Detecting outlying studies in network meta-analysis using Bayes factors

SILVIA METELLI[1], DIMITRIS MAVRIDIS[2], ANNA CHAIMANI[1]

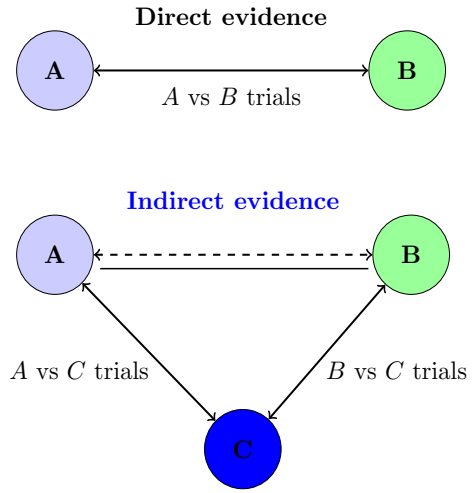
[1]Université de Paris, Inserm Research Center of Epidemiology and Statistics, France

[2]University of Ioannina, Department of Primary Education, Greece

41st Annual Conference of the International Society of Clinical Biostatistics, 23-27 August, 2020

Overview

- **Context**: systematic reviews/evidence synthesis
  - **Network meta-analysis** (NMA): simultaneous comparison of multiple treatments integrating **direct** with **indirect** evidence in a network of studies



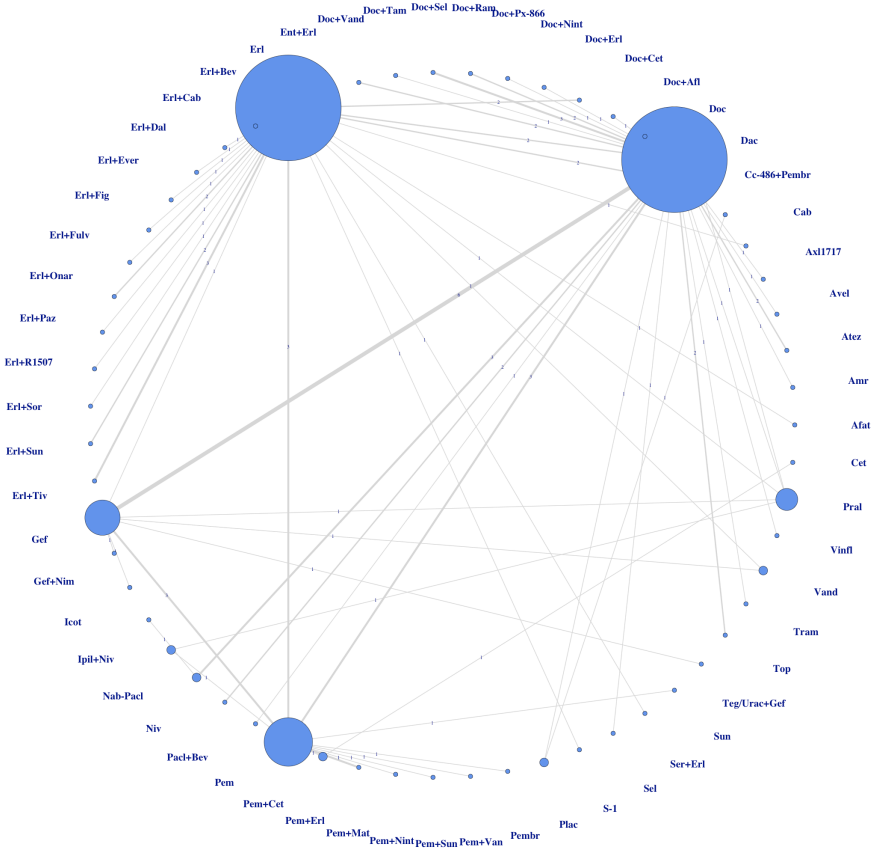
- **Outlier**: study with a clearly different effect estimate (e.g. extreme effect size)
- **Motivation**: outliers can bias NMA conclusions ⟶ wrong clinical decisions
- **Objective:** to develop (Bayesian) approach to detect outliers and to explore their influence on NMA results

note: despite flexible & commonly use, limited work so far on Bayesian outlier-detection methods in NMAs

