UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG

COMBINING RANDOMIZED CONTROLLED TRIALS AND REAL WORLD DATA

Kolloquium Berlin, 17 OCT 2023

Tim Friede

Department of Medical Statistics

University Medical Center Göttingen, Germany

ACKNOWLEDGEMENTS

- The talk draws on **joint work** with a number of colleagues including
  - Christian Röver
  - Sarah Friedrich (Augsburg)
  - Tim Mathes

- Grateful for **funding** by DFG (German Research Foundation; FR 3070/3-1)

DFG Deutsche Forschungsgemeinschaft

INNOVATIVE CLINICAL TRIALS

UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG

NEJM Evidence

Published May 24, 2022
NEJM Evid 2022; 1 (6)
DOI: 10.1056/EVIDctw2200060

CLINICAL TRIALS WORKSHOP | INNOVATION MINI-SERIES

Clinical Trials in the 21st Century — Promising Avenues for Better Studies

Michael J. Pencina, Ph.D.,[1] and B. Taylor Thompson, M.D.[2]

nature reviews drug discovery

COMMENT | 27 June 2022

Advancing innovative clinical trials to efficiently deliver medicines to patients

Complex innovative designs in clinical trials have the potential to increase efficiency and lower the cost of drug development, improving patient access to therapies. This article highlights designs and approaches based on a meeting linked to an ongoing FDA pilot program in the field.

INNOVATIVE CLINICAL TRIALS

UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG

NEJM Evidence

Published May 24, 2022
NEJM Evid 2022; 1 (6)
DOI: 10.1056/EVIDctw2200060

CLINICAL TRIALS WORKSHOP | INNOVATION MINI-SERIES

Clinical Trials in the 21st Century — Promising Avenues for Better Studies

Michael J. Pencina, Ph.D.,[1] and B. Taylor Thompson, M.D.[2]

Go Bayesian!

Bayesian approaches

- Borrowing in master protocols / basket trials
- Utilizing external data
- Adaptive designs

nature reviews drug discovery

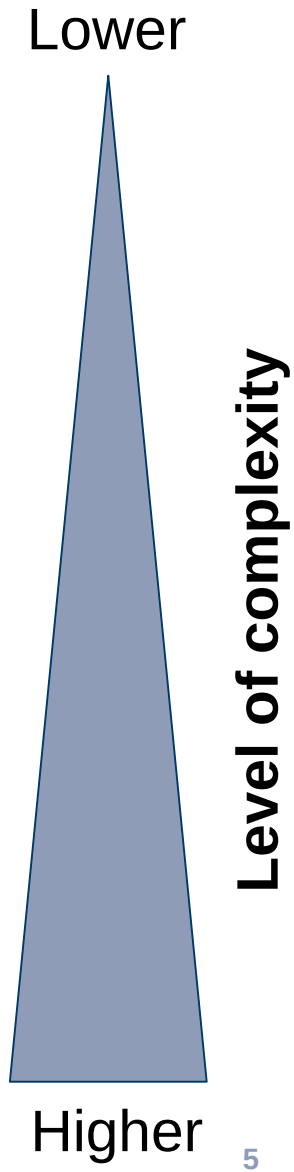
COMMENT | 27 June 2022

Advancing innovative clinical trials to efficiently deliver medicines to patients

Complex innovative designs in clinical trials have the potential to increase efficiency and lower the cost of drug development, improving patient access to therapies. This article highlights designs and approaches based on a meeting linked to an ongoing FDA pilot program in the field.

EVIDENCE SYNTHESIS

- **Pairwise meta-analysis**
  - comparing two treatments
- **Meta-regression**
  - including study-level covariates
- **Network meta-analysis**
  - comparing multiple treatments indirectly
- **RCT with historical controls**
  - integrating control group data from previous trials
- **Generalized (or cross design) synthesis**
  - combining data from different types of studies

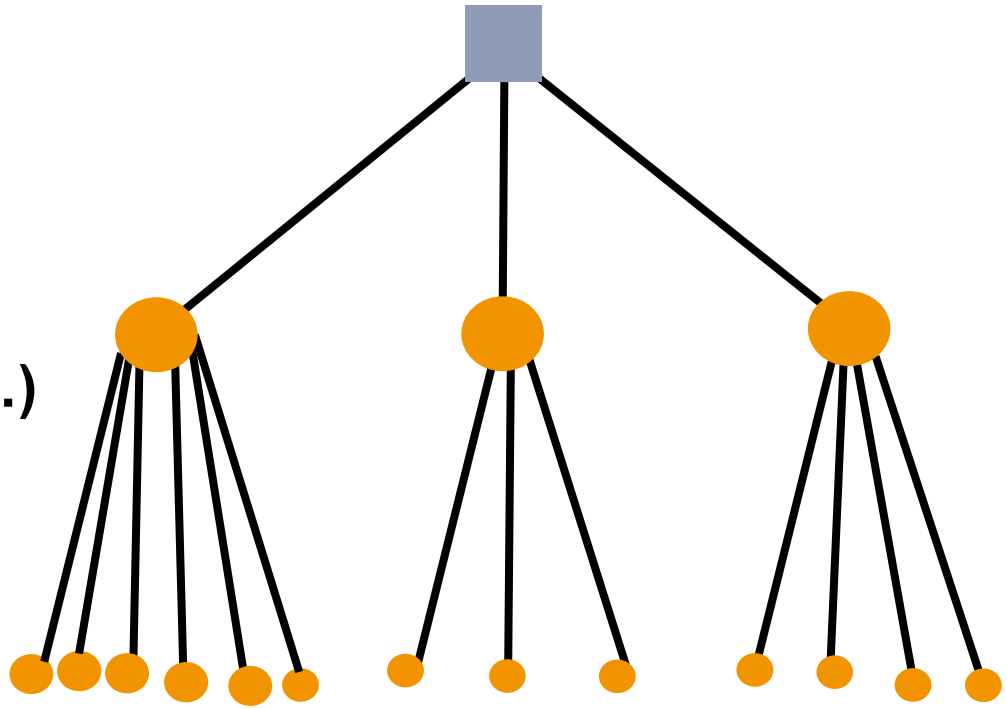


HIERARCHICAL MODELS

Meta-analysis

Studies (RCT, registry, …)

Patients



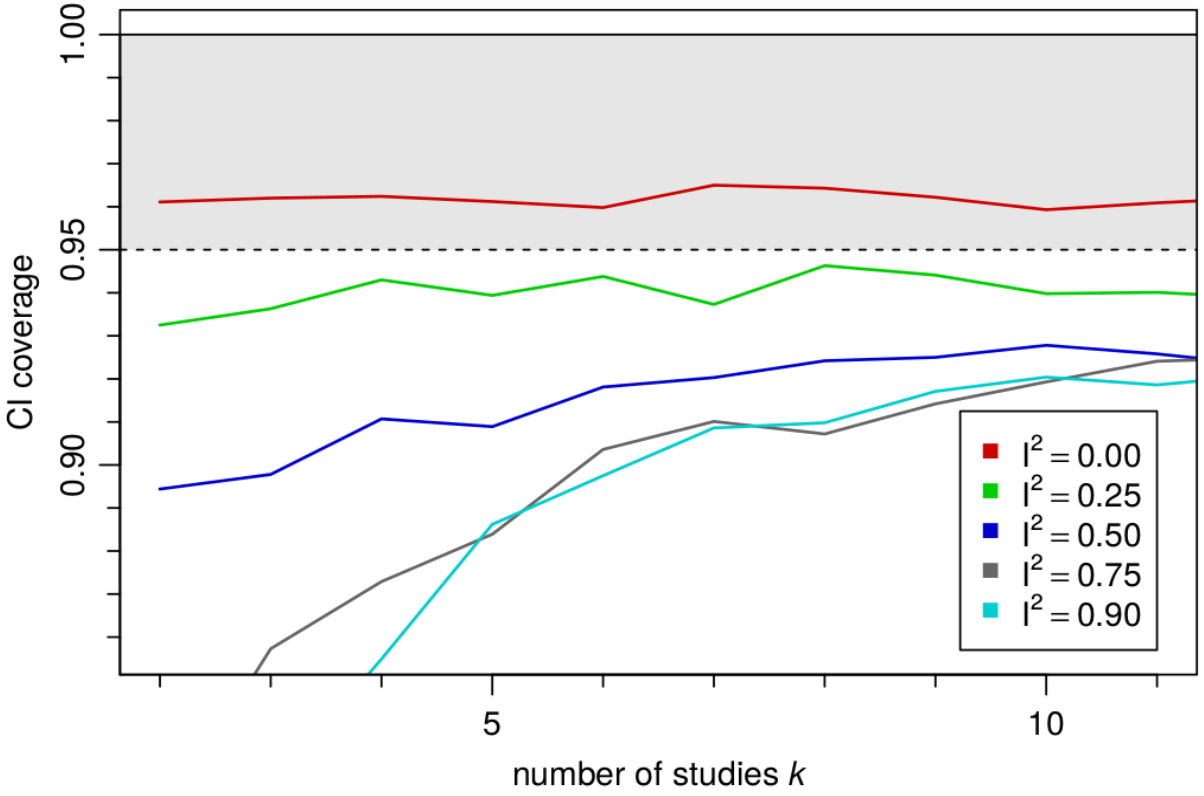
Example: **Normal-normal hierarchical model (NNHM)** for random-effects meta-analysis

$$y_i|\theta_i \sim N(\theta_i, \sigma_i^2), \qquad \theta_i|\mu,\tau \sim N(\mu,\tau^2)$$

