Forschungs- und Lehreinheit Hebammenwissenschaft
Medizinische Hochschule Hannover

Funded by DFG Deutsche Forschungsgemeinschaft German Research Foundation

MHH Medizinische Hochschule Hannover

Local inconsistency detection using the Kullback–Leibler divergence measure*

Loukia M Spineli[1]
[1]Forschungs- und Lehreinheit Hebammenwissenschaft Medizinische Hochschule Hannover

***From the published article:**
Spineli LM. Local inconsistency detection using the Kullback-Leibler divergence measure. Syst Rev. 2024;13(1):261.
doi: 10.1186/s13643-024-02680-4.

Corresponding Author: PD Dr Phil Loukia Spineli
Forschungs- und Lehreinheit Hebammenwissenschaft
Medizinische Hochschule Hannover
Carl-Neuberg-Straße 1, 30625 Hannover
spineli.Loukia@mh-hannover.de
https://www.mhh.de/hebammenwissenschaft

Background

Consistency in the different evidence sources constitutes the cornerstone of network meta-analysis (NMA). Several methods have been proposed to evaluate consistency locally and globally. Local consistency evaluation has received relatively more methodological attention for being intuitively more appealing and long-established since the NMA introduction. Local evaluation targets closed loops of evidence in the network, where there is direct and indirect evidence for the involved treatment comparisons. The difference between these two evidence sources, termed inconsistency, is usually tested using a two-sided Z-test. However, testing for inconsistency has been criticised for having low statistical power and misinterpreting a statistically non-significant inconsistency as evidence of consistency.

Methods

Interpretation index of inconsistency

We drew inspiration from the Kullback-Leibler divergence (KLD) measure [1] to propose a novel, straightforward ***interpretation index*** for ***local inconsistency evaluation***. The proposed index, D, considers the whole distribution of the estimated direct and indirect effects to quantify the *average* information loss when indirect effect replaces the direct effect, and vice versa, for a comparison in the closed loop. Minimum information loss coincides with an index close to 0, which would imply low inconsistency that may not threaten the validity of NMA results.

- KLD approximating direct effect: $D_{D,I} = \frac{1}{2}\left[\left(\frac{\hat{s}_D}{\hat{s}_I}\right)^2 + \frac{(\hat{\mu}_D - \hat{\mu}_I)^2}{\hat{s}_I^2} - 1 + ln\left(\frac{\hat{s}_I^2}{\hat{s}_D^2}\right)\right]$
- KLD approximating indirect effect: $D_{I,D} = \frac{1}{2}\left[\left(\frac{\hat{s}_I}{\hat{s}_D}\right)^2 + \frac{(\hat{\mu}_D - \hat{\mu}_I)^2}{\hat{s}_D^2} - 1 + ln\left(\frac{\hat{s}_D^2}{\hat{s}_I^2}\right)\right]$

$$\boldsymbol{D = \frac{D_{D,I} + D_{I,D}}{2}}$$

$\hat{\mu}_D$, direct estimate; $\hat{\mu}_I$, indirect estimate; $\hat{s}_D$, standard error of direct estimate; $\hat{s}_I$, standard error of indirect estimate

Threshold of acceptably low inconsistency

Based on the *opinion elicitation framework* of Spiegelhalter et al. [2], we assume that the direct and indirect effects of a comparison follow the same normal distribution but with different variances:

$$\theta_D \sim N(\mu, \tau^2) \text{ and } \theta_I \sim N(\mu, 2\tau^2)$$

Then $\theta_D - \theta_I$ would follow a normal distribution with variance $3\tau^2$. Subsequently, $|\theta_D - \theta_I|$ follows a half-normal distribution with scale parameter $\sqrt{3}\tau$ and median $\Phi^{-1}(0.75) \times \sqrt{3}\tau \cong 1.17\tau$.