Motivating data

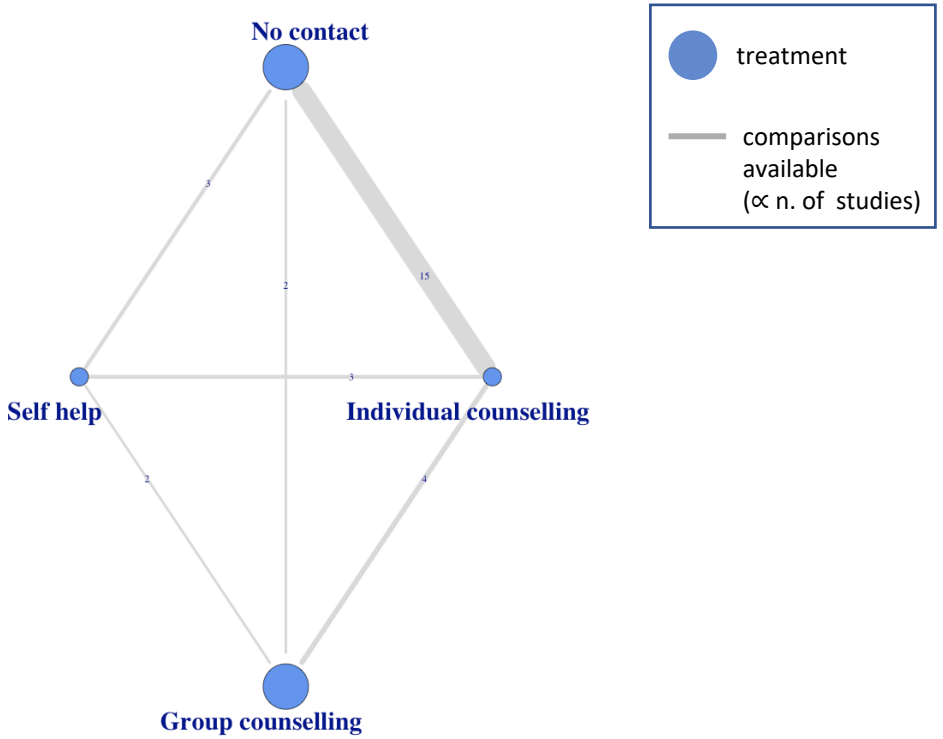
Non small cell lung cancer network

- 112 trials, 62 treatments
- many 'weak' links: outliers?



Smoking cessation network

- 24 trials, 4 treatments (types of counseling)
- fully connected, well studied network



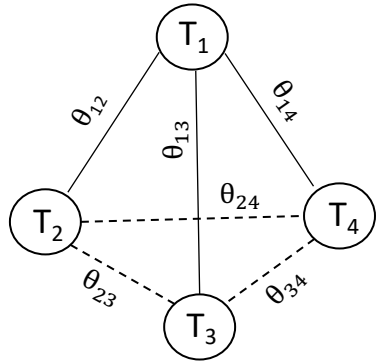
Standard NMA model: (Bayesian) random effects model

N studies, T treatments

For each study *i*: $y_{i,XY}$ observed relative treatment effect (with standard error)

$$y_{i,XY} \sim N\left(\theta_{XY} + \delta_{i,XY}, \sigma^2_{i,XY}\right),$$

$$\delta_{iXY} \sim N(0, \tau_{XY})$$



- $\forall\,(X,Y)$: summary relative effect $\theta_{XY} = \theta_X - \theta_Y$, ($\theta_X, \theta_Y$ basic parameters)
- common heterogeneity $\tau_{XY} = \tau$ across studies
- account for multi-arm trials: $\delta_i \sim N\left(0, \Psi_i^2\right)$, Ψ_i^2 between-study covariance matrix

- **Bayesian approach**: need to specify priors for parameters to estimate (basic parameters, heterogeneity):

  $\theta = (\theta_1, \ldots, \theta_{T-1})^T \sim P(\theta)$, $\tau \sim P(\tau)$ (typically vague priors are assigned)

