Complementing evidence from a small scale RCT by registry data in a rare disease setting

Christian Röver and Tim Friede

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

September 25, 2018

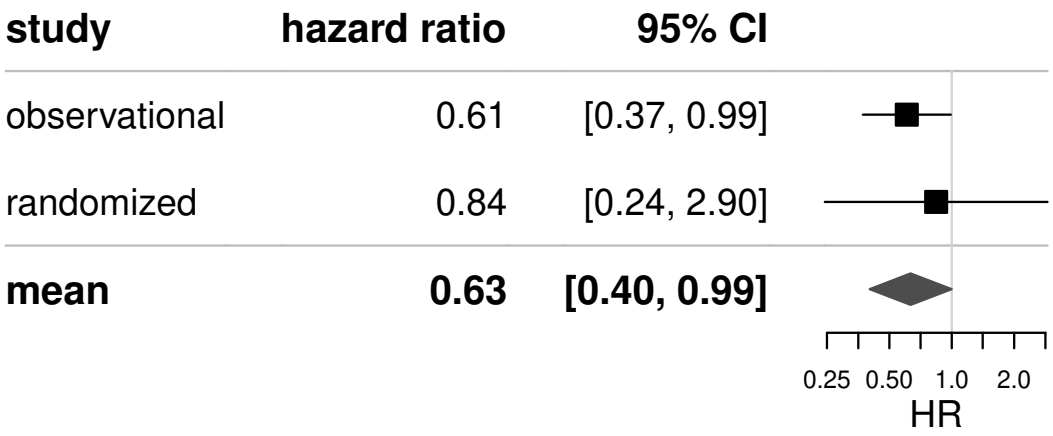
InSPiRe
Innovative methodology for small populations research

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement number FP HEALTH 2013-602144.

Introduction

The Creutzfeld-Jakob disease (CJD) example

- Creutzfeld-Jakob disease (CJD): a (very) **rare disease**
- A small **randomized trial** ($N = 12$) on the use of Doxycycline was conducted (endpoint: survival),
  registry data ($N = 88$) was considered in addition
  (analysis stratified by propensity scores)
- heterogeneity anticipated
- both estimates were combined
  (using standard random-effects meta-analysis)[1]



[1] D. Varges et al. Doxycycline in early CJD – a double-blinded randomised phase II and observational study. *Journal of Neurology, Neurosurgery and Psychiatry* 88(2):119–125, 2017.

- **normal-normal hierarchical model (NNHM)**:

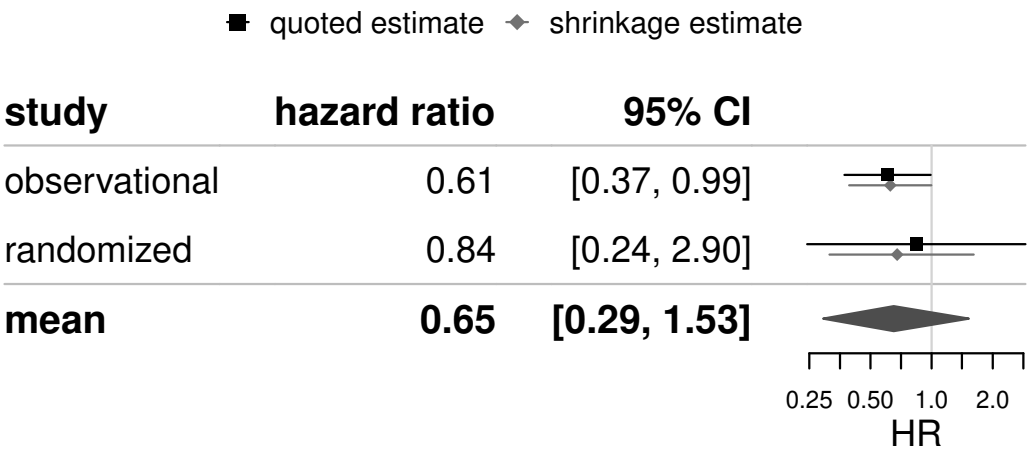
$$\begin{aligned} y_i|\theta_i &\sim \text{Normal}(\theta_i,\ {\sigma_i}^2), \\ \theta_i|\mu,\tau &\sim \text{Normal}(\mu,\ \tau^2) \qquad (\text{for } i = 1,\ldots,k) \end{aligned}$$

Data:

- estimates y_i
- standard errors σ_i

Parameters:

- study-specific effects θ_i
- overall effect μ
- heterogeneity τ



- (Bayesian approach:
  prior specification for μ and τ)

- **normal-normal hierarchical model (NNHM)**:

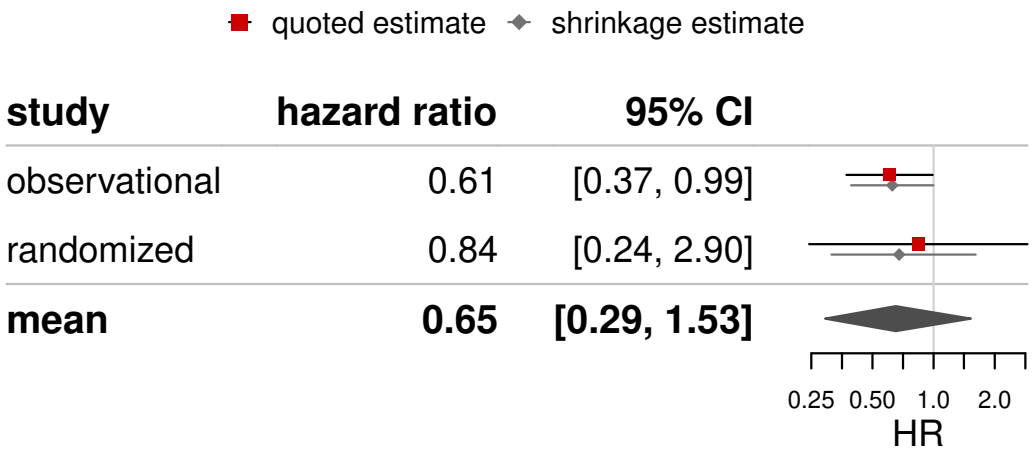
$$
\begin{aligned}
y_i|\theta_i &\sim \text{Normal}(\theta_i,\ \sigma_i{}^2), \\
\theta_i|\mu,\tau &\sim \text{Normal}(\mu,\ \tau^2) \qquad (\text{for } i = 1,\ldots,k)
\end{aligned}
$$

Data:

- estimates y_i
- standard errors σ_i

Parameters:

- study-specific effects θ_i
- overall effect μ
- heterogeneity τ



- (Bayesian approach:
  prior specification for μ and τ)

- **normal-normal hierarchical model (NNHM)**:

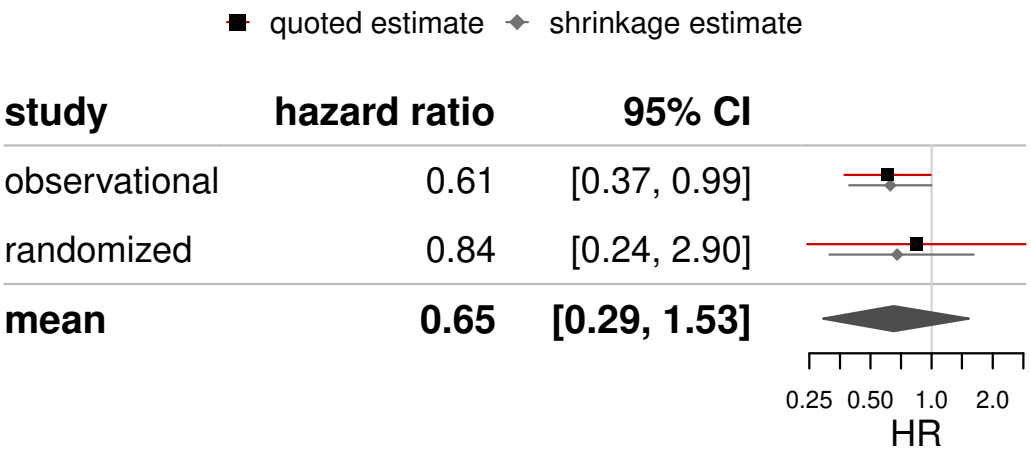
$$\begin{aligned} y_i|\theta_i &\sim \text{Normal}(\theta_i,\ \sigma_i{}^2), \\ \theta_i|\mu,\tau &\sim \text{Normal}(\mu,\ \tau^2) \qquad (\text{for } i = 1,\ldots,k) \end{aligned}$$

Data:

