The University Of Sheffield.

Evidence synthesis with limited studies

EFSPI
Latest Trends in HTA, 15th February 2019

Dr. Shijie (Kate) Ren, Senior Research Fellow
School of Health and Related Research, University of Sheffield, Sheffield, UK.

s. ren@sheffield.ac.uk

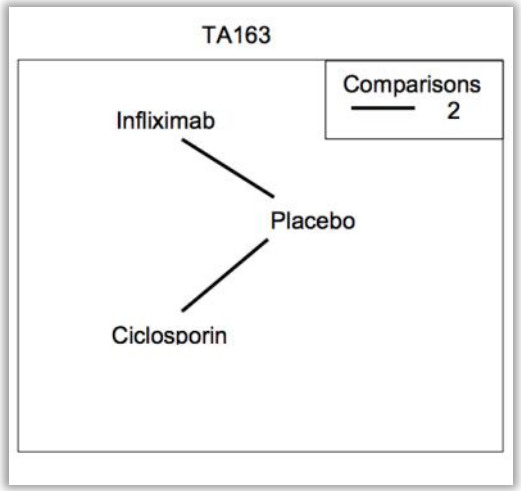
A WORLD TOP 100 UNIVERSITY

Motivation

- Evidence can arise from multiple sources in HTA
  - Pairwise meta-analysis (MA), network meta-analysis (NMA)
  - Fixed effect (FE) and random effects (RE) model
- RE model generally preferred
  - Allow for heterogeneity
  - Generalisable beyond included studies
- Problem: limited number of studies
  - Insufficient data to estimate the heterogeneity reliably
- Solution: Bayesian approach
  - What prior?

© The University of Sheffield

NICE TA163



- Disease: ulcerative colitis
- Outcome measure: colectomy rate at 3 months

No heterogeneity
FE model
0.72 (0.18, 2.70)

High heterogeneity
RE (vague prior)
0.70 (0.01, 84.6)

Variability of treatment effects among studies

© The University of Sheffield

Common scenario

NICE STAs

- The company
  - "very few studies…to support the estimation of a random effects model"
  - "instability in the WinBUGS model"
  - "random effects models did not converge"

→ Default to the use of a fixed effect model

- The expert review group (ERG)
  - "too few studies…not a valid reason…"
  - "external information should be used…plausible posterior uncertainty"

© The University of Sheffield

Justification of model choice

NICE STAs (2005-2016)

Method (number of submissions)		Justification	N(%)
Pairwise meta-analysis (38*)	FE model only (7)	No justification	5(71%)
		Check heterogeneity	2(29%)
Network meta-analysis (71*)	FE model only (24)	Insufficient data	17(71%)
		No justification	6(25%)
		Check heterogeneity	1(4%)

*: Multiple analyses and analyses for multiple outcomes may have been conducted in one submission.

- Uncertainty: underestimate/overestimate
- How to overcome the problem?
  - Incorporating external evidence

