

# The Network meta-analysis with Missing Outcome data (NEMO) project: a combination of a systematic overview, empirical and simulation study

Loukia Spineli

*AG Hebammenwissenschaft, Klinik für Frauenheilkunde und Geburtshilfe*



Medizinische Hochschule  
Hannover

Gefördert durch

**DFG** Deutsche  
Forschungsgemeinschaft

[Spineli.Loukia@mh-hannover.de](mailto:Spineli.Loukia@mh-hannover.de)

# External collaborators



## **Professor Holger Schünemann**

Department of Health Research Methods, Evidence and Impact (HEI),  
and Department of Medicine, McMaster University, Canada.



## **Dr. Juan Yepes-Nuñez**

Department of Health Research Methods, Evidence and Impact (HEI),  
McMaster University, Canada.

School of Medicine, University of Antioquia, Medellín, Colombia.



## **Konstantinos Pateras, PhD cand.**

Department of Biostatistics and Research Support, Julius Center for  
Health Sciences and Primary Care, University Medical Center Utrecht,  
Utrecht, The Netherlands.

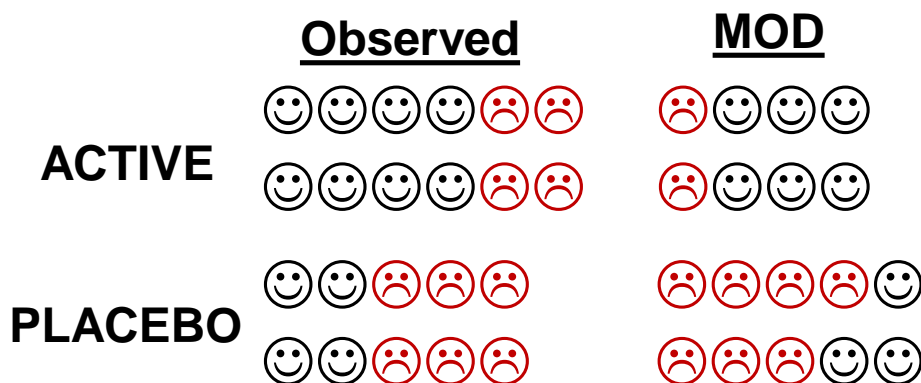


## **Chrysostomos Kalyvas, M.Sc.**

Department of Biostatistics and Research Decision Sciences, Merck  
Sharp and Dohme Europe Inc., Brussels, Belgium.

# Why the NEMO project was proposed? (1)

- **Missing outcome data (MOD)** are ubiquitous in clinical trials...
- ... and in systematic reviews, inevitably!
- Implications of MOD on results:
  - **loss of precision**, if MOD are excluded;
  - **risk of bias**, if MOD are handled inappropriately  
(i.e. exclusion or imputation before analysis).



%MOD: 40% vs 50% (active vs placebo)

Truth	14/20	7/20	4.3 (1.2, 16)
Exclude	8/ <u>12</u>	4/ <u>10</u>	3.0 (0.5, 17)
Impute ☹	8/20	4/20	2.7 (0.6, 11)
Impute ☺	16/20	14/20	1.7 (0.4, 7.3)

# Why the NEMO project was proposed? (2)

- Several strategies to address MOD in systematic reviews\* **but without evidence on their performance:**
  - exclude before analysis with & without uncertainty due to MOD;
  - impute before analysis & ignore uncertainty due to MOD;
  - model using pattern-mixture or selection model.

\*Higgins JPT, Deeks JJ, Altman DG on behalf of the Cochrane Statistical Methods. Chapter 16: Special topics in statistics. In: Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration; 2011

# Why the NEMO project was proposed? (3)

- **Intention-to-treat (ITT)** principle preferred in the presence of MOD.
- Appropriate application of ITT in systematic reviews requires:
  - Distinction between observed outcomes and MOD;  
*(observation-carried-forward methods hinder such an accurate distinction)*
  - Information on the reasons for MOD;  
*(transparent and detailed reporting of trials is required)*
  - Sensitivity analyses that include all randomised individuals under plausible scenarios about MOD.  
*(to evaluate the robustness of the results of primary analysis)*

# Why the NEMO project was proposed? (4)

- Reviewers fail to provide an accurate extraction of outcome data from the included trials.

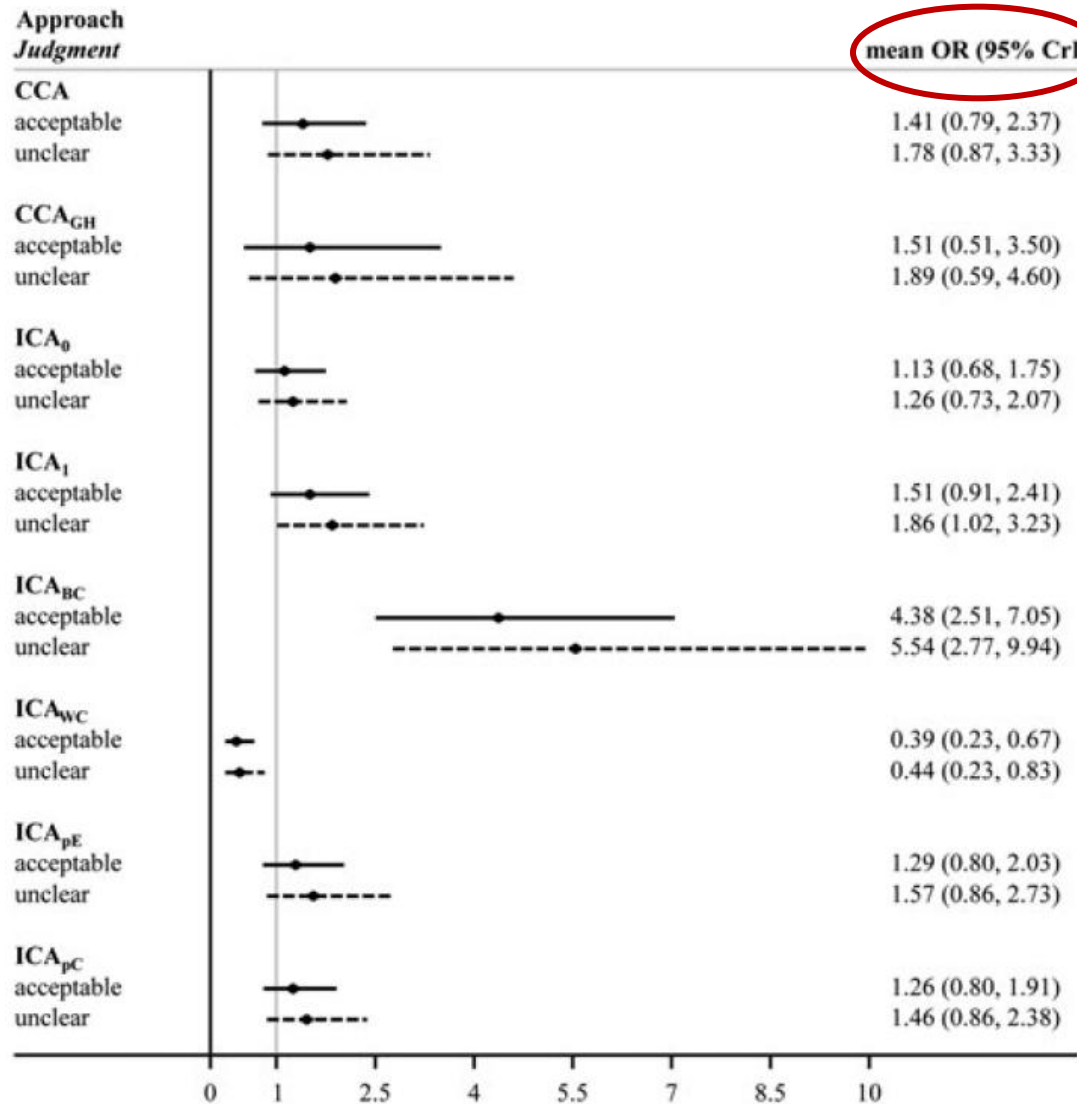
Item	Aspects reported in the systematic review	Extraction evaluation		
		Acceptable	Unclear	Unacceptable <sup>a</sup>
i.	Information on outcome of completers for all included studies	✓		
ii.	Information on outcome “withdrawal” for all included studies	✓	✓	✓
iii.	Explicit description of how missing outcome data in meta-analyses is handled	✓	✓	
iv.	Information on outcome of completers and/or outcome “withdrawal” but only for a minority of included studies			✓

<sup>a</sup>Meta-analyses that included at least one study with calculated negative non-events were immediately considered to be “unacceptably” extracted.

Cochrane Mental Health Group	Extraction Is Judged		
	Acceptable	Unclear	Unacceptable
Depression, anxiety and neurosis	10	17 <sup>a</sup>	25
Schizophrenia	0	42 <sup>a</sup>	16
Developmental, psychosocial and learning problems	1	1	1
<b>Total</b>	<b>11</b>	<b>60</b>	<b>42</b>

<sup>a</sup>Three systematic reviews in the Schizophrenia Group and one in the Depression, Anxiety and Neurosis Group did not provide adequate information for indicate “acceptable” extraction but reported attrition rates below 5%, which were too low to raise accuracy concerns.

# Why the NEMO project was proposed? (5)

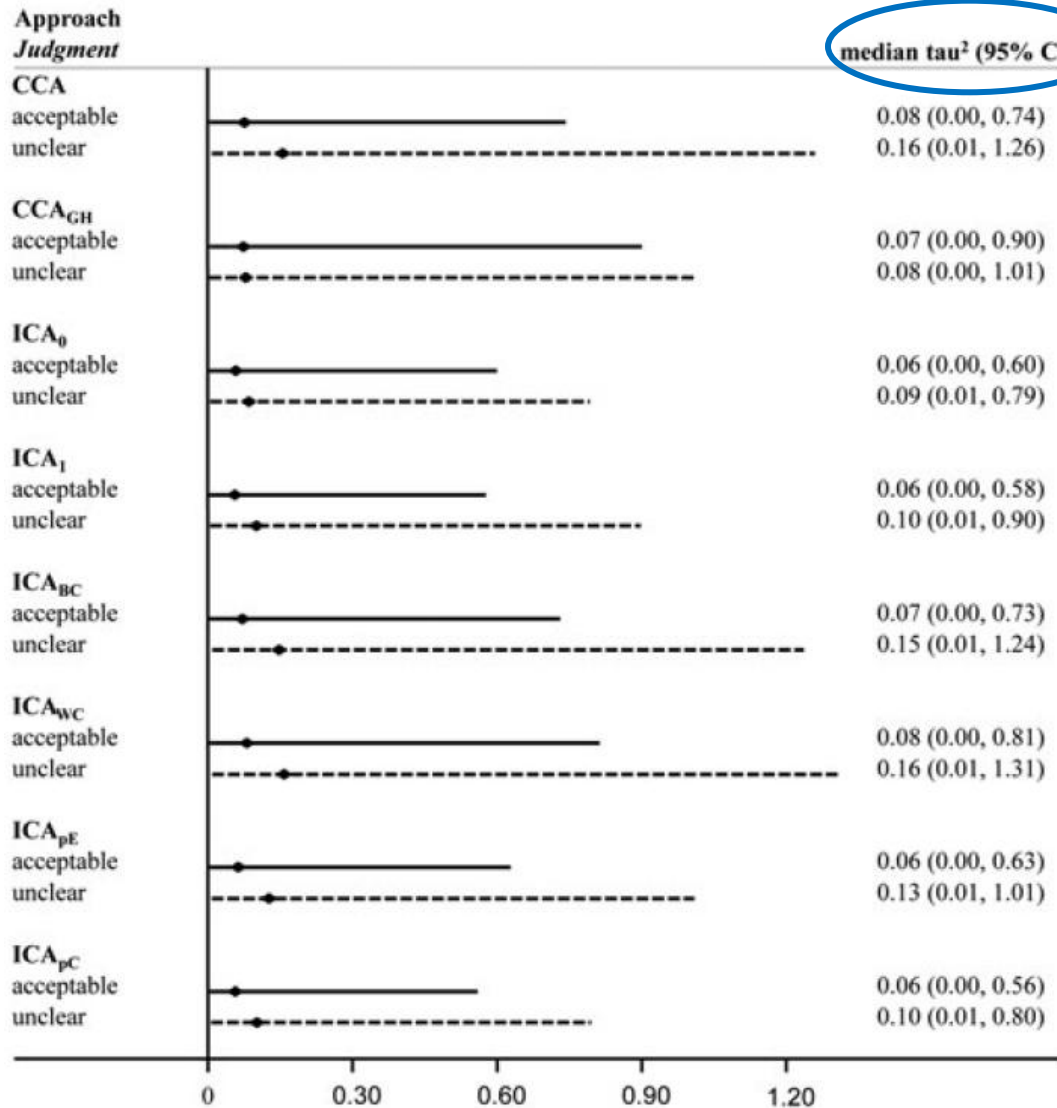


➤ Inaccurate extraction of trial data may :

- lead to **biased meta-analysis results**;
- **hinder the performance** of methods to handle MOD.

