The University Of Sheffield.

Evidence synthesis with limited studies

Incorporating genuine prior information about between-study heterogeneity

PSI conference 2018

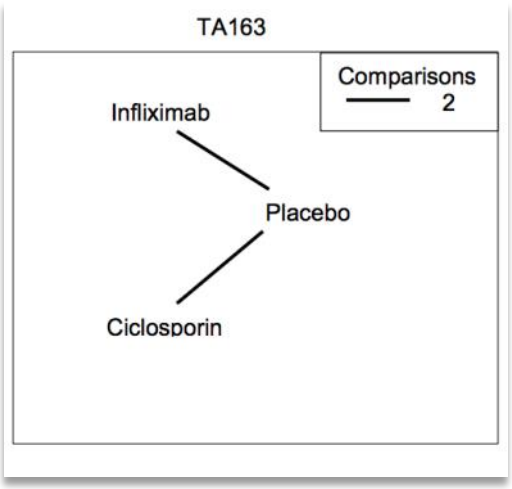
Dr. Shijie (Kate) Ren, Research Fellow, School of Health and Related Research, University of Sheffield, Sheffield, UK

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
  - National Institution for Health and Care Excellence (NICE) single technology appraisal (STA)

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

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"

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

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

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 could be reasonable?

Technical challenges

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

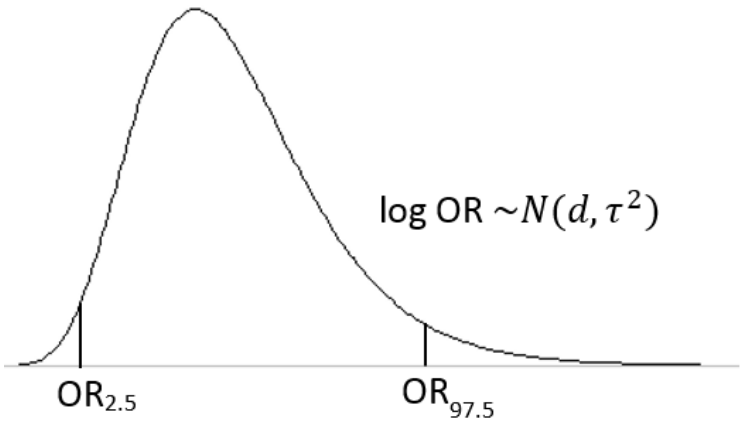
- Elicit the 'range' of treatment effects
- Transform the prior distribution for the 'range' to obtain the prior for the heterogeneity

Notation

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

What quantity to elicit?

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



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 not be 10 times more than the smallest OR in a population of studies.)

What if ?

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

Three-stage procedure for elicitation

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

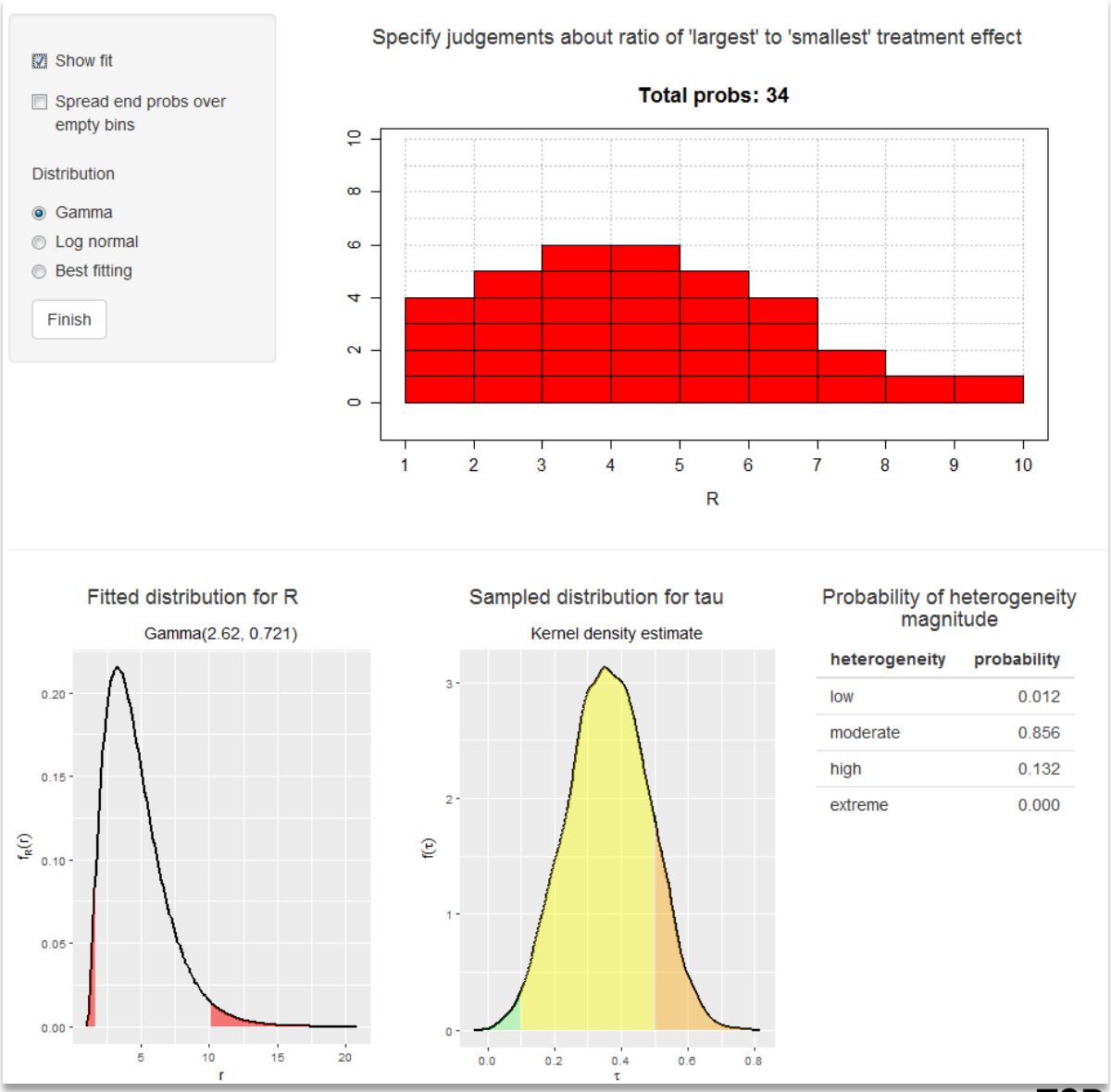
Express some values in $[1, R_{max}]$ as more likely than others → Informative Prior 2: Truncated Turner et al. (2012)

