Network Meta-Interpolation

Fast and accurate NMA with effect modification

Ofir Harari ofir@coreclinicalsciences.com

Joint work with Mohsen Soltanifar, Andre Verhoek, Joseph Cappelleri, Mario Ouwens, Caitlin Daly and Bart Heeg

June 3, 2023

Effect modification in Network Meta-Analysis

Effect modification

- When the magnitude of the effect of treatment on an outcome varies along the value of a variable, X, we call X an *effect modifier* (EM).
- In statistical terms, this means that X interacts with the treatment assignment.
  - Different treatment effects for different levels of X.
- Example: in the OAK study[1] (2017, Atezolizumab vs. Docetaxel for advanced NSCLC patients), the ITT hazard ratio for overall survival was 0.66 [CI: 0.51–0.86] for males and 0.81 [0.66–0.99] for females
  - Is sex an effect modifier?
- In network meta-analysis (NMA), an evidence network of trials with unbalanced EM distributions could lead to considerable bias.

Effect modification (continued)

- Consider the following hypothetical evidence network –

Study	Trt1	Trt2	% age > 65	% high severity	TE	SE
1	A	B	62.8	45.7	1.387	0.188
2	A	B	84.3	61.8	1.475	0.183
3	A	B	78.7	59.2	1.612	0.187
4	A	C	73.3	50.3	2.757	0.204
5	A	C	70.5	46.8	2.526	0.198
6	A	C	60.0	40.5	2.134	0.191
7	A	D	53.3	30.3	0.667	0.187

- Here the proportion of patients aged 65+ ranges from 53% to 84%, while the proportion of severe cases (at baseline) varies between 30% and 62%.
- If age and/or disease severity are effect modifiers, our analysis must take into account this study heterogeneity.

NMA methods and shared effect modification

Existing methods

The following methods assume(in their basic form) *shared effect modification* (SEM), that is: that all EMs affect outcome in the exact same way across different treatments.

1. **Network Meta Regression (NMR)**
   - Aggregate-level data (AgD) network meta-regression (NMR).[2]
   - The version the allows non-SEM is over-parameterized for the typical network size.
2. **Multi-level Network Meta Regression (ML-NMR)**
   - A recent approach[3] implemented in the *multinma* R package[4].
   - Resolves study imbalance by learning the EM distribution from IPD studies and sampling random pseudo-populations from the AgD studies.
   - In practice, requires one IPD per treatment class to fit a non-SEM model.

The SEM assumption is extremely limiting, and is probably rarely reasonable to make.

- Fitting a SEM model to non-SEM data only exacerbates bias.

Existing methods (continued)

3. **Matching-Adjusted Indirect Comparison (MAIC) and Simulated Treatment Comparison (STC)**

- MAIC uses propensity-score weighting, while STC is based on regression adjustments.
- Both methods are limited to two studies: one IPD and one AgD.
- ITC when using these methods is restricted to the EM conditions of the AgD study.
- Even within said restrictions, their performance is inadequate[5].

Network Meta-Interpolation (NMI)

Data requirements

We assume that we have the following information available for all studies -

Trt1	Trt2	# patients	Prop. age > 65	Prop. high severity	TE	SE
A	B	600	0.628	0.457	1.387	0.188
A	B		1		1.829	0.242
A	B		0		0.492	0.319
A	B			1	1.363	0.267
A	B			0	1.464	0.270

- 1^{st} row: treatment effect estimate and standard error for the full trial (62.8% 65 or older, 45.7% severe cases at baseline).
- 2^{nd} and 3^{rd} rows: TE and SE for the >= 65 and < 65 subgroups, respectively.
- 4^{th} and 5^{th} rows: TE and SE data on the high and low severity subgroups, respectively.
- This kind of information is often reported, but is hardly extracted - and never used.
- Empty cells in the age and severity columns stand contain missing data.

