Quantifying the robustness of primary analysis results: A case study on missing outcome data in pairwise and network meta-analysis§

Loukia M Spineli[1], Chrysostomos Kalyvas[2], Katerina Papadimitropoulou[3]

[1]Midwifery Research and Education Unit, Hannover Medical School, Hannover, Germany
[2]Biostatistics and Research Decision Sciences, MSD Europe Inc., Brussels, Belgium
[3]Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands

Gefördert durch DFG Deutsche Forschungsgemeinschaft

Hannover Medical School

§Published in Research Synthesis Methods (doi: 10.1002/jrsm.1478)

Sensitivity Analysis (SA)

- Integral part of the systematic reviews.
- Explore the sensitivity of primary analysis results to different re-analyses of the same dataset or part thereof.
- Popular in the analysis of missing participant outcome data (MOD)
  - Do different reasons for MOD affect the primary analysis results?

Patsopoulos et al. (2008); National Research Council (2010)

SA to address MOD in meta-analyses

- Most systematic reviews addressing MOD **do not** perform SA.
- Those few systematic reviews conducting SA:
  - consider a few clinically implausible scenarios for MOD;
  - do not provide a definition of similar results;
  - emphasise changes in the statistical significance (for significance level 5%) of the summary results.
- Conclusions are sensitive to the selected significance level!

Spineli et al. (2015); Kahale et al. (2018); Spineli et al. (2018)

Proposed framework for SA

- **Progressively stringent** yet clinically relevant scenarios for the MOD mechanism.
- **Objective measure of 'divergence'** of each re-analysis from the primary analysis results.
- **Objective definition of** a total **minimally accepted 'divergence'** between the primary analysis and all re-analysis results for a comparison.
- **Comprehensive illustration** of the summary **results under all different scenarios**.

Informative Missingness Parameter

- Measure the departure from the MAR assumption via a pattern-mixture model:

Informative Missingness Odds Ratio (IMOR)

$$\frac{\text{odds of event given MOD}}{\text{odds of event given completers}}$$

Informative Missingness Difference of Means (IMDoM)

Mean outcome given MOD – Mean outcome given completers

MAR assumption: IMOR = 1 and IMDoM = 0

White et al. (2008); Mavridis et al. (2015)

Clinically plausible scenarios (1)

Defined for an intervention arm

Progressively stringent scenarios

log IMOR: $-ln(3), -ln(2),$ **0**, $ln(2)$, and $ln(3)$

IMDoM: $-2, -1,$ **0**, 1, and 2

MAR assumption

Mavridis et al. (2015); Turner et al. (2015); Spineli (2019a)

Clinically plausible scenarios (2)

5 assumed values → **5^2** possible pairs of values

… for a pairwise meta-analysis or a star-shaped network

Scenario	log IMOR		IMDoM	
	Active	Control	Active	Control
1	$-ln(3)$	$-ln(3)$	-2	-2
2	$-ln(3)$	$-ln(2)$	-2	-1
…	…	…	…	…
13 (MAR)	$\boldsymbol{ln(1)}$	$\boldsymbol{ln(1)}$	**0**	**0**
…	…	…	…	…
24	$ln(3)$	$ln(2)$	2	1
25	$ln(3)$	$ln(3)$	2	2

Clinically plausible scenarios (3)

5 assumed values → **5^2** possible pairs of values

… for a complex network of four interventions

log IMOR (similarly for IMDoM)

