ij}^{\mathrm{review}} \sim \mathrm{N}(\alpha_{ij}, \mathrm{var}_{ij})$$

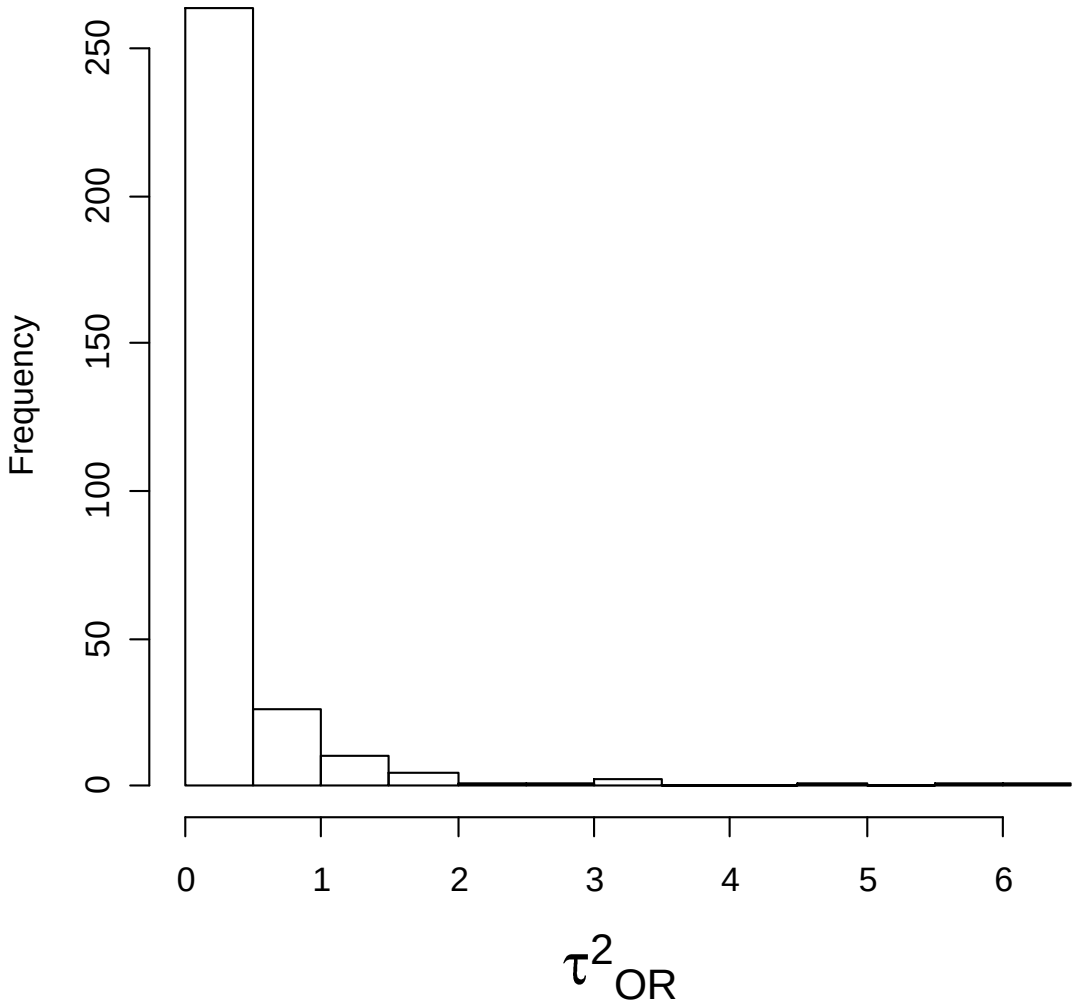
$$\alpha_{ij} \sim \mathrm{N}(\alpha_j, \tau_j)$$

$$\tau_j^2 = \exp(\log \mathrm{tau}_j)$$

$$\log \mathrm{tau}_j \sim \mathrm{N}(\mu, \sigma^2)$$

Results

Histogram of methods-of-moments estimates of τ^2_{OR}



Summary Stats: Method-of-moments estimators

Parameter	Prop=0	1st quartile	Median	3rd quartile
Pooled OR		0.56	0.95	1.54
Pooled RR		0.68	0.97	1.30
Pooled log OR		-0.58	-0.05	0.43
Pooled log RR		-0.39	-0.03	0.27
τ^2_{OR}	0.49	0	0.0084	0.25
τ^2_{RR}	0.51	0	0	0.075