- estimates y_i
- standard errors σ_i

Parameters:

- study-specific effects θ_i
- overall effect μ
- heterogeneity τ



- (Bayesian approach:
  prior specification for μ and τ)

- **normal-normal hierarchical model (NNHM)**:

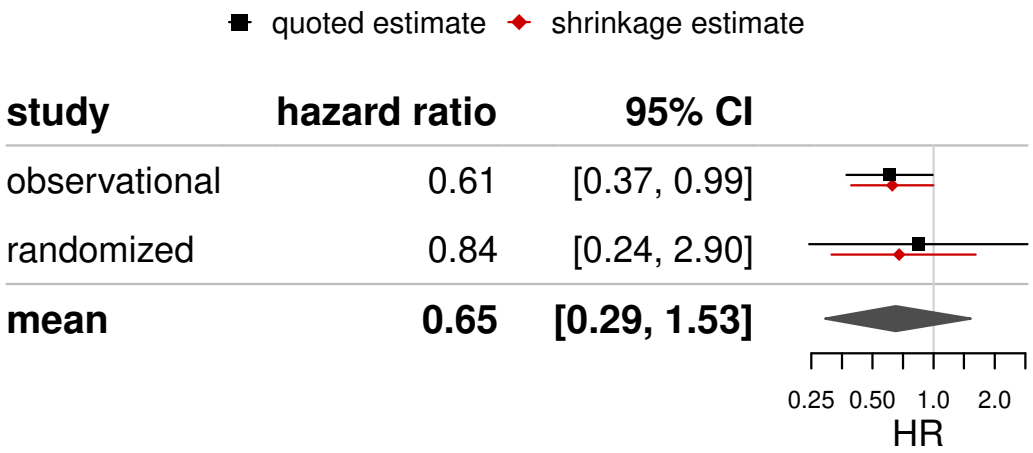
$$
\begin{aligned}
y_i|\theta_i &\sim \text{Normal}(\theta_i,\ {\sigma_i}^2), \\
\theta_i|\mu,\tau &\sim \text{Normal}(\mu,\ \tau^2) \qquad (\text{for } i = 1,\ldots,k)
\end{aligned}
$$

Data:

- estimates y_i
- standard errors σ_i

Parameters:

- study-specific effects θ_i
- overall effect μ
- heterogeneity τ



- (Bayesian approach:
  prior specification for μ and τ)

- **normal-normal hierarchical model (NNHM)**:

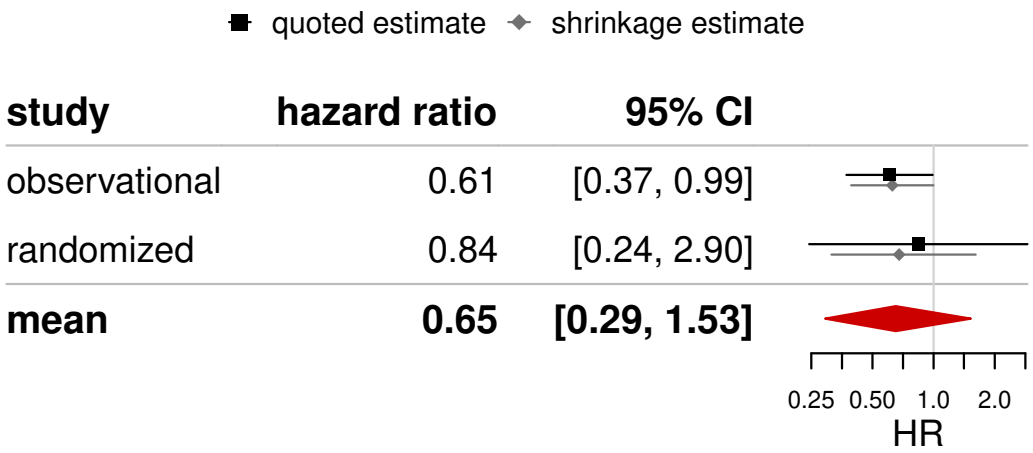
$$
\begin{aligned}
y_i|\theta_i &\sim \text{Normal}(\theta_i,\ {\sigma_i}^2), \\
\theta_i|\mu,\tau &\sim \text{Normal}(\mu,\ \tau^2) \qquad (\text{for } i = 1,\ldots,k)
\end{aligned}
$$

Data:

- estimates y_i
- standard errors σ_i

Parameters:

- study-specific effects θ_i
- overall effect μ
- heterogeneity τ



- (Bayesian approach:
  prior specification for μ and τ)

- **normal-normal hierarchical model (NNHM)**:

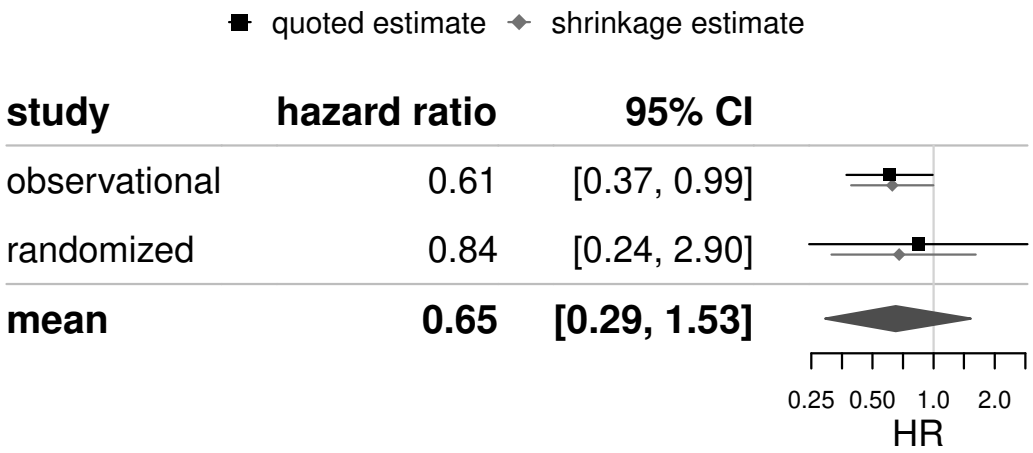
$$\begin{aligned} y_i|\theta_i &\sim \text{Normal}(\theta_i, \sigma_i^2), \\ \theta_i|\mu,\tau &\sim \text{Normal}(\mu, \tau^2) \qquad (\text{for } i = 1,\ldots,k) \end{aligned}$$

Data:

- estimates y_i
- standard errors σ_i

Parameters:

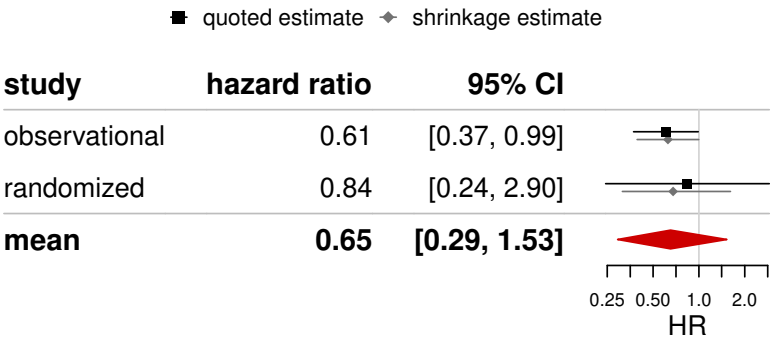
- study-specific effects θ_i
- overall effect μ
- heterogeneity τ



- (Bayesian approach: prior specification for μ and τ)

commonly:

- main interest in overall effect μ



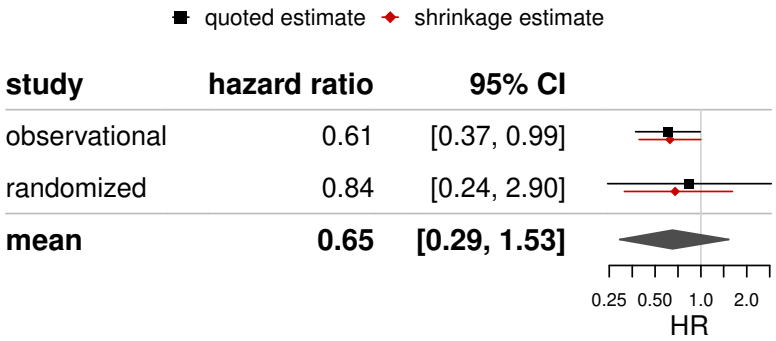
shrinkage estimation:

- (updated) estimate of study's specific effect θ_i
- based on all estimates $(y_1, \ldots, y_k, \sigma_1, \ldots, \sigma_k)$
- more or less "shrunk" towards the overall mean μ, (depending on heterogeneity)
- a.k.a. *best linear unbiased prediction (BLUP)* [2]

[2] S.W. Raudenbush, A.S. Bryk. Empirical Bayes meta-analysis. *Journal of Educational Statistics* 10(2):75–98, 1985.
G.K. Robinson. That BLUP is a good thing: The estimation of random effects. *Statistical Science* 6(1):15–51, 1991.

commonly:

- main interest in overall effect μ



shrinkage estimation:

- (updated) estimate of study's specific effect θ_i
- based on all estimates $(y_1, \ldots, y_k, \sigma_1, \ldots, \sigma_k)$
- more or less "shrunk" towards the overall mean μ, (depending on heterogeneity)
- a.k.a. *best linear unbiased prediction (BLUP)* [2]

[2] S.W. Raudenbush, A.S. Bryk. Empirical Bayes meta-analysis. *Journal of Educational Statistics* 10(2):75–98, 1985.
G.K. Robinson. That BLUP is a good thing: The estimation of random effects. *Statistical Science* 6(1):15–51, 1991.