© The University of Sheffield

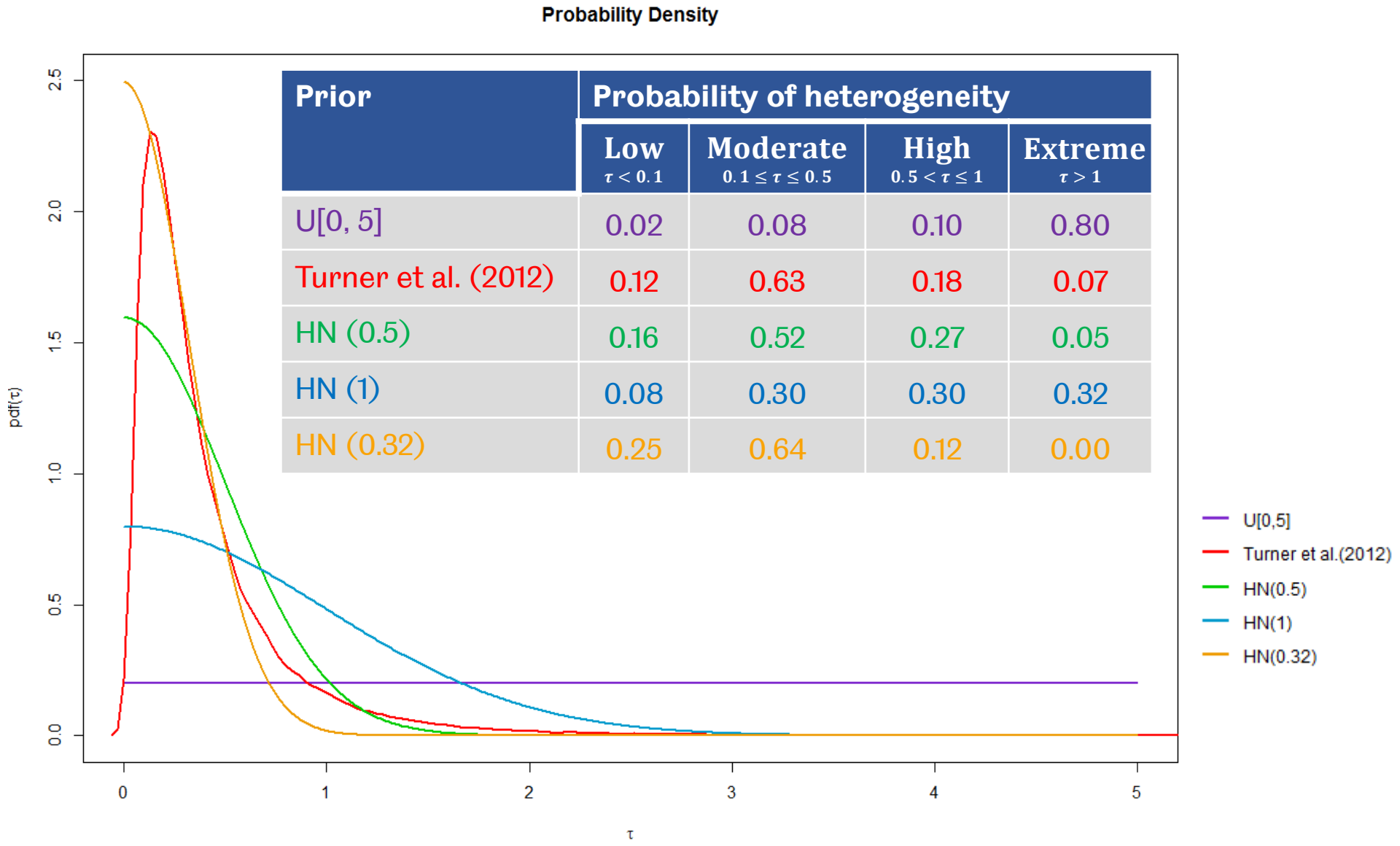
Currently commonly used external evidence

- Empirical evidence
  - Turner et al. (2012), Rhodes et al. (2015): predictive distributions for the heterogeneity parameter in various settings using studies from the Cochrane Database of Systematic Reviews
  - Outcomes: binary outcome [Turner et al. (2012)]; standardised mean difference [Rhodes et al. (2015)]
- Other suggested priors
  - Spiegelhalter et al. (2004): Half-Normal (0.51)
  - Friede et al. (2016): Half-Normal (0.5) and Half-Normal (1)
  - NICE DSU TSD 3: Half-Normal (0.32)
  - Outcomes: binary

© The University of Sheffield

Which one to choose?

Prior	Probability of heterogeneity			
	Low $\tau < 0.1$	Moderate $0.1 \leq \tau \leq 0.5$	High $0.5 < \tau \leq 1$	Extreme $\tau > 1$
U[0, 5]	0.02	0.08	0.10	0.80
Turner et al. (2012)	0.12	0.63	0.18	0.07
HN (0.5)	0.16	0.52	0.27	0.05
HN (1)	0.08	0.30	0.30	0.32
HN (0.32)	0.25	0.64	0.12	0.00



© The University of Sheffield

Experts' beliefs

- Subjective; controversial; possible to quantify?
- Motivating example 1

TA 163: Colectomy rate at 3 months	OR , median (95% CrI/PrI) infliximab vs. placebo	
RE model $\tau_{OR} \sim uniform\,[0,5]$	0.70 (0.01, 84.59)	Posterior distribution
	0.69 (0, 2498.82)	Predictive distribution

Don't believe the results (implausible upper limit). ➔ RE model should not be used.

So what do you believe *would* be reasonable?