Outline of the method

- NMI is a “population adjustment” heuristic that aims to balance on study EM distribution.
- The idea: put effort into data pre-processing and keep modelling simple.
- The method consists of the following three steps:

  1. Enrich the data by imputing the missing cells in the table for all studies.
  2. Interpolate treatment effect estimates and standard errors for all studies at “neutral” EM values.
  3. Perform standard NMA[6] on the resultant study-level data, now that all studies are balanced.

Modelling assumptions (two EMs)

For each study, and assume that the data are generated by the generalized linear model

$$\mathrm{g}\left(\theta;\boldsymbol{\beta}\right)=\beta_{00}+\beta_{01}x_{1i}+\beta_{02}x_{2i}+\beta_{12}x_{1i}x_{2i}+\beta_{03}T_i+\beta_{13}x_{1i}T_i+\beta_{23}x_{2i}T_i,$$

- Here x_{1i} and x_{2i} are two binary covariate values for patient i.
- T_i is the treatment indicator (1 for treatment, 0 for baseline).
- The method is applicable to any number of binary EMs.

Modelling assumptions (continued)

- The relative treatment effect at the aggregate level is given by

$$\hat{\Delta}\left(\bar{x}_1, \bar{x}_2\right) := \mathrm{g}\left(\theta; \hat{\boldsymbol{\beta}}, \bar{x}_1, \bar{x}_2, T = 1\right) - \mathrm{g}\left(\theta; \hat{\boldsymbol{\beta}}, \bar{x}_1, \bar{x}_2, T = 0\right)$$

$$= \hat{\beta}_{03} + \hat{\beta}_{13}\bar{x}_1 + \hat{\beta}_{23}\bar{x}_2.$$

- The associated variance is then

$$\hat{\sigma}^2_{\hat{\Delta}}\left(\bar{x}_1, \bar{x}_2\right) = \hat{\sigma}_{00} + \bar{x}_1^2\hat{\sigma}_{11} + \bar{x}_2^2\hat{\sigma}_{22} + 2\bar{x}_1\hat{\sigma}_{01} + 2\bar{x}_2\hat{\sigma}_{02} + 2\bar{x}_1\bar{x}_2\hat{\sigma}_{12},$$

  where

$$\hat{\sigma}_{ij} = \mathrm{Cov}\left(\hat{\beta}_{i3}, \hat{\beta}_{j3}\right), \quad 0 \leq i \leq j \leq 2.$$

- Both equations are linear in their respective parameters (given the $\bar{x}$-s).

Step 1

- Recall that our study-level data is of the form

$$\begin{bmatrix} \bar{x}_1 & \bar{x}_2 & \widehat{\mathrm{TE}}_1 & \widehat{\mathrm{SE}}_1 \\ 1 & & \widehat{\mathrm{TE}}_2 & \widehat{\mathrm{SE}}_2 \\ 0 & & \widehat{\mathrm{TE}}_3 & \widehat{\mathrm{SE}}_3 \\ & 0 & \widehat{\mathrm{TE}}_4 & \widehat{\mathrm{SE}}_4 \\ & 1 & \widehat{\mathrm{TE}}_5 & \widehat{\mathrm{SE}}_5 \end{bmatrix}$$

- First, we take advantage of any IPD we have to estimate the Pearson correlation ρ_{X_1,X_2}.
- Then we use it to replace the missing values with the best linear unbiased prediction (BLUP)

$$\bar{x}_j^* \left(x_i\right) = \hat{\rho}_{X_1,X_2} \hat{\sigma}_{X_1} \hat{\sigma}_{X_2} \left(x_i - \bar{x}_j\right) + \bar{x}_j,\ 1 \leq i \neq j \leq 2,\ x_i = 0, 1.$$

Step 1: an example

- The original study

Trt1	Trt2	x1	x2	TE	se
A	B	0.628	0.457	1.387	0.188
A	B	1		1.829	0.242
A	B	0		0.492	0.319
A	B		1	1.363	0.267
A	B		0	1.464	0.270

- …and the imputed one

Trt1	Trt2	x1	x2	TE	se
A	B	0.628	0.457	1.387	0.188
A	B	1.000	0.538	1.829	0.242
A	B	0.000	0.321	0.492	0.319
A	B	0.739	1.000	1.363	0.267
A	B	0.535	0.000	1.464	0.270

