Data-based lumping and splitting of treatments in network meta-analysis in the presence of heterogeneity – a flexible non-parametric Bayesian approach

Tim Disher, PhD, RN | Sandpiper Analytics

MSR126

Background

- A critical challenge in Network Meta-analysis (NMA) is defining treatment nodes which profoundly impacts results (Xing & Lin, 2020).
- Lumping distinct treatments can obscure differences, while splitting similar ones reduces precision and increases the risk of spurious findings due to multiplicity (Barrientos, Page, & Lin, 2024; Efthimiou & White, 2020).
- While most lumping/splitting decisions are qualitative (Dhippayom, Saldanha, Chaiyakunapruk, & Devine, 2022), data-driven solutions exist and demonstrate that the choice can substantively affect NMA results (Xing & Lin, 2020).
- Proposed data-driven approaches include a Bayesian non-parametric method allowing for uncertainty in classification (Barrientos et al., 2024) and a frequentist method using generalized fused lasso to penalize differences (Kong, Daly, & Béliveau, 2024).
- Limitations of current data-driven methods include implementation difficulties (e.g., custom MCMC), potential interpretation issues with specific priors, and the failure to account for treatment effect heterogeneity, particularly adjustment for baseline risk (Barrientos et al., 2024; Cameron et al., 2018).
- A need exists for a flexible framework that facilitates implementation, clarifies interpretation, accounts for heterogeneity (like baseline risk), and allows incorporating varying degrees of domain-specific knowledge into the clustering process.

Methods

- We adapted and extended a Bayesian non-parametric approach, replacing the Dirichlet process prior with a spike-and-slab base measure with a regularized horseshoe prior. This facilitates implementation with generic samplers (e.g., JAGS), avoids mixing issues, and frames treatment effects in terms of regularization.
- We incorporated meta-regression on baseline risk to account for heterogeneity, allowing for clustering under varying baseline risks. We also developed methods to integrate domain knowledge by limiting the number of clusters or specifying informative priors. We illustrate these methods in an application to the certolizumab baseline risk adjustment example from NICE TSD 3 (Dias, Sutton, Welton, & Ades, 2011).

Results

All model implementations showed good mixing and were well calibrated under simulation from the known data-generating process. Node lumping differed under the unadjusted and adjusted models, with two groups in the former and three in the latter. Model fit statistics were best under the baseline risk adjusted clustering model (DIC: 152) and second best under the standard non-clustering adjusted model (DIC: 154). This finding suggests that out of sample predictive performance can be maintained or improved while reducing dimensionality of the network by more than 50%. Effective sample size in the DP BLR NMA was 2-4x that of the unclustered NMA.

A

Tocilizumab					
1.07 (0.56 to 2.09)	Adalimumab				
1.28 (0.45 to 3.60)	1.19 (0.46 to 3.16)	Etanercept			
1.54 (0.78 to 3.27)	1.42 (0.81 to 2.83)	1.21 (0.48 to 3.18)	CZP		
1.72 (0.86 to 3.39)	1.60 (0.88 to 2.93)	1.35 (0.53 to 3.54)	1.13 (0.58 to 2.05)	Infliximab	
6.55 (1.95 to 23.24)	6.10 (1.82 to 20.63)	5.14 (1.24 to 22.61)	4.25 (1.29 to 15.08)	3.79 (1.20 to 13.94)	Rituximab
9.46 (5.78 to 16.41)	8.80 (6.11 to 13.90)	7.39 (3.29 to 18.27)	6.17 (3.93 to 9.82)	5.48 (3.59 to 8.75)	1.45 (0.45 to 4.43)	Placebo