© The University of Sheffield

Experts' beliefs

- Motivating example 2
  - Spiegelhalter et al. (2004): Half-Normal (0.51)
  - NICE DSU TSD: Half-Normal (0.32)

Published opinion

DJ Spiegehalter, KR Abrams and JP Myles

Dias S, Sutton AJ, Welton NJ and Ades AE

© The University of Sheffield

Elicitation framework

- Aim: Construct a genuine prior distribution for the heterogeneity parameter using external information
  - Empirical evidence: Turner et al. (2012), Rhodes et al. (2015)
  - Experts' beliefs: experts, published opinion

Technical challenges

How to construct probability distributions for abstract model parameters from judgements about interpretable observable quantities

- Elicit the 'range' of treatment effects
- Transform the prior distribution for the 'range' to obtain the prior for the heterogeneity

© The University of Sheffield

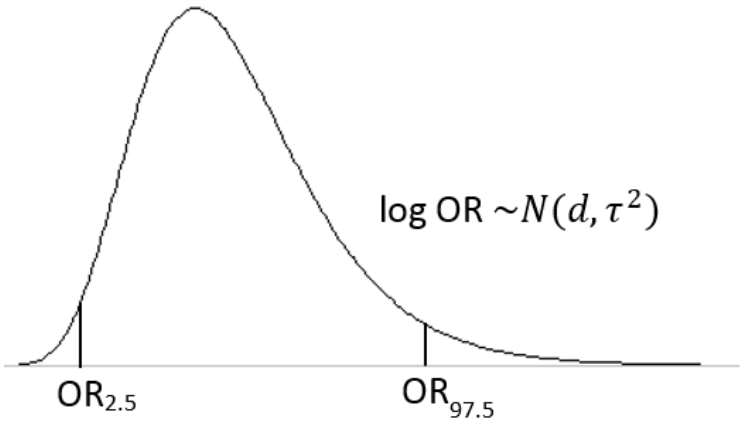
Notation

- δ_i : treatment effect in study i, for $i = 1, \ldots, S$
  - log OR
  - log HR
  - mean difference
- $\delta_1, \ldots, \delta_S \sim \mathrm{N}(\mathrm{d}, \tau^2)$

© The University of Sheffield

What quantity to elicit?[12]

- Assume δ_i is log OR
- Quantity of interest
  - Heterogeneity parameter, τ
- Interpretable observable quantities
  - Study-specific treatment effect, OR
- Propose to elicit: $R = \frac{OR_{97.5}}{OR_{2.5}}$



© The University of Sheffield

How does it work?

$$\delta_{97.5} - \delta_{2.5} = 2 \times 1.96\tau = 3.92\tau$$
$$\rightarrow \log(R) = 3.92\tau$$
$$\rightarrow \tau = \frac{\log(R)}{3.92} \quad (1)$$

- Make judgements about R, then judgements about τ can be inferred using equation (1)
- Less formal definition of R
  - The ratio of the largest to the smallest OR that could arise over a set of studies
  - The 'range' of treatment effects
  - TA163: $R = 10$ (The largest OR of having colectomy when comparing infliximab to placebo could be 10 times more than the smallest OR in a population of studies.)

© The University of Sheffield

What if ?

- Expert is only able to specify a point estimate of R
  - No probabilistic distribution
- Expert is not able to say anything about R

Three-stage procedure for elicitation

© The University of Sheffield

Three-stage procedure

1: Confirm need for RE model → FE model

2: Upper bound for the ratio

$R_{max} = 10$. OR in one study no more than 10 times that of another → Informative Prior 1: Turner et al. (2012)

3: Full distribution for the ratio

Express some values in $[1, R_{max}]$ as more likely than others → Informative Prior 2: Truncated Turner et al. (2012)