Step 1 output

- Now that we have a completed study-level matrix (all 5 rows)

$$\begin{bmatrix} \bar{x}_1 & \bar{x}_2 & \widehat{\mathrm{TE}}_1 & \widehat{\mathrm{SE}}_1 \\ 1 & \bar{x}_2^*(1) & \widehat{\mathrm{TE}}_2 & \widehat{\mathrm{SE}}_2 \\ 0 & \bar{x}_2^*(0) & \widehat{\mathrm{TE}}_3 & \widehat{\mathrm{SE}}_3 \\ \bar{x}_1^*(1) & 0 & \widehat{\mathrm{TE}}_4 & \widehat{\mathrm{SE}}_4 \\ \bar{x}_1^*(0) & 1 & \widehat{\mathrm{TE}}_5 & \widehat{\mathrm{SE}}_5 \end{bmatrix} = \begin{bmatrix} | & | & | & | \\ \bar{\boldsymbol{x}}_1^* & \bar{\boldsymbol{x}}_2^* & \hat{\boldsymbol{\Delta}} & \hat{\boldsymbol{\sigma}}_{\hat{\Delta}} \\ | & | & | & | \end{bmatrix},$$

we may proceed to Step 2 of NMI and go back to our equations.

Step 2

- The first equation

$$\hat{\Delta}\left(\bar{x}_1, \bar{x}_2\right) = \hat{\beta}_{03} + \hat{\beta}_{13}\bar{x}_1 + \hat{\beta}_{23}\bar{x}_2,$$

  induces the system

$$\boldsymbol{M}_1 \hat{\boldsymbol{\beta}}_{3\times 1} = \hat{\boldsymbol{\Delta}}_{5\times 1}$$

  for

$$\boldsymbol{M}_1 = \begin{bmatrix} | & | & | \\ \mathbf{1} & \bar{\boldsymbol{x}}_1^* & \bar{\boldsymbol{x}}_2^* \\ | & | & | \end{bmatrix}_{5\times 3}$$

- This overdetermined system gives rise to the least square solution

$$\tilde{\hat{\boldsymbol{\beta}}} = \left(\boldsymbol{M}_1^\mathsf{T}\boldsymbol{M}_1\right)^{-1}\boldsymbol{M}_1^\mathsf{T}\hat{\boldsymbol{\Delta}}$$

Step 2 (continued)

- The second equation,

$$\hat{\sigma}^2_{\hat{\Delta}}\left(\bar{x}_1, \bar{x}_2\right) = \hat{\sigma}_{00} + \bar{x}_1^2\hat{\sigma}_{11} + \bar{x}_2^2\hat{\sigma}_{22} + 2\bar{x}_1\hat{\sigma}_{01} + 2\bar{x}_2\hat{\sigma}_{02} + 2\bar{x}_1\bar{x}_2\hat{\sigma}_{12},$$

  takes the form

$$\boldsymbol{M}_2\hat{\boldsymbol{\sigma}}_{6\times 1} = \hat{\boldsymbol{\sigma}}^2_{\hat{\boldsymbol{\Delta}},5\times 1}$$

  for

$$\boldsymbol{M}_2 = \begin{bmatrix} | & | & | & | & | & | \\ \mathbf{1} & \bar{\boldsymbol{x}}_1^{*\,2} & \bar{\boldsymbol{x}}_2^{*\,2} & 2\bar{\boldsymbol{x}}_1^* & 2\bar{\boldsymbol{x}}_2^* & 2\bar{\boldsymbol{x}}_1^*\bar{\boldsymbol{x}}_2^* \\ | & | & | & | & | & | \end{bmatrix}_{5\times 6}$$

- Here we use the Moore-Penrose pseudo-inverse for underdetermined systems

$$\tilde{\hat{\boldsymbol{\sigma}}} = \boldsymbol{M}_2^{\mathsf{T}}\left(\boldsymbol{M}_2\boldsymbol{M}_2^{\mathsf{T}}\right)^{-1}\hat{\boldsymbol{\sigma}}^2_{\hat{\boldsymbol{\Delta}}}$$