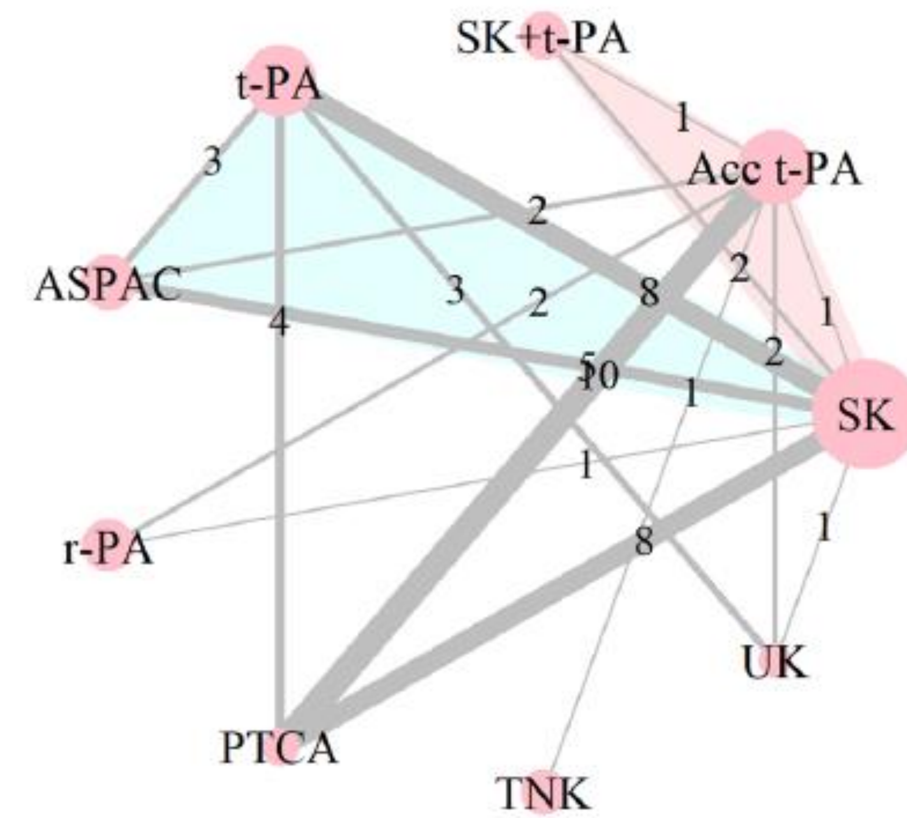
Replacing $\hat{\mu}_D - \hat{\mu}_I = 1.17\tau$, $\hat{s}_D = \tau$ and $\hat{s}_I = \sqrt{2}\tau$ in $D_{D,I}$ and $D_{I,D}$, we obtain $D = 0.64$.

$\boldsymbol{D < 0.64}$ for a comparison implies **acceptably low inconsistency**; otherwise, inconsistency may be material