Informative Prior 2 → Expert; Published opinion*

Informative Prior 3: Elicited prior

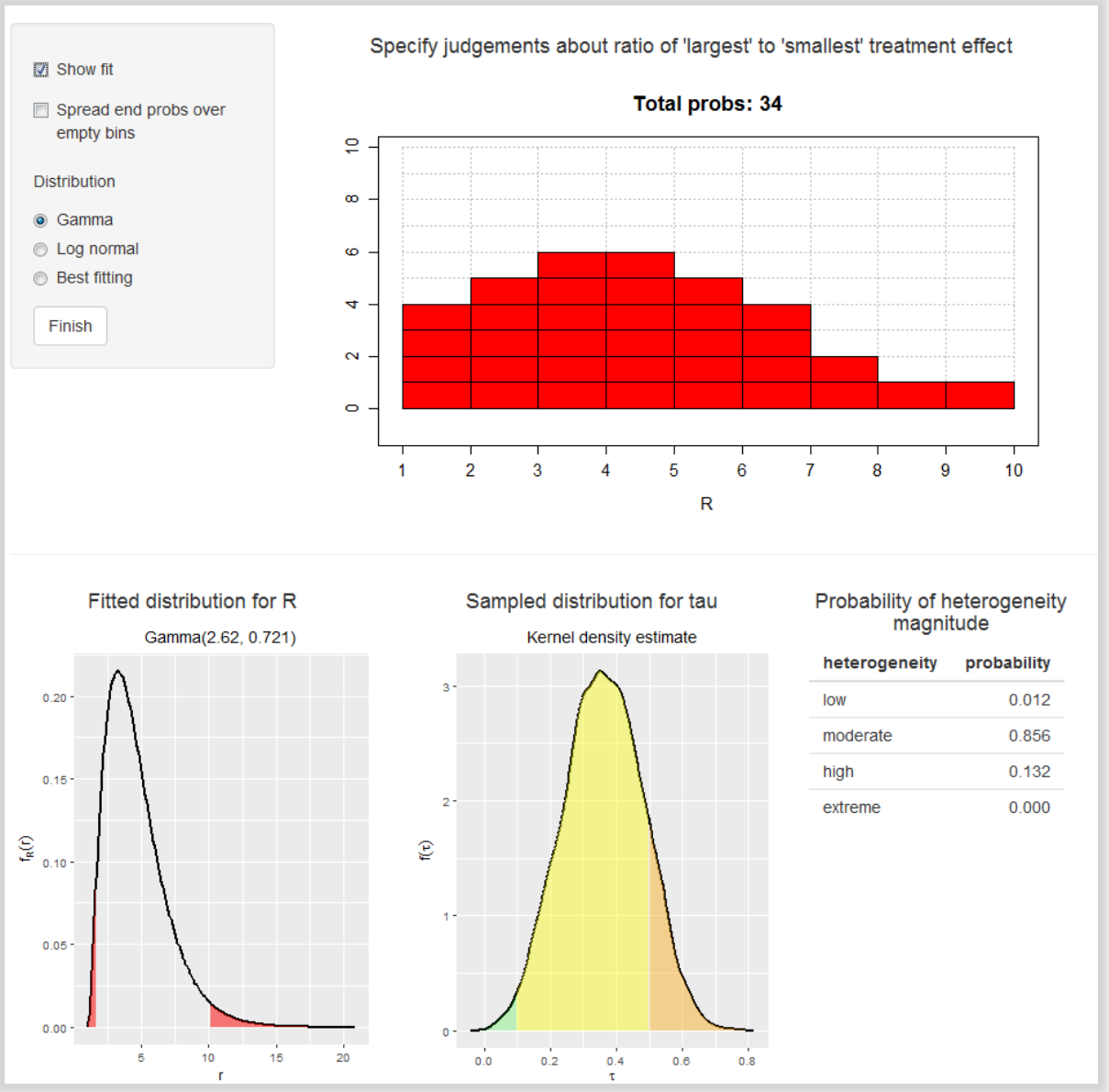
*: Smith et al. (1995): $R_{max} = 10$

Spiegelhalter et al. (2004): $R_{max} = 50$

© The University of Sheffield

Implementation

- R package: `SHELF`
- function: `elicitHeterogen()`
- See Ren et al. (2018) for details



© The University of Sheffield

Extensions

- NMA
  - Homogeneous variance model
  - Elicit the “range” for a pairwise comparison [most comfortable about expression beliefs]
  - Feedback on if agree with the elicited probabilities for other pairwise comparisons in the network