Outlier NMA model: mean-shift model

- Outlier: study with a 'shifted' mean - different definitions (e.g. shifted variance) lead to different models
- Assumption: effect size of outlying study i shifted by a factor $v \in R$

$$y_{i,XY} \sim N\left(\theta_{XY} + vi + \delta_{i,XY}, \sigma^2_{i,XY}\right), \quad \delta_i \sim N\left(0, \Psi_i^2\right)$$

Testing for outliers:

- For each study i, we test if $y_{i,XY}$ has a shifted effect:

$$H_0: \nu_i = 0; \; H_1: \nu_i \neq 0$$

$if\ \nu_i \neq 0$ mean-shift model looks more plausible: study i is a potential outlier

Bayes factors offer a principled way of testing such hypotheses

Outlier detection: cross-validatory Bayes factors

- We have two models: standard model M_0 (H_0), mean-shift model M_1 (H_1)
- from Bayes theorem we have:

$$\underbrace{\frac{P(M_1|y)}{P(M_0|y)}}_{\text{posterior odds}} = \underbrace{\frac{p(\theta_1)}{p(\theta_0)}}_{\text{prior odds}} \times \underbrace{\frac{\int L(y|\theta_1, M_1)P(\theta_1|M_1)d\theta_1}{\int L(y|\theta_0, M_0)P(\theta_0|M_0)d\theta_0}}_{\text{Bayes factor}}$$

- Bayes factor (BF): change from prior odds to posterior odds
  - → provides a **measure of plausibility of H_1 over H_0**
- We test this hypothesis for each study: leave-one-out cross validation (LOO-CV) scheme
  - standard backward search
  - modified search (restricted to groups of studies comparing the same treatments)

Outlier detection: Bayesian p-value (alternative method)

- Posterior predictive p-values (Bayesian p-values): **measure departure from the assumed (NMA) model**

allows to quantify uncertainty of being outlier

$$p_{D_i} = P(D_i(y^*_{XY}, \theta) \geq D_i(y_{XY}, \theta)|y)$$

where:

- y^* hypothetical future values generated from predictive distribution
- $\mathrm{D}(\cdot)$ some function measuring the discrepancy of model vs. data y
- $\mathrm{P}(\cdot|y)$ posterior distribution of (θ, y^*) given y

- **we propose two choices for D:**
    1. model likelihood
    2. tailored measure of 'outlyingness' for each study

Down-weighting scheme (post-detection step)

- ***What to do once outliers are identified?*** **Down-weight via informative priors.**

Simple idea:

1. NMA model **with additional variance weight** $(0 < w_i < 1)$ for outliers only
2. perform the NMA analysis again
3. compare results

	non-outlying studies	outlying studies
Model	$y_{i,XY} \sim N\left(\theta_{XY} + \delta_{i,XY}, \sigma^2_{i,XY}\right)$	$y_{i,XY} \sim N\left(\theta_{XY} + \delta_{i,XY}, \sigma^2_{i,XY}/w_i\right)$
Informative prior	None	$w_i \sim \text{Beta}(a, b)$ a, b centered at values <0.5

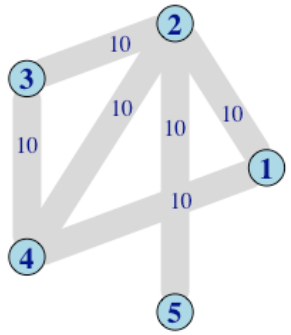
Applications

- **Simulation study:**
  - assess performance of the methods proposed

- **Two real networks of interventions:**
  - demonstrate the methods in practice

Simulation study: networks geometry and settings

- **Balanced design**
  - 10 studies per comparison
  - fairly well connected



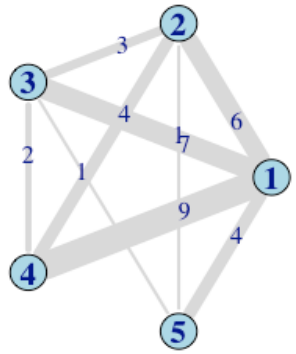
Simulated scenarios:

- network geometries
- $\tau \in \{0, 0.032, 0.096, 0.287\}$ (according to Turner 2012)
- contaminate with 1 or 3 outliers