Shrinkage estimation

Why shrinkage?

- often of primary interest: a particular study (-outcome)
- here:
  randomized study
  additional data
- aim: to infer the randomized study's outcome
  (a **shrinkage estimate**, *not* an overall mean[3])
- NNHM (meta-analysis) model provides framework
- useful when data are sparse (e.g., rare diseases)

[3] 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.

Shrinkage estimation

The MAP / MAC connection

- two ways to analyze *i*th estimate:
  - **Meta-analytic-combined (MAC)** approach:
    perform joint meta-analyis of all studies,
    determine *i*th shrinkage estimate
  - **Meta-analytic-predictive (MAP)** approach:
    meta-analyze all but *i*th study;
    resulting posterior yields *meta-analytic predictive (MAP)* prior,
    use MAP prior and data y_i to infer θ_i
- both approaches yield identical results[4]
- MAP approach
  - additional motivation
  - quantification of information contributed by additional studies

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

Shrinkage estimation

Two-study scenario

- consider: primary interest in randomized trial outcome
  (no “breaking of randomization” by pooled analysis)
- does it make sense to consider shrinkage estimates from a 2-study meta-analysis?
- how do shrinkage estimates behave in general?

Shrinkage estimation

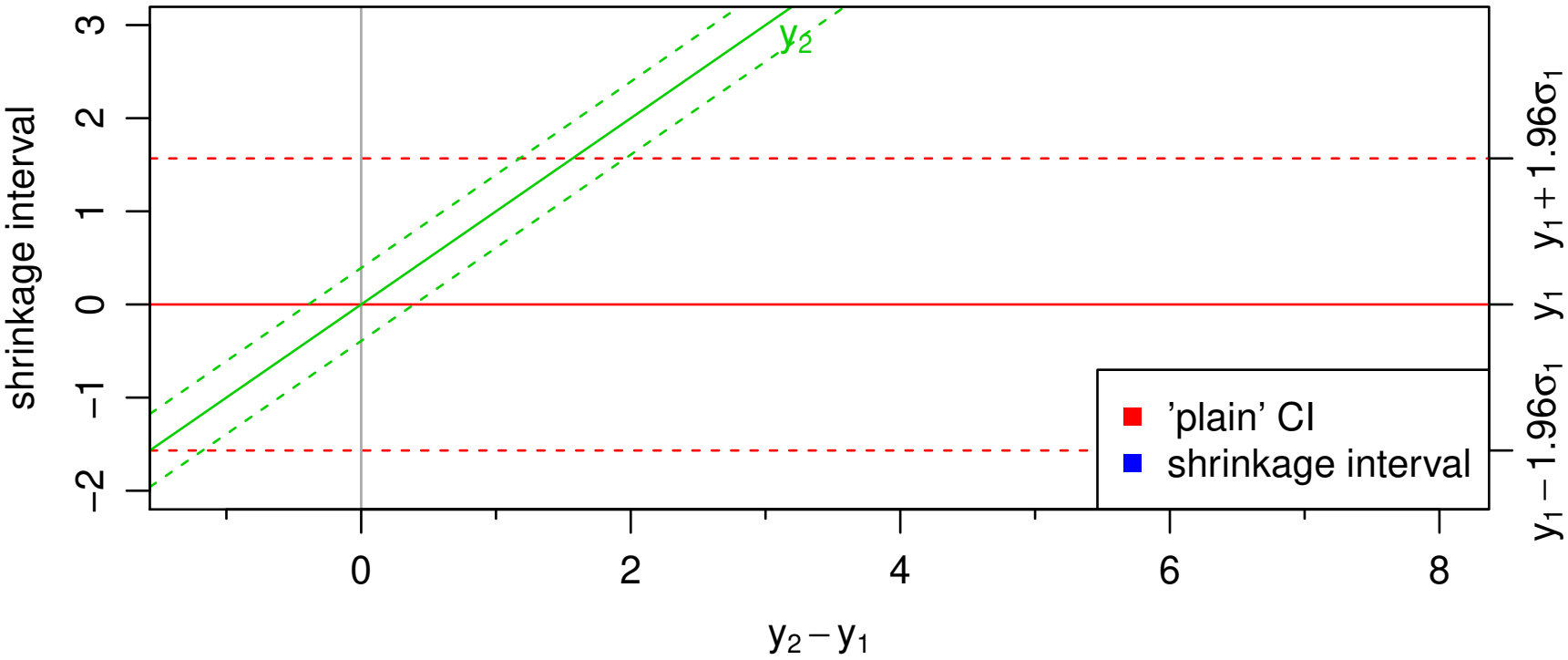
Two-study scenario

- consider: primary interest in randomized trial outcome
  (no “breaking of randomization” by pooled analysis)
- does it make sense to consider shrinkage estimates from a 2-study meta-analysis?
- how do shrinkage estimates behave in general?

- investigate example cases
- consider pair of studies, binary endpoint (log-OR);
  $n_1 = 25$, $n_2 = 400 \quad \rightarrow \quad$ approx. $\sigma_1 = 0.8$, $\sigma_2 = 0.2$
  effect prior: $p(\mu) =$ uniform
  heterogeneity prior: $p(\tau) =$ half-Normal(0.5)