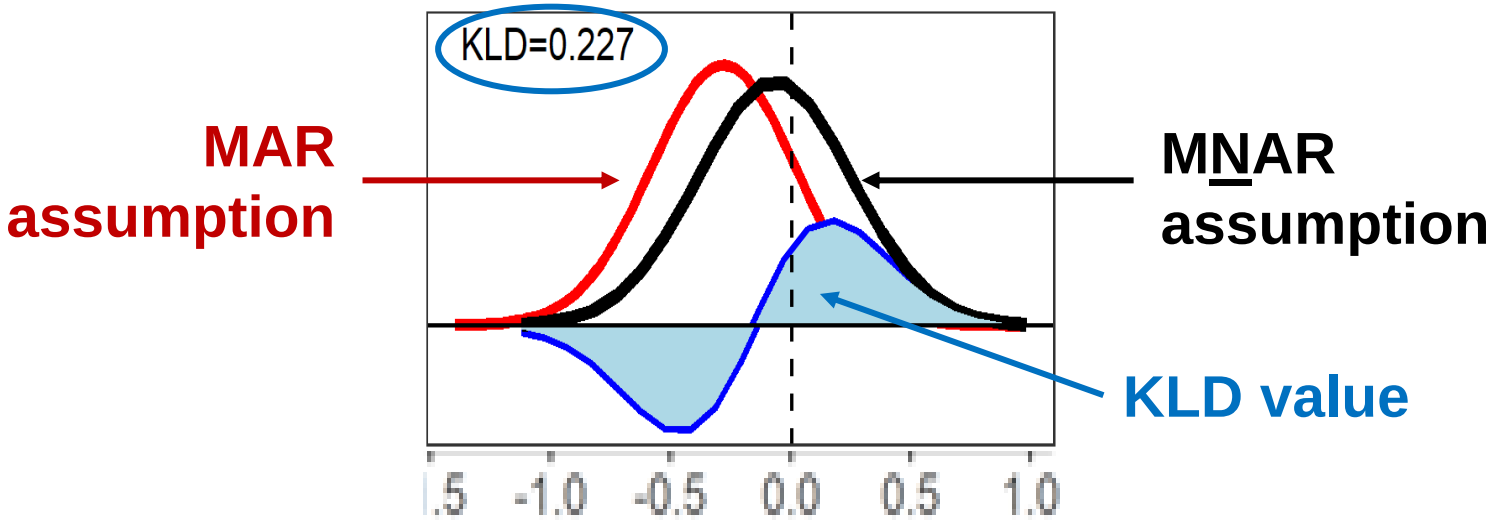
Scenario	A	B	C	D (ref.)
1	$-ln(3)$	$-ln(3)$	$-ln(3)$	$-ln(3)$
2	$-ln(3)$	$-ln(3)$	$-ln(3)$	$-ln(2)$
…	…	…	…	…
13 (MAR)	$\boldsymbol{ln(1)}$	$\boldsymbol{ln(1)}$	$\boldsymbol{ln(1)}$	$\boldsymbol{ln(1)}$
…	…	…	…	…
24	$ln(3)$	$ln(3)$	$ln(3)$	$ln(2)$
25	$ln(3)$	$ln(3)$	$ln(3)$	$ln(3)$

Spineli (2019b)

The Kullback–Leibler divergence (KLD)

How much two probability distributions differ from each other?

(e.g., normal posterior distribution of log OR under MAR and MNAR assumptions)



Kullback & Leibler (1951)

The robustness index (RI)

- **24 informative scenarios**

versus

MAR assumption

⇨ **24 KLD values**

➢ What is the ***total divergence*** from the primary analysis (MAR)?

$$RI = \sqrt{\sum_{i \in S} {D_{13,i}}^2}$$

$$S = \{1, 2, \dots, 12, 14, \dots, 25\}$$

- Pairwise meta-analysis ⇨ **One** RI
- NMA with T interventions ⇨ $\binom{T}{2}$ RIs

The threshold of robustness (1)

- No universally accepted definition of similar results.