*From a systematic review with a meta-analysis with ,**acceptable**' extraction and the same meta-analysis on LOCF data (,**unclear**').*

# Why the NEMO project was proposed? (6)



➤ Inaccurate extraction of trial data may :

- lead to **biased meta-analysis results**;
- **hinder the performance** of methods to handle MOD.

*From a systematic review with a meta-analysis with ,**acceptable**' extraction and the same meta-analysis on LOCF data (,**unclear**').*



**‘What is the reporting quality of systematic reviews of multiple interventions in terms of MOD?’**

# Search strategy

## Published databases considered:

- Zarin et al. → **inception – 14/04/2015;  $\geq 4$  interventions**
- Tan et al. } → **1997 – 07/2012;  $\geq 3$  interventions**
- Bafeta et al. }
- Nikolakopoulou et al. → **inception – 12/2012;  $\geq 4$  interventions**

## Our own search:

**08/2012 – 03/2017;  $\geq 3$  interventions; using **Petropoulou et al. search strategy** in MEDLINE, EMBASE and the CDSRs.**

# Eligibility strategy (1)

## Following Zarin et al.:

- ✓ Systematic reviews of randomized controlled trials (RCTs);
- ✓ No language restriction;
- ✗ Diagnostic test accuracy studies;
- ✗ Genetic studies;
- ✗ Observational studies;
- ✗ Mixture of RCTs and observational studies;
- ✗ Number of included trials  $<$  number of interventions.

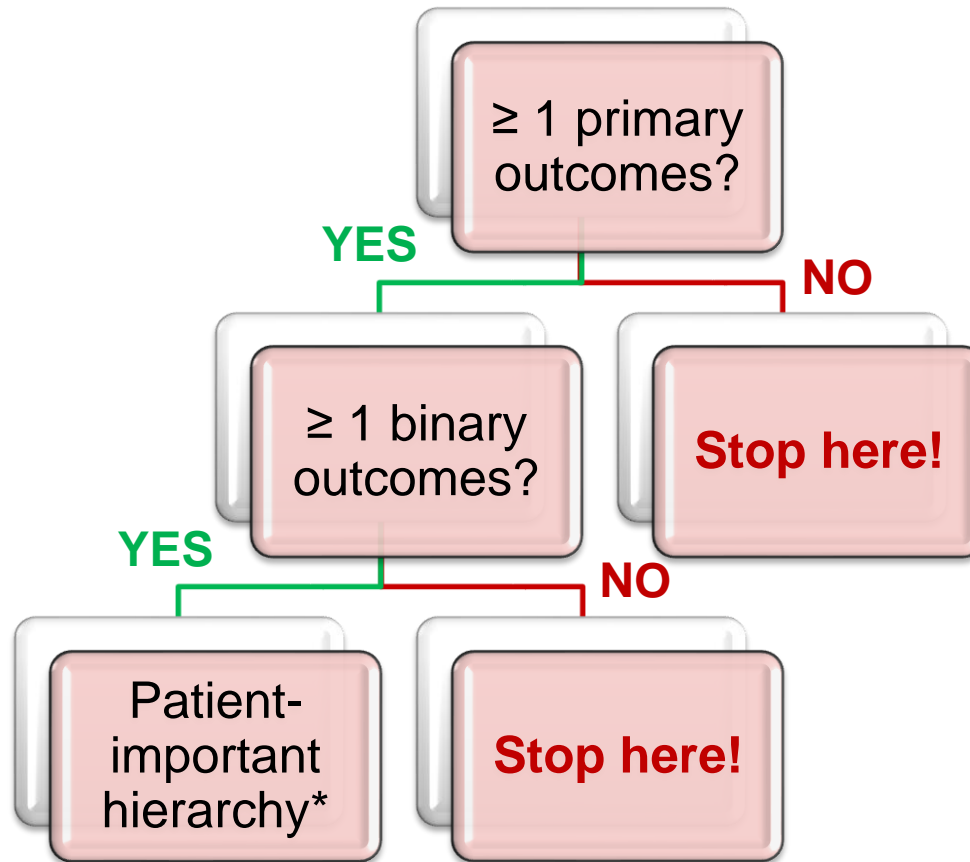
# Eligibility strategy (2)

## Our own criteria:

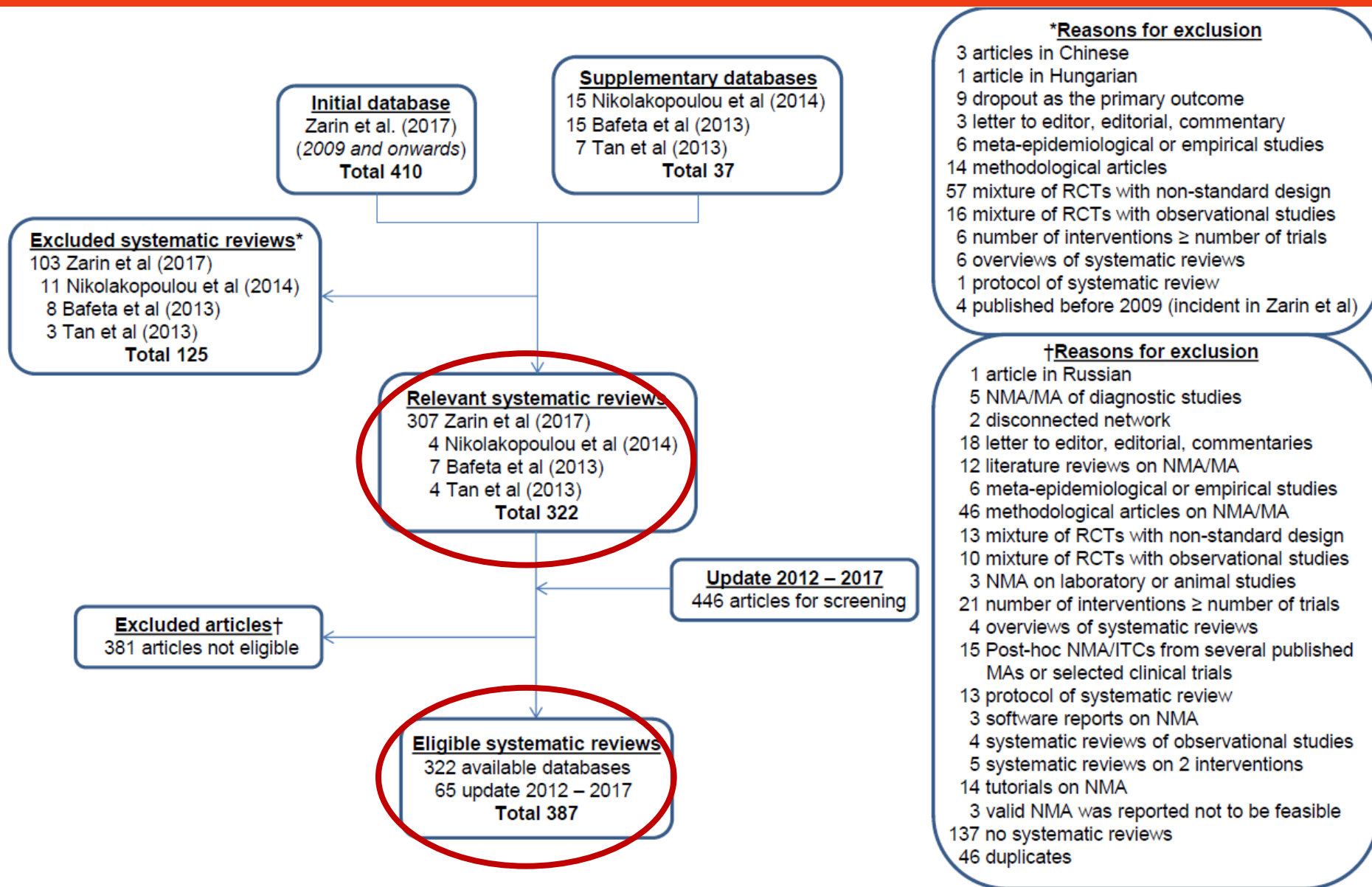
- ✓ At least 3 interventions are investigated;
- ✓ Systematic reviews published from **2009 and onwards**;
  - ❖ *the new Cochrane risk of bias tool was published during 2009*
- ✗ RCTs with non-standard design;
  - ❖ *e.g. quasi, crossover, factorial, cluster, split-mouth.*
- ✗ Overviews of systematic reviews;
- ✗ MOD investigated as single primary outcome.