Mean $\tau^2_{OR} = 0.27$

Mean $\tau^2_{RR} = 0.11$

Option 1: OR

Range for τ	Observed	Expected counts		
		Inverse Gamma (0.511, 0.031)	Gamma (0.65, 1.218)	Log-Normal (-1.563, 1.494)
(0,0.1]	4	2.13	10.08	3.36
(0.1, 0.5]	78	98.89	63.31	84.71
(0.5, 1]	58	29.52	60.55	49.13
>1	21	27.54	27.07	23.80
χ^2 goodness-of-fit p-value		< 0.0001	0.036	0.4603
Kolmogorov-Smirnov p-value		0.0000340	0.364	0.5388

Option 1: RR

Range for τ	Observed	Expected counts		
		Inverse Gamma (0.416, 0.00616)	Gamma (0.524, 2.217)	Log-Normal (-2.645, 1.710)
(0,0.1]	22	33.42	23.40	19.26
(0.1, 0.5]	90	82.82	82.92	98.44
(0.5, 1]	36	15.98	40.99	25.99
>1	5	17.70	5.69	9.32
χ^2 goodness-of-fit p-value		< 0.0001	0.7104	0.073
Kolmogorov-Smirnov p-value		0.00022	0.1277	0.7329

Option 2 – Goodness-of-fit

Deviance Information Criterion (DIC) for the three distributions (lower is better)

Distribution	OR	RR
Inverse Gamma	157	-135
Gamma	101	-187
Log-Normal	87	-207

Option 2: Estimated parameters

	Shape (inverse-gamma & gamma); mean (log-Normal)		Rate (inverse gamma & gamma); precision (log-Normal)	
OR	ML	Post. median & 95% CrI	ML	Post. median & 95% CrI
Inverse gamma	0.511	0.51 (0.42, 0.61)	0.031	0.03 (0.023, 0.04)
Gamma	0.65	0.65 (0.53, 0.77)	1.218	1.21 (0.91, 1.56)
Log-Normal	-1.56	-1.56 (-1.80, -1.33)	0.45	0.44 (0.35, 0.55)
RR				
Inverse gamma	0.416	0.41 (0.34, 0.49)	0.0061	0.0061 (0.004,0.0081)
Gamma	0.524	0.52 (0.43, 0.62)	2.217	2.19 (1.61, 2.89)
Log-Normal	-2.645	-2.65 (-2.92, 2.37)	0.34	0.34 (0.27, 0.42)

Option 3: Model fit statistics based on RCT-level data

Odds ratios scale

Distribution	Posterior Median & 95% CrI	DIC
Log-Normal	Mean: -3.16 (-3.84, -2.68) Precision: 0.35 (0.20, 0.61)	4706
Inverse Gamma	Shape: 0.95 (0.67, 1.40) Rate: 0.03 (0.01, 0.06)	4716