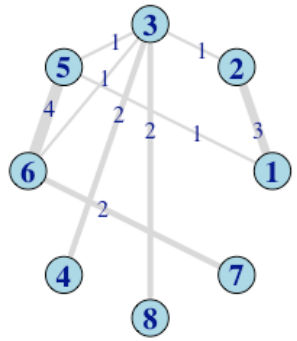
32 scenarios in total

- **Unbalanced design**

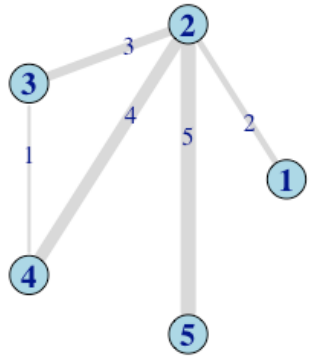
1. Well connected



2. Fairly connected



3. Poorly connected



Simulation study results (detection)

- Unbalanced case (fairly connected network), 3 artificial outliers

Bayes Factor (BF)*

* BF>3 moderate, 3<BF<10 substantial
10<BF<150 strong, BF>150 decisive

	$\tau = 0$	$\tau = 0.032$	$\tau = 0.096$	$\tau = 0.287$
Outlier 1	**511.1**	**287.1**	**9.1**	2.5
Outlier 2	**118.2**	**1540.1**	2.7	1.3
Outlier 3	**9284.1**	**32.1**	3.2	0.98
No outliers	None	None	1 BF ~ 2	1 BF ~ 3