# Eligibility strategy (4)

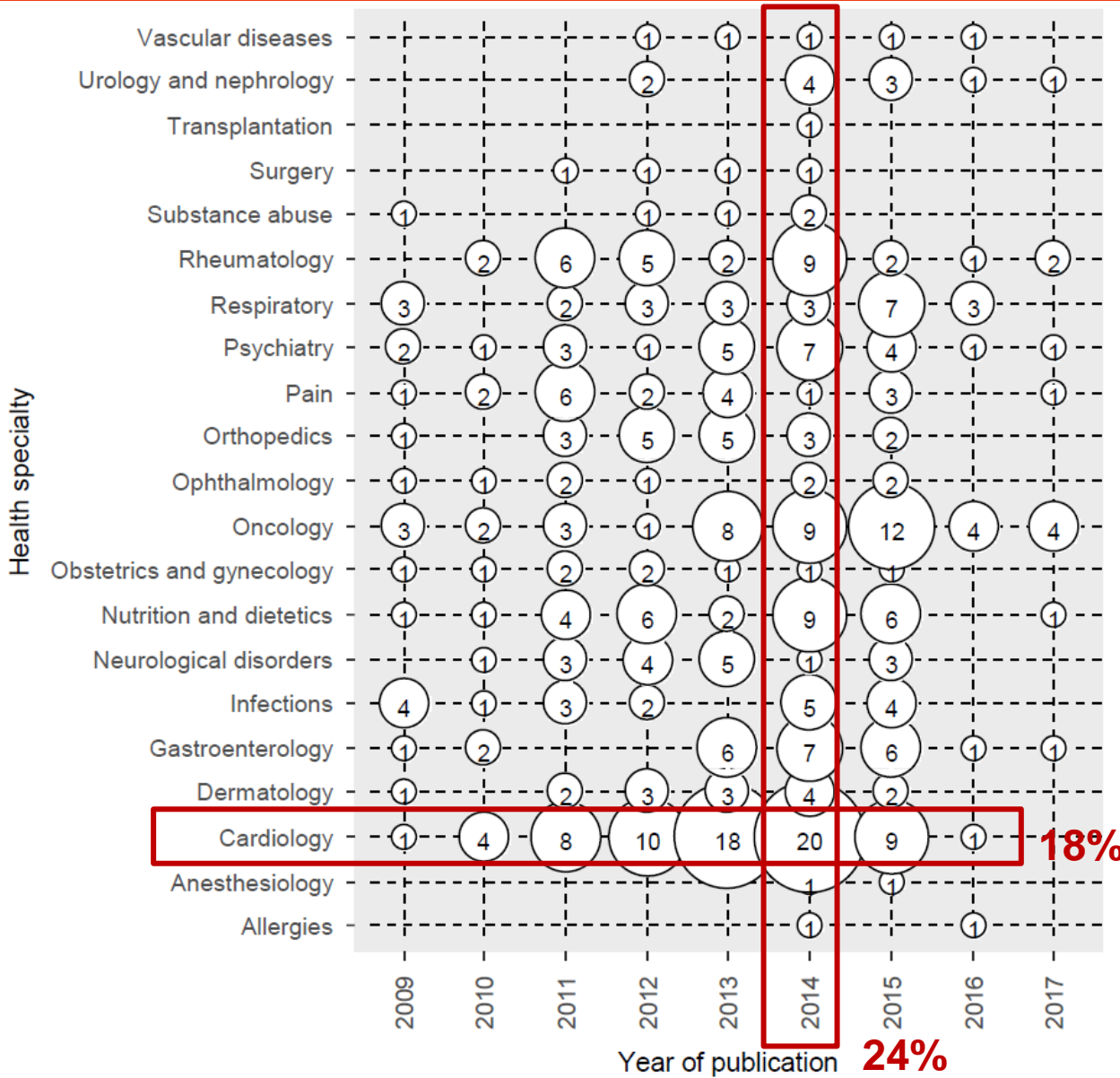
For each eligible Systematic review:



\*Akl EA, Kahale LA, Agarwal A, et al. Impact of missing participant data for dichotomous outcomes on pooled effect estimates in systematic reviews: a protocol for a methodological study. Syst Rev 2014;3:137