Proposed intuitive rule

Low divergence → **RI < 1** SD of low heterogeneity

Substantial divergence → **RI ≥ 1** SD of low heterogeneity

Low statistical heterogeneity*

Scale	Prior	Median	95% interval
log OR	log-normal	0.08	0.003 – 2.18
SMD	log-t (5 df)	0.03	0.0002 – 5.16

*Empirically-based prior for τ^2 in a general health-care setting

Turner et al. (2015); Rhodes et al. (2015)

The threshold of robustness (2)

Decision rule in NMA

Robustness in the network

low divergence in **all** possible comparisons

$(RI < \sqrt{median})$

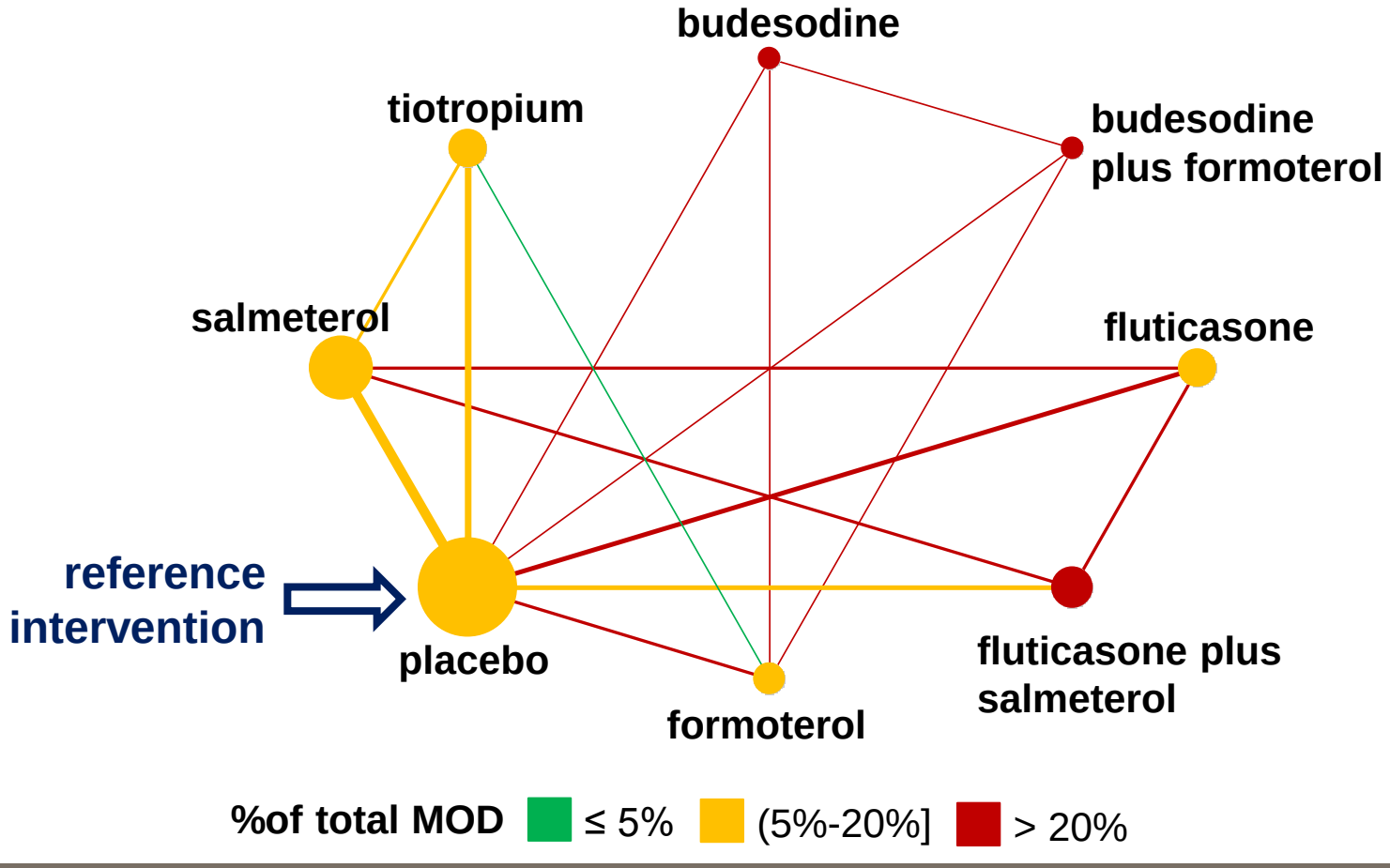
Lack of robustness in the network

substantial divergence in **at least one** possible comparison

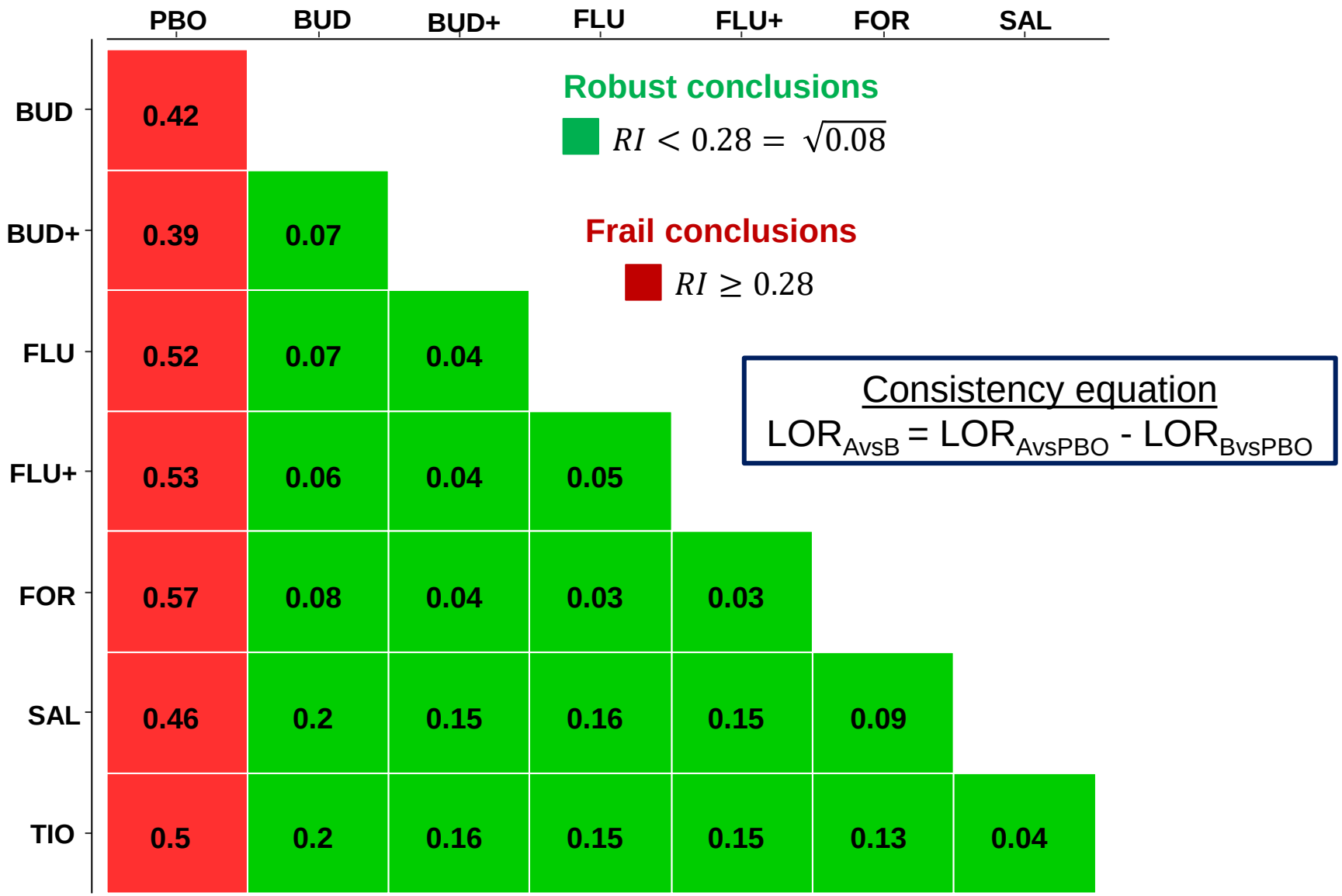
$(RI \geq \sqrt{median})$

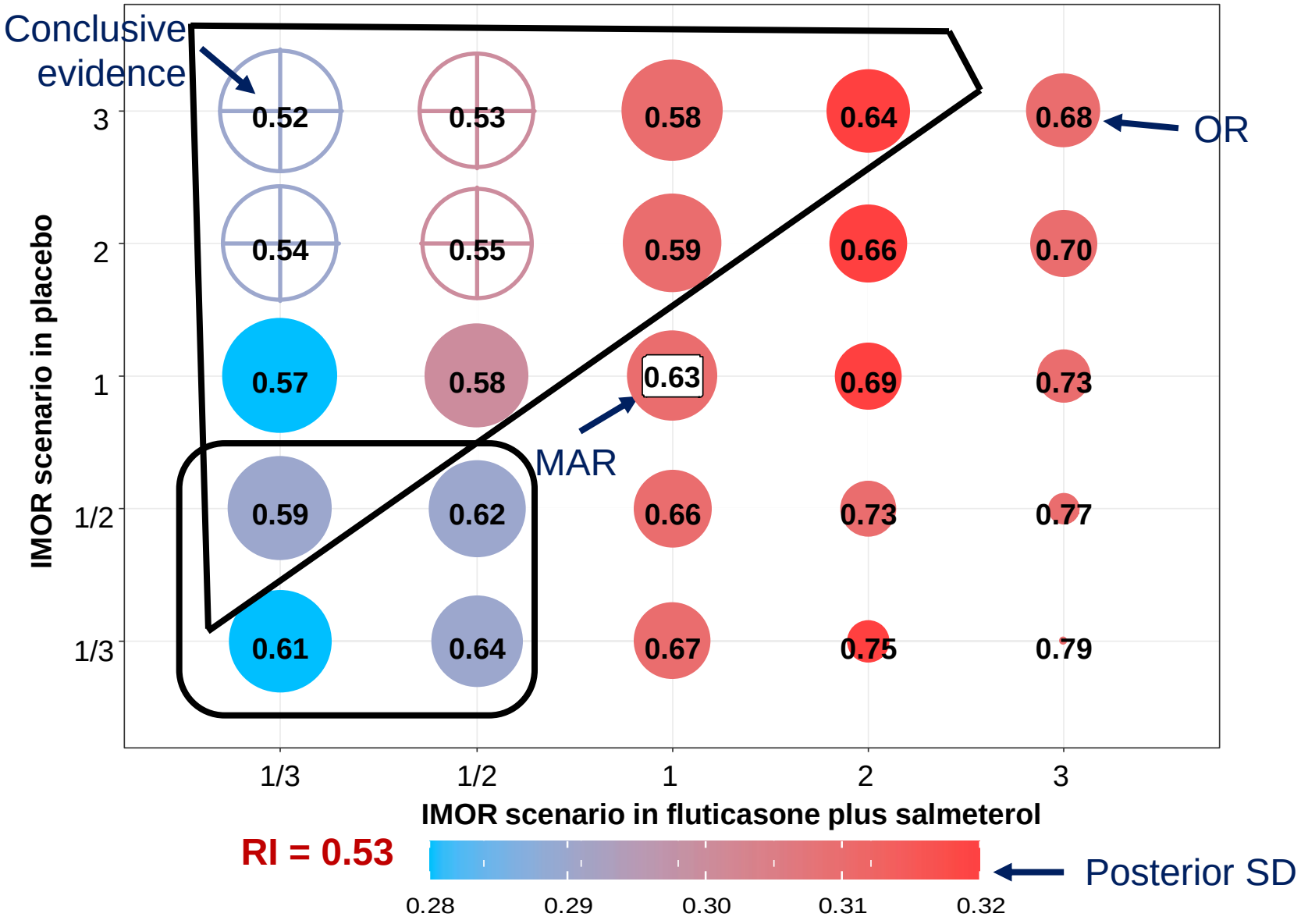
Illustrative example (1)