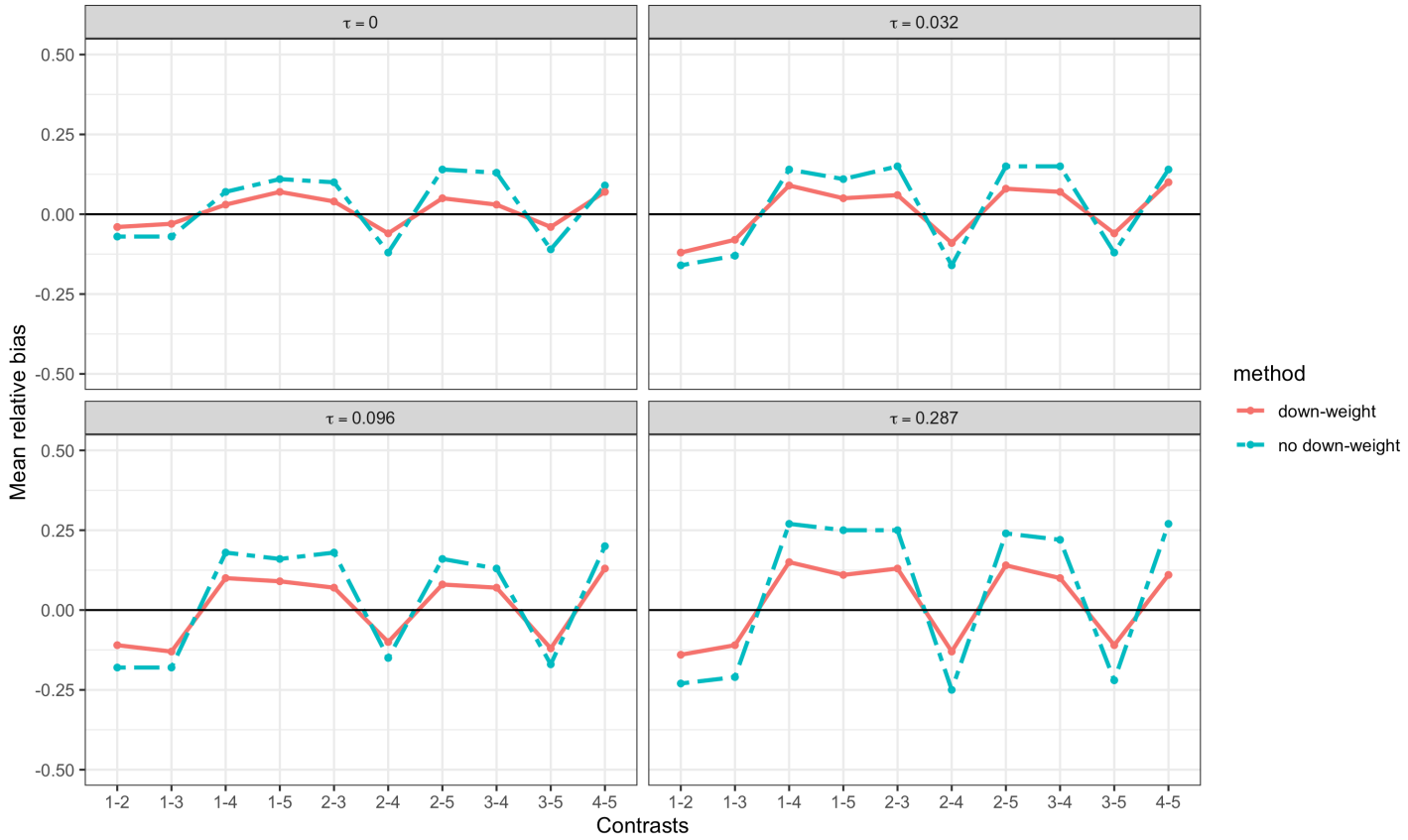
Bayesian p-value*

* D based on *outlyingness* measure

	$\tau = 0$	$\tau = 0.032$	$\tau = 0.096$	$\tau = 0.287$
Outlier 1	**0.001**	**0.001**	**0.01**	0.12
Outlier 2	**<0.01**	**<0.0001**	**0.001**	0.07
Outlier 3	**<0.0001**	**0.01**	0.07	0.22
No outliers	None	None	None	0.05

Simulation study results (down-weight)

- Unbalanced case (poorly connected network), 3 artificial outliers
- Relative bias= $(\hat{\theta}^{MC} - \theta^{\text{true}})/\theta^{\text{true}}$, with $\hat{\theta}^{MC}$: Monte Carlo average of estimated effects



Real data

Lung cancer network	Smoking cessation network
Study 67: BF=1353.2 p-value=<0.001	Study 3: BF=287.8 p-value=<0.01
Study 42: BF=876.1 p-value=<0.01	--------
Study 7 : BF=10.2 p-value=<0.01	--------

Results (compared with forward search (FS) algorithm for NMA, Petropoulou 2019)**:**

- **3 potential outliers in lung cancer data**
- **1 potential outlier in smoking cessation data**
- FS detection method: similar results
- sensitivity analyses and down-weighting suggest first two studies in lung cancer and study 3 in smoking data are influential

Final Remarks

Conclusions

- outlying studies need attention when synthesizing evidence
- two outlier-detection methods proposed (model-based Bayesian)
- promising results:
  - in simulations, good performance of both methods, down-weight improves estimates precision
  - in real data, some detected studies proved influential
- amount of heterogeneity and number of studies play crucial role

Future directions

- extend to multiple outcomes
- use external data to inform down-weight (e.g. informative *power priors*)... but external data not easy to get!

Some useful references

1. Lu AG and Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. *Journal of the American Statistical Association*. 2006;474:447-459.

2. Salanti G, Higgins JP, Ades AE, and Ioannidis JP. Evaluation of networks of randomized trials. *Statistical Methods in Medical Research*. 2008;17(3):279–300

3. Mavridis D, Moustaki I, Wall M, and Salanti G. Detecting outlying studies in meta-regression models using a forward search algorithm. *Research Synthesis Methods*. 2017;8(2):199–211.

4. Petropoulou M, Salanti G, Rucker G, Schwarzer G, Moustaki I, and Mavridis D. A forward search algorithm for detection of extreme study effects in network meta-analysis. Manuscript, 2019.

5. Zhang J, Fu H, and Carlin BP. Detecting outlying trials in network meta-analysis. *Statistics in Medicine*. 2015;34(19):2695–2707.

6. Jeffreys H. Theory of probability. Oxford University Press, 1998.

7. Raftery A. Bayesian model selection in social research (with discussion by Andrew Gelman & Donald b. Rubin, and Robert Hauser, and a rejoinder), 1995.

8. Kass R, Raftery A. Bayes Factors. *Journal of the American Statistical Association.* 1995; 90(430): 773-795.

THANK YOU!