Network meta-analysis with integrated nested Laplace approximations (INLA)

Rafael Sauter[1], Leonhard Held[1]

[1]Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Switzerland

University of Zurich[UZH]

Network meta-analysis (NMA)

Motivation

- A systematic review collects trials comparing not only treatment A vs. B ...
  but also A vs. C, A vs. D and D vs. B, etc.
- A network meta-analysis (NMA) aggregates the evidence combining direct and indirect comparisons instead of using pairwise direct comparisons only.

Challenges of a NMA

- Like in pairwise meta-analysis the treatment effects can show excessive between-trial variability called heterogeneity.
- Combining direct and indirect evidence may produce contradictory estimates, called network inconsistency.

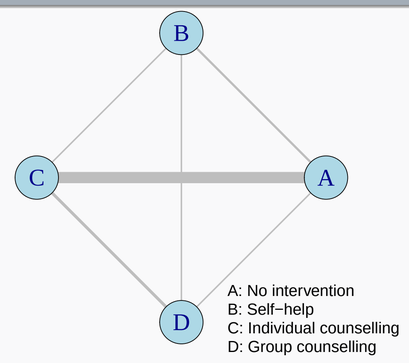
Advantages of a NMA:

- Comprehensive use of all available data (direct and indirect evidence).
- Comparison of interventions which have not been directly compared in any experiment.
- Improved precision for each comparison.

$\Longrightarrow$ **We implement established NMA models in INLA**, instead of computationally intensive MCMC sampling!

Application: a network about smoking cessation (Hasselblad, 1998)

- The binomial outcome $y_{iX} \sim \text{Bin}(n_{iX}, \pi_{iX})$ for treatment X and trial i is successful smoking cessation after 6 to 12 months.
- The network compares interventions A, B, C, D.
- There are 24 (22 two-arm and 2 three-arm) trials.
- An edge represents an observed direct treatment comparison d_{XY}.

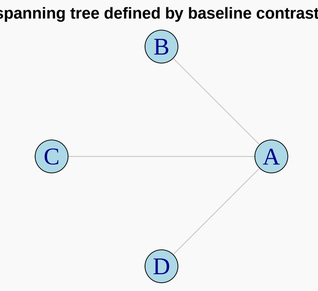


- Network consistency implies...

$$d_{BC} = d_{AC} - d_{AB}$$
$$d_{BD} = d_{AD} - d_{AB}$$
$$d_{CD} = d_{AD} - d_{AC}$$

... thus only 3 baseline contrasts $\boldsymbol{d}_b = (d_{AB}, d_{AC}, d_{AD})^\top$ are needed to describe the network.

- The remaining contrasts $\boldsymbol{d}_r$ are linear combinations of $\boldsymbol{d}_b$.



NMA model

- A logistic regression model describes the smoking cessation network

$$\text{logit}(\pi_{iX}) = a_{iX}$$
$$\text{logit}(\pi_{iY}) = a_{iX} + d_{XY} + \gamma_{iXY}.$$

- The baseline treatment effect a_{iX} is a nuisance parameter. Our interest is in d_{XY}.
- Heterogeneity is modeled by trial-specific random effects $\gamma_{iXY} \sim \text{N}(0, \tau^2)$.

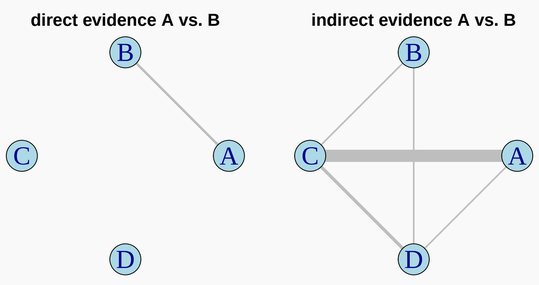
2 approaches to include inconsistency:

Random effects inconsistency

- Lu and Ades (2006) introduce inconsistency by cycle-specific random effects $\xi_{XYZ} \sim \text{N}(0, \kappa^2)$ for $\boldsymbol{d}_r$ such that e.g. $d_{YZ} = d_{XZ} - d_{XY} + \xi_{XYZ}$.
- The model for treatment π_{iY} with baseline treatment X is extended:
  $\text{logit}(\pi_{iY}) = a_{iX} + d_{XY} + \gamma_{iXY} + \xi_{XYZ}$.

Node-splitting inconsistency

- Dias et al. (2010) suggest to estimate d_{XY} based on direct $d_{XY}^{\text{dir.}}$ and based on indirect $d_{XY}^{\text{ind.}}$ evidence only. This is a node-split for X, Y.



- Inconsistency for d_{XY} is defined by the linear combination

$$i_{XY} = d_{XY}^{\text{dir.}} - d_{XY}^{\text{ind.}}.$$

Bayesian estimation with INLA

- This NMA model can be estimated with integrated nested Laplace approximations (INLA) (Rue et al., 2009).
- INLA approximates the marginal posterior distributions and is a fast and accurate alternative to MCMC sampling.

Multi-arm trials

- If a trial i compares more than 2 treatments then $\boldsymbol{\gamma}_i$ is not assumed to be independent.
- (Lu and Ades, 2006) assume a homogeneous correlation structure

$$\boldsymbol{\gamma}_i \sim \text{N}(\mathbf{0}, \boldsymbol{T})$$

- The variances on the diagonal of $\boldsymbol{T}$ are all equal to τ^2.
- Correlations in $\boldsymbol{T}$ are under consistency equal to $\rho = 1/2$.
- In INLA $\boldsymbol{T}$ can be written as Kronecker product

$$\boldsymbol{T} = \mathbf{C} \otimes \tau^2$$

where $\mathbf{C} = (1-\rho)\mathbf{I} + \rho\mathbf{J}$, $\mathbf{I}$ is the identity matrix and $\mathbf{J}$ is a matrix of ones.