B

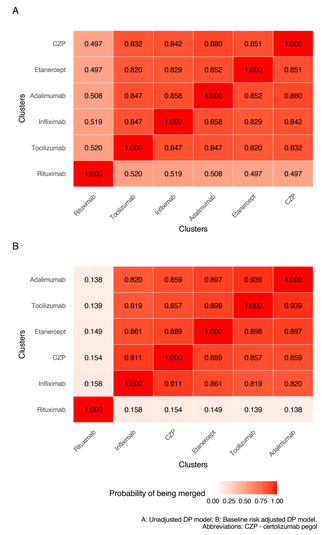
Tocilizumab					
1.00 (0.90 to 1.00)	Adalimumab				
1.00 (1.00 to 1.62)	1.00 (1.00 to 1.60)	Etanercept			
1.00 (1.00 to 1.65)	1.00 (1.00 to 1.63)	1.00 (1.00 to 1.63)	CZP		
1.00 (1.00 to 1.70)	1.00 (1.00 to 1.68)	1.00 (1.00 to 1.65)	1.00 (1.00 to 1.39)	Infliximab	
5.58 (1.00 to 12.04)	5.59 (1.00 to 12.01)	5.44 (1.00 to 11.90)	5.32 (1.00 to 11.50)	5.23 (1.00 to 11.45)	Rituximab
7.24 (5.55 to 10.05)	7.24 (5.62 to 10.03)	7.11 (5.26 to 9.97)	7.00 (5.11 to 9.13)	6.94 (4.91 to 9.13)	1.30 (0.60 to 8.16)	Placebo

Credible interval contains 1 No Yes

A: Baseline risk adjusted standard model; B: Baseline risk adjusted DP model. Abbreviations: CZP - certolizumab pegol

Results

In both the unadjusted and adjusted DP models (Figure 2 Panel A and B), active therapies other than rituximab showed similarly high probability of being grouped into the same cluster. In the unadjusted model, however, rituximab has an approximately 50% probability of being included in this cluster whereas that probability dropped to approximately 15% when heterogeneity in baseline risk was accounted for. This difference was also found in the modal network, which combined all therapies in the unadjusted model but with large uncertainty (41%) versus combining all therapies except rituximab after adjustment with increased confidence (64%).



A: Unadjusted DP model; B: Baseline risk adjusted DP model. Abbreviations: CZP - certolizumab pegol

Conclusions

Data-based lumping/splitting decisions informed by prior expectation of unique treatments is feasible to implement within existing NICE TSD code. Lumped treatments can differ significantly in adjusted compared to unadjusted models. Data-based lumping of treatments can have large effects on effective sample size on comparisons. Future research should consider to ability to leverage prior information in treatments that are more likely to be lumped (eg, minor dose differences, treatment in same class).

References

Dhippayom, T., Saldanha, I. J., Chaiyakunapruk, N., & Devine, B. (2022). Characterizing health services interventions for evidence syntheses by using the theme, intensity, provider/platform (TIP) framework. *Journal of Evidence-Based Medicine, 15*(4), 333–337.

Xing, A., & Lin, L. (2020). Effects of treatment classifications in network meta-analysis. Research Methods in Medicine & Health Sciences, 1(1), 12–24.

Barrientos, A. F., Page, G. L., & Lin, L. (2024). Non-parametric bayesian approach to multiple treatment comparisons in network meta-analysis with application to comparisons of anti-depressants. Journal of the Royal Statistical Society Series C: Applied Statistics, 73(5), 1333–1354.

Cameron, C., Hutton, B., Druchok, C., McElligott, S., Nair, S., Schubert, A., … Villacorta, R. (2018). Importance of assessing and adjusting for cross-study heterogeneity in network meta-analysis: A case study of psoriasis. Journal of Comparative Effectiveness Research, 7(11), 1037–1051.

Dias, S., Sutton, A. J., Welton, N. J., & Ades, A. (2011). NICE DSU technical support document 3: Heterogeneity: Subgroups, meta-regression, bias and bias-adjustment.

Kong, X., Daly, C. H., & Béliveau, A. (2024). Generalized fused lasso for treatment pooling in network meta-analysis. Statistics in Medicine.

SANDPIPER ANALYTICS

sandpiper-analytics.com

tim.disher@sandpiper-analytics.com

linkedin.com/in/tim-disher/