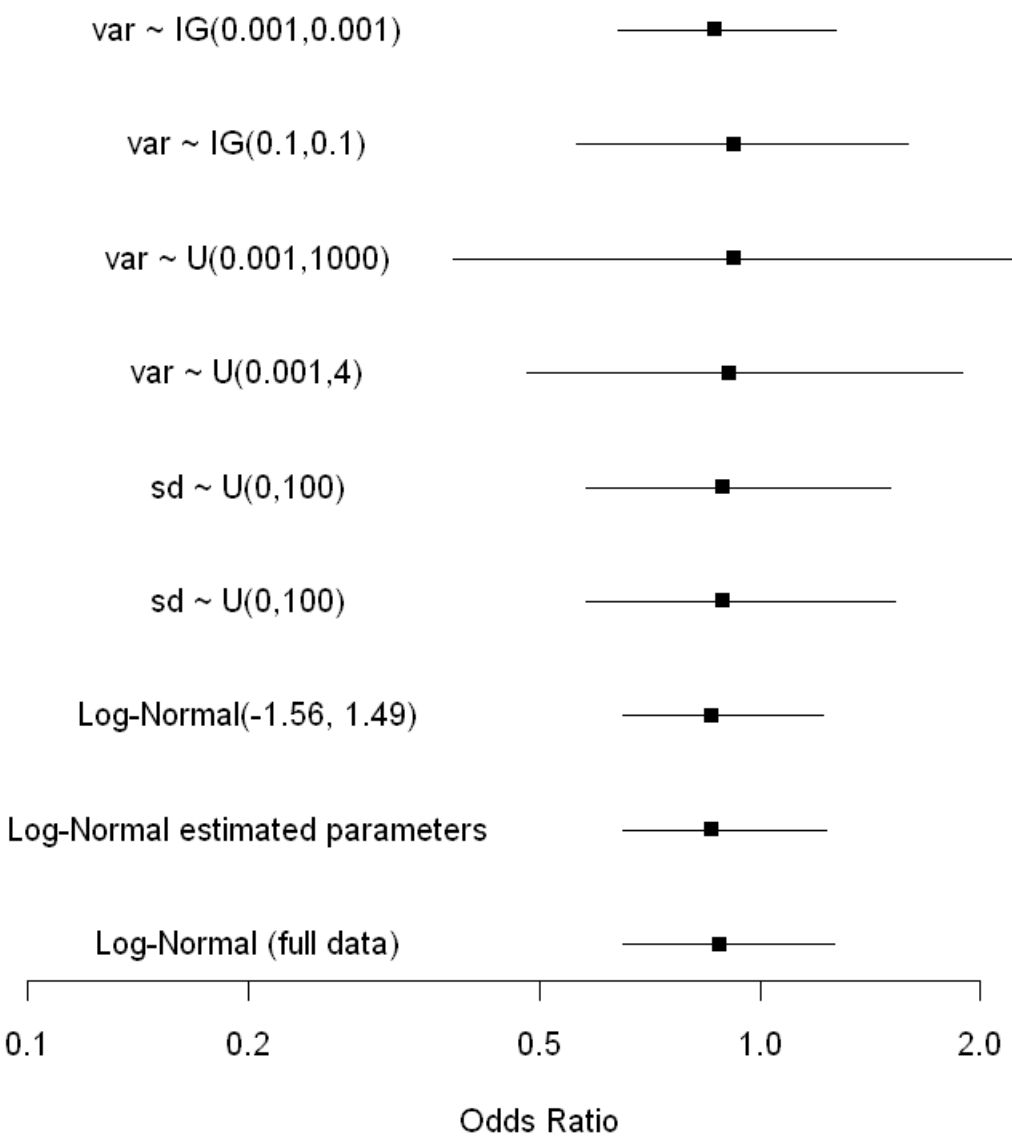
Relative Risk scale

Distribution	Posterior Median & 95% CrI	DIC
Log-Normal	Mean: -4.41 (-5.16, -3.89) Precision: 0.32 (0.18, 0.56)	2733
Inverse Gamma	Shape: 0.87 (0.63, 1.24) Rate: 0.0063 (0.0024, 0.014)	2743