# 387 selected Systematic Reviews

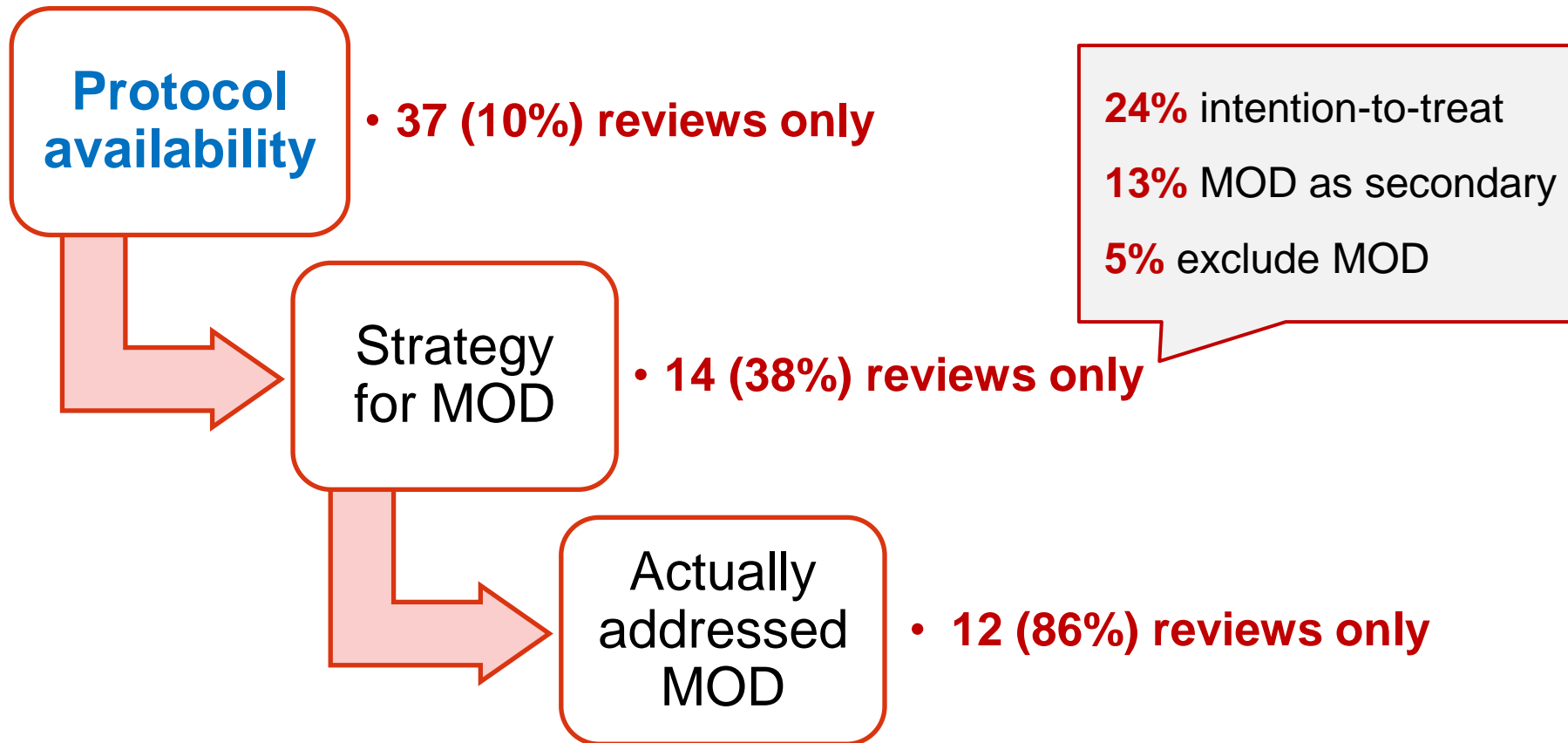


**Non-Cochrane  
reviews: 98%**

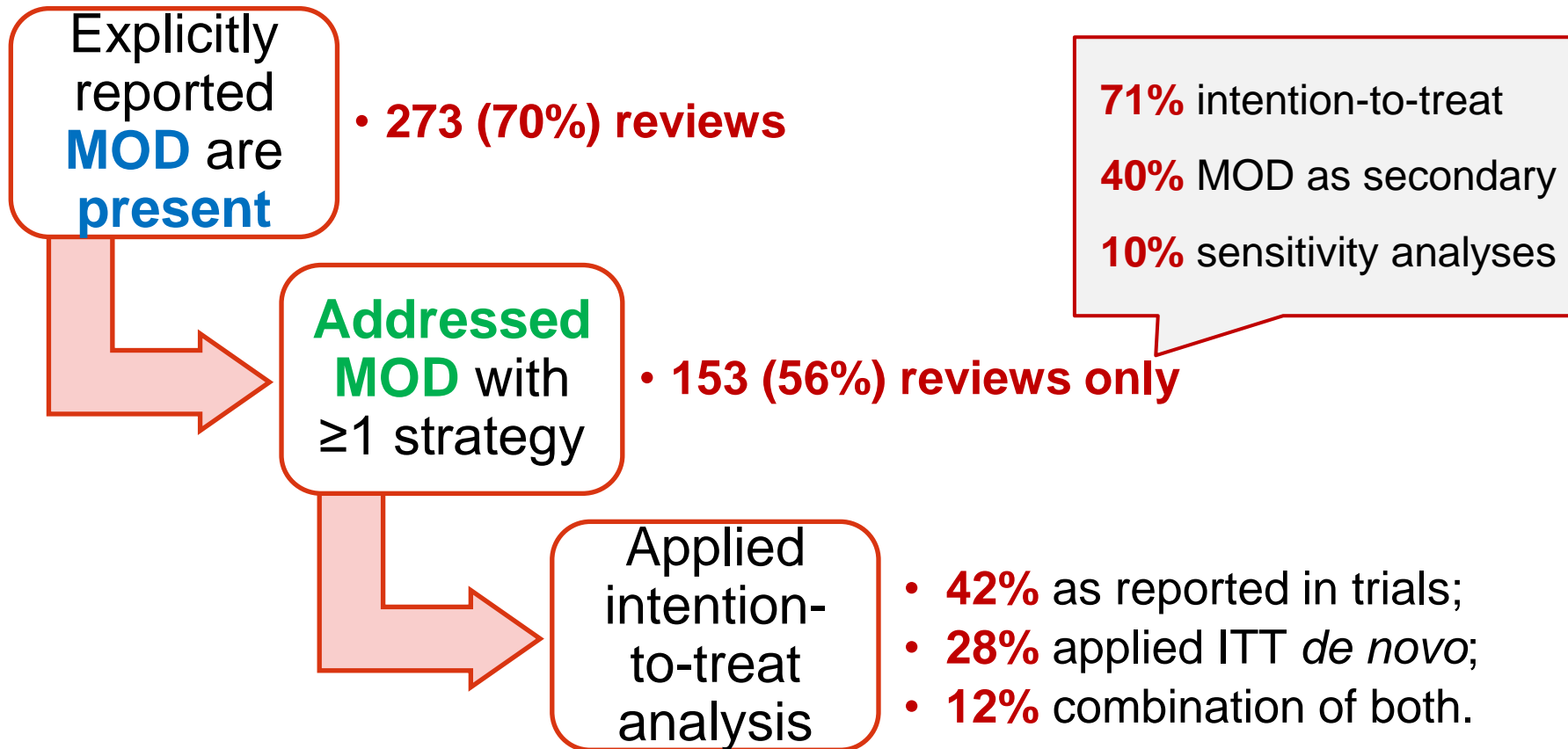


# Addressing MOD in the

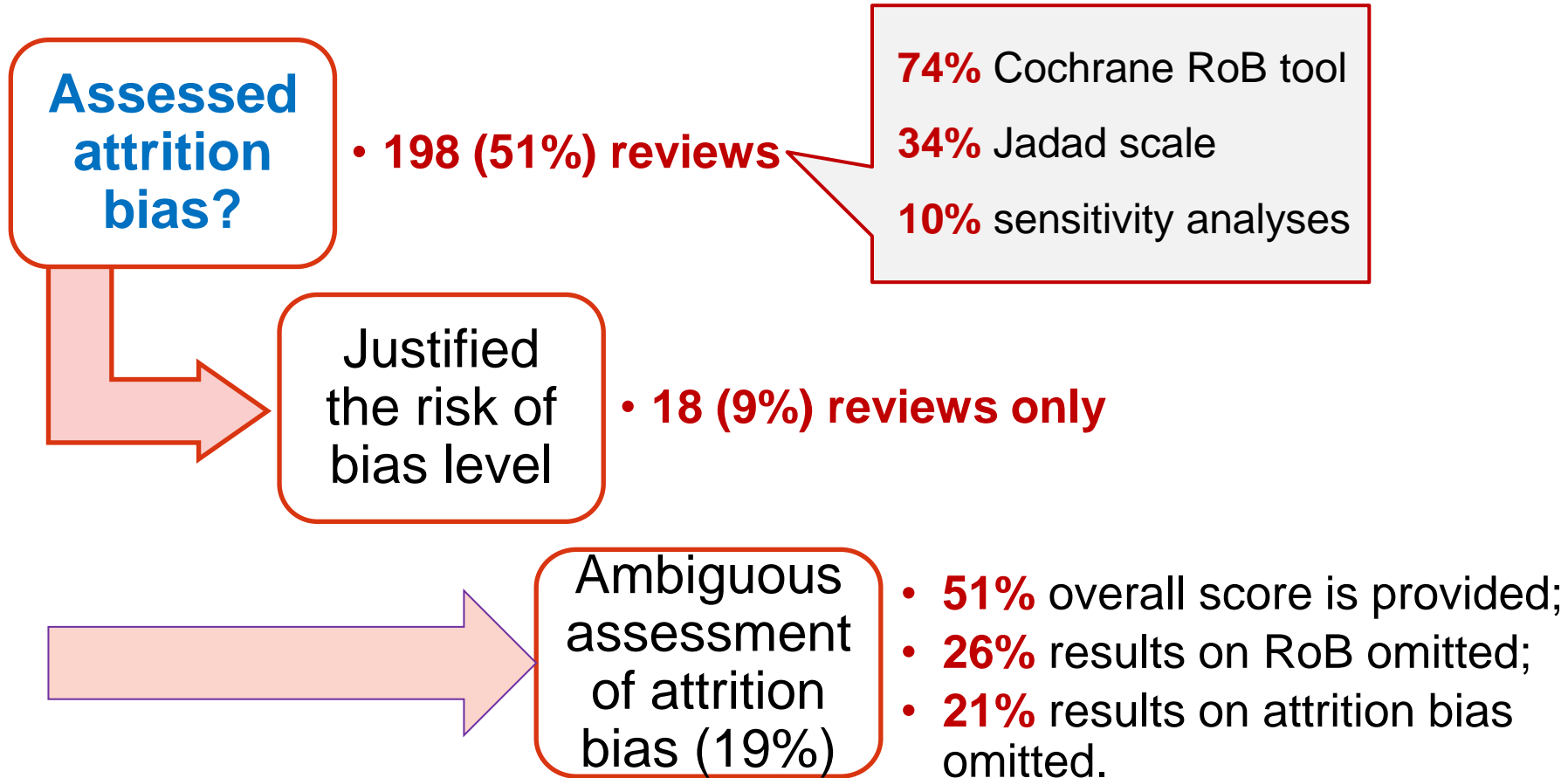
**PROTOCOL**



# Addressing MOD in the **REVIEW**



# Addressing MOD in the **REVIEW**



# Addressing MOD in the **REVIEW**

## Implications of MOD?

- **88 (32%) out of 273 reviews**

**66%** MOD as secondary  
**18%** sensitivity analysis  
**12%** MOD prevalence

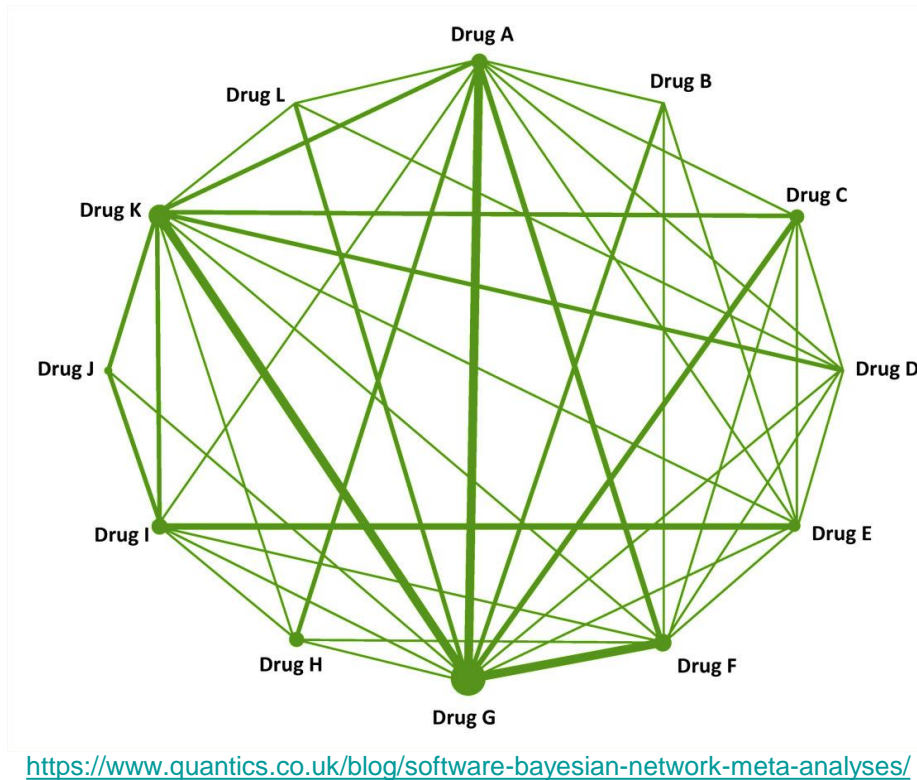
## Where reported mainly?

- **74 (84%) in Discussion**
- **46 (52%) in Abstract**

## Which NMA components mainly?

- **84 (95%) NMA effects;**
- **13 (15%) intervention ranking.**

# 273 NMAs with present MOD



# Addressing MOD in the **NMA**

Strategy  
explicitly  
reported

- in 113 (40%) reviews only
  - 92% *claimed intention-to-treat*

60% as reported in trials  
22% exclusion of MOD  
17% intention-to-treat

Actual  
strategy  
judged

- agreed with 14 (12%) reviews
- able to judge in 95 (35%) reviews

Intention-to-  
treat analysis  
(16; 17%)

- 8 all MOD as non-events;
- 1 all MOD as events;
- 7 no scenario is provided.

# Addressing MOD in the **NMA**

**Applied  
sensitivity  
analysis**

• **16 (6%) reviews only**

**7** trial exclusion

**5** available case analysis

**3** imputation

Reported  
any  
changes

• **2 reviews only**

‘To carry out a clinically sound analysis, we used a conservative approach and imputed outcomes for the missing participants assuming that they did not respond to treatment’

Justified strategy  
used in primary &  
sensitivity analysis

• **12 out of 273** reviews

Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. Lancet. 2009;373(9665):746-58.

# Conclusions

- ❑ The **quality of reporting and handling MOD** in systematic reviews with NMAs is **particularly inadequate!**
- ❑ Reviewers remain **unaware of the presence and importance of MOD** in systematic reviews of multiple interventions.
- ❑ Poor handling of MOD attests to **limited knowledge of the reviewers** regarding the existing relevant methodology.
- ❑ **Education** amongst reviewers, peer reviewers and journal editors is deemed **necessary!**



**‘Do different models for binary MOD agree in  
terms of core components of NMA ?’**

# Informative Missingness Odds Ratio (IMOR)

Pattern-mixture model (White et al., 2008)

$$IMOR = \frac{p^m / (1 - p^m)}{p^o / (1 - p^o)} \quad \begin{array}{l} \text{'Odds of an event being missing to} \\ \text{odds of an event being observed'} \end{array}$$

$$\log(IMOR) = \delta_{ik} \sim N(\Delta_{ik}, \sigma_{ik}^2) \quad \text{arm } k \text{ of trial } i$$

$\Delta$	interpretation
$\Delta > 0$	more likely that a missing case to be an event
$\Delta < 0$	less likely that a missing case to be an event
$\Delta = 0$	Missing at random (on average)

$$\sigma_{ik}^2 = \sigma^2 = 1$$

# Prior structures of logIMOR (1)

## Identical structure\*

... log IMORs are considered identical depending on further assumptions on whether missingness mechanisms are:

- **common in the whole network**
- **intervention-related**
- **trial-related**

Assumption	Prior specification
Common-within-network	$\delta_{ik} = \delta, \delta \sim N(0,1)$
Intervention-specific	$\delta_{ik} = \delta_{t_{ik}}, \delta_{t_{ik}} \sim N(0,1)$
Trial-specific	$\delta_{ik} = \delta_i, \delta_i \sim N(0,1)$

\*Turner NL, Dias S, Ades AE, Welton NJ. A Bayesian framework to account for uncertainty due to missing binary outcome data in pairwise meta-analysis. Stat Med 2015;34:2062-80

# Prior structures of logIMOR (2)

## Hierarchical structure (Turner et al. 2015)

... log IMORs are considered different yet related to each other by borrowing information either across:

- the whole network
- different trials for the same intervention
- different interventions in the same trial

Assumption	Prior specification
Common-within-network	$\delta_{ik} \sim N(\Delta, \sigma^2)$ $\Delta \sim N(0, 1), \sigma \sim U(0, 1)$
Intervention-specific	$\delta_{ik} \sim N(\Delta_{tik}, \sigma_{tik}^2)$ $\Delta_{tik} \sim N(0, 1), \sigma_{tik} \sim U(0, 1)$
Trial-specific	$\delta_{ik} \sim N(\Delta_i, \sigma_i^2)$ $\Delta_i \sim N(0, 1), \sigma_i \sim U(0, 1)$

# Missingness scenarios

$$\delta_{tik} \sim N(\Delta_{tik}, 1)$$

Scenario*	Interpretation	Abbrev.
$e^{\Delta_{ik}} = 1$	Missing at random more likely	<b>MAR</b>
$e^{\Delta_{tik}} = 2$	More missing are events	<b>MME</b>
$e^{\Delta_{tik}} = 1/2$	More missing are non-events	<b>MMNE</b>
$e^{\Delta_{tik}} = 2$ & $e^{\Delta_R} = 1/2$	<b>MME</b> for non-references; <b>MMNE</b> for reference	<b>BC</b>
$e^{\Delta_{tik}} = 1/2$ & $e^{\Delta_R} = 2$	<b>MMNE</b> for non-references; <b>MME</b> for reference	<b>WC</b>



\*Spineli LM, Higgins JP, Cipriani A, Leucht S, Salanti G. Evaluating the impact of imputations for missing participant outcome data in a network meta-analysis. Clin Trials 2013;10:378-88

# Classification of networks based on MOD (1)

**Prevalence: low, moderate, large**

- ❑ **Low** MOD → 12/29 (41%) networks
  - ❑ **Moderate** and **balance** MOD → 14/29 (48%) networks
  - ❑ **Moderate** and **imbalance** MOD → 2/29 networks
  - ❑ **Large and imbalance** MOD → 1/29 networks
- Characterising networks**