STANDARD METHOD FAILS

UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG

- **Standard method for random-effects meta-analysis** (DerSimonian-Laird) **<u>with (very) few studies</u>**
  - Underestimates between-study heterogeneity
  - Fails to account for uncertainty in estimation of heterogeneity



IntHout et al, 2014; Röver et al, 2015

BAYESIAN META-ANALYSIS

- **Idea**: Weakly informative prior on between-trial heterogeneity τ for meta-analysis with few studies (Spiegelhalter et al, 2004), with uninformative prior on treatment effect μ
  - Avoids zero estimates of between-trial heterogeneity
  - Accounts for uncertainty in the estimation of the heterogeneity
- **Easy to compute**
  - Application of DIRECT algorithm (Röver & Friede, 2017) (which is faster than MCMC sampling and does not require inspection of convergence diagnostics)
  - R package **`bayesmeta`** by Christian Röver (available from CRAN)

"WHERE DOES THE PRIOR COME FROM?"

- Theoretical arguments, simulations, empirical data

Received: 16 July 2020 | Revised: 13 January 2021 | Accepted: 16 January 2021
DOI: 10.1002/jrsm.1475

RESEARCH ARTICLE

Research Synthesis Methods WILEY

On weakly informative prior distributions for the heterogeneity parameter in Bayesian random-effects meta-analysis

Christian Röver[1] | Ralf Bender[2] | Sofia Dias[3] | Christopher H. Schmid[4] | Heinz Schmidli[5] | Sibylle Sturtz[2] | Sebastian Weber[6] | Tim Friede[1]

Received: 25 February 2022 | Revised: 16 November 2022 | Accepted: 18 March 2023
DOI: 10.1002/sim.9731

RESEARCH ARTICLE

Statistics in Medicine WILEY

Summarizing empirical information on between-study heterogeneity for Bayesian random-effects meta-analysis

Christian Röver[1] | Sibylle Sturtz[2] | Jona Lilienthal[2] | Ralf Bender[2] | Tim Friede[1]

PRIORS COVERING SMALL TO LARGE HETEROGENEITY ON LOG-ODDS RATIO SCALE

Table 1. Between-trial heterogeneity for log-odds ratios: τ values representing small to very large heterogeneity, with 95% intervals for across-trial odds ratios (exp (θ_j)).

Heterogeneity		95% interval
Small:	$\tau = 0.125$	0.783–1.28
Moderate:	$\tau = 0.25$	0.613–1.63
Substantial:	$\tau = 0.5$	0.325–2.66
Large:	$\tau = 1$	0.141–7.10
Very large:	$\tau = 2$	0.020–50.4

Table 2. Between-trial heterogeneity for log-odds ratios: three priors covering small to large heterogeneity.

Prior distribution	Median	95% interval
Half normal (scale = 0.5)	0.337	(0.016, 1.12)
Half normal (scale = 1.0)	0.674	(0.031, 2.24)
Uniform (0, 4)	2.0	(0.1, 3.9)

Friede et al. (2017) RSM

AN EXAMPLE IN HEART FAILURE

ESC European Society of Cardiology

European Journal of Heart Failure (2023) **25**, 1080–1090
doi:10.1002/ejhf.2860

RESEARCH ARTICLE

Effect of intravenous iron replacement on recurrent heart failure hospitalizations and cardiovascular mortality in patients with heart failure and iron deficiency: A Bayesian meta-analysis

Stefan D. Anker[1]*, Muhammad Shahzeb Khan[2], Javed Butler[3,4], Stephan von Haehling[5], Ewa A. Jankowska[6], Piotr Ponikowski[6], and Tim Friede[7]

AN EXAMPLE IN HEART FAILURE

UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG

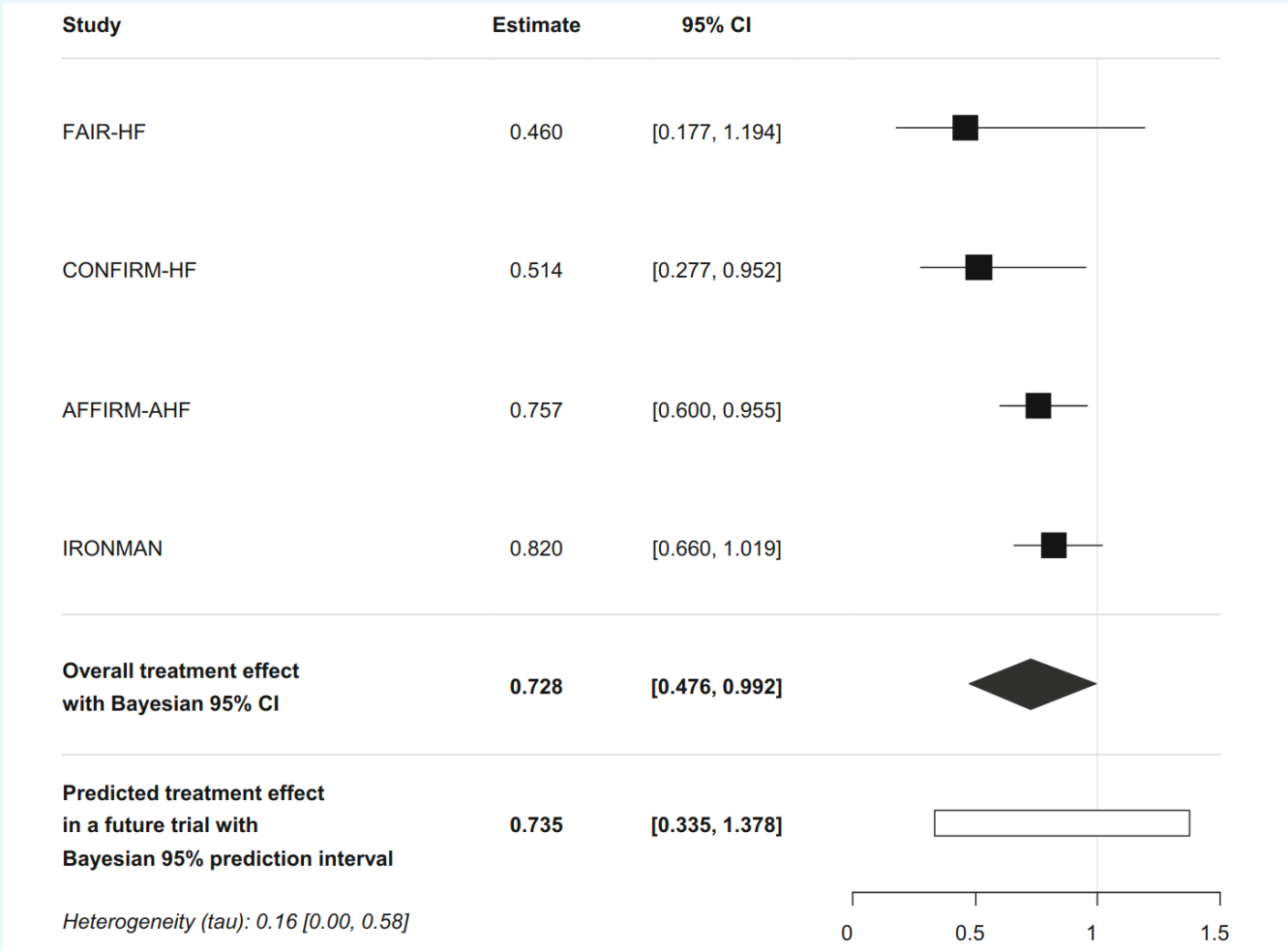


Figure 1 Forest plot for the effect of intravenous iron on the composite endpoint of total heart failure hospitalizations and cardiovascular mortality using Bayesian random-effects meta-analysis. The black squares represent the effects of the individual studies. The prediction estimate indicates what the effect in a future trial maybe. CI, credible interval.