Shrinkage estimation

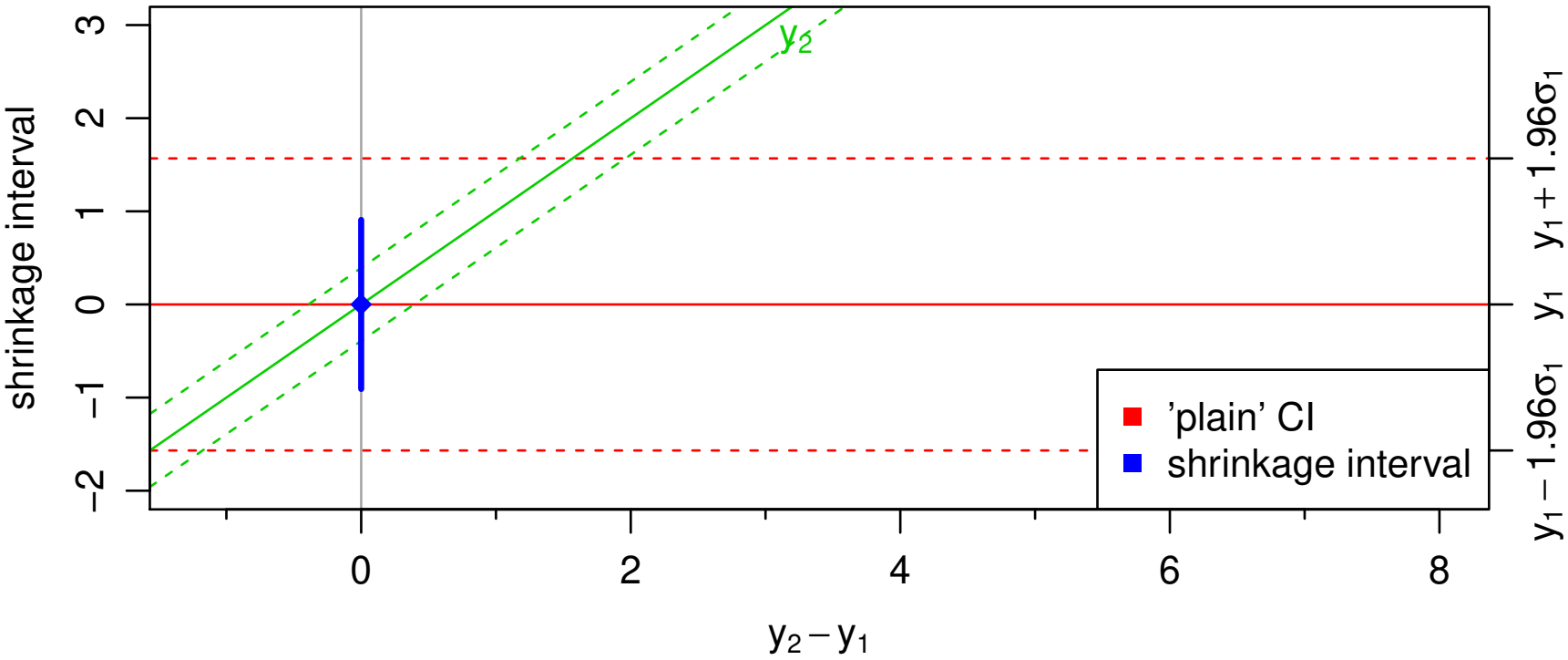
Two-study scenario



- $\sigma_1 = 0.8,\ \sigma_2 = 0.2$, interested in θ_1

Shrinkage estimation

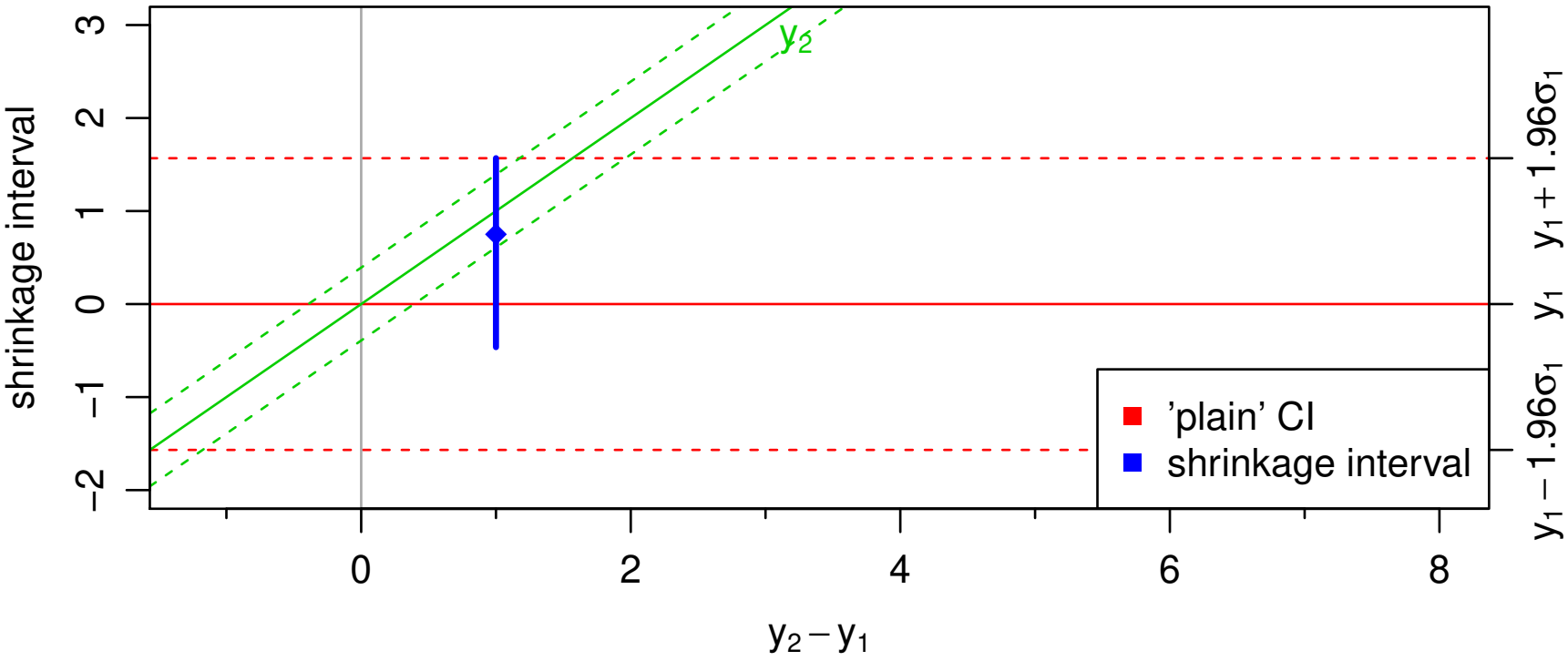
Two-study scenario



- $\sigma_1 = 0.8,\ \sigma_2 = 0.2$, interested in θ_1

Shrinkage estimation

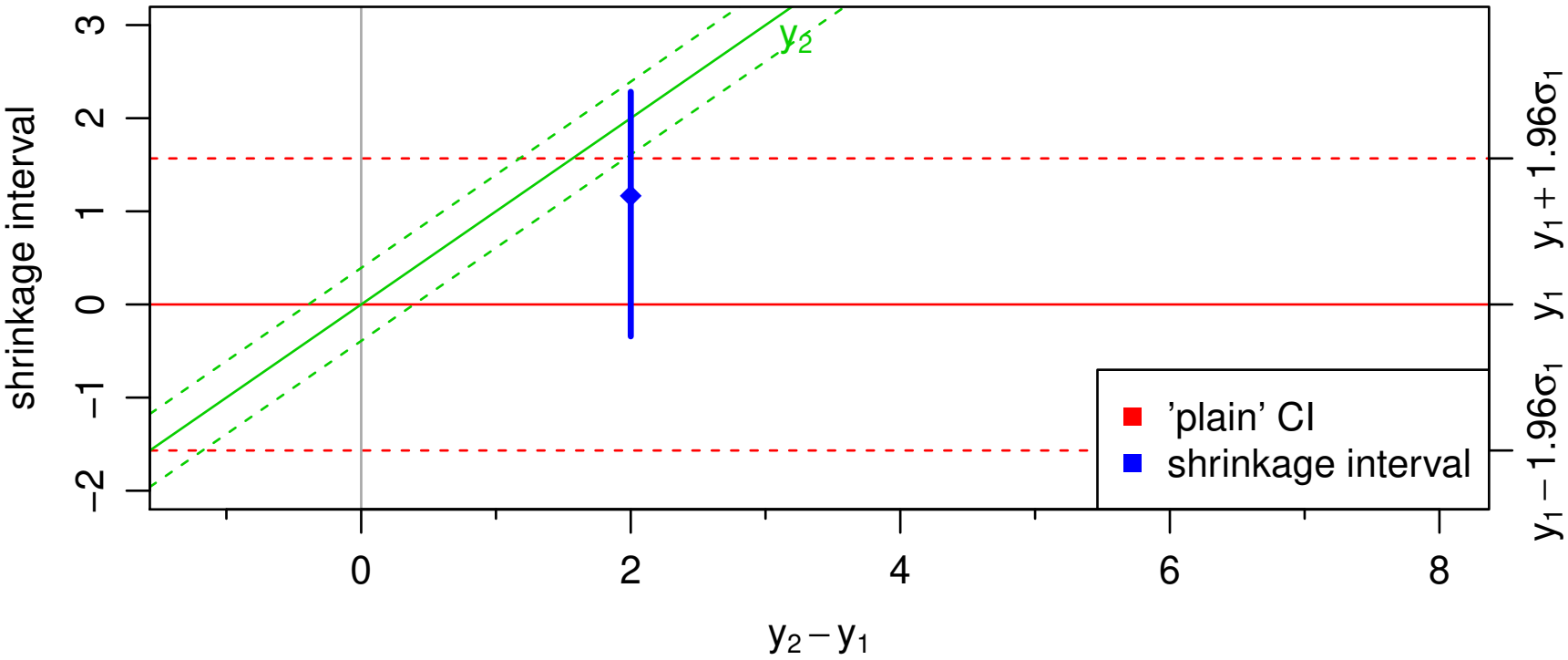
Two-study scenario



- $\sigma_1 = 0.8,\ \sigma_2 = 0.2$, interested in θ_1

Shrinkage estimation

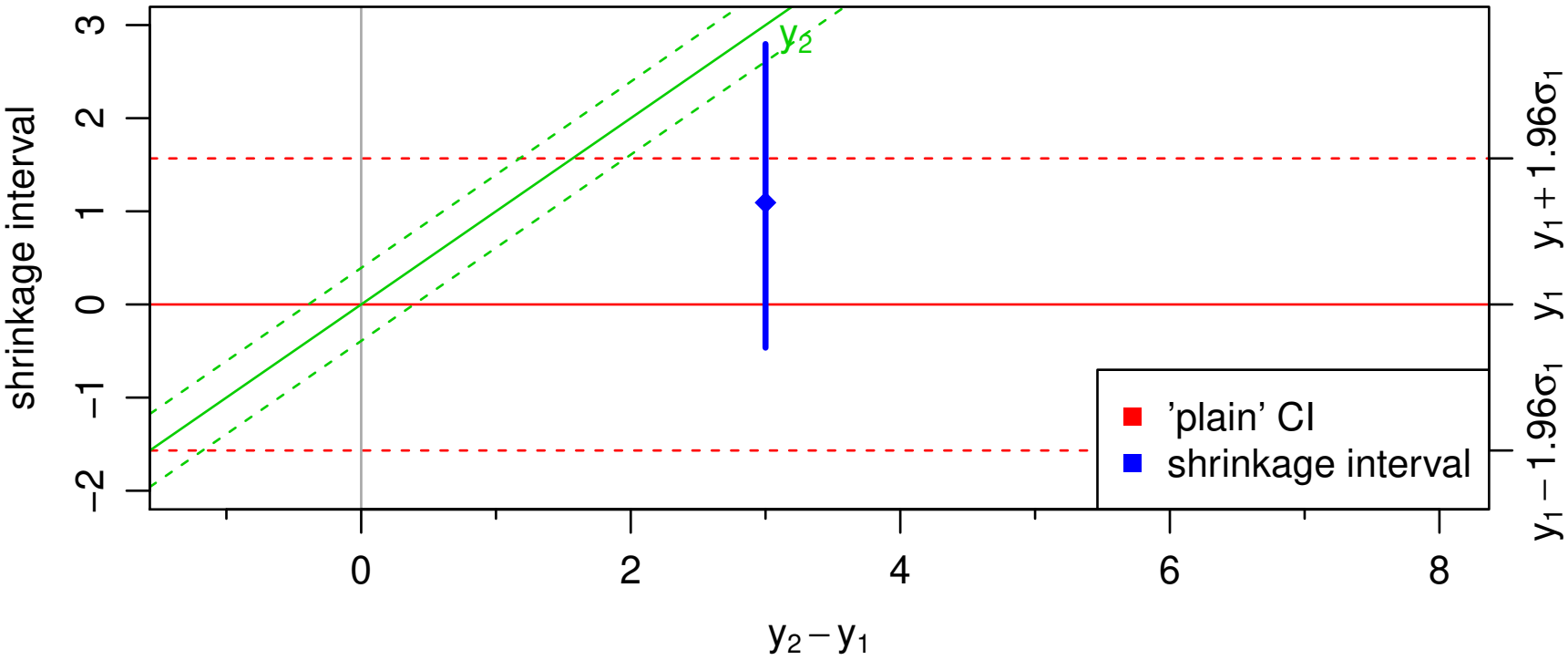
Two-study scenario



- $\sigma_1 = 0.8,\ \sigma_2 = 0.2$, interested in θ_1

Shrinkage estimation

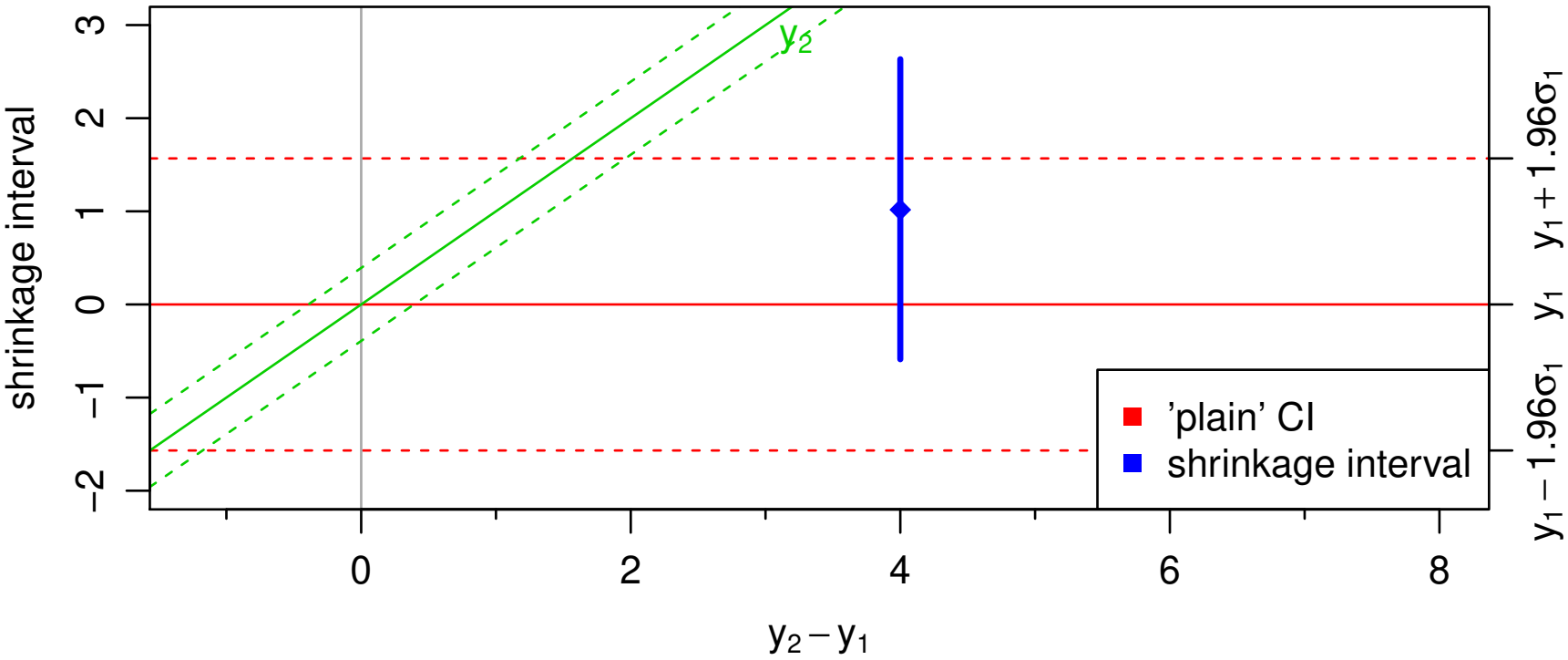
Two-study scenario



- $\sigma_1 = 0.8,\ \sigma_2 = 0.2$, interested in θ_1

Shrinkage estimation

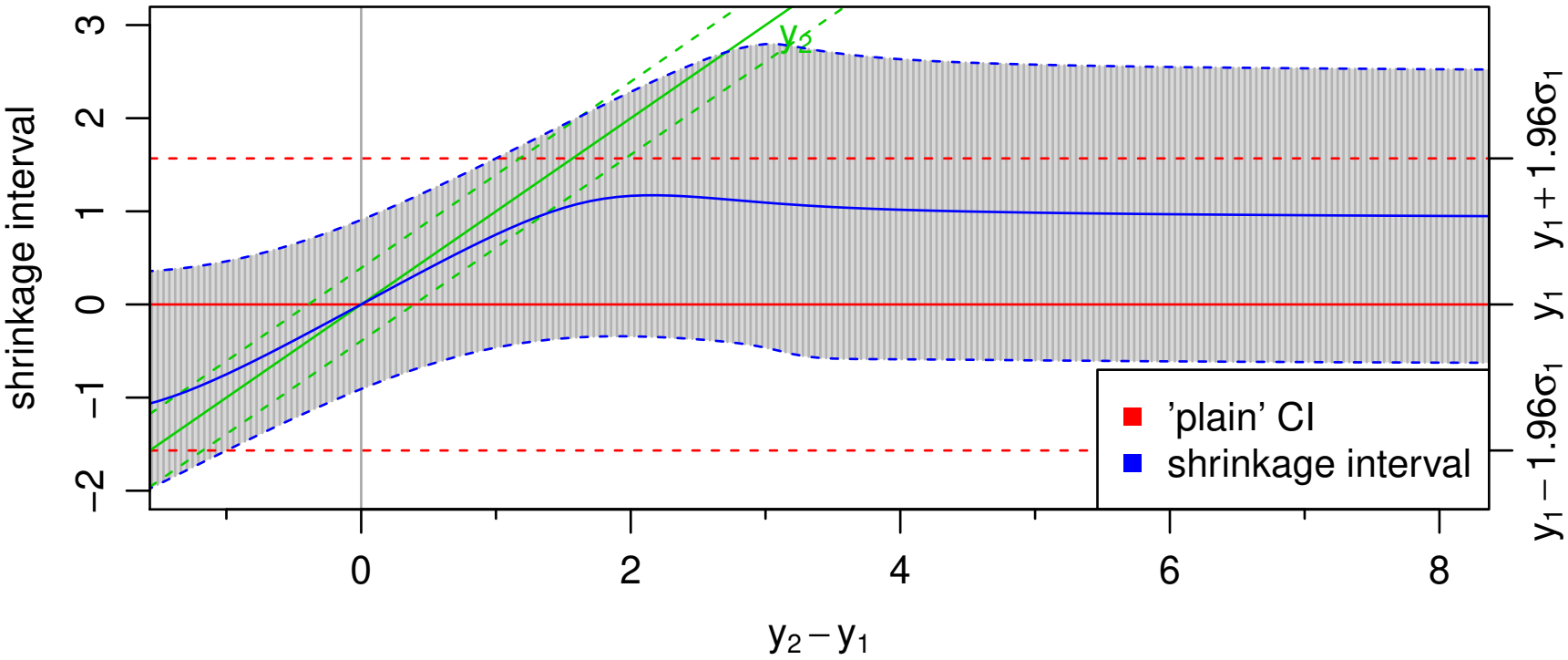
Two-study scenario



- $\sigma_1 = 0.8,\ \sigma_2 = 0.2$, interested in θ_1

Shrinkage estimation

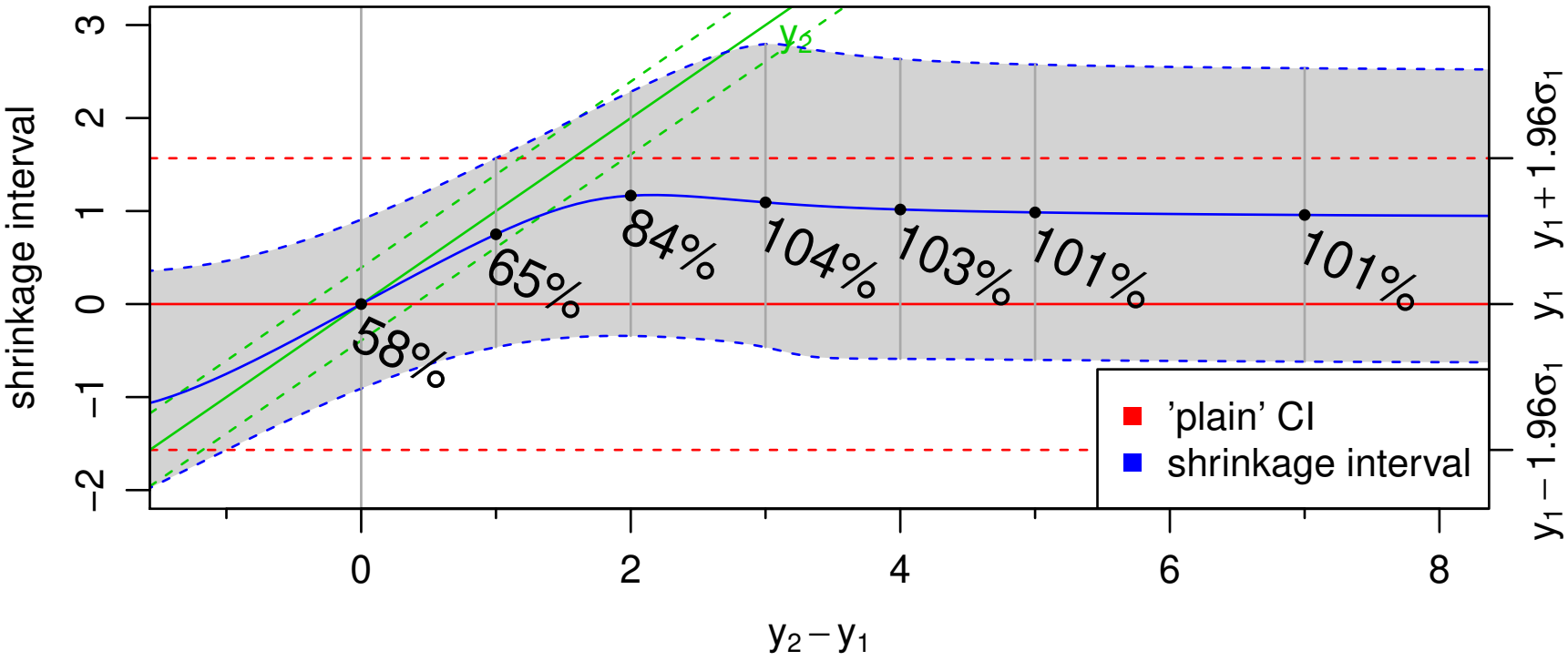
Two-study scenario



- $\sigma_1 = 0.8,\ \sigma_2 = 0.2$, interested in θ_1
- **robust** behaviour

Shrinkage estimation

Two-study scenario



- $\sigma_1 = 0.8,\ \sigma_2 = 0.2$, interested in θ_1
- **robust** behaviour