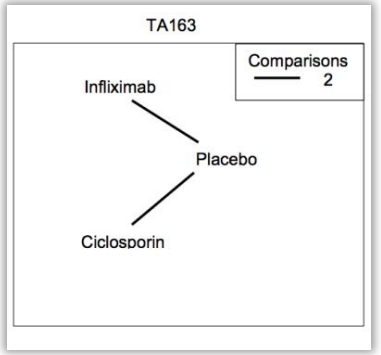
© The University of Sheffield

Extensions

- Scale-free: hazard ratio, relative risk, ratio of means
  - Three-stage procedure [check if Turner et al. (2012) is also applicable]
- Continuous: Three-stage procedure with modification
  - Dichotomise the response using some appropriate cut-off to define a new treatment effect on the OR scale
  - Three-stage procedure considering ORs
  - Given a prior for τ_{OR} (variability in log ORs), convert it to a prior for τ_{MD} on the original scale via $\tau_{MD} = \omega\tau_{OR}$
  - $\omega = \sigma\frac{\sqrt{3}}{\pi}$
  - σ: an estimate of an individual level SD
  - See Ren et al. (2018) for details

© The University of Sheffield

NICE TA163



- Disease: ulcerative colitis
- Outcome measure: colectomy rate at 3 months
- Fixed effect model was used
- Re-analyse using a random effects with
  - $\tau_{OR} \sim$ U[0, 5]
  - $\tau_{OR} \sim$ HN(0.5)
  - Turner et al. (2012) prior: $\tau^2_{OR} \sim lognormal(-2.56, 1.74^2)$
  - Truncated Turner et al. (2012): $R_{max} = 10$
  - Elicited prior: $(R_{OR} - 1) \sim gamma(2.62, 0.721)$ and $\tau_{OR} = \log(R_{OR})/3.96$

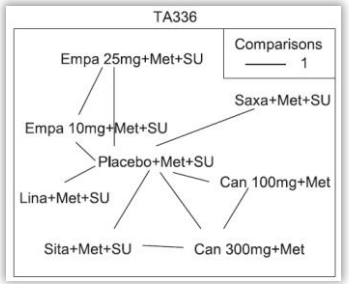
© The University of Sheffield

Results (TA163)

Method	OR median (95% CrI) infliximab vs. placebo	OR median (95% CrI) ciclosporin vs. placebo	Probability of heterogeneity			
			Low	Moderate	High	Extreme
FE	0.72 (0.18, 2.70)	0.13 (0.03, 0.44)	0	0	0	0
RE (vague prior)	0.70 (0.01, 84.59)	0.02 (0, 1.46)	0.01	0.05	0.07	0.87
RE (HN(0.5))	0.71 (0.15, 3.40)	0.12 (0.02, 0.50)	0.15	0.51	0.29	0.05
RE (Turner prior)	0.71 (0.14, 3.25)	0.11 (0.01, 0.48)	0.11	0.62	0.18	0.08
RE (truncated Turner prior)	0.69 (0.17, 2.77)	0.12 (0.03, 0.48)	0.15	0.78	0.07	0
RE (elicited prior)	0.71 (0.17, 2.97)	0.12 (0.03, 0.47)	0.01	0.85	0.14	0

© The University of Sheffield

NICE TA336



- Disease: type 2 diabetes mellitus
- Outcome measure: CFB in body weight (kg) at 24 wks
- Fixed effect model was used
- Re-analyse using a random effects with
  - $\tau_{MD} \sim \text{U}[0, 5]$
  - Rhodes et al. (2015) prior: $\log(\tau^2_{SMD}) \sim t(-3.44, 2.59^2, 5)$
  - Turner et al. (2012) prior: $\tau^2_{OR} \sim lognormal(-2.56, 1.74^2)$ and $\tau_{MD} = 2.61 \times \tau_{OR}/1.81$
  - Truncated Turner et al. (2012): $R_{max} = 10$
  - Elicited prior: $(R_{OR} - 1) \sim gamma(1.94, 0.741)$ and $\tau_{MD} = 2.61 \times \log(R_{OR})/(3.96 \times 1.81)$

© The University of Sheffield

Results (TA336)