Anker et al (2023) EJHF

AN EXAMPLE IN HEART FAILURE

- **Implementation in R with bayesmeta**

```r
library("bayesmeta")

# HF hospitalisations or CV death (LWYY)
# Bayesian random-effects meta-analysis

taupriordensity <- function(t){dhalfnormal(t, scale=0.5)}
taupriordensity_sens <- function(t){dhalfnormal(t, scale=1)}

# Overall
y=c(-0.77613, -0.66598, -0.27842, log(0.82))
se=c(0.48667, 0.31459, 0.11881, (log(1.02)-log(0.66))/(2*qnorm(0.975)))
study = c("FAIR-HF",  "CONFIRM-HF", "AFFIRM-AHF", "IRONMAN")

bma <- bayesmeta(y = y, sigma = se, label = study, tau.prior=taupriordensity)
summary(bma)
forestplot(bma, exponentiate=TRUE)
```

AN EXAMPLE IN HEART FAILURE

- **For comparison, a standard (frequentist) meta-analysis** of the same for studies as in Anker et al (2023) EJHF …

```
y=c(-0.77613, -0.66598, -0.27842, log(0.82))
se=c(0.48667, 0.31459, 0.11881, (log(1.02)-log(0.66))/(2*qnorm(0.975)))
study = c("FAIR-HF",  "CONFIRM-HF", "AFFIRM-AHF", "IRONMAN")

# classical (frequentist) random-effects meta-analysis
ma01 <- rma.uni(yi=y,vi=se^2)
print(ma01)
forest(ma01, transf=exp, slab=study)
```

AN EXAMPLE IN HEART FAILURE

- **For comparison, a standard (frequentist) meta-analysis** of the same for studies as in Anker et al (2023) EJHF …

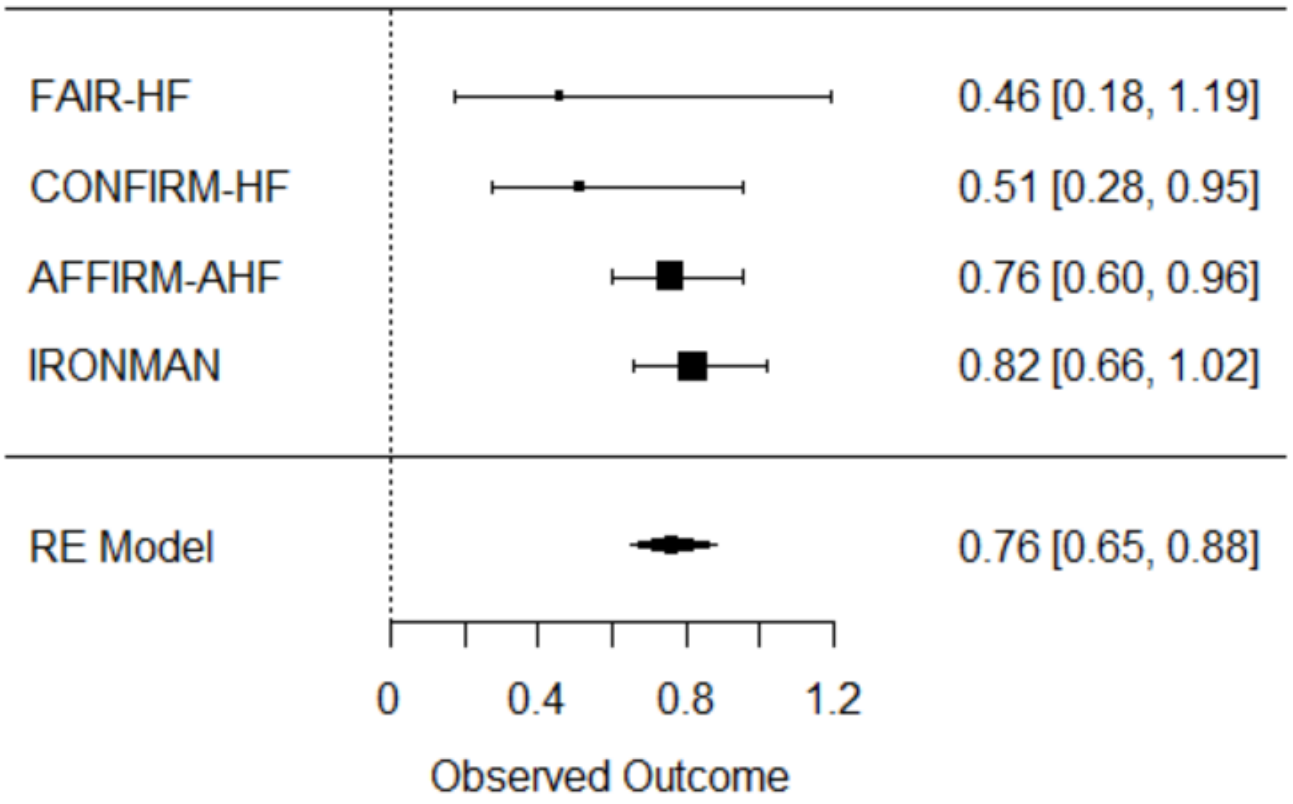
```
Random-Effects Model (k = 4; tau^2 estimator: REML)

tau^2 (estimated amount of total heterogeneity): 0 (SE = 0.0183)
tau (square root of estimated tau^2 value):      0
I^2 (total heterogeneity / total variability):   0.00%
H^2 (total variability / sampling variability):  1.00

Test for Heterogeneity:
Q(df = 3) = 3.0808, p-val = 0.3793
```

- Standard (frequentist) analysis estimates tau to be zero, i.e. no between-study heterogeneity.
- In my view very unlikely to be true …

AN EXAMPLE IN HEART FAILURE

UNIVERSITÄTSMEDIZIN GÖTTINGEN : UMG

- **For comparison, a standard (frequentist) meta-analysis** of the same for studies as in Anker et al (2023) EJHF …



- No between-trial heterogeneity results in (a) common-effect estimate and (b) shorter confidence interval (likely too short)

EXAMPLE: DOXYCYCLINE IN EARLY CREUTZFELDT-JAKOB DISEASE (CJD)

Neurodegeneration

OPEN ACCESS

RESEARCH PAPER

Doxycycline in early CJD: a double-blinded randomised phase II and observational study

Daniela Varges,[1] Henrike Manthey,[1] Uta Heinemann,[1] Claudia Ponto,[1] Matthias Schmitz,[1] Walter J Schulz-Schaeffer,[2] Anna Krasnianski,[1] Maren Breithaupt,[1] Fabian Fincke,[1] Katharina Kramer,[3] Tim Friede,[3] Inga Zerr[1]

http://dx.doi.org/10.1136/jnnp-2016-313541 (open access)

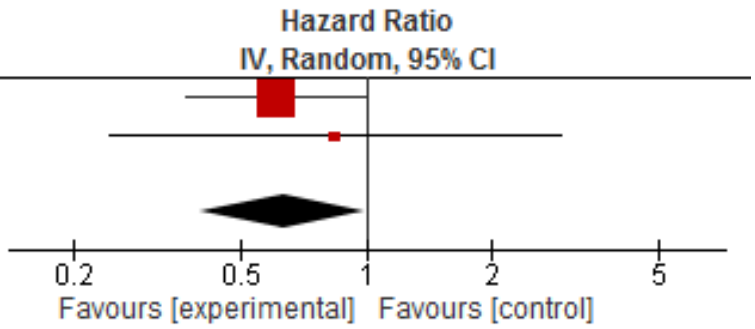
EXAMPLE: DOXYCYCLINE IN EARLY CREUTZFELDT-JAKOB DISEASE (CJD)