Step 2 (continued)

Once we have obtained $\tilde{\hat{\boldsymbol{\beta}}}$ and $\tilde{\hat{\boldsymbol{\sigma}}}$, the relative treatment effect at any new vector $\left(\bar{x}_1^{\text{new}}, \bar{x}_2^{\text{new}}\right)$ and its associated variance are approximated by

$$\tilde{\hat{\Delta}}\left(\bar{x}_1^{\text{new}}, \bar{x}_2^{\text{new}}\right) = \tilde{\hat{\beta}}_{03} + \tilde{\hat{\beta}}_{13}\bar{x}_1^{\text{new}} + \tilde{\hat{\beta}}_{23}\bar{x}_2^{\text{new}}$$

and

$$\tilde{\hat{\sigma}}^2_{\hat{\Delta}}\left(\bar{x}_1^{\text{new}}, \bar{x}_2^{\text{new}}\right) = \tilde{\hat{\sigma}}_{00} + \left(\bar{x}_1^{\text{new}}\right)^2\tilde{\hat{\sigma}} + \left(\bar{x}_2^{\text{new}}\right)^2\tilde{\hat{\sigma}}_{22} + 2\bar{x}_1^{\text{new}}\tilde{\hat{\sigma}}_{01} + 2\bar{x}_2^{\text{new}}\tilde{\hat{\sigma}}_{02} + 2\bar{x}_1^{\text{new}}\bar{x}_2^{\text{new}}\tilde{\hat{\sigma}}_{12},$$

respectively.

Step 2: an example

- Suppose that we wish to perform ITC at $(\bar{x}_1^{\text{new}}, \bar{x}_2^{\text{new}}) = (0.675, 0.475)$
- Step 1 output

Trt1	Trt2	x1	x2	TE	se
A	B	0.628	0.457	1.387	0.188
A	B	1.000	0.538	1.829	0.242
A	B	0.000	0.321	0.492	0.319
A	B	0.739	1.000	1.363	0.267
A	B	0.535	0.000	1.464	0.270

is now collapsed into

Trt1	Trt2	x1	x2	TE	se
A	B	0.675	0.475	1.44	0.189

Step 3

- Repeat steps 1-2 for each AgD study in the evidence network.
- The final dataset will contain exactly one row per study.
- Perform standard (fixed or random effect) AgD NMA on that resultant dataset.

Summary of NMI

1. Impute all missing sub-population means in the AgD subgroup analyses using BLUP.
2. Solve two linear systems (per study) to obtain relative TEs and SEs at neutral EM values.
3. Perform standard AgD NMA.

A worked example

The AgD data sets used

```
#Displaying NMR/NMA AgD
disp_tab(NMR_AgD)
```

Study	Trt1	Trt2	x1	x2	TE	se
1	A	B	0.628	0.457	1.387	0.188
2	A	B	0.843	0.618	1.475	0.183
3	A	B	0.787	0.592	1.612	0.187
4	A	C	0.733	0.503	2.757	0.204
5	A	C	0.705	0.468	2.526	0.198
6	A	C	0.600	0.405	2.134	0.191
7	A	D	0.533	0.303	0.667	0.187

```
#Displaying ML-NMR AgD
disp_tab(ML_NMR_AgD[1:10,])
```

Tr	n	r	X1	X2
A	318	58	0.597	0.472
B	282	133	0.663	0.440
A	317	65	0.842	0.625
B	283	150	0.845	0.611
A	292	57	0.767	0.589
B	308	169	0.805	0.594
A	301	56	0.751	0.495
C	299	234	0.716	0.512
A	280	58	0.682	0.461
C	320	245	0.725	0.475