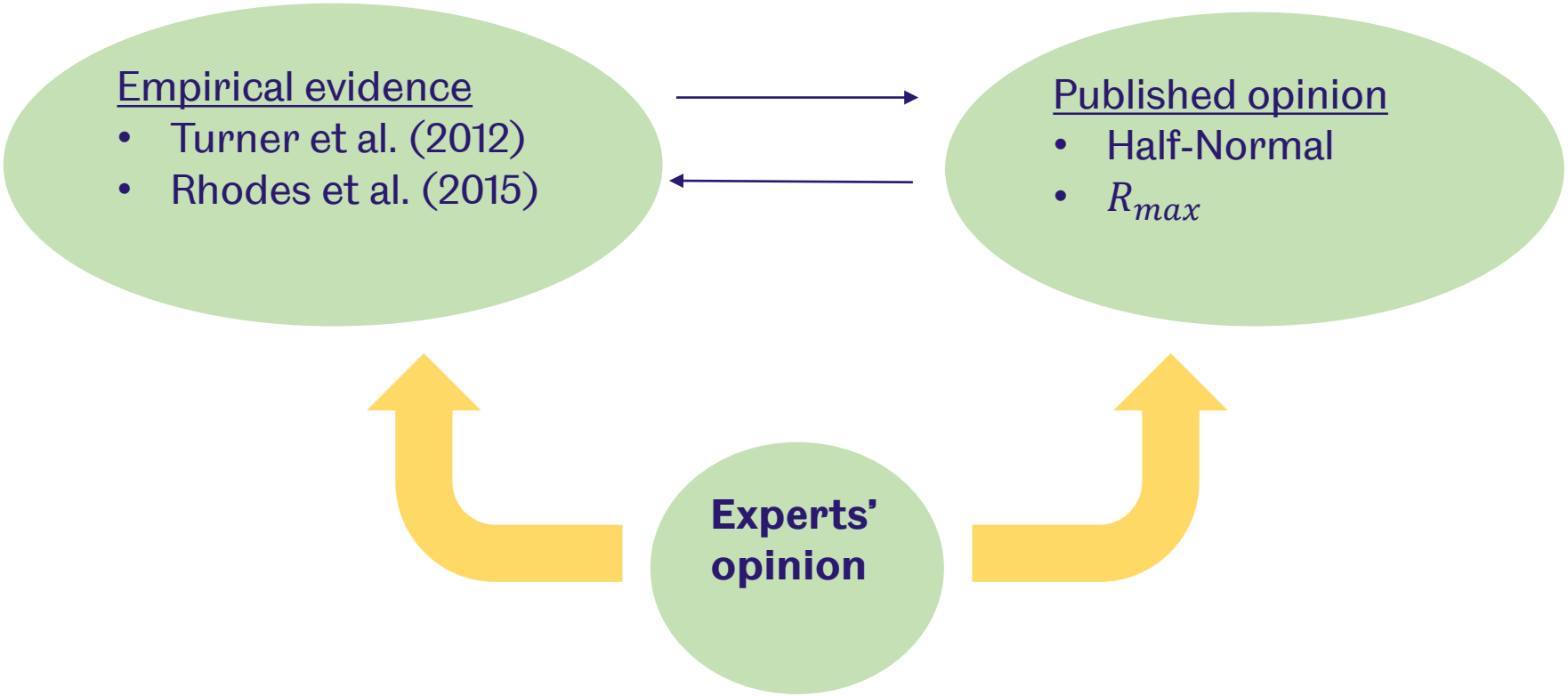
Method	MD median [95% CrI] Empa vs. placebo	MD median [95% CrI] Empa vs. Linagliptin	Probability of heterogeneity			
			Low	Moderate	High	Extreme
FE	-1.77 [-2.18, -1.35]	-2.10 [-2.64, -1.54]	0	0	0	0
RE (vague prior)	-1.76 [-6.10, 2.70]	–2.08 [–8.12, 4.08]	0.14	0.32	0.19	0.35
RE (Rhodes prior)	-1.79 [-2.54, -0.99]	-2.10 [-3.18, -1.01]	0.45	0.45	0.06	0.04
RE (Rhodes prior*)	-1.77 [-3.08, -0.45]	-2.10 [-3.94, -0.27]	0.27	0.50	0.14	0.09
RE (Turner prior*)	-1.77 [-2.88, -0.63]	-2.10 [-3.65, -0.51]	0.18	0.70	0.10	0.02
RE (truncated Turner prior*)	-1.77 [-2.62, -0.93]	-2.10 [-3.30, -0.93]	0.21	0.75	0.04	0
RE (elicited prior*)	-1.78 [-2.76, -0.80]	-2.10 [-3.47, -0.72]	0.08	0.88	0.03	0

*: with modification (see Ren et al. (2018))