UNIVERSITÄTSMEDIZIN GÖTTINGEN : UMG

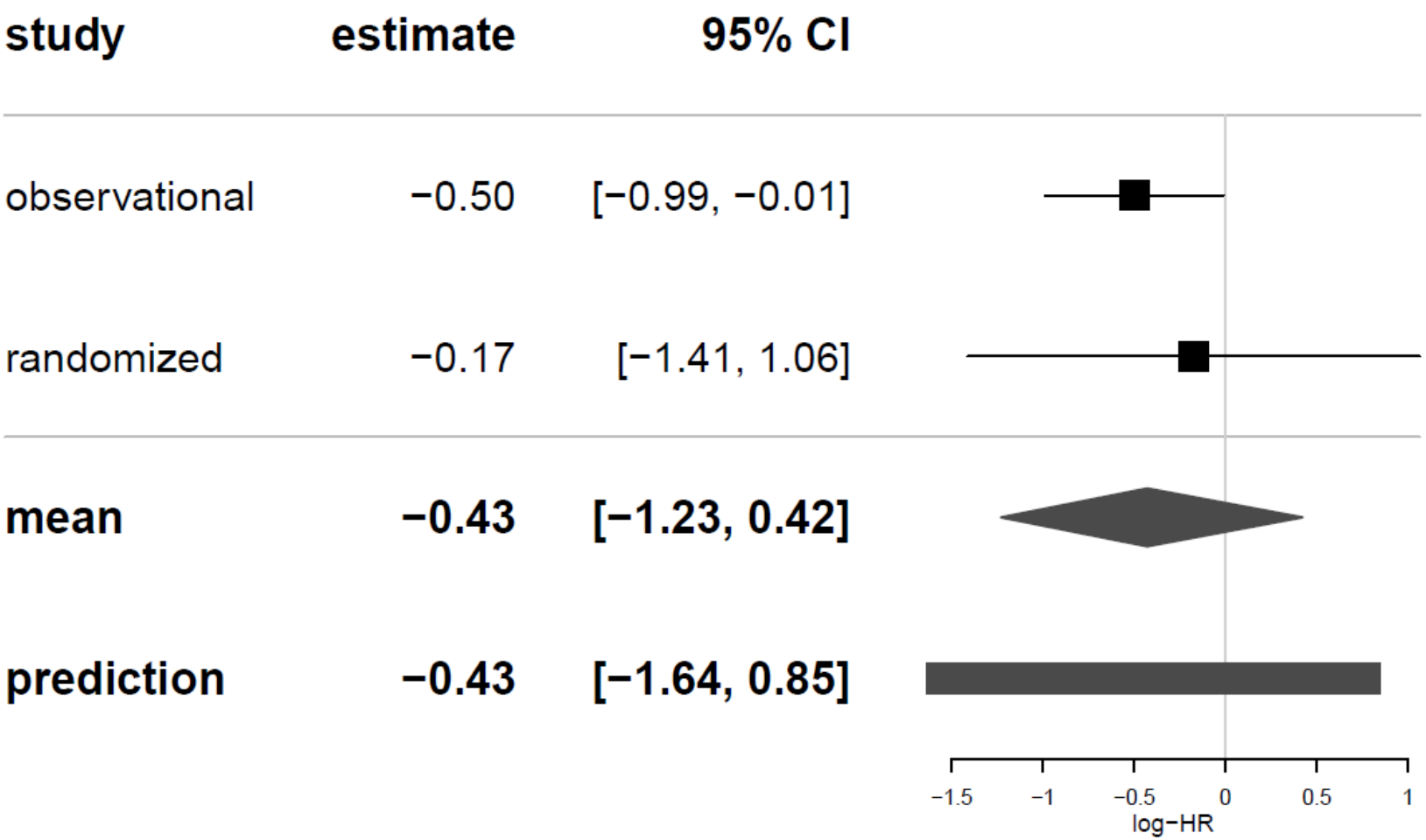
- **Creutzfeldt-Jakob disease**
  - prevalence of 1–9 cases per 1,000,000 people
  - qualifies as **rare disease** (EU: less than 5 in 10,000)
- **Varges et al (2017)** conducted:
  - double-blinded randomized phase II trial (n=12)
  - observational study (n=88) (Cox regression stratified by terciles of the propensity scores)
  - survival time as primary outcome

Study or Subgroup	log[Hazard Ratio]	SE	Weight	Hazard Ratio IV, Random, 95% CI
Observational study	-0.49948	0.2493	86.5%	0.61 [0.37, 0.99]
RCT	-0.17344	0.6312	13.5%	0.84 [0.24, 2.90]
Total (95% CI)			**100.0%**	**0.63 [0.40, 1.00]**

Heterogeneity: Tau² = 0.00; Chi² = 0.23, df = 1 (P = 0.63); I² = 0%
Test for overall effect: Z = 1.96 (P = 0.05)



EXAMPLE IN CJD: BAYESIAN RANDOM-EFFECTS META-ANALYSIS



Computed with **`bayesmeta`**; HN(0.5) prior for τ

QUANTITIES OF INTEREST

Different quantities of interest in hierarchical models

- **average effect** (μ) across studies
  - standard (pairwise) meta-analysis
- effect (θ_{k+1}) of a future study
  - **prediction / extrapolation**: e.g. adult to children; bridging
- effect (θ_i) of an individual study in the light of the other studies (**shrinkage estimator**)
  - e.g. **small RCT with borrowing from registry**; borrowing between subgroups in a basket trial; bridging study

EXAMPLE IN CJD: SHRINKAGE ESTIMATOR

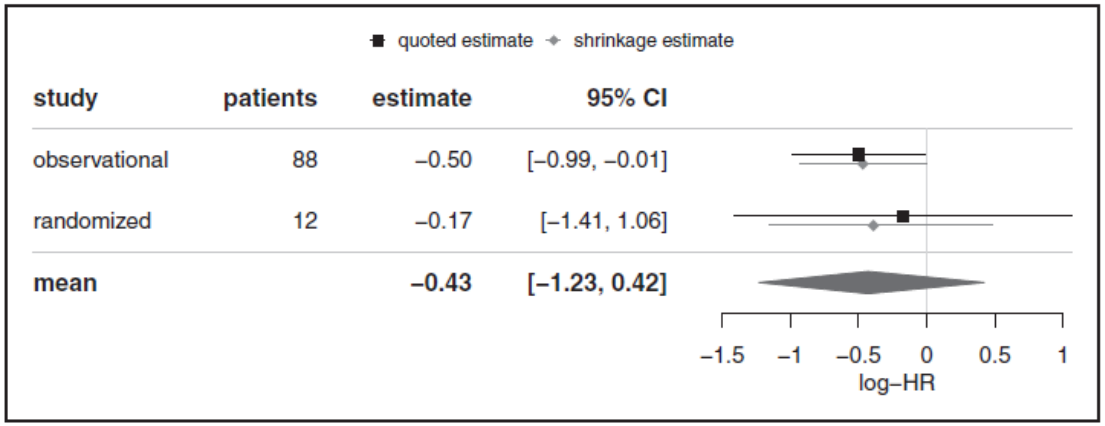


Figure 2. Forest plot for the CJD example (log-HR outcome). The shrinkage interval for the log-HR based on randomized evidence here is [−1.16, 0.48], spanning only two-thirds of the original confidence interval width.

- RCT shrinkage interval width: 66% of original CI width
- Translates into 129% gain in sample size (about 27 instead of 12 patients)