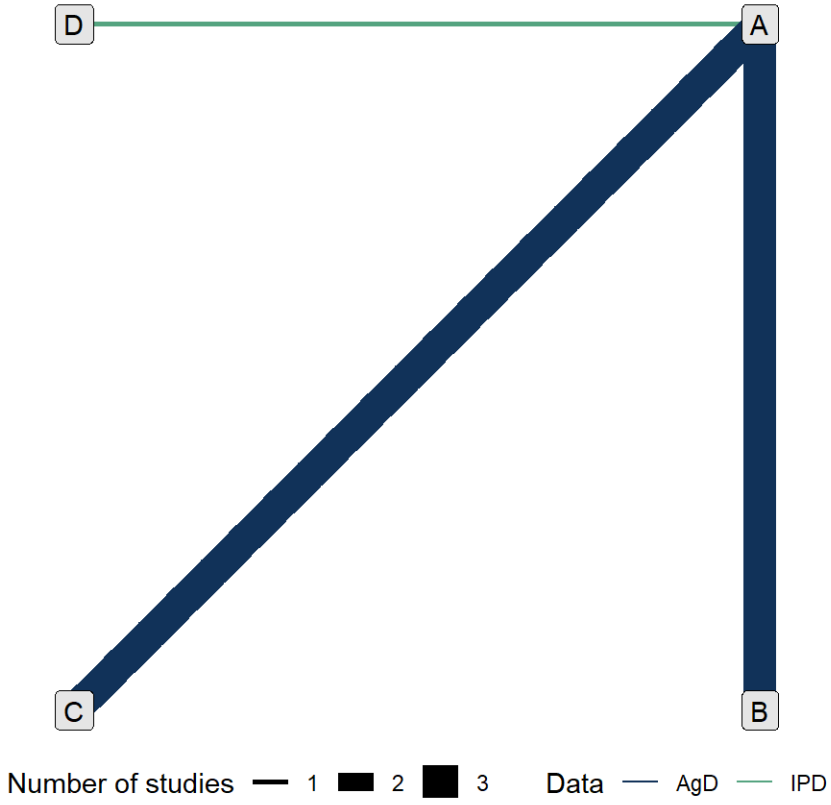
```
#Displaying NMI AgD
disp_tab(NMI_AgD[6:15,])
```

Study	Trt1	Trt2	x1	x2	TE	se
2	A	B	0.843	0.618	1.475	0.183
2	A	B	1		1.577	0.198
2	A	B	0		0.946	0.525
2	A	B		1	1.517	0.232
2	A	B		0	1.412	0.298
3	A	B	0.787	0.592	1.612	0.187
3	A	B	1		1.858	0.215
3	A	B	0		0.643	0.398
3	A	B		1	1.662	0.241
3	A	B		0	1.542	0.296

The network

```
net = combine_network(
  set_ipd(IPD %>%
            mutate(TrC = Tr),
          study = Study,
          trt = Tr,
          trt_class = TrtClass,
          r = Y),
  set_agd_arm(ML_NMR_AgD %>%
                mutate(TrC = Tr),
              study = Study,
              trt = Tr,
              trt_class = TrtClass,
              r = r,
              n = n),
  trt_ref = "A")

plot(net)
```



NMI (Step 1-2)