- relative **shrinkage interval width**: may be substantially shorter

Shrinkage estimation

Two-study simulations

- how do shrinkage intervals behave **on average**?
- what gain can we expect (if any)?

- investigate:
  - coverage
  - interval width

- consider again pairs of studies (binary endpoint);
  $n_1, n_2 \in \{25, 100, 400\}$,
  $\sigma_1, \sigma_2 \in \{0.8, 0.4, 0.2\}$
- prior: uniform prior for μ, half-Normal(0.5) for heterogeneity τ
  (sensitivity analysis with half-Normal(1.0))
- derive estimate for θ_1

Shrinkage estimation

Two-study simulations: **coverage** (%)

	τ						
		small	*moderate*	*substantial*	*large*	*very large*	
n_1/n_2	0.0	0.1	0.2	0.5	1.0	2.0	$*$
25/400	99.8	99.5	99.0	93.4	84.1	79.4	94.7
25/100	98.7	98.8	98.3	93.6	86.1	79.9	95.1
100/400	98.5	98.1	97.2	93.3	90.7	90.6	94.9
25/25	96.7	96.8	96.1	94.6	90.4	84.5	95.0
100/100	96.8	96.7	96.4	94.0	91.3	91.0	95.7
400/400	96.9	96.7	95.0	93.9	93.9	94.1	95.0
100/25	96.0	95.8	95.1	94.8	93.9	92.6	94.7
400/100	95.2	95.8	95.2	94.8	93.7	93.8	95.1
400/25	95.2	94.9	95.3	94.7	94.8	94.5	95.3

$*$: heterogeneity τ drawn from prior distribution

Shrinkage estimation

Two-study simulations: **coverage** (%)