Röver & Friede (2020) SMMR

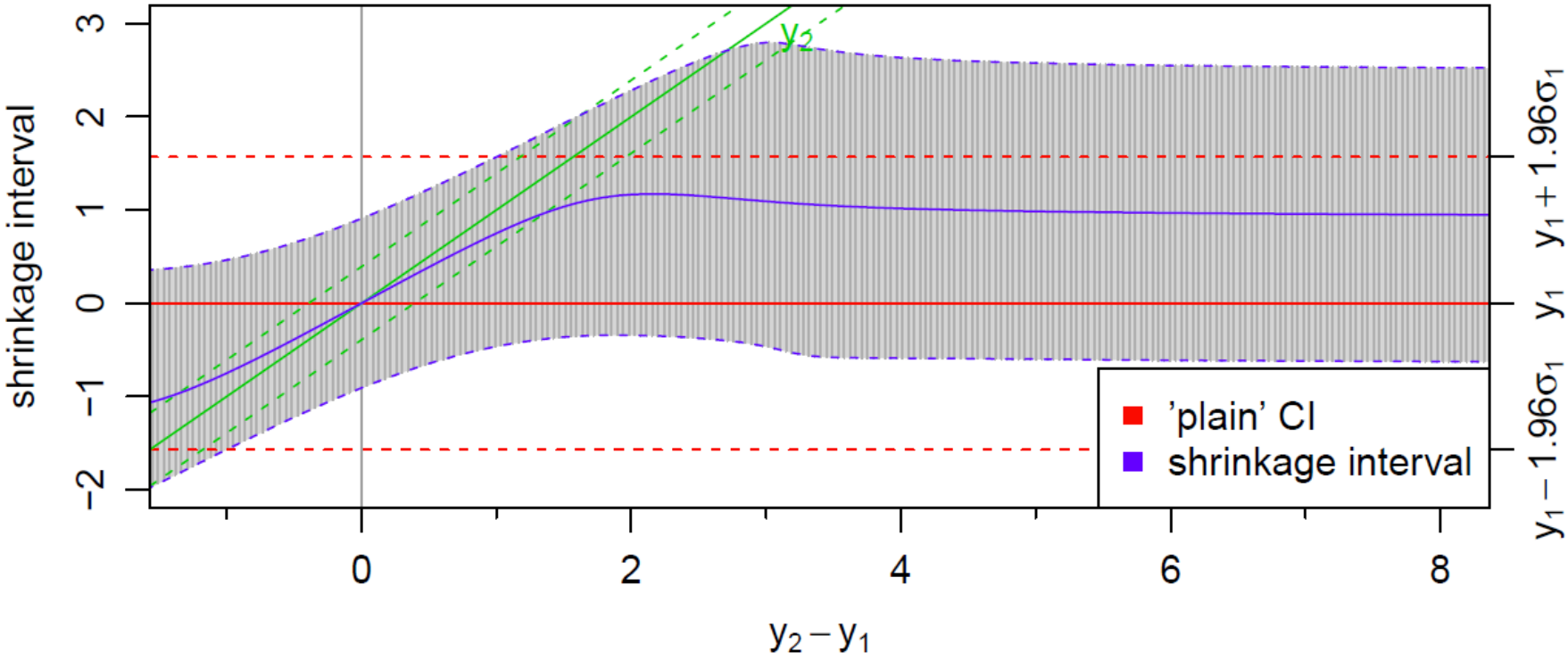
SHRINKAGE ESTIMATION IN R

UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG

```
# specify the data:
cjd <- cbind.data.frame("study" = c("observational", "randomized"),
                        "logHR" = c(-0.49948, -0.17344),
                        "logHR.se" = c(0.2493, 0.6312),
                        stringsAsFactors=FALSE)
# analyze:
require("bayesmeta")
bm <- bayesmeta(y         = cjd$logHR,
                sigma     = cjd$logHR.se,
                labels    = cjd$study,
                tau.prior = function(t){dhalfnormal(t, scale=0.5)})
# show results:
bm
forestplot(bm)
# show shrinkage estimates:
bm$theta
# interval length ratio (66%);
(q <- diff(bm$theta[7:8,"randomized"])
      / (2*qnorm(0.975)*bm$theta[2,"randomized"]))

# effective sample size gain (129%);
(1/q)^2 - 1
```

Appendix in Röver & Friede (2020) SRSM

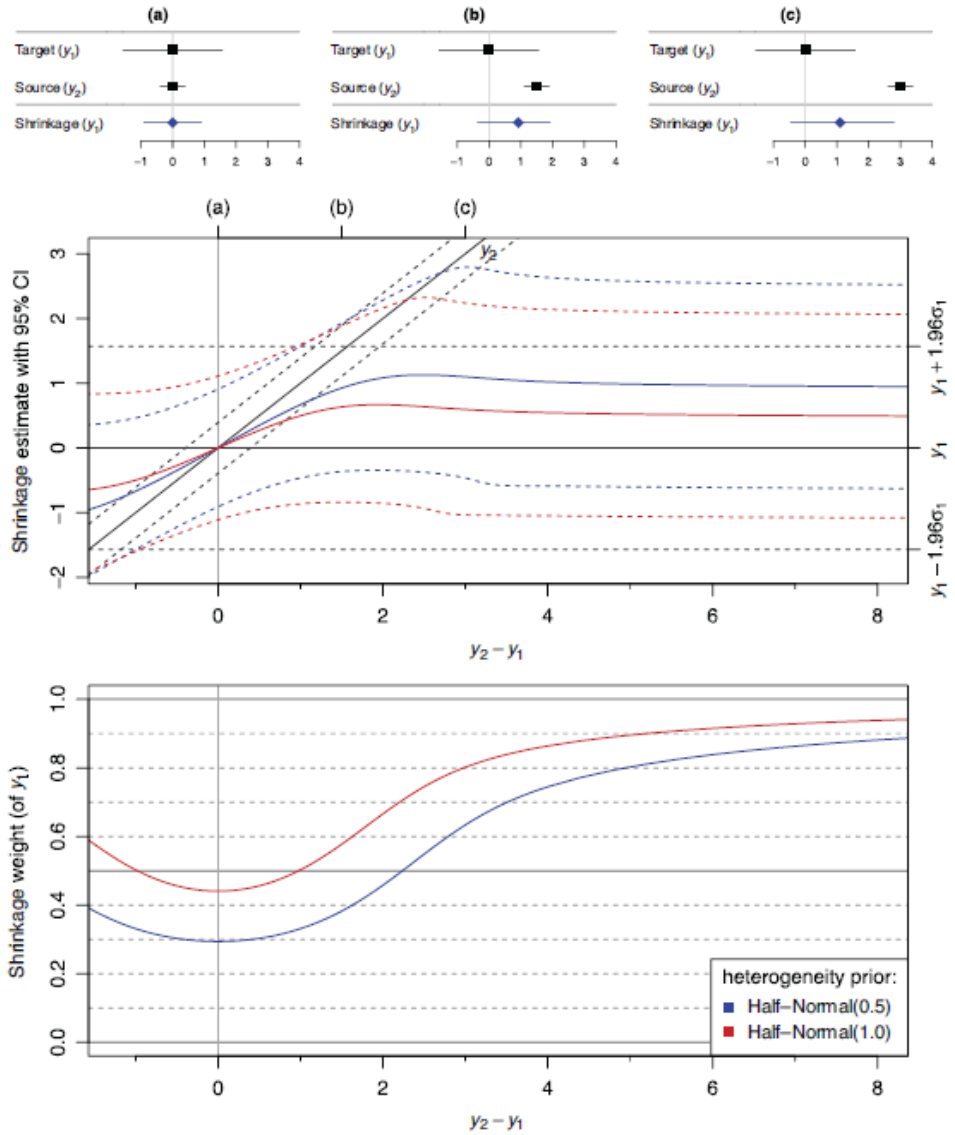
DYNAMIC BORROWING



- $n_1 = 25,\ n_2 = 400,\ p(\tau) = \text{HN}(0.5)$, interested in θ_1

Röver & Friede (2020) SMMR

BOUNDS FOR THE WEIGHTS



- Lower bound on the target's weight for any data realization (y_1, y_2) or any heterogeneity prior given by common-effect (CE) weight $\sigma_2^{-2}/(\sigma_1^{-2} + \sigma_2^{-2})$
- In this example, $\sigma_1 = 0.8$ and $\sigma_2 = 0.2$ resulting in CE weight 1/17 (5.9%)
- Minimum where $y_1 = y_2$
- Min. weight 29% for HN(0.5)
- Larger weight for larger scale of heterogeneity prior

Röver & Friede (2021) Biom J

BOUNDS FOR THE WEIGHTS: CJD EXAMPLE

- Lower bound on the RCT's weight for any data realization (y_1, y_2) or any heterogeneity prior: $\sigma_2^{-2}/(\sigma_1^{-2}+\sigma_2^{-2})$=13.5%

TABLE 2 Data from Varges et al. (2017) on an observational and a randomized study investigating the effect of doxycycline on survival in CJD

		Patients		log(HR)	
i	Study	Treatment	Control	y_i	σ_i
1	Observational	55	33	−0.499	0.249
2	Randomized	7	5	−0.173	0.631

TABLE 3 Estimates for the CJD example

	Mean weight		Effect estimate θ_2	
τ prior	Minimum	Actual	Mean	95% CI
HN(0.5)	38.9%	39.5%	−0.370	[−1.157, 0.477]
HN(1.0)	52.1%	53.1%	−0.326	[−1.232, 0.664]
		(100.0%	−0.173	[−1.410, 1.064])

For different heterogeneity priors (HN(0.5) or HN(1.0)), the corresponding minimum (coincidence) weight is given, as well as the resulting weight for the actual data along with the corresponding shrinkage estimates. The very last line shows the estimate based only on y_2 and σ_2 for comparison.

Röver & Friede (2021) Biom J

BAYESIAN BORROWING AND TYPE I ERROR RATE CONTROL

UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG

- Type I error rate control (a frequentist property) cannot be guaranteed with Bayesian borrowing (Kopp-Schneider et al, 2019)
- Computer simulations used to explore impact of Bayesian borrowing on (frequentist) type I error rate

Table 1. Coverage (%) of shrinkage intervals for estimation of the first study's mean parameter (θ_1).