↓

Informative Prior 3: Elicited prior

Implementation

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



NICE TA163

- Disease: ulcerative colitis
- Outcome measure: colectomy rate at 3 months
- Fixed effect model was used
- Re-analyse using a random effects with
  - A vague prior U[0, 5]
  - Informative prior 1: Turner et al. (2012)
  - Informative prior 2: Truncated Turner el al. (2012)
  - Informative prior 3: Elicited prior gamma (2.62, 0.721)

Results

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 (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

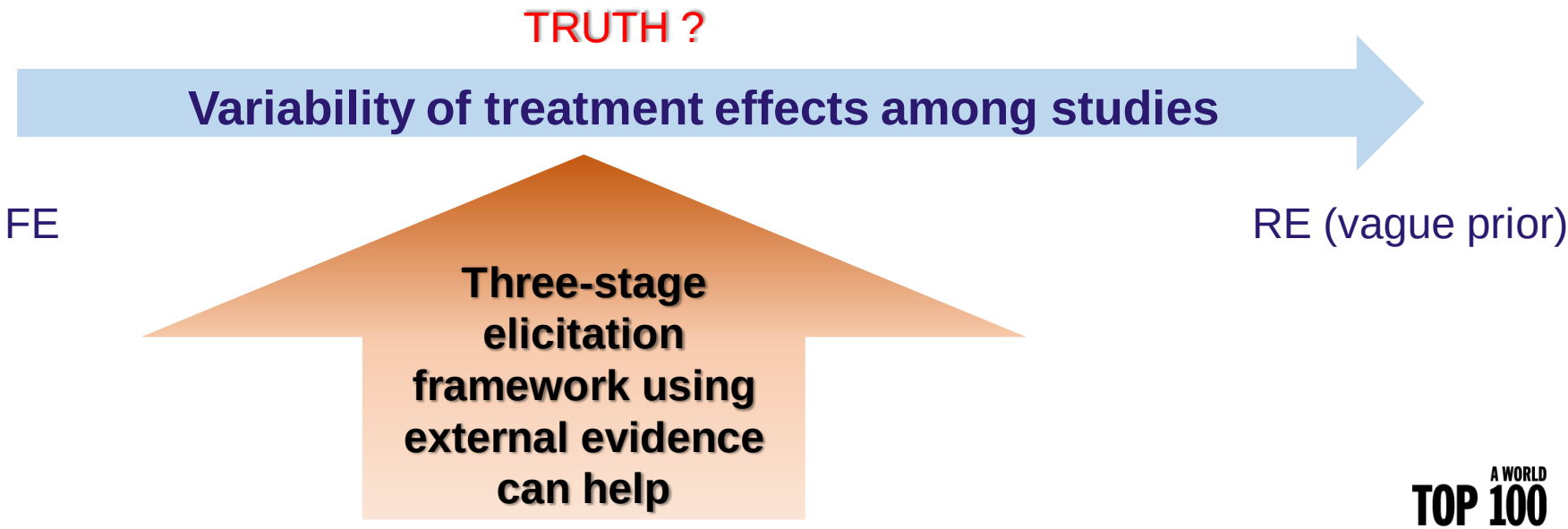
Extensions

- Elicitation framework for other types of outcome measures
  - Ordered categorical
  - Continuous
- See Ren et al. (2018) for details

Summary

Evidence synthesis with limited studies

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



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). 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.

Thank you.