	τ						
		small	*moderate*	*substantial*	*large*	*very large*	
n_1/n_2	0.0	0.1	0.2	0.5	1.0	2.0	$*$
25/400	99.8	99.5	99.0	93.4	84.1	79.4	94.7
25/100	98.7	98.8	98.3	93.6	86.1	79.9	95.1
100/400	98.5	98.1	97.2	93.3	90.7	90.6	94.9
25/25	96.7	96.8	96.1	94.6	90.4	84.5	95.0
100/100	96.8	96.7	96.4	94.0	91.3	91.0	95.7
400/400	96.9	96.7	95.0	93.9	93.9	94.1	95.0
100/25	96.0	95.8	95.1	94.8	93.9	92.6	94.7
400/100	95.2	95.8	95.2	94.8	93.7	93.8	95.1
400/25	95.2	94.9	95.3	94.7	94.8	94.5	95.3

$*$: heterogeneity τ drawn from prior distribution

- **good coverage** for non-extreme heterogeneity

Shrinkage estimation

Two-study simulations: **relative interval width** (%)

	τ						
		small	*moderate*	*substantial*	*large*	*very large*	
n_1/n_2	0.0	0.1	0.2	0.5	1.0	2.0	$*$
25/400	62.3	62.7	63.0	65.6	72.1	83.1	65.1
25/100	67.5	67.4	67.9	69.8	75.2	84.2	69.5
100/400	78.5	78.7	79.9	85.2	91.4	95.9	83.4
25/25	78.9	79.0	79.0	79.7	81.8	86.8	79.7
100/100	85.1	85.4	85.7	88.5	92.5	96.2	87.5
400/400	89.9	90.5	91.9	95.5	97.8	99.0	93.7
100/25	92.9	92.9	93.0	93.4	94.6	96.6	93.3
400/100	95.0	95.1	95.4	96.7	98.1	99.1	96.2
400/25	98.0	98.0	98.1	98.2	98.6	99.2	98.2