τ prior:		HN (0.5)							HN (1.0)						
n_1/n_2	τ:	0.0	0.1	0.2	0.5	1.0	2.0	*	0.0	0.1	0.2	0.5	1.0	2.0	*
25/400		99.7	99.6	98.9	93.4	84.0	79.0	94.7	99.3	99.3	99.0	96.7	92.5	90.5	95.1
25/100		98.7	98.7	98.1	93.9	86.1	80.0	95.1	98.4	98.6	98.5	96.5	93.2	90.8	94.4
100/400		98.7	98.2	97.1	93.2	90.9	90.4	94.9	98.1	97.7	97.2	94.8	93.7	93.5	95.3
25/25		96.6	96.7	96.1	94.5	90.5	84.6	95.0	97.0	97.2	96.6	95.7	94.0	92.1	94.9
100/100		96.7	96.5	96.3	94.0	91.1	90.7	95.7	96.7	96.4	96.6	95.3	93.7	93.6	94.9
400/400		96.7	96.6	95.0	94.0	94.0	93.9	95.0	96.4	96.4	95.0	94.9	94.9	94.8	95.0
100/25		96.0	95.6	95.3	94.8	93.8	92.3	94.7	96.0	95.8	95.6	95.2	94.7	94.3	94.8
400/100		95.5	95.6	95.4	94.7	93.7	93.8	95.1	95.6	95.5	95.5	94.9	94.3	94.5	95.1
400/25		95.1	95.1	95.2	94.7	94.9	94.5	95.3	95.0	95.2	95.2	94.8	95.0	95.0	95.2

Note: Sample sizes (n_1 and n_2) as well as settings for the heterogeneity prior ($p(\tau)$) and actual heterogeneity values (τ) are varied. The columns labelled by an asterisk (*) correspond to drawing the heterogeneity from its corresponding prior distribution.

Röver & Friede (2020) SMMR

EXAMPLE: EARLY PRO-TECT

www.kidney-international.org clinical trial

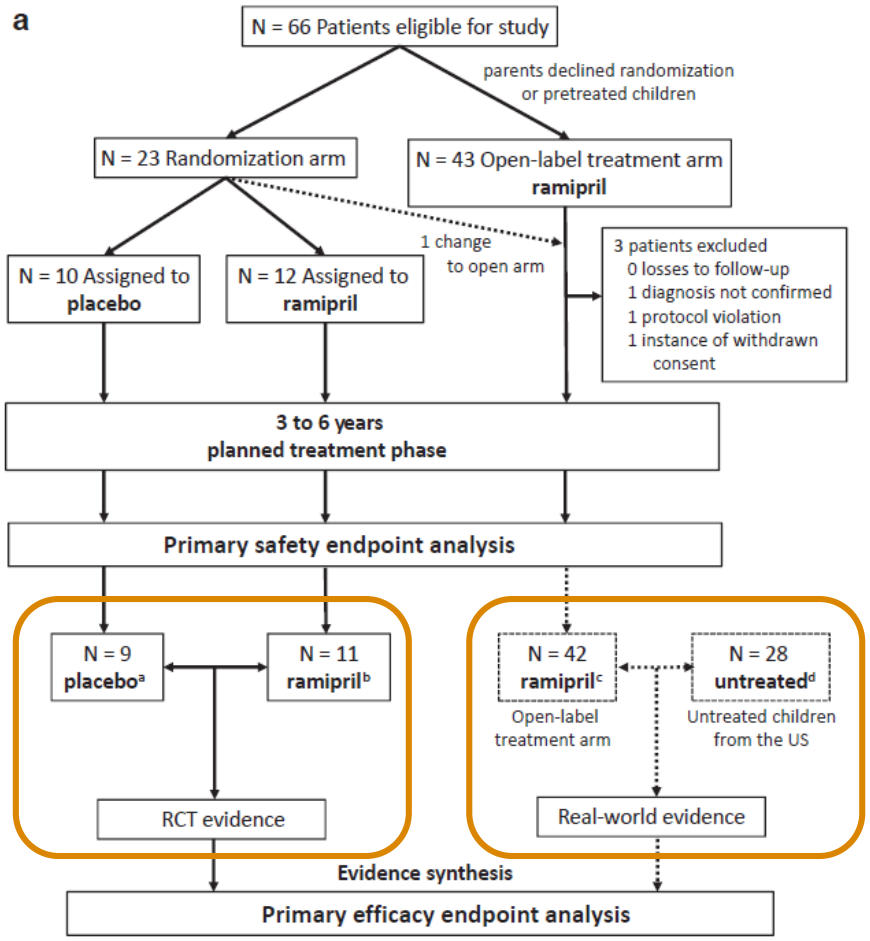
A multicenter, randomized, placebo-controlled, double-blind phase 3 trial with open-arm comparison indicates safety and efficacy of nephroprotective therapy with ramipril in children with Alport's syndrome

see commentary on page 1104

OPEN

Oliver Gross[1], Burkhard Tönshoff[2], Lutz T. Weber[3], Lars Pape[4], Kay Latta[5], Henry Fehrenbach[6], Baerbel Lange-Sperandio[7], Hildegard Zappel[8], Peter Hoyer[9], Hagen Staude[10], Sabine König[11], Ulrike John[12], Jutta Gellermann[13], Bernd Hoppe[14], Matthias Galiano[15], Britta Hoecker[2], Rasmus Ehren[3], Christian Lerch[4], Clifford E. Kashtan[16], Markus Harden[17], Jan Boeckhaus[1] and Tim Friede[17]; for the German Pediatric Nephrology (GPN) Study Group and EARLY PRO-TECT Alport Investigators[18,19]

EXAMPLE: EARLY PRO-TECT TRIAL

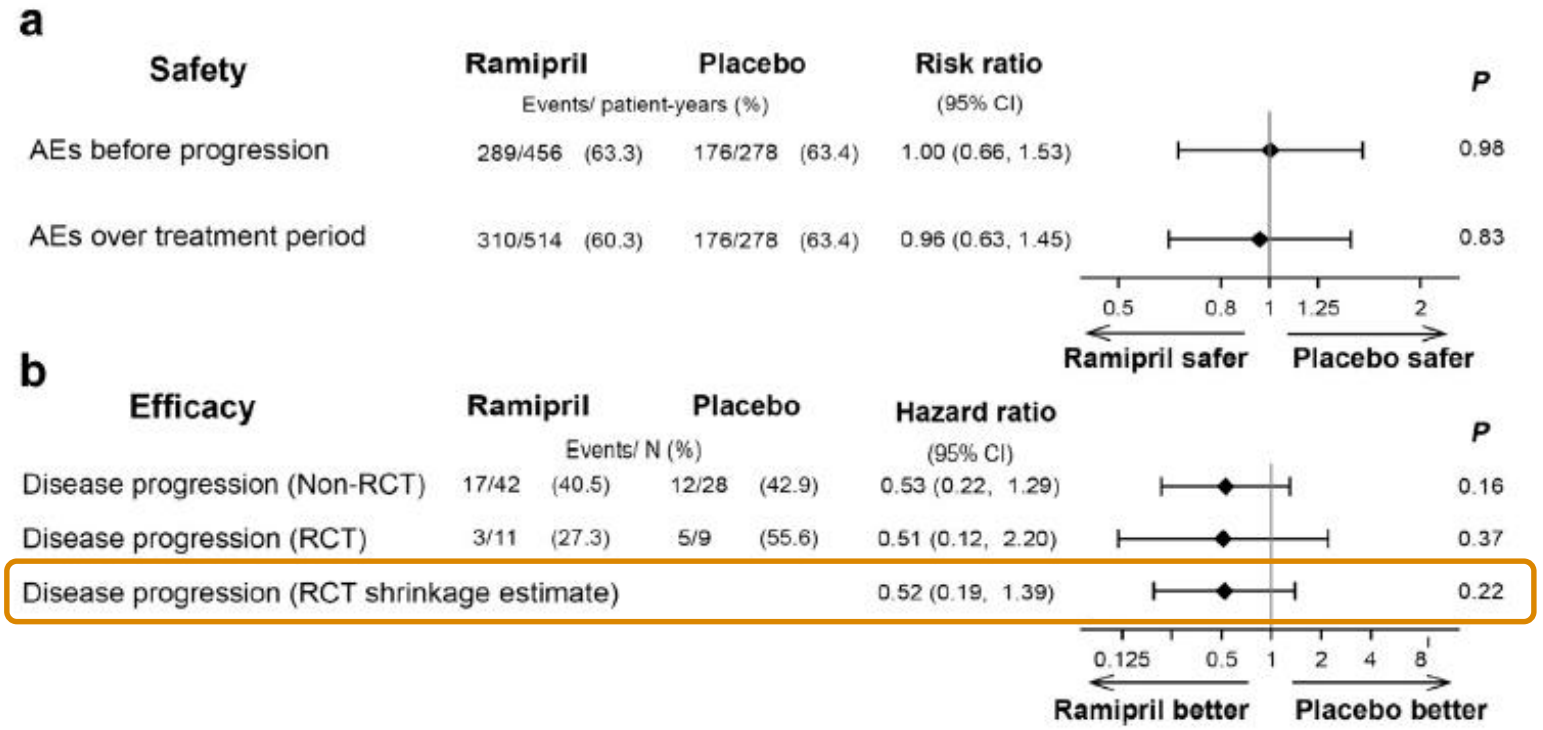


- **Randomised controlled trial in children with Alport's syndrome** (rare genetic disorder leading to end-stage kidney disease)
- **Observational data**
  - Open-label treatment arm
  - Natural disease cohort (registry)