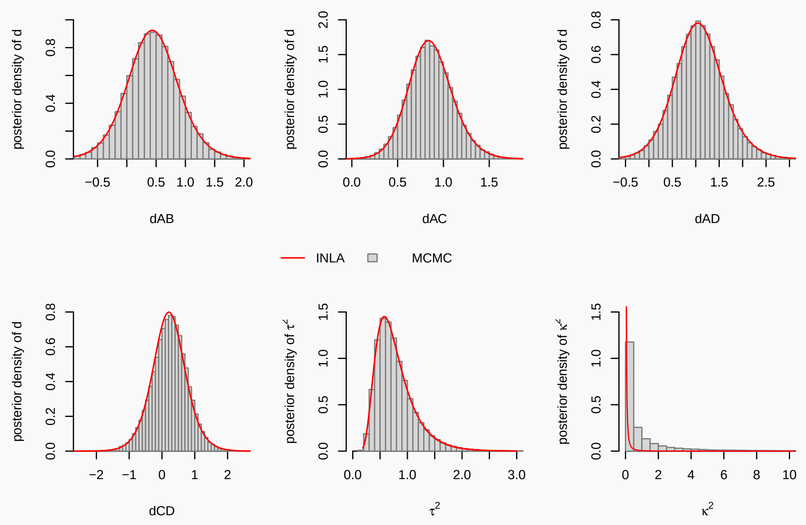
- This correlation structure is implemented in INLA with the `group` statement (see Riebler et al., 2012).

Node-splitting

- Node-splitting is implemented in INLA by using two likelihoods. One for $d_{XY}^{\text{dir.}}$ and one for $d_{XY}^{\text{ind.}}$ (see Martino et al., 2011).
- The heterogeneity hyperparameter τ^2 is the same for both likelihoods and estimated based on all data.

Application

Compare MCMC and INLA for smoking cessation NMA with random effects inconsistency



Marginal posterior distributions for $\boldsymbol{d}_b$ and one linear combination (d_{CD}) by INLA (red line) and MCMC (histogram) for the NMA model with random effects for heterogeneity τ^2 and incoherence κ^2.

Compare MCMC and INLA for smoking cessation NMA with node-splitting inconsistency

	MCMC			INLA		
	Mean	Stdev.	p-value	Mean	Stdev.	p-value
i_{AB}	-0.36	0.84	0.66	-0.42	0.77	0.58
i_{AC}	-0.24	0.80	0.77	0.22	0.69	0.74
i_{AD}	0.35	1.01	0.74	0.23	0.96	0.82
i_{BC}	-0.69	0.91	0.43	-0.65	0.83	0.42
i_{BD}	0.09	1.04	0.92	-0.03	0.96	0.97
i_{CD}	-1.85	1.10	0.08	-1.89	0.94	0.03

Inconsistency estimates i_{XY} based on the linear combination of the marginal posteriors $d_{XY}^{\text{dir.}}$ and $d_{XY}^{\text{ind.}}$.

Compare computation time for node-splitting by MCMC and INLA

Computation time for 6 node-splits in the smoking cessation network with a laptop (2.7GHz 4 core processor):

- **0.17 min.** with INLA.
- **16.73 min.** with MCMC using 150'000 iterations, done with `JAGS` through `R`-package `gemtc` (van Valkenhoef and Kuiper, 2014).

Summary

- INLA is much faster than MCMC sampling which is an advantage if node-splitting for a large network is needed.
- The approximations by INLA are deterministic and convergence checks are not needed.
- Also other NMA models can be implemented in INLA such as summary-level data models (Lumley, 2002).
- More details can be found in Sauter and Held (2015).

References

Dias, S., N. J. Welton, D. M. Caldwell, and A. E. Ades (2010). Checking consistency in mixed treatment comparison meta-analysis. *Statistics in Medicine 29*(7-8), 932–944.

Hasselblad, V. (1998). Meta-analysis of multitreatment studies. *Medical Decision Making 18*(1), 37–43.

Lu, G. and A. E. Ades (2006). Assessing evidence inconsistency in mixed treatment comparisons. *Journal of the American Statistical Association 101*(474), 447–459.

Lumley, T. (2002). Network meta-analysis for indirect treatment comparisons. *Statistics in Medicine 21*(16), 2313–2324.

Martino, S., R. Akerkar, and H. Rue (2011). Approximate Bayesian inference for survival models. *Scandinavian Journal of Statistics 38*(3), 514–528.

Riebler, A., L. Held, and H. Rue (2012, 03). Estimation and extrapolation of time trends in registry data - Borrowing strength from related populations. *The Annals of Applied Statistics 6*(1), 304–333.

Rue, H., S. Martino, and N. Chopin (2009, 2). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations (with discussion). *Journal of the Royal Statistical Society - Series B 71*, 319–392.

Sauter, R. and L. Held (2015). Network meta-analysis with integrated nested Laplace approximations. *Biometrical Journal*. Early view available online.

van Valkenhoef, G. and J. Kuiper (2014). *gemtc: GeMTC network meta-analysis*. R package version 0.6.

rafael.sauter@uzh.ch