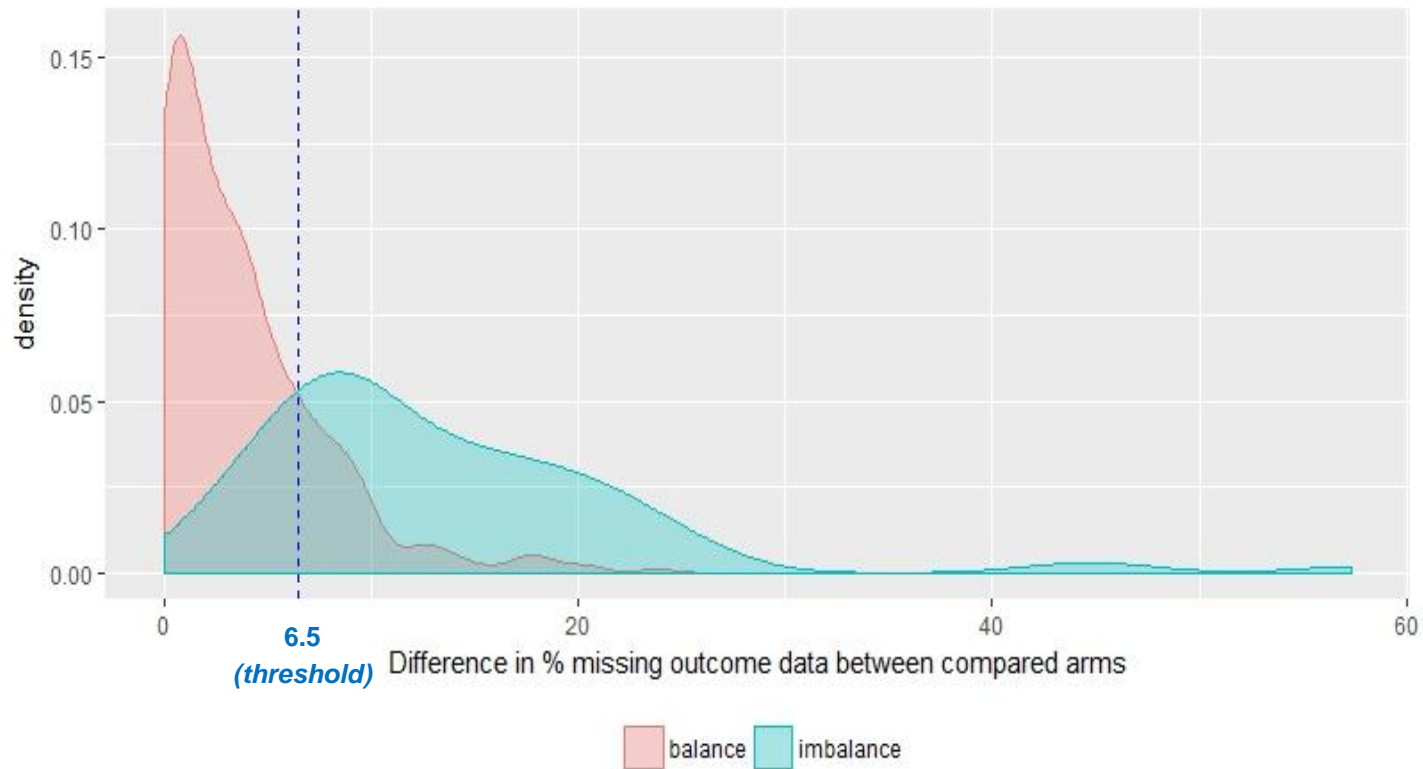
Use of the *‘five to twenty’ rule* based on Sackett et al (1997) ***in each network:***

- **low** → median of total %MOD across trials < 5%
- **high** → median of total %MOD across trials > 20%
- **moderate** → otherwise

# Classification of networks based on MOD (2)

## Balance vs. Imbalance in the compared arms

Calculate difference in %MOD between compared arms ***within each trial*** :



# Model specification & Presentation of results



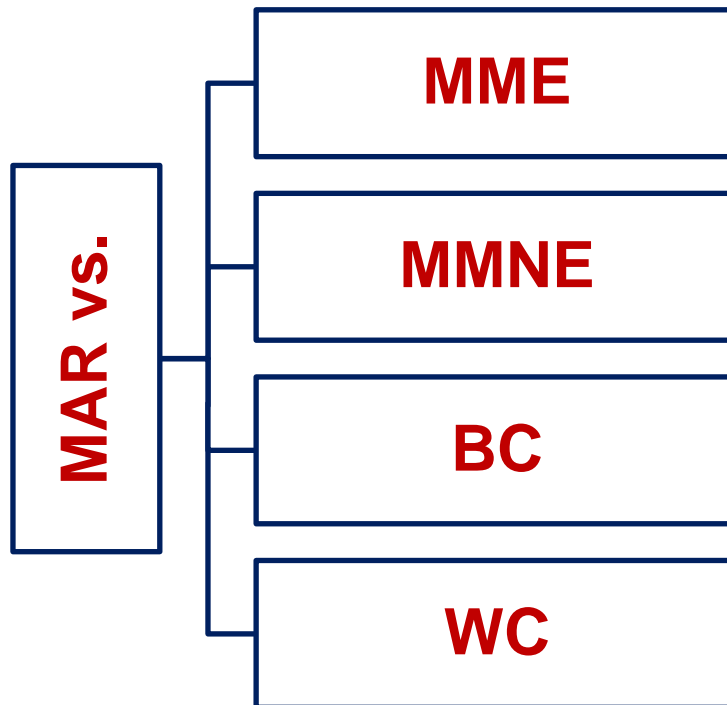
- Bayesian random-effects NMA model by Dias et al (2013);
- Incorporation of log IMOR as described in Turner et al (2015);
- For location parameters use of  $N(0, 10\,000)$ ;
- Use of empirical priors on between-trial variance based on Turner et al (2015);
- 3 chains for 10 000 updates and 1 000 burn-in.
- Bland-Altman plots to investigate the agreements in the compared methods



# Research questions (1)

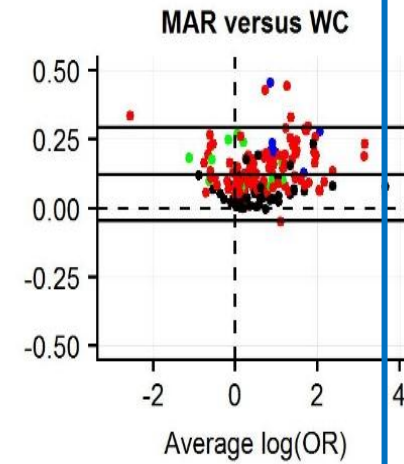
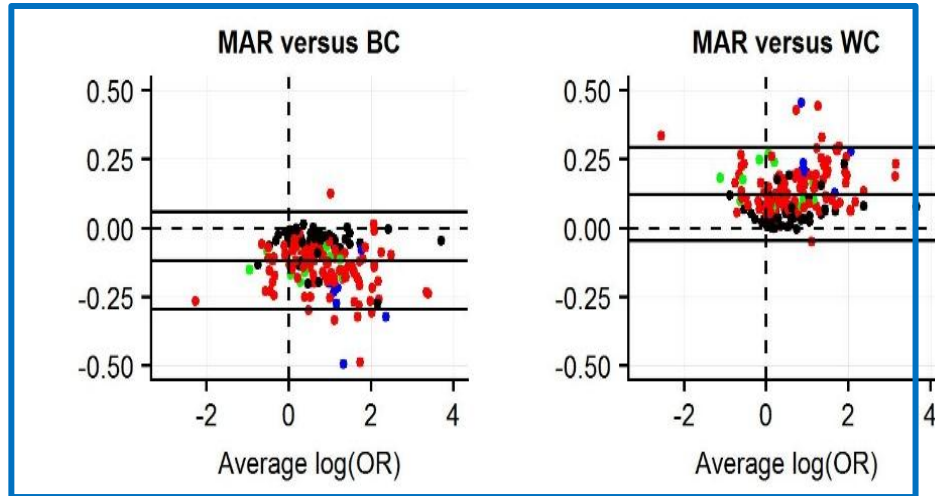
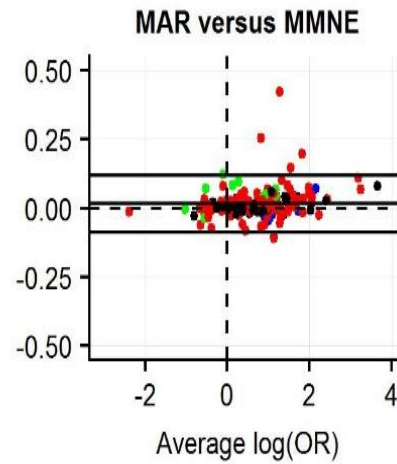
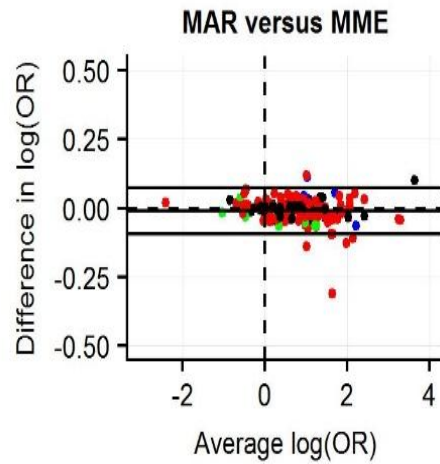
## Agreement between on average MAR and extreme scenarios

- Identical log IMORs with intervention-specific normal prior distribution



$$\delta_{ik} = \delta_{tik}, \delta_{tik} \sim N(\Delta_{tik}, 1)$$

# Posterior mean



# Research questions (2)

## Agreement between accountability & ignorance of uncertainty due to MOD

- Identical log IMORs with intervention-specific normal prior distribution

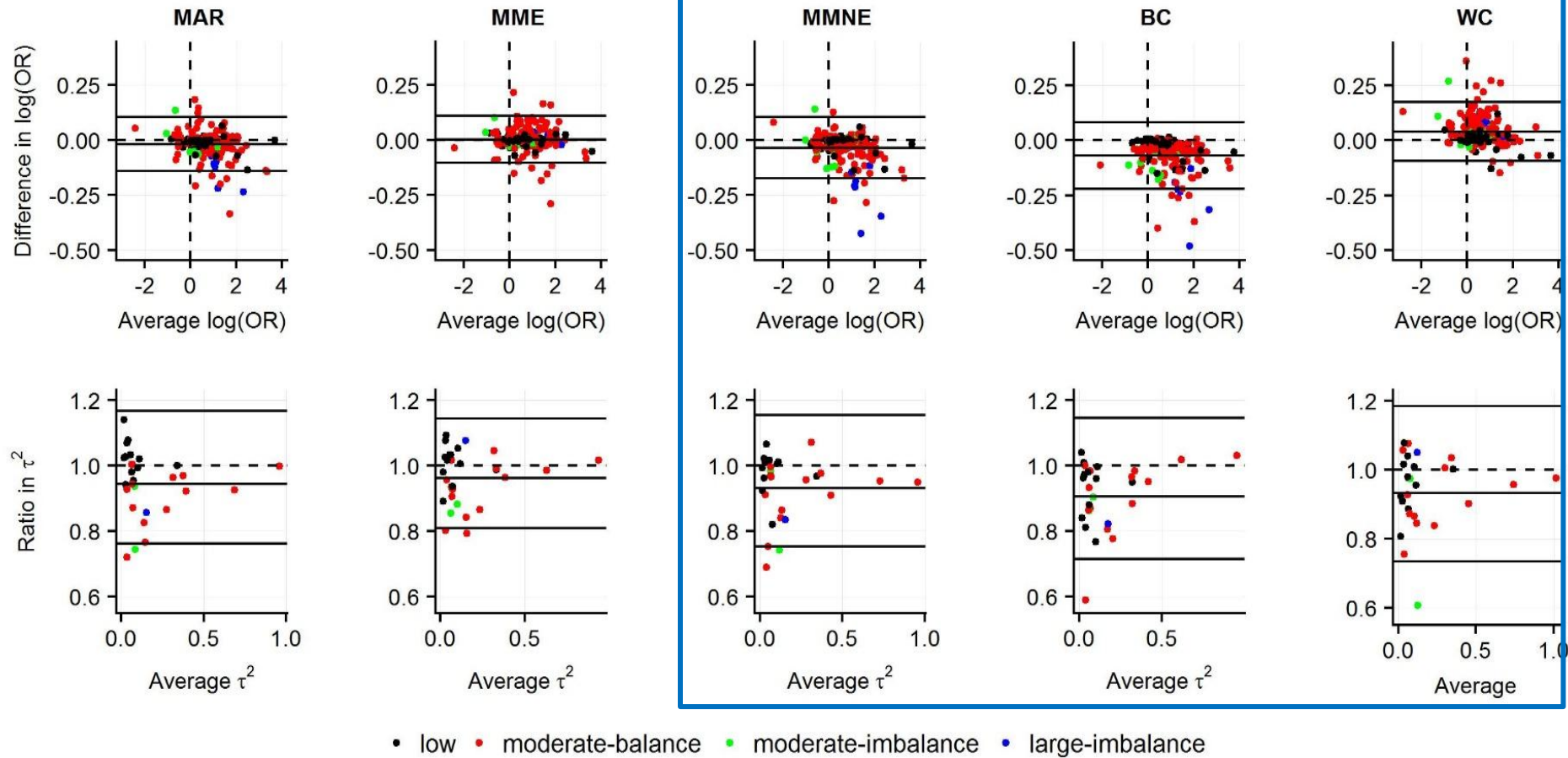
$$\delta_{ik} = \delta_{t_{ik}}, \delta_{t_{ik}} \sim N(\Delta_{t_{ik}}, 1)$$