Figure 1 in Gross et al (2020) Kidney International

EXAMPLE: EARLY PRO-TECT TRIAL

- Figure 2 in Gross et al (2020) Kidney International



- **Increased precision in estimating the treatment effect**: Interval shortened by 42%; equivalent to raising the sample size of the RCT from 20 to 43; i.e. 70 patients in RWE count as 23 RCT patients

COMPREHENSIVE COHORT STUDIES

- **Schmoor et al (1996) Stat Med**

STATISTICS IN MEDICINE, VOL. 15, 263–271 (1996)

RANDOMIZED AND NON-RANDOMIZED PATIENTS IN CLINICAL TRIALS: EXPERIENCES WITH COMPREHENSIVE COHORT STUDIES

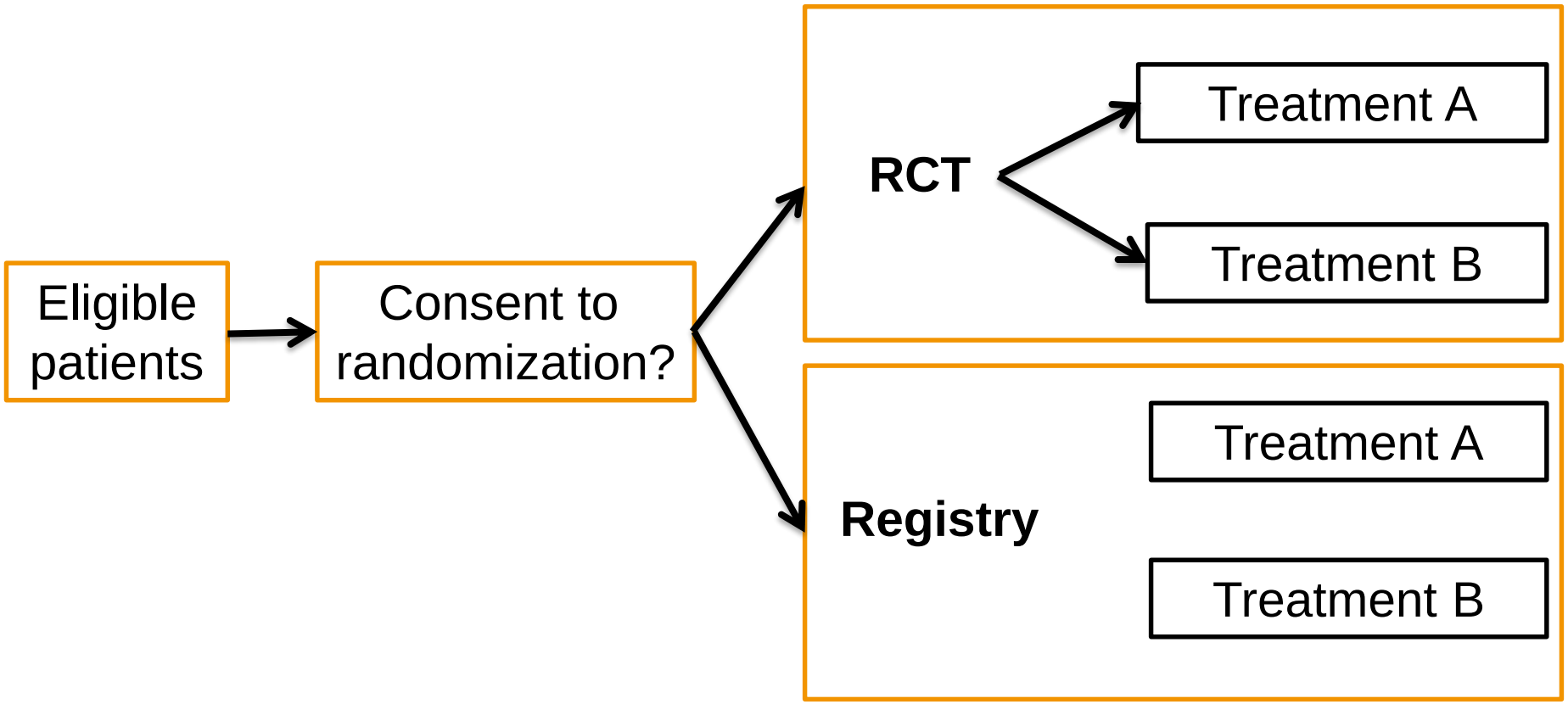
CLAUDIA SCHMOOR, MANFRED OLSCHEWSKI AND MARTIN SCHUMACHER

Institute of Medical Biometry and Informatics, University of Freiburg, Stefan-Meier-Str. 26, D-79104 Freiburg, Germany

SUMMARY

In clinical research, randomized trials are widely accepted as the definitive method of evaluating the efficacy of therapies. Random assignment of patients to treatment ensures internal validity of the comparison of new treatments with controls. An assessment of external validity can best be achieved by comparing the randomized study sample to the population of patients who met the eligibility criteria but did not consent to randomization. The Comprehensive Cohort Study (CCS) is designed to recruit all patients fulfilling the clinical eligibility criteria regardless of their consent to randomization. The CCS concept was adopted in the major clinical trials of the German Breast Cancer Study Group (GBSG) conducted between 1983 and 1989. In this period 124 centres recruited 2084 patients in three clinical trials. 734 (35 per cent) of these patients accepted being randomized, while 1350 (65 per cent) chose one of the treatments under study; the randomization rates differed remarkably between trials. In this paper we examine the representativeness of the randomized patients in the three trials. Based on a median follow-up of about 5 years we present results on the external validity of the treatment effects estimated in the randomized patients by means of Cox's proportional hazards model and compare them between trials. We discuss advantages and disadvantages of the CCS design and conclude that its use is only justified under extraordinary circumstances.

COMPREHENSIVE COHORT DESIGN



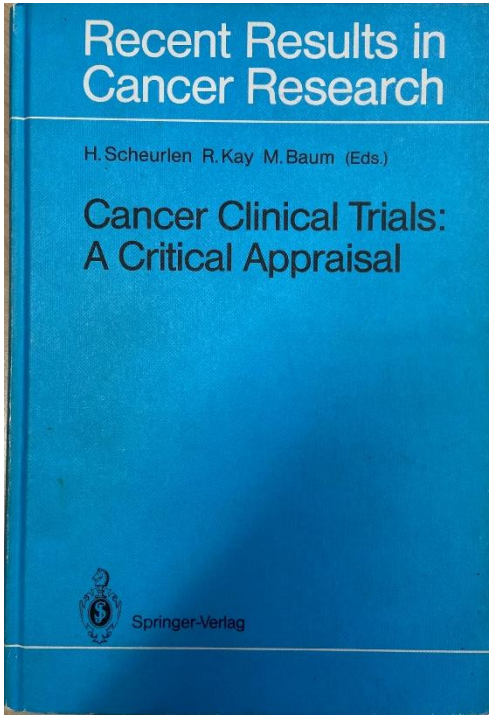
Adapted from Figure 1 in Schmoor et al (1996)

ORIGINS OF THE CCS DESIGN

Some references from Schmoor et al (1996)

REFERENCES

1. Olschewski, M. and Scheurlen, H. 'Comprehensive Cohort Study: An alternative to randomized consent design in a breast preservation trial', *Methods of Information in Medicine*, **24**, 131–134 (1985).
2. Principal Investigators of CASS and their Associates 'National Heart, Lung, and Blood Institute Coronary Artery Surgery Study', *Circulation*, **63 I**, 1–82 (1981).
3. Olschewski, M., Schumacher, M. and Davis, K. B. 'Analysis of randomized and non-randomized patients in clinical trials using the comprehensive cohort follow-up study design', *Controlled Clinical Trials*, **13**, 226–239 (1992).