© The University of Sheffield

Summary

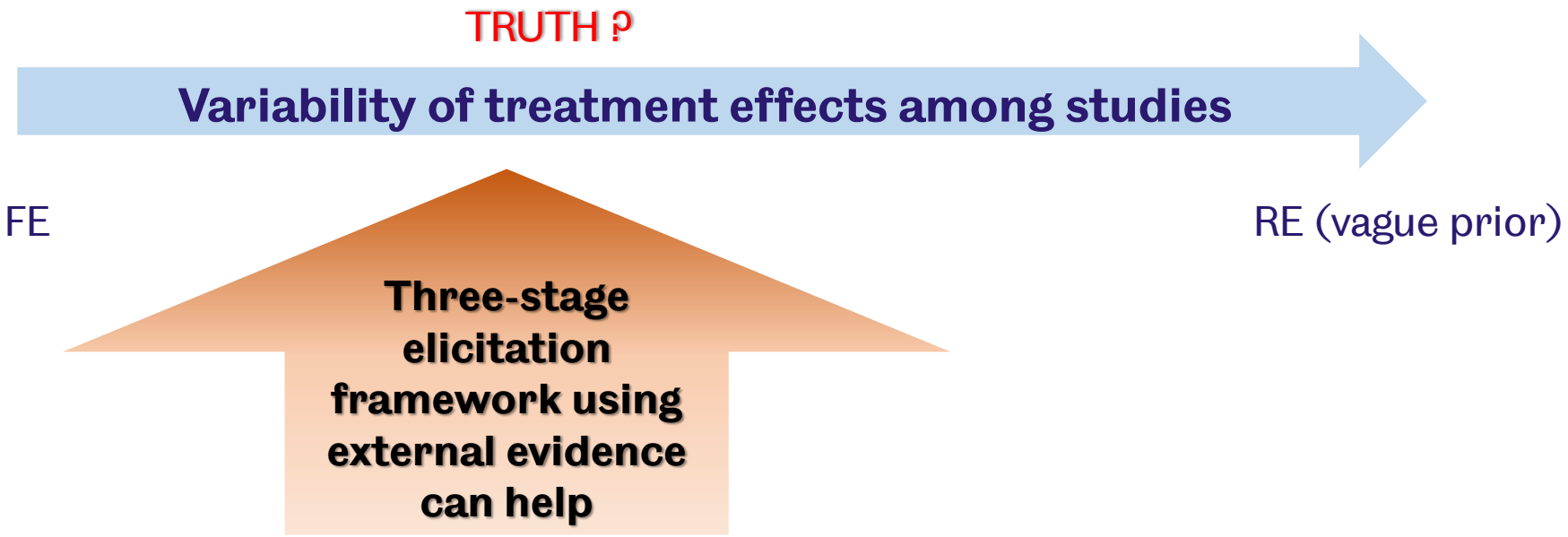
- Use external evidence for τ



- The scales matter: scale-free vs. continuous

© The University of Sheffield

Summary



- Use genuine prior distribution for τ
- Minimum requirement: the ratio of the largest to the smallest OR (the 'range' of treatment effects)

© The University of Sheffield

References

- Ren et al. (2018) Incorporating genuine prior information about between-study heterogeneity in random effects pairwise and network meta-analyses. Medical Decision Making 38;4:531-542. Open-access.
- Oakley JE. (2017) SHELF: Tools to Support the Sheffield Elicitation Framework. R package version 1.2.3. Available from: https://cran.r-project.org/package=SHELF
- Turner et al. (2012) Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. Int J Epidemiol. 41:818–827.
- Rhodes et al. (2015) Predictive distributions were developed for the extent of heterogeneity in meta-analyses of continuous outcome data. J Clin Epidemiol. 68:52–60.
- Smith TC, Spiegelhalter DJ, Thomas A. Bayesian approaches to random-effects meta-analysis: A comparative study. Stat Med. 1995;14:2685–2699.
- Spiegelhalter DJ, Abrams KR, Myles JP. Bayesian Approaches to Clinical Trials and Health-Care Evaluation. Chichester: John Wiley & Sons, Ltd; 2004.
- Friede T, et al. (2017). Meta-analysis of few small studies in orphan diseases. Research Synthesis Methods 8;1: 79-91.

© The University of Sheffield