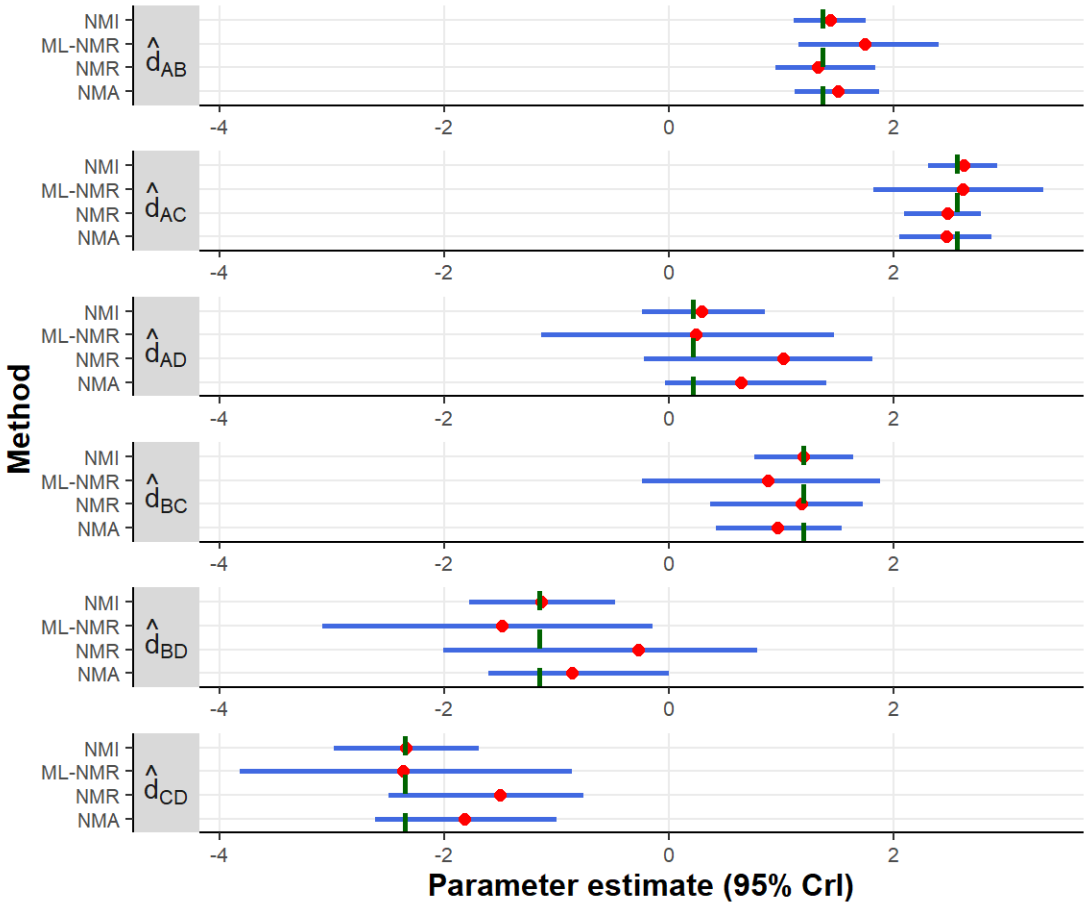
```
disp_tab(tail(NMI_object$Imputed, 10))
```

Study	Trt1	Trt2	x1	x2	TE	se
6	A	C	0.600	0.405	2.134	0.191
6	A	C	1.000	0.489	3.139	0.279
6	A	C	0.000	0.279	1.056	0.287
6	A	C	0.725	1.000	3.768	0.380
6	A	C	0.515	0.000	1.307	0.233
7	A	D	0.533	0.303	0.667	0.187
7	A	D	1.000	0.394	0.008	0.278
7	A	D	0.000	0.200	1.305	0.266
7	A	D	0.692	1.000	-0.286	0.374
7	A	D	0.464	0.000	0.997	0.223

```
disp_tab(NMI_object$Final) #data for NMA
```

Study	Trt1	Trt2	x1	x2	TE	se
1	A	B	0.675	0.475	1.439	0.189
2	A	B	0.675	0.475	1.370	0.216
3	A	B	0.675	0.475	1.480	0.196
4	A	C	0.675	0.475	2.786	0.200
5	A	C	0.675	0.475	2.571	0.195
6	A	C	0.675	0.475	2.540	0.209
7	A	D	0.675	0.475	0.289	0.221

Results


NMI
ML-NMR
NMR
NMA
$\hat{d}_{AB}$
$\hat{d}_{AC}$
$\hat{d}_{AD}$
$\hat{d}_{BC}$
$\hat{d}_{BD}$
$\hat{d}_{CD}$
-4
-2
0
2
Method
Parameter estimate (95% CrI)