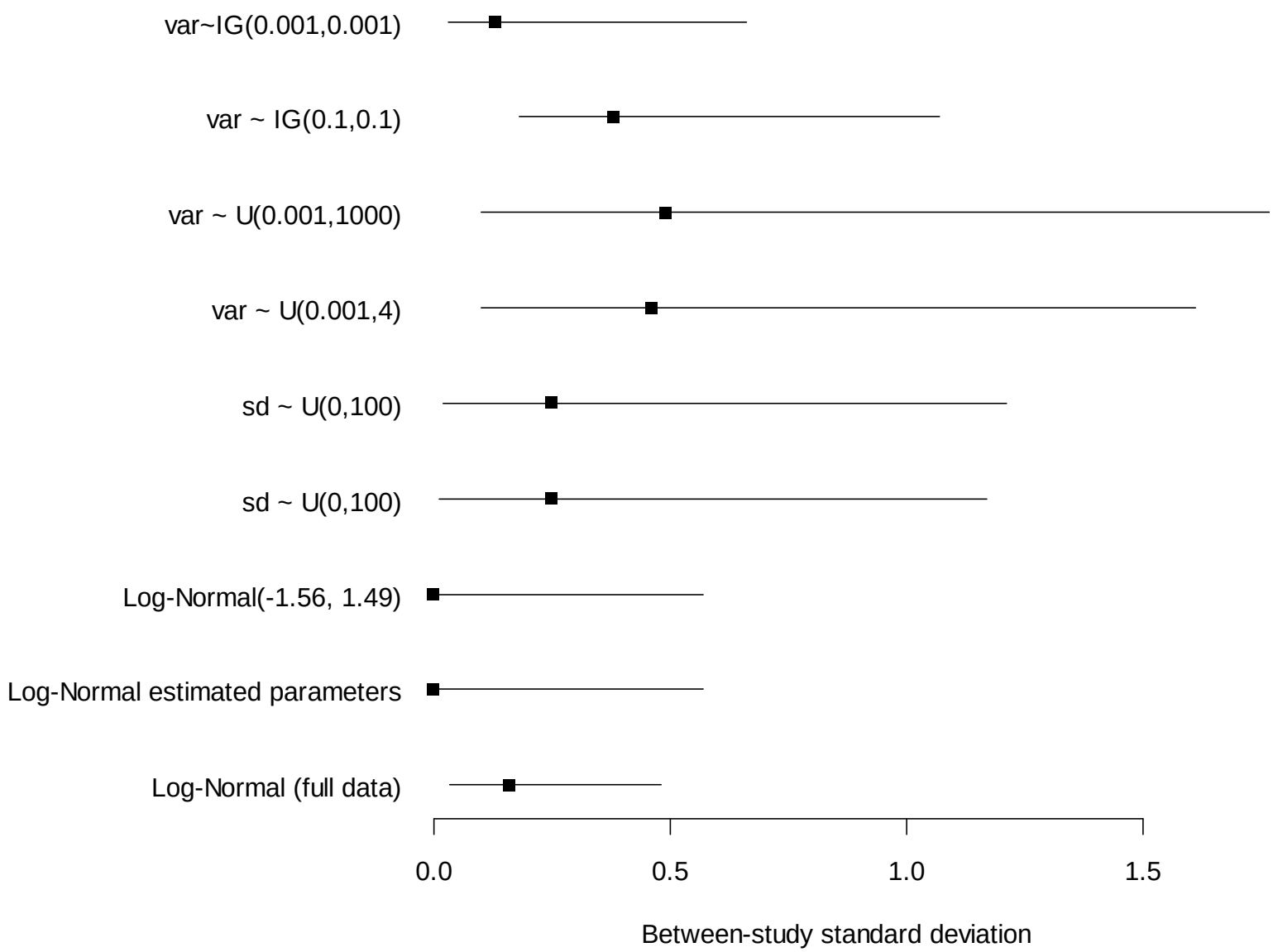
Summary

- Log-Normal fits best in all three approaches
- Less heterogeneity in the RRs than in the ORs

Application to example - OR



Posterior distribution for τ



Strengths

- Proposed priors describe reasonable beliefs about between-study heterogeneity
- If you do a new meta-analysis using this empirical prior
  - Provided you believe the prior, you should believe the posterior
  - Regardless of how many studies were in your meta-analysis

Limitations

- Have used Cochrane reviews only
- Have used the alternative, less desirable formulation of the Bayesian model
  - Study-level data is log OR and associated variance, not raw numbers of events, non-events
  - Does not account for uncertainty in variance of log ORs
- Have used just one review for illustration – impact of priors on results may be different for other examples

Questions

- Does it matter that the data was extracted by just one person?
- Are Cochrane reviews different from other reviews?
- Can heterogeneity be exactly equal to zero?
  - i.e. should I have a point mass at 0?
- Priors for pooled OR and heterogeneity assumed independent – does this matter?
- I used a continuity correction for trials with zero events – how much would the results change on using an alternative method?
- Exchangeability?