WEB MATERIAL

Introducing the Treatment Hierarchy Question in Network Meta-Analysis

Georgia Salanti, Adriani Nikolakopoulou, Orestis Efthimiou, Dimitris Mavridis, Matthias Egger, and Ian R. White

Contents

WEB APPENDIX 1

Framing the Treatment Hierarchy Question Using Absolute Estimands and Relative Treatment Effects

Treatment hierarchy questions may be expressed using *absolute estimands* (like means and probabilities) or *relative treatment effects* (like mean differences or risk ratios). A series of independent posterior distributions estimates the absolute estimands, while a joint posterior distribution estimates the relative treatment effects. In either case, ranking metrics are statistics that summarize these distributions to answer a treatment hierarchy question.

Any question that can be expressed using relative treatment effects can also be expressed using absolute estimands, but the opposite is not always possible. For example, treatment hierarchy question 1 can be expressed as "Which treatment has the largest reduction in estimated mean post-treatment LDL-C compared to treatment A?". This is the same question, since the treatment with the smallest value of M_i must also have the smallest value of D_{iA} (where we interpret $D_{AA} = 0$). However, treatment hierarchy question 2 cannot be expressed in terms of relative treatment effects δ_{ij}.

Treatment hierarchy questions that cannot be expressed in terms of relative treatment effects can result in treatment hierarchies that might seem inappropriate or counterintuitive. For example, in Figure 1, it might appear strange to prefer treatment C to treatment B based on its lower probability of having a mean LDL-C above 2.5 mmol/L (*treatment hierarchy question 2*), when it also has a larger estimated mean LDL-C (*treatment hierarchy question 1*). We can debate whether treatment hierarchy question 2 is of relevance to a particular decision-making context or unsuitable for producing treatment guidelines. But we cannot argue that the obtained hierarchy C, B, A is wrong, because it correctly answers a valid treatment hierarchy question.

In two-arm randomized trials and pairwise meta-analysis, conclusions are in practice based on the (single) relative treatment effect δ_{ij}. Because the uncertainty in the estimation of δ_{ij} is usually distributed equally around the centre of the distribution D_{ij} (a symmetric distribution), all treatment hierarchy questions usually give the same answer. For example, if $D_{ij} < 0$ indicating that treatment i has lower estimated LDL-C than treatment j, then also the probability that treatment i beats treatment j is greater than ½ and hence greater than the probability that treatment j beats treatment i.

WEB APPENDIX 2

Sources of Imprecision in the Estimation of Treatment Effects in Network Meta-Analysis

Most NMA models provide estimates of relative treatment effects δ_{AP}, δ_{BP}, δ_{CP} for A, B, C versus a reference treatment, say P. The absolute estimands μ_A, μ_B, μ_C and μ_P can be directly estimated from the model or obtained by combining the δ's with an estimate for μ_P. Each relative treatment effect or absolute estimand is estimated with uncertainty, which depends on the following three factors:

1. The amount of direct and indirect information for treatment and treatment comparison. Little information for an absolute estimand is available if the treatment features in few studies, if its studies have large sampling error (small sample size, few events or large standard deviations), if it is compared with few other interventions or is part of few closed loops of evidence. Similarly, the relative treatment effect is uncertain if few studies (or small studies) examine the comparison of interest and if the comparison is part of few closed loops.

2. The heterogeneity in relative treatment effects. If the relative treatment effects δ are heterogeneous across studies examining the same comparison, the uncertainty in their estimation will be larger. As absolute estimands are estimated from relative treatment effects, heterogeneity in δ will result in more uncertainty in the estimation of μ as well.

3. Residual incoherence. Evidence of large incoherence (disagreement between direct and indirect evidence) should prevent researchers from synthesizing the data. Even if a treatment is involved in evidence loops with small or moderate amounts of incoherence, the credibility of the estimated summary effect decreases. However, this produces estimates with less precision and larger credible intervals only when residual incoherence is explicitly modelled within the NMA.

The first two situations outlined above or their combination always result in increased uncertainty about δ and μ and this uncertainty plays a major role when estimating treatment hierarchy using ranking metrics.