Conclusions

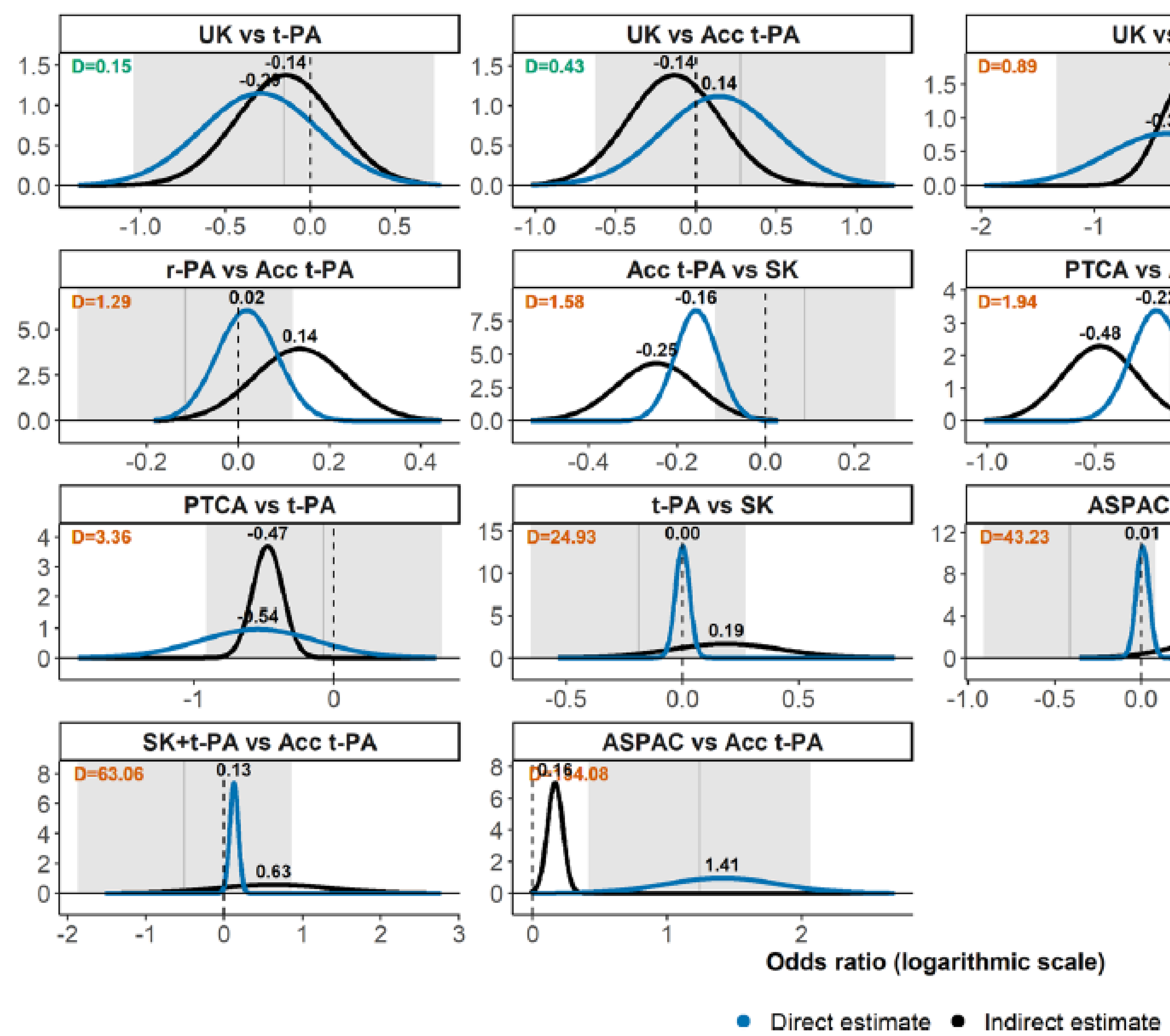
The interpretation index revealed material inconsistency ($D \geq 0.64$) for most split nodes, with some nodes being associated with very large D values as the distributions of direct and corresponding indirect effects hardly overlapped. To understand whether inconsistency may govern a network requires a thorough evaluation of the collated evidence, including model fit and outlier detection in addition to statistical testing (the current status quo). The proposed interpretation index is a valuable addition to the inconsistency evaluation toolkit and can uncover nodes with material inconsistency when statistical tests are inconclusive.

Results



Originally, the proposed interpretation index was applied to three networks with mutliple treatments. Here, we present the ***thrombolytics network*** [3,4], consisting of 48 two-arm studies and 2 three-arm studies comparing 8 thrombolytic treatments and angioplasty for acute myocardial infarction: streptokinase (SK), alteplase (t-PA), accelerated alteplase (Acc t-PA), streptokinase plus alteplase (SK + t-PA), reteplase (r-PA), tenecteplase (TNK), percutaneous transluminal coronary angioplasty (PTCA), urokinase (UK), and anistreplase (ASPAC).

The ***Bayesian node-splitting approach*** with automated generation of split nodes [5] was applied. The ***rnmamod R package*** was employed to run the node-splitting approach and calculate the interpretation index for each split node.



Posterior densities for 14 split nodes. The grey area and vertical line indicate the inconsistency's 95% interval (approximated using the reported posterior mean and standard deviation) and posterior mean. Green and orange D values (top left of each plot) indicate acceptably low and material inconsistency. The threshold of 0.64 was employed.

References

1. Kullback S, Leibler RA. On information and sufficiency. Ann Math Stat. 1951;22(1):79–86.
2. Spiegelhalter DJ, Abrams KR, Myles J. Hierarchical priors. In: Bayesian approaches to clinical trials and health-care evaluation. Chichester: John Wiley & Sons; 2004. p. 167–74.
3. Boland A, Dundar Y, Bagust A, Haycox A, Hill R, Mujica Mota R, et al. Early thrombolysis for the treatment of acute myocardial infarction: a systematic review and economic evaluation. Health Technol Assess. 2003;7(15):1–136.
4. Keeley EC, Boura JA, Grines CL. Primary angioplasty versus intravenous thrombolytic therapy for acute myocardial infarction: a quantitative review of 23 randomised trials. Lancet. 2003;361(9351):13–20.
5. 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(1):80–93.