A more comprehensive simulation study

Simulation description

- We generated 1000 random datasets from the following non-SEM model -

$$\log \frac{p_i}{1-p_i} = \varepsilon_i + \begin{cases} -1.39 + 0.69T_i + x_{1i}T_i & \text{A-B trials} \\ -1.39 + T_i + 1.61x_{1i}T_i + x_{2i}T_i & \text{A-C trials} \\ -1.39 + 1.5T_i - 1.2x_{1i}T_i - x_{2i}T_i & \text{A-D trial} \end{cases}$$

and 1000 more from the following SEM model -

$$\log \frac{p_i}{1-p_i} = \varepsilon_i + \begin{cases} -1.39 + 0.69T_i + 1.61x_{1i}T_i + x_{2i}T_i & \text{A-B trials} \\ -1.39 + T_i + 1.61x_{1i}T_i + x_{2i}T_i & \text{A-C trials} \\ -1.39 + 1.5T_i + 1.61x_{1i}T_i + x_{2i}T_i & \text{A-D trial} \end{cases}$$

where $\varepsilon_i \sim \mathcal{N}\left(0, 0.2^2\right)$.

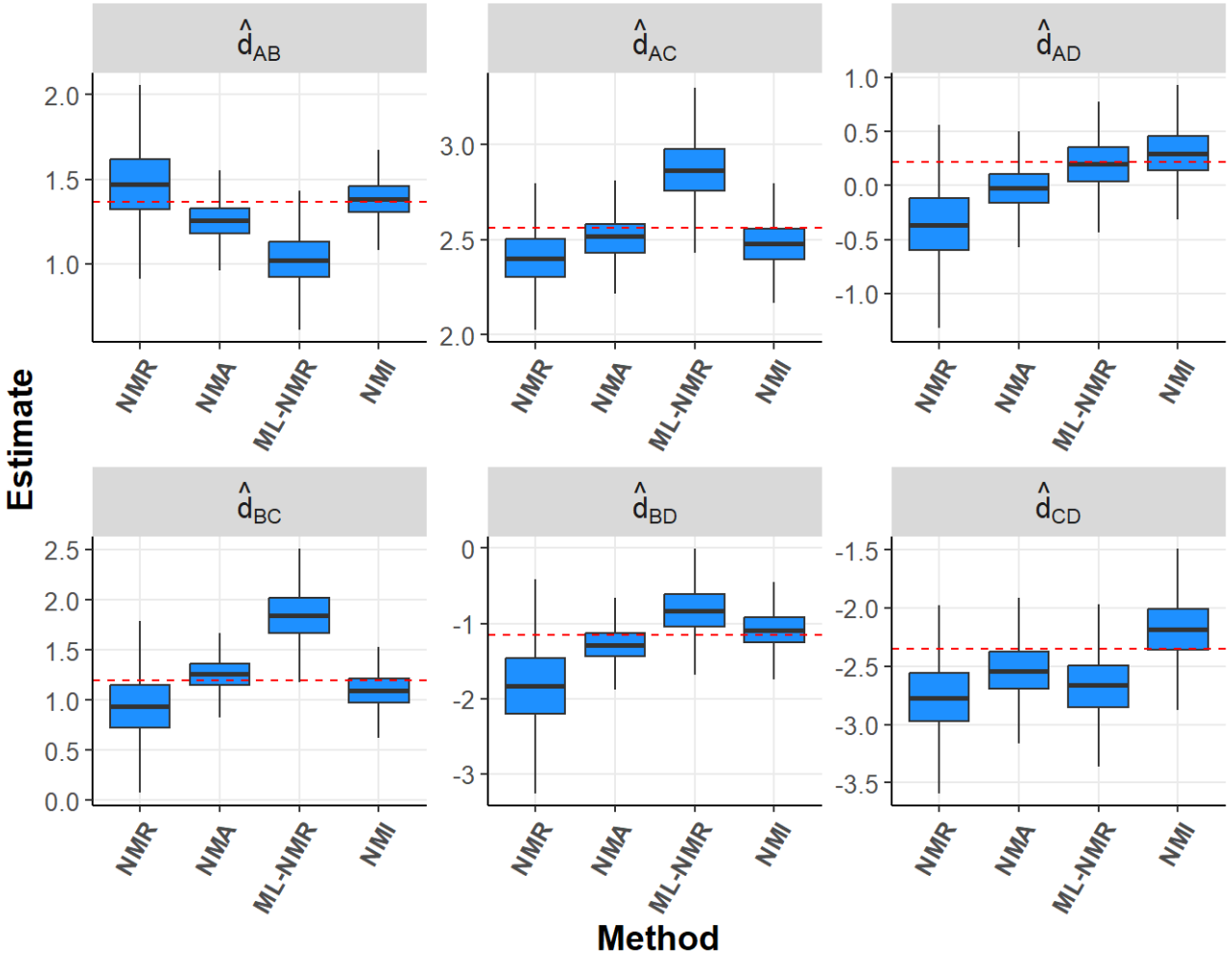
Simulation description (continued)