$*$: heterogeneity τ drawn from prior distribution

Shrinkage estimation

Two-study simulations: **relative interval width** (%)

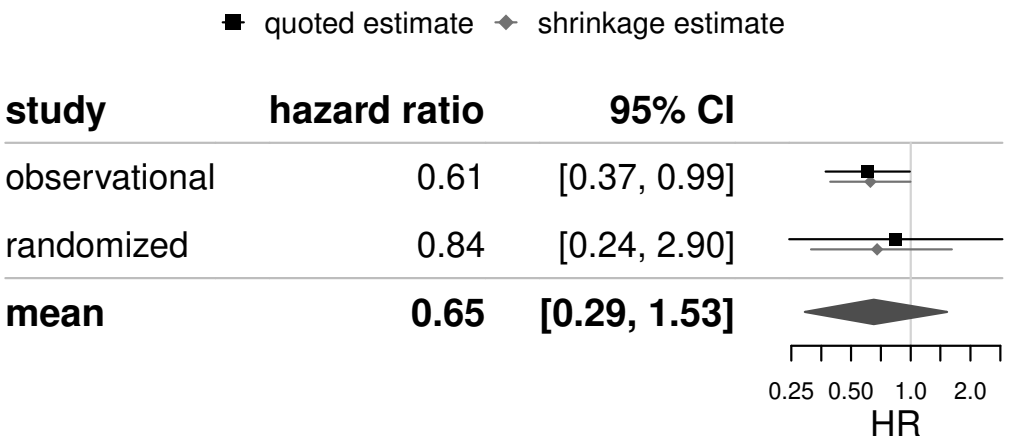
	τ						
		small	*moderate*	*substantial*	*large*	*very large*	
n_1/n_2	0.0	0.1	0.2	0.5	1.0	2.0	$*$
25/400	62.3	62.7	63.0	65.6	72.1	83.1	65.1
25/100	67.5	67.4	67.9	69.8	75.2	84.2	69.5
100/400	78.5	78.7	79.9	85.2	91.4	95.9	83.4
25/25	78.9	79.0	79.0	79.7	81.8	86.8	79.7
100/100	85.1	85.4	85.7	88.5	92.5	96.2	87.5
400/400	89.9	90.5	91.9	95.5	97.8	99.0	93.7
100/25	92.9	92.9	93.0	93.4	94.6	96.6	93.3
400/100	95.0	95.1	95.4	96.7	98.1	99.1	96.2
400/25	98.0	98.0	98.1	98.2	98.6	99.2	98.2

$*$: heterogeneity τ drawn from prior distribution

- substantial **precision gain** possible

Shrinkage estimation

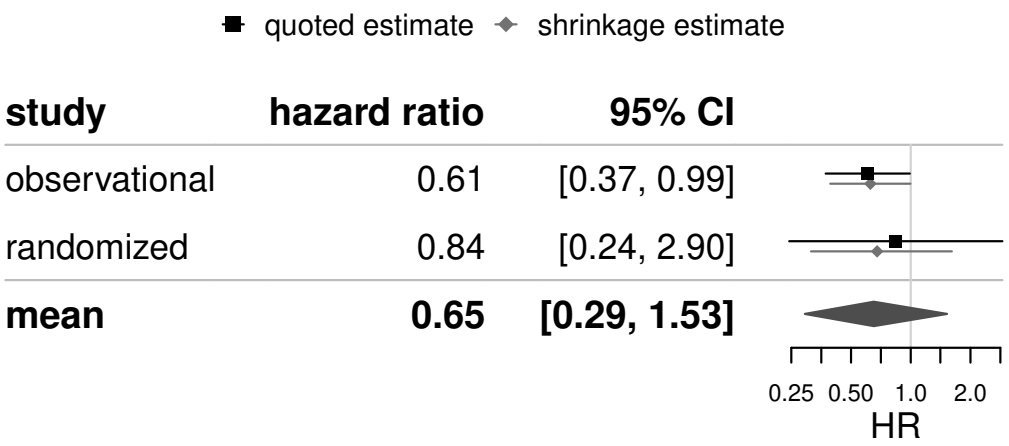
The Creutzfeld-Jakob disease (CJD) example



- shrinkage interval width: 66%
- 129% sample size gain (12 $\rightarrow$ $\approx$27 patients)
- results not dominated by external data (only $\approx$15 of 88 pts. contributed)

Shrinkage estimation

The Creutzfeld-Jakob disease (CJD) example



- shrinkage interval width: 66%
- 129% sample size gain (12 $\rightarrow$ $\approx$27 patients)
- results not dominated by external data (only $\approx$15 of 88 pts. contributed)

```
> require("bayesmeta")
> # perform analysis:
> bm <- bayesmeta(y = cjd$logHR, sigma = cjd$logHR.se,
+                 labels = cjd$study,
+                 tau.prior = function(t){dhalfnormal(t, scale=0.5)})
>
> # show shrinkage estimates:
> print(exp(bm$theta[c(7,4,8),"randomized"]))
95% lower    median 95% upper
0.3142006 0.6767489 1.6112718
```

Conclusions

Shrinkage estimation for 2 studies

- readily motivated, transparent
- valid (coverage close to nominal level)
- robust behaviour
- potentially substantial gain despite 'pathological' setting ($k = 2$)
- especially if external data come with great precision ($\sigma_2 \leq \sigma_1$)
- special "$k = 2$"-case: alternative parametrisation possible (reference to "overall mean" μ not necessary)
- article under review[5]
- computations quick & easy using **`bayesmeta`** R package[6]

[5] C. Röver, T. Friede. Dynamically borrowing strength from another study. *arXiv preprint 1806.01015* (submitted for publication), 2018.

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

+++ additional slides +++

CJD example

R code

```
cjd <- cbind.data.frame("study"    = c("observational", "randomized"),
                        "logHR"    = c(-0.49948, -0.17344),
                        "logHR.se" = c(0.2493, 0.6312))

# analyze:
require("bayesmeta")
bm <- bayesmeta(y         = cjd$logHR,
                sigma     = cjd$logHR.se,
                labels    = cjd$study,
                tau.prior = function(t){dhalfnormal(t, scale=0.5)})

# show results:
print(bm)

# show forest plot:
forestplot(bm, xlab="log-HR")

# show shrinkage estimates:
print(bm$theta)
print(exp(bm$theta[c(7,4,8),"randomized"]))
```

Alternative model parametrization

Details

- the **normal-normal hierarchical model (NNHM)**:

$$\begin{aligned} y_i|\theta_i &\sim \text{Normal}(\theta_i,\, \sigma_i^2), \\ \theta_i|\mu,\tau &\sim \text{Normal}(\mu,\, \tau^2) \qquad (\text{for } i = 1,\ldots,k) \end{aligned}$$

- the alternative **reference model**:

$$\begin{aligned} y_i|\vartheta_i &\sim \text{Normal}(\vartheta_i,\, \sigma_i^2), \\ \vartheta_1|\alpha,\beta &\sim \text{Normal}(\alpha,\, 0) \qquad (\text{i.e., } \vartheta_1 = \alpha), \\ \vartheta_2|\alpha,\beta &\sim \text{Normal}(\alpha,\, \beta^2) \end{aligned}$$

- both models yield **identical shrinkage estimates**[7] for $k = 2$ and
  - (improper) uniform priors for μ and α
  - (any) heterogeneity prior with density $p(\tau) = f_\star(\tau)$, and matching prior with density $p(\beta) = \frac{1}{\sqrt{2}} f_\star\left(\frac{\beta}{\sqrt{2}}\right)$ for β

[7] C. Röver, T. Friede. Dynamically borrowing strength from another study. *arXiv preprint 1806.01015* (submitted for publication), 2018.

Heterogeneity (τ)

Half-Normal prior: motivation (1)

- **recommended family**: half-t, half-Normal, half-Cauchy (*not* recommended: inverse-Gamma)[8]
- effect measure here: logarithmic **ratio** (odds ratio, hazard ratio,...)
- heterogeneity τ may be translated into implied **spread of effects** θ_i and $\exp(\theta_i)$
- Spiegelhalter et al. (2004)[9] proposed **categories**
  - "reasonable": $0.1 < \tau < 0.5$
  - "fairly high": $0.5 < \tau < 1.0$
  - "fairly extreme": $\tau > 1.0$
- Turner & al. (2015)[10] empirically investigated heterogeneity in meta-analyses archived in the Cochrane Library

---

[8] A. Gelman. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis* 1(3):515–534, 2006.