- Scenarios considered: MAR, MME, MMNE, BC and WC

uncertainty about the scenario considered	
Accounted for vs.	ignored
$\delta_{t_{ik}} \sim N(\Delta_{t_{ik}}, 1)$	$\delta_{t_{ik}} = \Delta_{t_{ik}}$

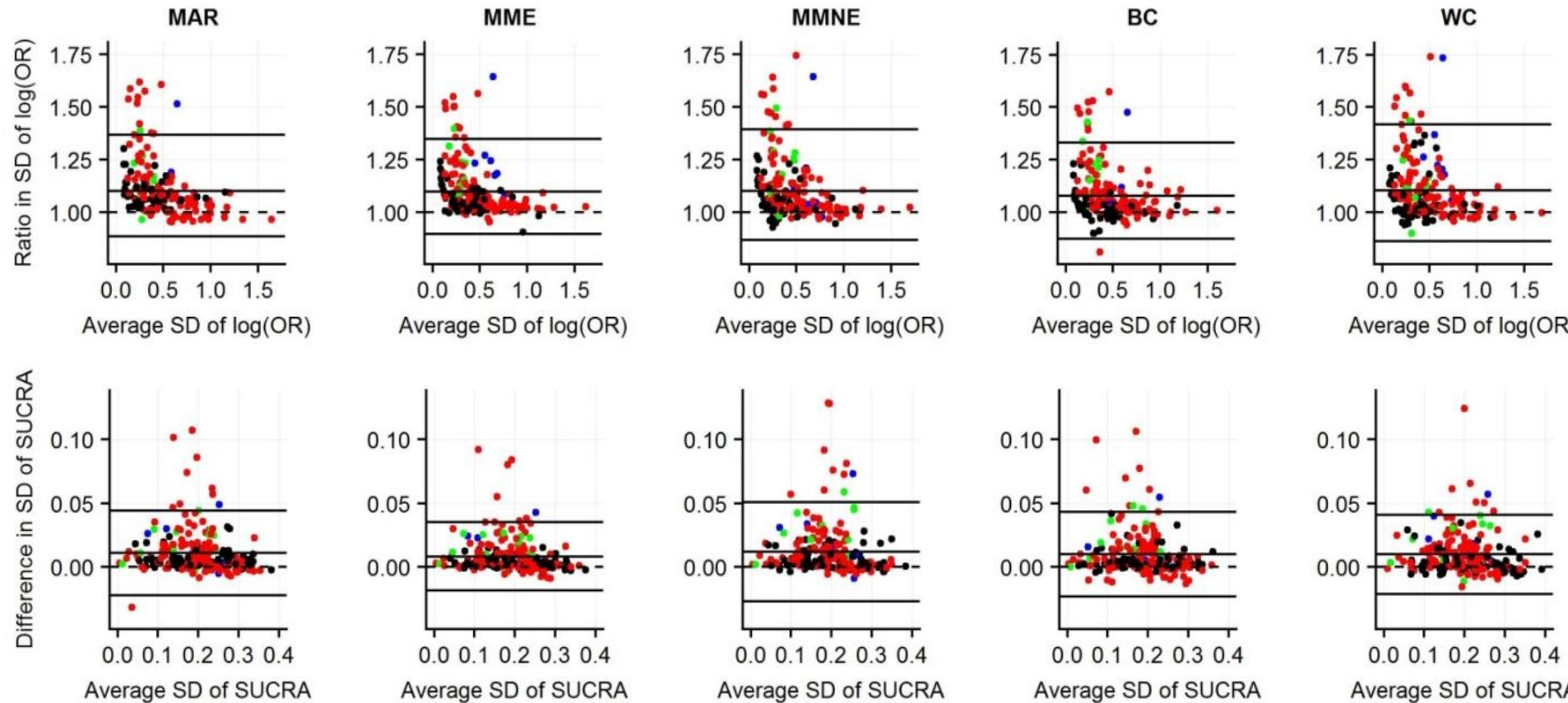
# Posterior mean/median

## *Accounting vs ignoring uncertainty due to MOD*



# Posterior standard deviation

## *Accounting vs ignoring uncertainty due to MOD*



• low • moderate-balance • moderate-imbalance • large-imbalance

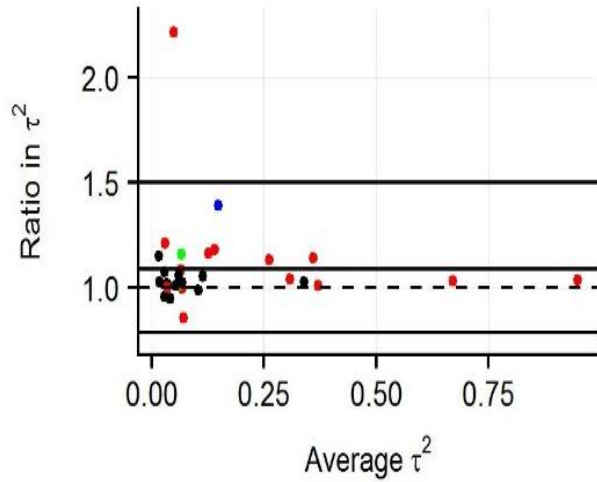
# Research questions (3)

## **Agreement between identical and hierarchical structure of log IMOR**

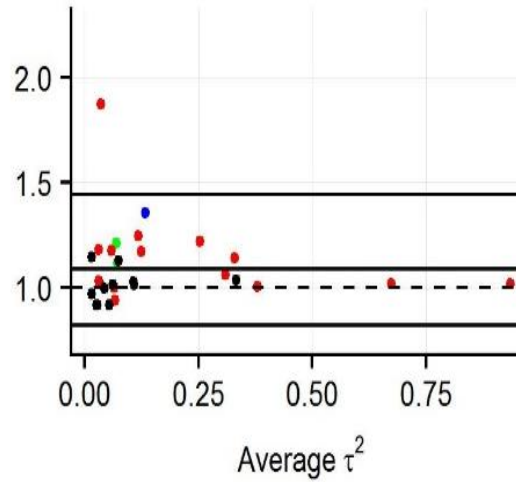
- MAR on average scenario
- In both structures use of normal prior distribution on log IMORs assumed:
  - ☐ Common-within network
  - ☐ Intervention-specific
  - ☐ Trial-specific

# Posterior median

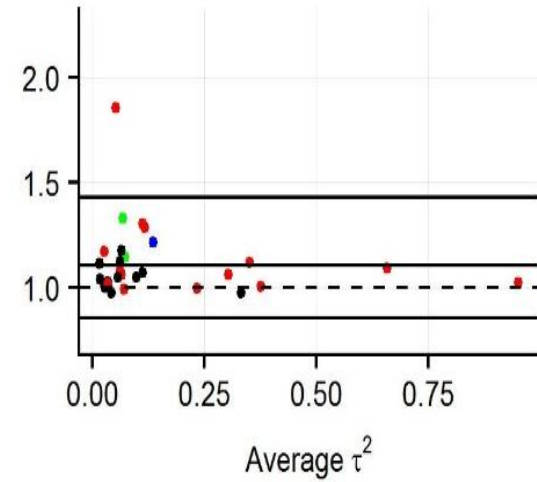
Common: Identical vs. Hierarchical



Trial: Identical vs. Hierarchical



Intervention: Identical vs. Hierarchical



• low • moderate-balance • moderate-imbalance • large-imbalance

# Take-home message

- ✓ Use **‘on average MAR’** for primary analysis and assumptions with clinical plausibility as sensitivity analyses.
- ✓ **Model the missingness mechanism via the IMOR parameter** in order to accommodate the uncertainty about the missingness scenarios considered.
- ✗ **Avoid imputing or excluding MOD either before analysis or by fixing the missingness parameter!**

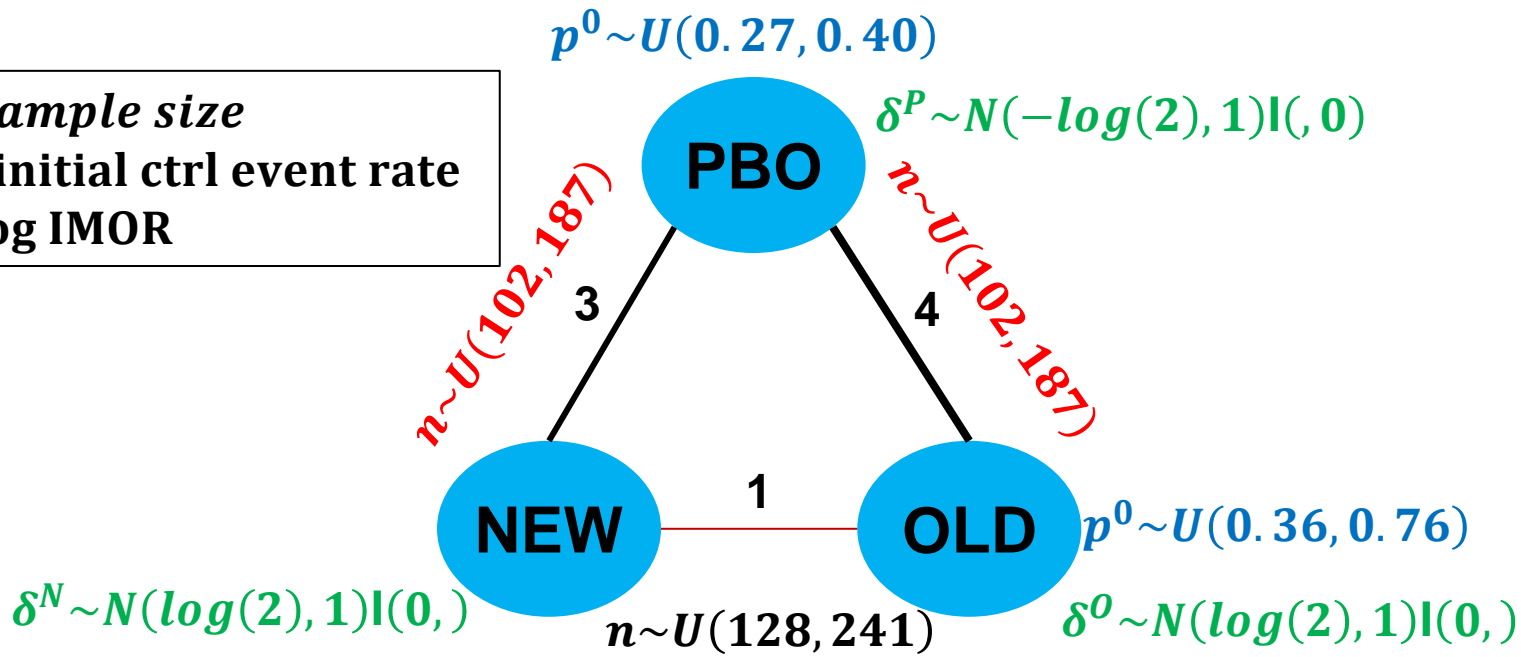
👉 Simulation study follows ...