- The network consisted of three A-B (AgD), three A-C (AgD) and one A-D (IPD) studies.
- The effect modifiers X_1 and X_2 where drawn from a bivariate distribution with correlation $\rho_{X_1X_2} = 0.244$ and marginal binary distributions with

	A-B trials			A-C trials			A-D trial
$\Pr(X_1 = 1)$	0.50	0.56	0.62	0.68	0.73	0.79	0.85
$\Pr(X_2 = 1)$	0.30	0.36	0.42	0.48	0.53	0.59	0.65

- Each time, the posterior means of all treatment effects were recorded along with their 95% CrIs for all methods.
- Root mean squared error (RMSE) was calculated for each treatment effect estimate, as well as the proportion of the time that the 95% CrI covered the true value.

Simulation results: non-SEM data



Simulation results (averaged over all effects)

- Simulation results: non-SEM data

Method	RMSE	Coverage
NMR	0.541	80.2%
NMA	0.241	98.3%
ML-NMR	0.447	92.6%
NMI	0.228	99.0%

- Simulation results: SEM data

Method	RMSE	Coverage
NMR	0.381	91.2%
NMA	0.381	88.3%
ML-NMR	0.255	98.7%
NMI	0.222	99.0%

- Under non-SEM data, standard NMA is preferable to NMR/ML-NMR.
- ML-NMR shines under a SEM data model.
- NMI performs well either way.
- Read the paper for more scenarios.

Concluding remarks

Summary

- SEM is an extremely restrictive assumption in ITC that is impossible to validate.
- Adjusting for effect modification under a misplaced SEM assumption is worse than not adjusting at all.
- NMI requires extraction of subgroup analyses, but serves as an insurance policy against non-SEM data.
- NMI handles any number of studies and does not limit ITC to the EM conditions of one study.
- Our *Research Synthesis Methods* paper is open access and includes code and datasets.

RESEARCH ARTICLE | Open Access

Network meta-interpolation: Effect modification adjustment in network meta-analysis using subgroup analyses

Ofir Harari, Mohsen Soltanifar, Joseph C. Cappelleri, Andre Verhoek, Mario Ouwens, Caitlin Daly, Bart Heeg

First published: 25 October 2022 | https://doi.org/10.1002/jrsm.1608

Bibliography

References

1. Rittmeyer, A. *et al.* Atezolizumab versus docetaxel in patients with previously treated non-small-cell lung cancer (OAK): A phase 3, open-label, multicentre randomised controlled trial. *Lancet* **389**, 255–265 (2017).

2. Dias, S., Sutton, A. J., Welton, N. J. & Ades, A. E. *NICE DSU technical support document 3: Heterogeneity: Subgroups, meta-regression, bias and bias-adjustment*. (National Institute for Health; Clinical Excellence, 2011).

3. Phillippo, D. M. *et al.* Multilevel network meta-regression for population-adjusted treatment comparisons. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **183**, 1189–1210 (2020).

4. Phillippo, D. M. *Multinma: Network meta-analysis of individual and aggregate data in stan*. (2021).

5. Phillippo, D. M., Dias, S., Ades, A. E. & Welton, N. J. Assessing the performance of population adjustment methods for anchored indirect comparisons: A simulation study. *Statistics in Medicine* **39**, 4885–4911 (2020).

6. Dias, S., Welton, N., Sutton, A. & Ades, A. *NICE DSU technical support document 2: A generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials*. (National Institute for Health; Clinical Excellence, 2011).