[9] D.J. Spiegelhalter, K.R. Abrams, J.P. Myles. *Bayesian approaches to clinical trials and health-care evaluation*. John Wiley & Sons, 2004. Sec. 5.7.

[10] 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.

Heterogeneity (τ)

Half-Normal prior: motivation (2)

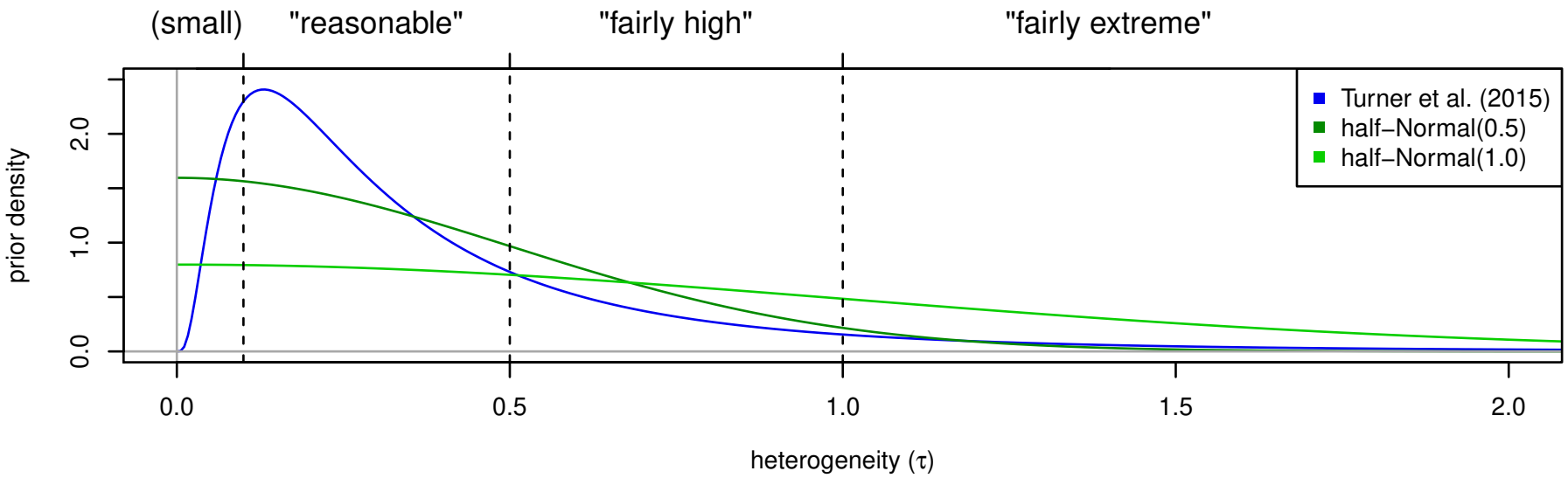
- proposed **categories**:
  - "reasonable": $0.1 < \tau < 0.5$
  - "fairly high": $0.5 < \tau < 1.0$
  - "fairly extreme": $\tau > 1.0$

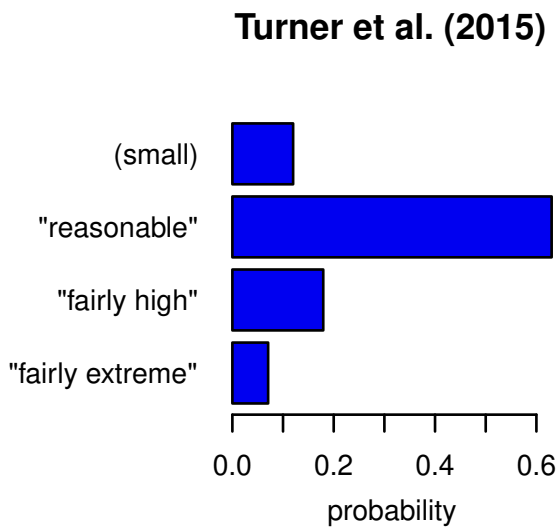
- **Implications** of certain τ values:
  95% range of effects $\exp(\theta_i)$
  spans a range of $\exp(3.92\tau)$
  (ratio largest / smallest)

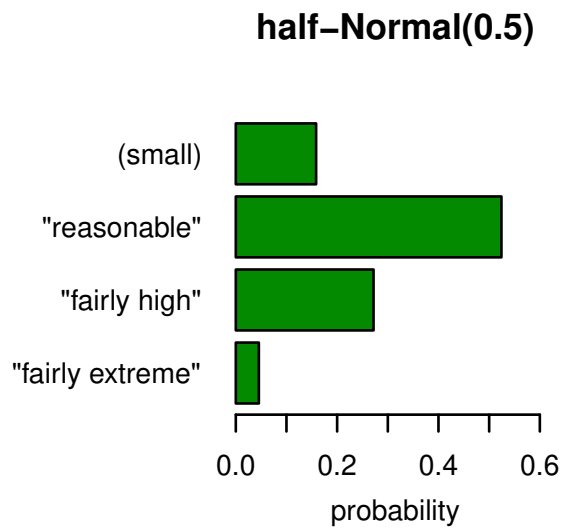
τ	$\exp(3.92\tau)$
0.0	1.00
0.1	1.48
0.2	2.19
0.5	7.10
1.0	50.4
2.0	2540

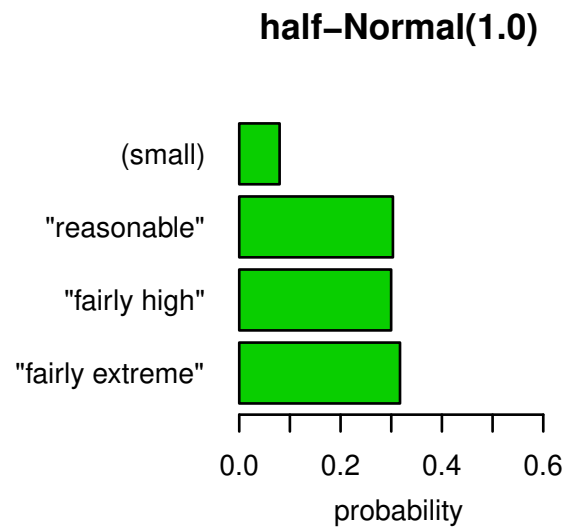
Heterogeneity (τ)

Half-Normal prior: motivation (3)









Standard errors and sample sizes

Heuristics

- assume: standard errors scale with $\frac{1}{\sqrt{N}}$
- doubling the sample size ($N = 2 \times N_0$) means a shorter s.e., shorter by a factor of $\frac{1}{\sqrt{2}} = 71\%$

N	$\frac{1}{\sqrt{N/N_0}}$
N_0	100 %
$2N_0$	71 %
$3N_0$	58 %
$4N_0$	50 %
$\vdots$	$\vdots$

Standard errors and sample sizes

Heuristics

- assume: standard errors scale with $\frac{1}{\sqrt{N}}$
- doubling the sample size ($N = 2 \times N_0$) means a shorter s.e., shorter by a factor of $\frac{1}{\sqrt{2}} = 71\%$

N	$\frac{1}{\sqrt{N/N_0}}$
N_0	100 %
$2N_0$	71 %
$3N_0$	58 %
$4N_0$	50 %
⋮	⋮

- inversely: a SE only $\frac{\sigma}{\sigma_0} = 71\,\%$ as wide implies a 100% gain in sample size
- generally: *effective sample size gain* $(\frac{\sigma}{\sigma_0})^{-2} - 1$

σ/σ_0	gain
100 %	0 %
90 %	23 %
80 %	56 %
70 %	104 %
50 %	300 %
⋮	⋮

Shrinkage estimation

Two-study simulations: **relative sample size gain** (%)

	τ						
n_1/n_2	0.0	0.1	0.2	0.5	1.0	2.0	$*$
25/400	162	160	158	144	113	68.4	147
25/100	123	123	121	111	89.6	56.3	113
100/400	64.5	64.0	60.0	43.8	25.7	12.7	49.4
25/25	61.2	60.9	60.7	58.4	51.8	36.9	58.7
100/100	38.8	38.1	37.1	29.6	19.4	10.1	32.3
400/400	24.2	22.9	19.4	11.0	5.5	2.4	15.1
100/25	15.9	16.0	15.8	14.8	11.9	7.5	14.9
400/100	11.0	10.7	10.0	7.3	4.2	2.0	8.3
400/25	4.1	4.1	4.0	3.7	2.9	1.7	3.7

$*$: heterogeneity τ drawn from prior distribution