**Which factors may affect the performance of different models for binary MOD in terms of core components of NMA?**

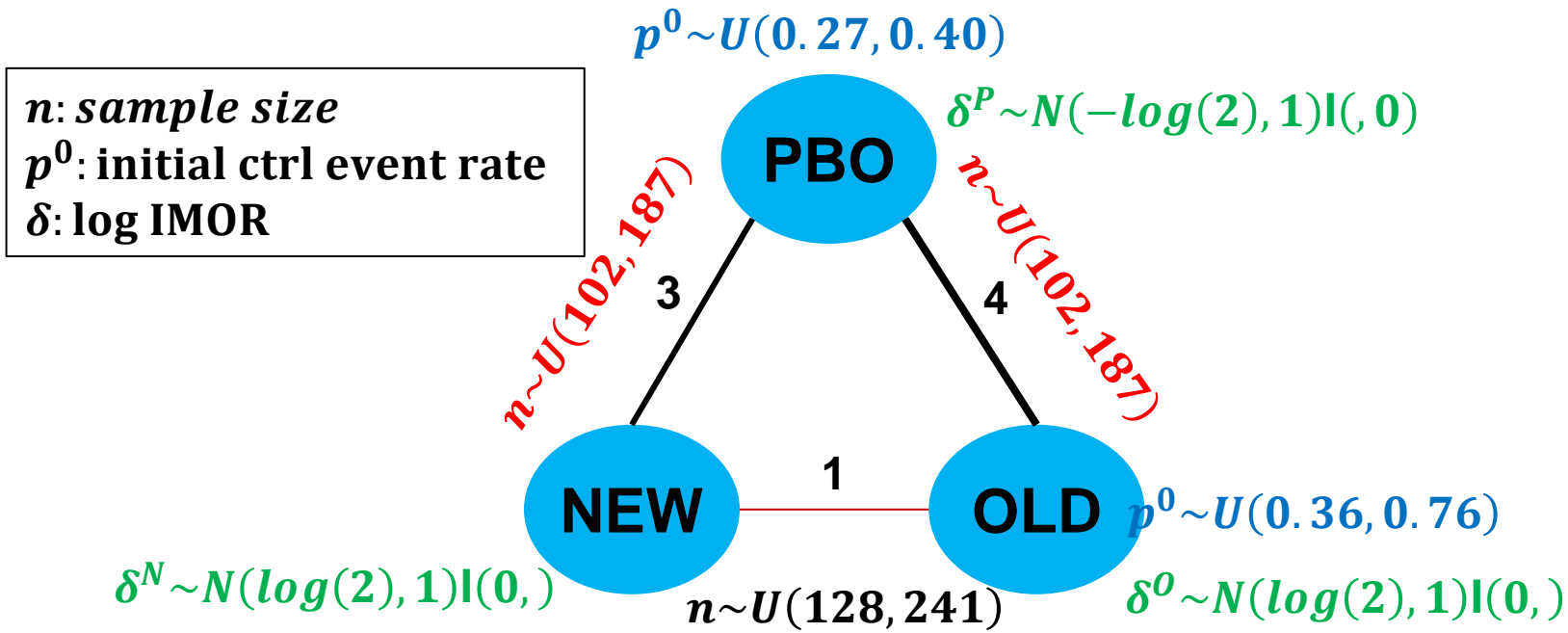
# Simulation set-up (1)

**$n$** : sample size  
 **$p^0$** : initial ctrl event rate  
 **$\delta$** : log IMOR



Comparison		Moderate %MOD		Large %MOD	
NEW	PBO	$U(0.05, 0.10)$	$U(0.11, 0.20)$	$U(0.21, 0.30)$	$U(0.31, 0.40)$
OLD	PBO				
NEW	OLD				

# Simulation set-up (2)



	small/low	substantial	$LOR_{NP} = \log(2)$ $LOR_{OP} = \log(1.5)$ $LOR_{NO} = LOR_{NP} - LOR_{OP} + IF$
$\tau^2$	all-cause mortality	generic health	
IF	$t(0, 0.44, df = 3)$	$t(1, 0.44, df = 3)$	

# Hartung & Knapp data generating model

1. Obtain **initial** event risks for the experimental arms,  $p^E$ :

$$p^{E,0} = \frac{p^{C,0} \cdot \exp(LOR)}{1 - p^{C,0} + p^{C,0} \cdot \exp(LOR)}$$

2. Obtain **initial** log odds for experimental and control arms, respectively:

$$\text{logit}^{E,0} = \log\left(\frac{p^{E,0}}{1-p^{E,0}}\right) \text{ and } \text{logit}^{C,0} = \log\left(\frac{p^{C,0}}{1-p^{C,0}}\right)$$

3. Generate **true** log odds for experimental and control arms, respectively:

$$\text{logit}^E \sim N\left(\text{logit}^{E,0}, \frac{2\tau^2}{3}\right) \text{ and } \text{logit}^C \sim N\left(\text{logit}^{C,0}, \frac{\tau^2}{3}\right)$$

4. Finally, obtain **true** event risks for experimental and control arms, respectively:

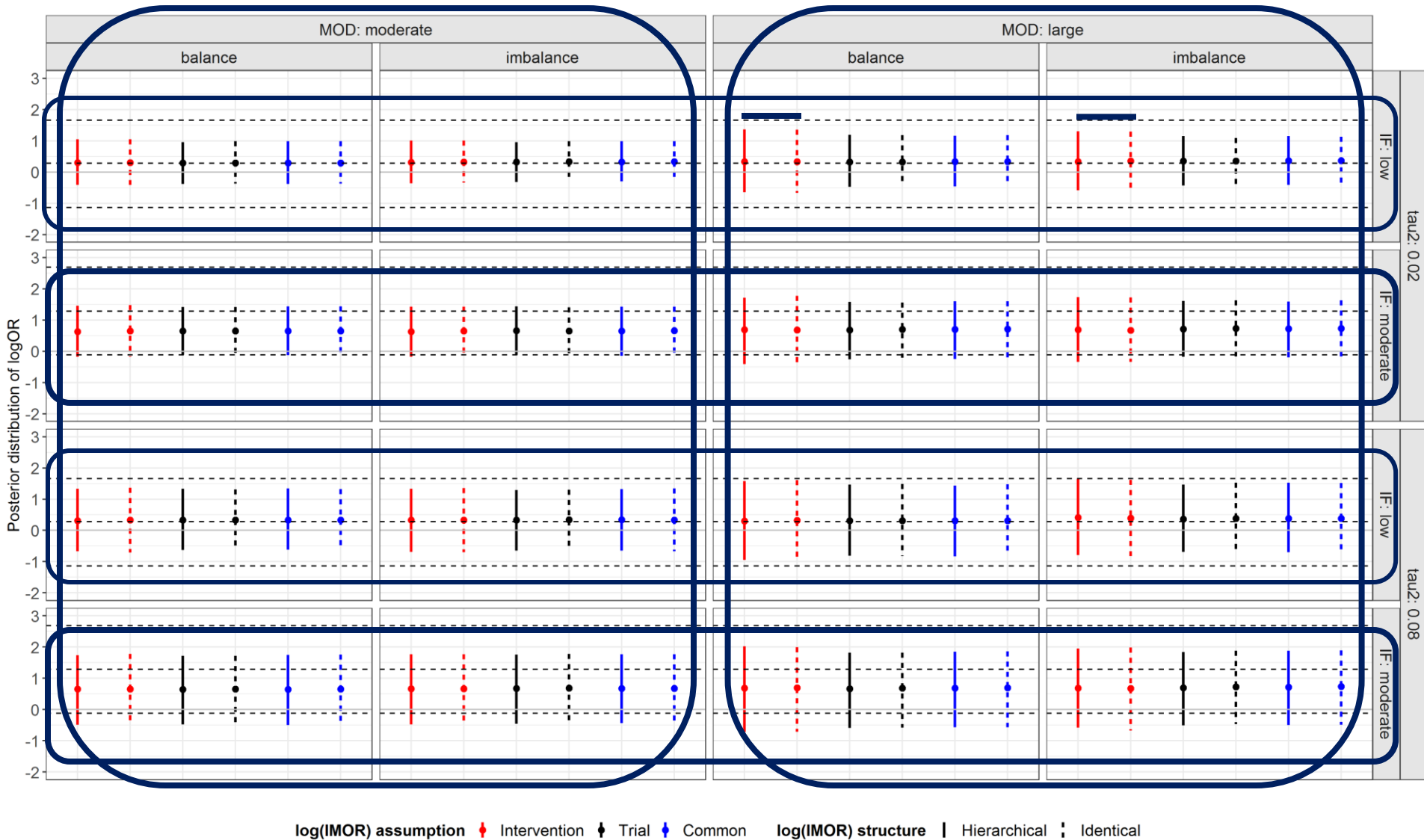
$$p^E = \frac{1}{1 + \exp(-\text{logit}^E)} \text{ and } p^C = \frac{1}{1 + \exp(-\text{logit}^C)}$$