- Exacerbation (yes/no) of chronic obstructive pulmonary disease



Baker et al. (2009)


PBO
BUD
BUD+
FLU
FLU+
FOR
SAL
BUD
BUD+
FLU
FLU+
FOR
SAL
TIO
0.42
0.39
0.07
0.52
0.07
0.04
0.53
0.06
0.04
0.05
0.57
0.08
0.04
0.03
0.03
0.46
0.2
0.15
0.16
0.15
0.09
0.5
0.2
0.16
0.15
0.15
0.13
0.04
Robust conclusions
$RI < 0.28 = \sqrt{0.08}$
Frail conclusions
$RI \geq 0.28$
Consistency equation
$LOR_{AvsB} = LOR_{AvsPBO} - LOR_{BvsPBO}$


Conclusive evidence
OR
MAR
IMOR scenario in placebo
IMOR scenario in fluticasone plus salmeterol
3
2
1
1/2
1/3
0.52
0.53
0.58
0.64
0.68
0.54
0.55
0.59
0.66
0.70
0.57
0.58
0.63
0.69
0.73
0.59
0.62
0.66
0.73
0.77
0.61
0.64
0.67
0.75
0.79
RI = 0.53
Posterior SD
0.28
0.29
0.30
0.31
0.32

Discussion points

- The RI is useful to broader set of sensitivity analyses.
  - Comparing different priors for the heterogeneity parameter.
- The researchers can increase or decrease the RI threshold.
  - Decrease for HTAs; increase for ‘exploratory’ reviews.
- The RI thresholds are not tailored to the research question.
  - Expert opinion to define similarity in treatment effects.
- The scenarios are based on relevant empirical studies.
  - Ideally, define the scenarios with an expert assistance.

References

1. Patsopoulos NA, Evangelou E, Ioannidis JP. Sensitivity of between-study heterogeneity in meta-analysis: proposed metrics and empirical evaluation. Int J Epidemiol. 2008;37(5):1148-1157.
2. National Research Council. The Prevention and Treatment of Missing Data in Clinical Trials Panel on Handling Missing Data in Clinical Trials. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press; 2010. Retrieved from www.nap.edu.
3. Kahale LA, Diab B, Brignardello-Petersen R, et al. Systematic reviews do not adequately report or address missing outcome data in their analyses: a methodological survey. J Clin Epidemiol. 2018;99:14-23.
4. 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(2):175-187.
5. Spineli LM, Yepes-Nuñez JJ, Schünemann HJ. 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.
6. 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(5):711-727.
7. 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(12):2062-2080.
8. Mavridis D, White IR, Higgins JP, Cipriani A, Salanti G. Allowing for uncertainty due to missing continuous outcome data in pairwise and network meta-analysis. Stat Med. 2015;34(5): 721-741.
9. Spineli LM. An empirical comparison of Bayesian modelling strategies for missing binary outcome data in network meta-analysis. BMC Med Res Methodol. 2019a;19(1):86.
10. Spineli LM. Modeling missing binary outcome data while preserving transitivity assumption yielded more credible network meta-analysis results. J Clin Epidemiol. 2019b;105:19-26.
11. Kullback S, Leibler RA. On information and sufficiency. Ann Math Stat. 1951;22(1):79-86.
12. Turner RM, Jackson D, Wei Y, Thompson SG, Higgins JPT. Predictive distributions for between-study heterogeneity and simple methods for their application in Bayesian meta-analysis. Stat Med. 2015;34(6):984-998.
13. Rhodes KM, Turner RM, Higgins JP. Predictive distributions were developed for the extent of heterogeneity in meta-analyses of continuous outcome data. J Clin Epidemiol. 2015; 68(1):52-60.
14. Baker WL, Baker EL, Coleman CI. Pharmacologic treatments for chronic obstructive pulmonary disease: a mixed-treatment comparison meta-analysis. Pharmacotherapy. 2009;29(8):891-905.

Brussels

Leiden

Thank you for your attention!

Hannover