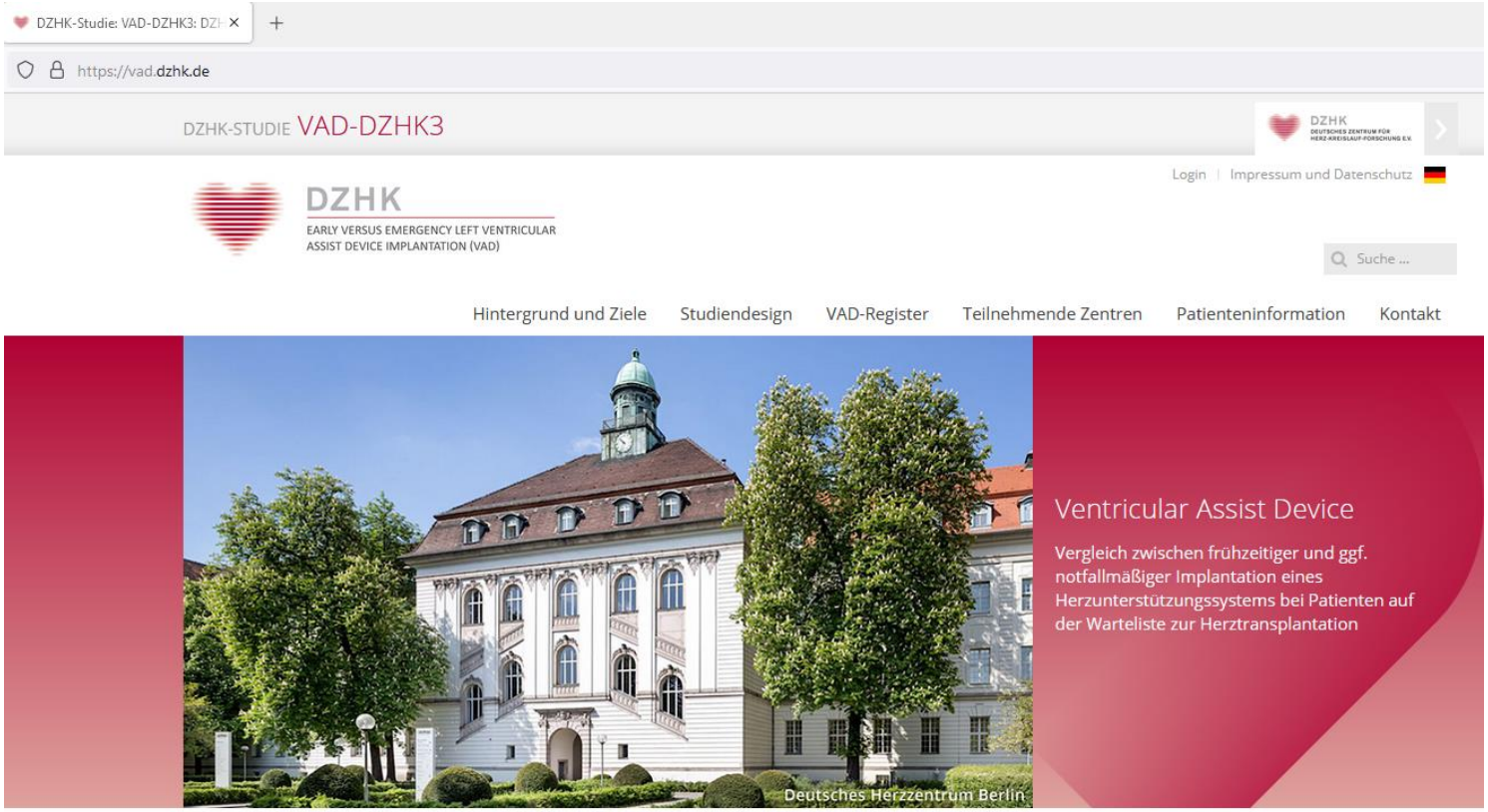
INTERNAL AND EXTERNAL VALIDITY

- **Randomized controlled trial**
  - Internal validity through randomisation
- **Assessment of external validity in comprehensive cohort studies** (Schmoor et al, 1996)
  - Comparisons of RCT and registry with regard to
    - baseline characteristics
    - follow-up / outcome
    - treatment effects

EXTENSION OF CCS APPROACH

- **Randomized controlled trial**
  - Internal validity through randomisation
- **Assessment of external validity in comprehensive cohort studies**
  - Comparisons of RCT and registry with regard to baseline characteristics and follow-up (Schmoor et al, 1996)
- **Data integration**
  - Meta-analytic framework to integrate data from RCT and registry (using appropriate causal inference approach) accounting for heterogeneity (Röver and Friede, 2020)

EXAMPLE: VAD-DZHK3



Kurzinfo VAD-Studie

Für Patienten mit Herzschwäche im Endstadium (terminale Herzinsuffizienz), die auf eine Transplantation warten, ist der Einsatz eines mechanischen Herzunterstützungssystems (Ventricular Assist Device, VAD) häufig die einzige Möglichkeit, die Wartezeit auf ein Spenderorgan zu überbrücken. Bisher gibt es jedoch keinen allgemein anerkannten Standard für den optimalen Zeitpunkt des Einsetzens (Implantation) eines VAD.

In der VAD-Studie wird jetzt eine frühzeitige mit einer gegebenenfalls notfallmäßigen VAD-Implantation bei Patienten auf der Warteliste zur Herztransplantation verglichen. Dadurch sollen leitlinienrelevante Erkenntnisse für die zukünftige Behandlung dieser Patienten gewonnen werden und damit das Überleben und die Lebensqualität der Betroffenen verbessert werden.

CAUSAL INFERENCE IN SMALL OBSERVATIONAL STUDIES

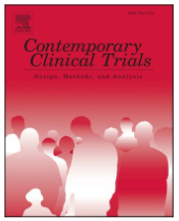
- **Data requirements**: characterization of patients, granularity of follow-up
- Do **causal inference methods** (e.g. propensity score based approach, g-computation) work with small sample sizes?

Contemporary Clinical Trials 99 (2020) 106213

Contents lists available at ScienceDirect

Contemporary Clinical Trials

journal homepage: www.elsevier.com/locate/conclintrial



Causal inference methods for small non-randomized studies: Methods and an application in COVID-19

Sarah Friedrich *, Tim Friede

Department of Medical Statistics, University Medical Center Göttingen, Humboldtallee 32, 37073 Göttingen, Germany

RECOMMENDATIONS FOR THE ANALYSIS OF SMALL NON-RANDOMIZED STUDIES

Based on (limited) simulations with binary outcome, binary treatment and covariates (Friedrich & Friede, 2020)

1. Unmeasured confounder rendered the methods useless. Therefore, careful clinical characterization of patients important
2. Effect measure: risk difference preferred over odds ratio
3. For small sample sizes, the best performance observed for covariate adjustment, PS covariate and doubly robust g-computation (based on quintiles)
4. IPTW performed well regarding bias and RMSE, but coverage of confidence intervals very low (and therefore not recommended)
5. Conduct simulations to explore properties of the methods in scenarios similar to the one at hand (R code available)

CONCLUSIONS AND DISCUSSION

- **Hierarchical models**
  - flexible statistical framework for evidence synthesis
- **Bayesian inference**: advantages over traditional methods in the presence of heterogeneity and only (very) few studies
  - easy to apply using R package **`bayesmeta`**
- **Cross-design synthesis of available evidence**
  - Promising in rare diseases
  - more practical (and regulatory) experience needed
- **Bounds for weights:** concerns of evidence being easily overwhelmed by external data are largely unwarranted
- **Alternative approaches** including power prior model

ANY QUESTIONS?

E-Mail: tim.friede@med.uni-goettingen.de
Homepage: https://medstat.umg.eu/
Twitter: @tim_friede

SOME REFERENCES

- Friede T, Röver C, Mathes T (2023) Verknüpfung von randomisierten kontrollierten Studien und Real World Data (Combining randomized controlled trials and real-world data). Prävention und Gesundheitsförderung (in press).
- Friede T, Röver C, Wandel S, Neuenschwander B (2017) Meta-analysis of few small studies in orphan diseases. Research Synthesis Methods 8: 79–91.
- Friedrich S, Friede T (2020) Causal inference methods for small non-randomized studies: Methods and an application in COVID-19. Contemporary Clinical Trials 99: 106213.
- Röver C, Friede T (2020) Dynamically borrowing strength from another study through shrinkage estimation. Statistical Methods in Medical Research 29: 293–308.
- Röver C, Friede T (2021) Bounds for the weight of external data in shrinkage estimation. Biometrical Journal 63: 1131–1143.