# Model specification & Presentation of results

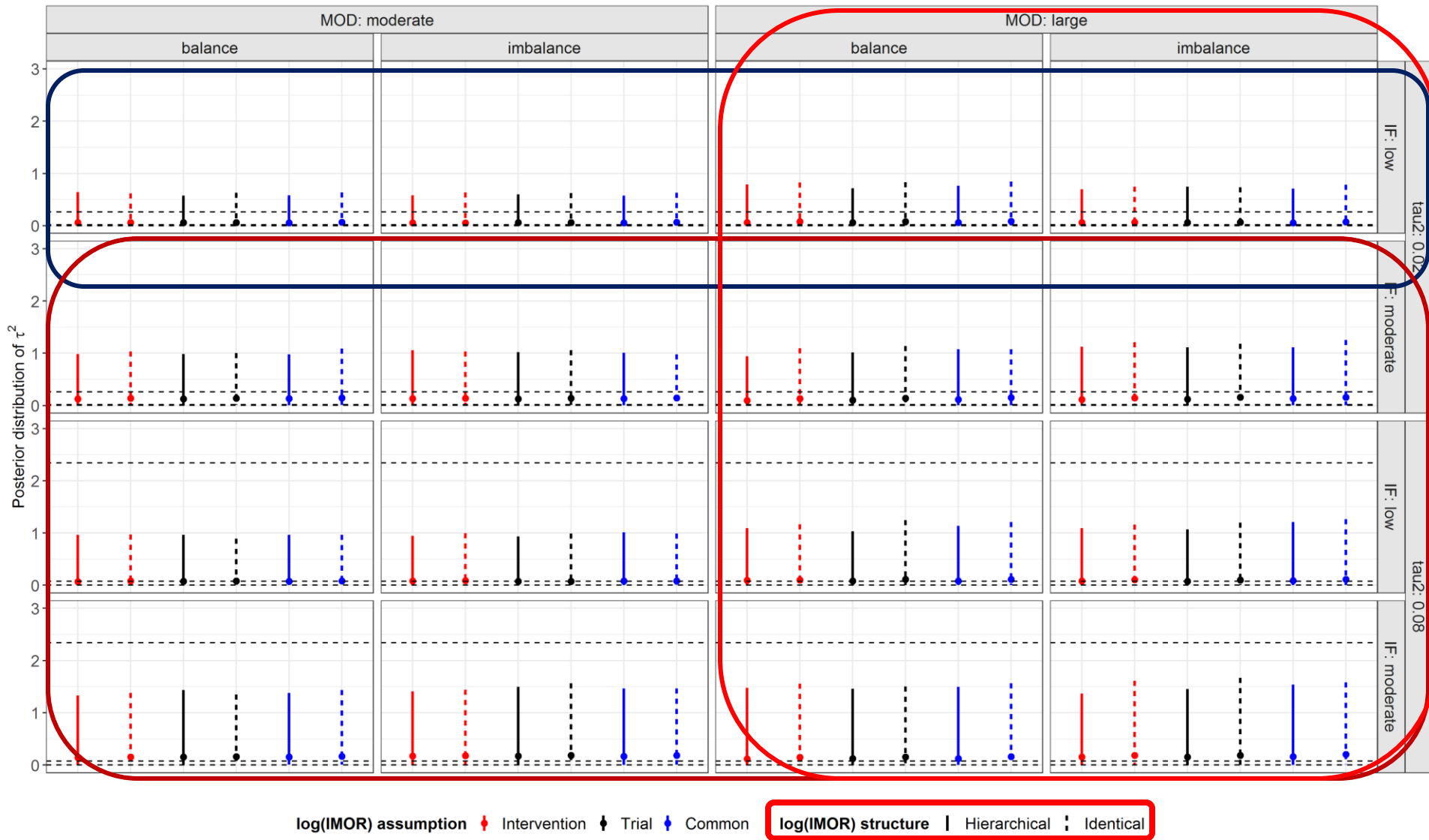
- **Bayesian random-effects NMA model** by Dias et al (2013);
- Application of the **node-splitting model** as developed by Dias et al (2010);
- Incorporation of log IMOR as described in Turner et al (2015);
- **Pattern-mixture model** for **identical** and **hierarchical logIMOR** (*common-within-network, trial-specific, and intervention-specific*).
- For **location parameters**, we used  $N(0, 10\ 000)$ ;
- For **between-trial variance**, we used  $N(0, 1)I(0, )$ ;
- For **logIMOR**, we used  $N(0,1)$  (on average MAR) in **all arms**;
- **1 000 triangles**, 3 chains for 20 000 updates and 2 000 burn-in;
- **Interval plots** for LOR (new vs old) and between-trial variance to correctly present results from Bayesian methods (prior distribution is presented with dotted parallel lines).



# Posterior distribution of LOR (new vs old)



# Posterior distribution of between-trial variance



# Limitations of the study

- ❑ We considered a simple network of **3 interventions and 2-arm trials**;
  - Simulating **a complex network with multi-arm trials\*** will shed more light on implications of different prior structures of log IMOR on NMA estimates;
- ❑ We investigated only **frequent events**;
  - Carpenter et al (2007): *'if an event is rare, missing data on very few patients can markedly alter estimated event rates'* → **affect NMA estimates.**
- ❑ **The degree of unbalanced MOD** was much smaller than the total extent of MOD in each trial → **We observed that much in the empirical study!**
  - **Larger imbalance of MOD** may have resulted in more imprecise log OR **under common-within-network and trial-specific prior structures.**

\*Seide SE, Jensen K, Kieser M. Simulation and data-generation for random-effects network meta-analysis of binary outcome. Stat Med. 2019 May 9. doi: 10.1002/sim.8193. [Epub ahead of print]

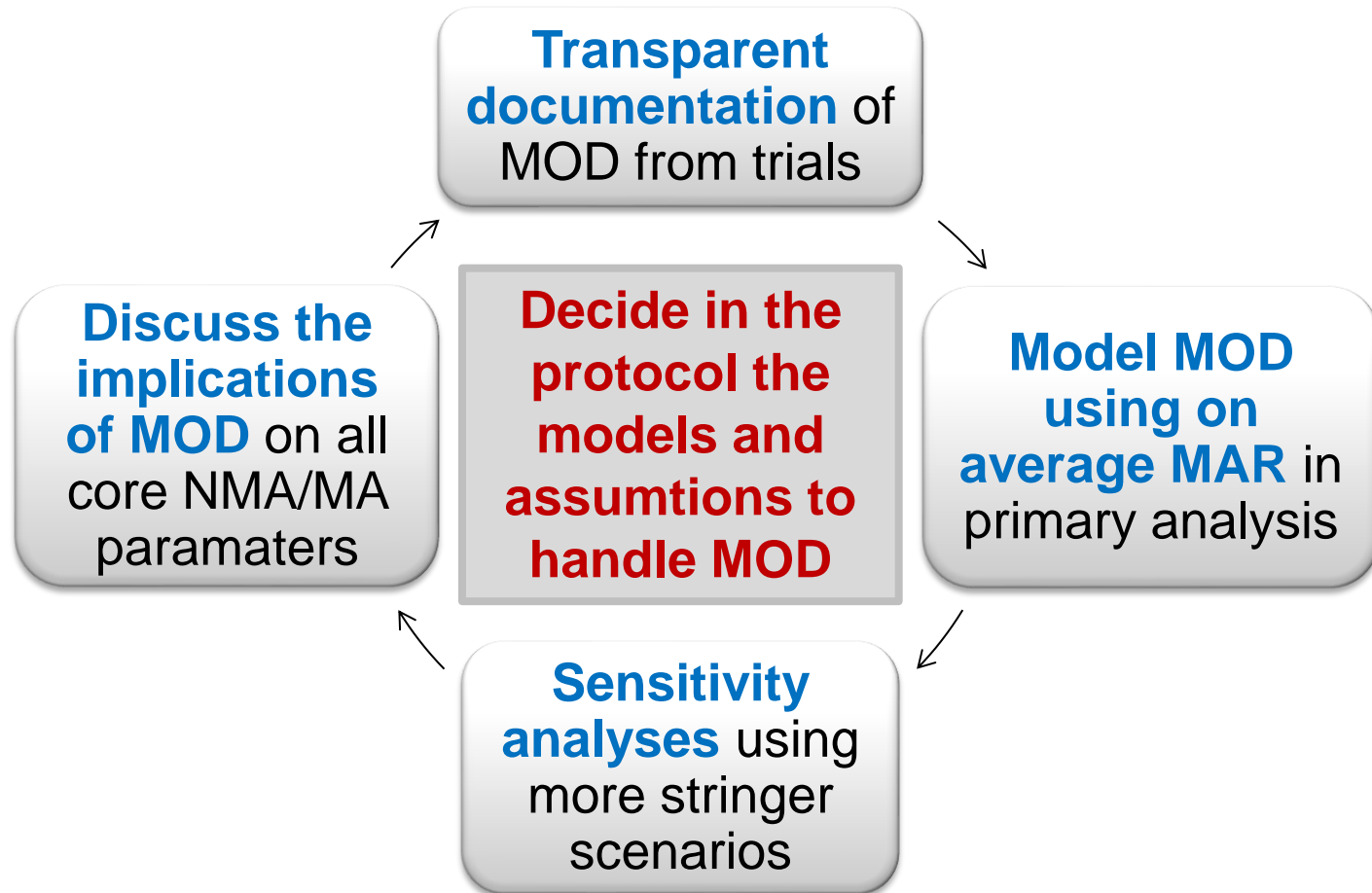


# Take-home message

- ❑ **Always decide at the protocol** on the proper prior structure of log IMOR that best aligns with the condition and interventions investigated;
- ❑ **Misspecification of the prior structure** may lead to spurious estimation of the uncertainty around log OR with implications for the conclusions;
- ❑ **Both identical and hierarchical structure may be considered** in the context of a sensitivity analysis;
- ❑ Though, we regard **hierarchical structure to be more plausible** in practice.
- ❑ **Results may be also generalized to conventional meta-analyses** with binary outcome.
- **Inferences are greatly restricted by the scenarios considered!!!**

# Future work in NEMO

- ❖ To assess and refine **methods to address continuous MOD in one-stage** (empirical and simulation study);
- ❖ Develop a **graphical tool for comprehensive sensitivity analyses** on plausible scenarios for MOD;
- ❖ **Develop an R-package** to ,accommodate‘ all strategies that have been investigated within the project;
- ❖ **Submit Habilitation** on ,NMA with MOD‘.



# Publications within NEMO

1. Spinelis LM, Kalyvas C, Pateras K. Participants' outcomes gone missing within a network of interventions: Bayesian modeling strategies. *Stat Med* 2019. DOI: 10.1002/sim.8207 [in press]
2. Spinelis LM. An empirical comparison of Bayesian modelling strategies for missing binary outcome data in network meta-analysis. *BMC Med Res Methodol* 2019;19(1):86.
3. Spinelis LM. Modeling missing binary outcome data while preserving transitivity assumption yielded more credible network meta-analysis results. *J Clin Epidemiol* 2019;105:19-26.
4. Spinelis LM, Yepes-Nuñez JJ, Schünemann H. A systematic survey shows that reporting and handling of missing outcome data in networks of interventions is poor. *BMC Med Res Methodol* 2018;18(1):115.
5. Spinelis LM. Missing binary data extraction challenges from Cochrane reviews in mental health and Campbell reviews with implications for empirical research. *Res Synth Methods* 2017;8(4):514-525.

## References (1)

- Akl EA, Carrasco-Labra A, Brignardello-Petersen R, et al. Reporting, handling and assessing the risk of bias associated with missing participant data in systematic reviews: a methodological survey. *BMJ Open* 2015;5:e009368.
- Bafeta A, Trinquart L, Seror R, Ravaud P. Analysis of the systematic reviews process in reports of network meta-analyses: methodological systematic review. *BMJ* 2013;347:f3675.
- Carpenter JR, Kenward MG. Missing data in randomised controlled trials—a practical guide. Health Technology Assessment Methodology Programme. Birmingham; 2007.
- Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med Decis Making* 2013;33:607–617.
- Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010;29:932-44.
- Nikolakopoulou A, Chaimani A, Veroniki AA, et al. Characteristics of networks of interventions: A description of a database of 186 published networks. *PLoS One* 2014;9:1–10.
- Petropoulou M, Nikolakopoulou A, Veroniki AA, et al. Bibliographic study showed improving statistical methodology of network meta-analyses published between 1999 and 2015. *J Clin Epidemiol* 2017;82:20–28.

## References (2)

- Sackett DL, Richardson WS, Rosenberg W, Haynes RB. Evidence-based medicine: how to practice and teach EBM. New York: Churchill Livingstone, 1997
- Spineli LM, Pandis N, Salanti G. Reporting and handling missing outcome data in mental health: A systematic review of cochrane systematic reviews and meta-analyses. Res Synth Methods 2015;6:175–87.
- Tan SH, Bujkiewicz S, Sutton A, et al. Presentational approaches used in the UK for reporting evidence synthesis using indirect and mixed treatment comparisons. J Health Serv Res Policy 2013;18:224–32.
- Turner NL, Dias S, Ades AE, Welton NJ. A Bayesian framework to account for uncertainty due to missing binary outcome data in pairwise meta-analysis. Stat Med 2015;34:2062-80.
- van Valkenhoef G, Dias S, Ades AE, Welton NJ. Automated generation of node-splitting models for assessment of inconsistency in network meta-analysis. Res Synth Methods 2016;7:80-93.
- White IR, Higgins JP, Wood AM. Allowing for uncertainty due to missing data in meta-analysis--part 1: two-stage methods. Stat Med 2008;27:711-727
- Zarin W, Veroniki AA, Nincic V, et al. Characteristics and knowledge synthesis approach for 456 network meta-analyses: a scoping review. BMC Med 2017;15:3.

**Thank you for your attention!**