# Use of novel methods for population-adjusted indirect comparisons in recent submissions to the National Institute for Health and Care Excellence (NICE)

Mariam Besada[1] and Greta Lozano-Ortega[1]

[1]Broadstreet HEOR, 300 – 177 West 7th Ave, Vancouver BC Canada

SA33

## BACKGROUND

- Health Technology Assessment (HTA) bodies often require comparative effectiveness data for reimbursement decisions. Standard network meta-analysis (NMA) is frequently used, but it requires an assumption of no important differences with respect to effect modifiers across studies. Bias can be introduced if populations differ significantly.
- Population-adjusted indirect comparisons (PAICs), such as matching-adjusted indirect comparisons (MAICs) and simulated treatment comparisons (STCs), address imbalances in baseline characteristics that influence treatment effect. MAIC is the most commonly used PAIC method due to its simplicity and transparency. For unanchored comparisons (i.e., indirect comparisons without a common comparator, often from single-arm trials), MAIC (and STCs) are suitable.
- MAICs require individual patient data (IPD) from at least one study. They are limited to pairwise comparisons and work by reweighting the IPD trial to match the characteristics of the comparator study, which often differs from the population of interest.
- In 2020, Phillippo *et al.* introduced a new PAIC method, multilevel network meta-regression (ML-NMR), an extension of Bayesian NMA. ML-NMR relies on randomized controlled trial (RCT) data and integrates IPD from some trials with aggregate data from others. This method overcomes limitations of existing approaches of adjustment in anchored comparisons (noting that ML-NMR, like NMA, is not relevant to unanchored comparisons), enabling estimates for the population of interest and accommodating larger network structures beyond pairwise indirect comparisons.[1]

## OBJECTIVE

**To identify any use of ML-NMR in NICE submissions to date, as well as challenges with other PAIC methods that may be improved upon with ML-NMR.**

## METHODS

- The study design is summarized in **Figure 1**.

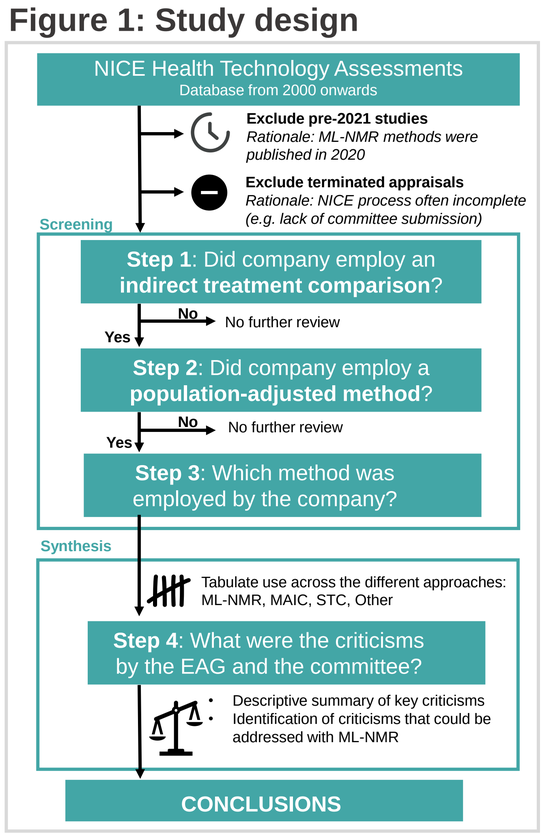
### DATA COLLECTION

- The NICE website was searched to identify completed oncology submissions from January 2021 onwards.
- Each appraisal was assessed by a reviewer and quality checked by a second reviewer.

### SYNTHESIS OF DATA

- The use of any indirect comparison was extracted from the company submission; amongst submissions with an anchored comparison, the evidence assessment group (EAG) report was reviewed, and criticisms were summarized.

**Figure 1: Study design**

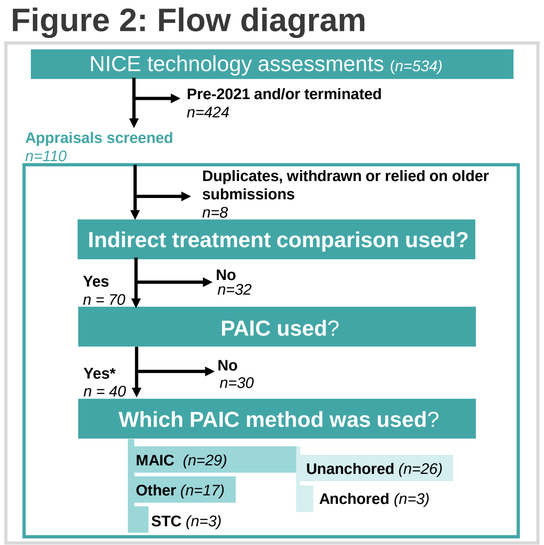



**Abbreviations**: EAG, Evidence Assessment Group; MAIC, Matching-adjusted indirect comparison; ML-NMR, Multilevel network meta-regression; NICE, National Institute for Health and Care Excellence; STC, Simulated treatment comparison

## RESULTS

- The search, conducted on May 1st, 2024, yielded 110 submissions (**Figure 2**); eight were excluded as they were either duplicates, withdrawn or relied on older company submissions.
- Of the 102 remaining, 40 (39%) included at least one population-adjusted method but none employed an ML-NMR.
- MAIC was the most employed method (n =29, 75%).
- Of the 29 MAICs conducted, only three were anchored.

**Figure 2: Flow diagram**




*Some submissions applied more than one method
**Abbreviations**: MAIC, Matching-adjusted indirect comparison; NICE, National Institute for Health and Care Excellence; PAIC, Population-adjusted indirect comparison; STC, Simulated treatment comparison

## RESULTS CONT.

- Five key criticisms across submissions that used anchored MAICs are summarized in **Figure 3**.

**Figure 3: Key criticisms of company submissions involving anchored MAICs**

**#1 APPLICABILITY:**
In situations where population characteristics in the trials differed from those in the UK, treatment effects estimated by the MAIC were scrutinized as they may not accurately reflect expected outcomes for UK patients.

> While an anchored MAIC was conducted between Trial A and Trial B, an unanchored MAIC was requested using a real world evidence dataset given the *"[real world evidence] dataset is more representative of patients in England"*

**#2 EFFECTIVE SAMPLE SIZE (ESS):**
MAICs often suffered from significantly reduced ESS. Exclusion of important covariates for ESS preservation raised concerns about the validity of the relative effects.

> *"The ESS estimate for the [treatment X] arm in Scenario 1 is approximately 54% of the original sample size after weighting, showing a substantial reduction in the number of patients informing the analysis. The ESS estimate for the [treatment X] arm was higher in Scenario 2 [93%]; however, fewer factors were included in the matching process meaning that important effect modifiers may not have been accounted for and therefore, residual confounding may be present"*

**#3 COMPARABILITY:**
Doubts were raised about the comparability of the anchor arms; comparability ensures that differences in outcomes reflect treatment effects rather than baseline population differences.

> *"There are concerns with regard to the comparability of the placebo arms across both [trial X] and [trial Y] trials. Specifically, the company evaluated the treatment effect between the placebo arms of both trials and found that the HR for OS was different from 1.0, although this result was statistically non-significant (HR=0.87; 95% CI [0.67-1.15]; p=0.326)."*

**#4 MISSING COVARIATE DATA:**
Missing covariate data weakened the reliability of estimates as it limits the ability to fully adjust for population differences between trials. Attempts should be made at imputing missing data to allow for potential adjustments.

> *"No attempt was made by the company to impute missing covariate data in the [trial C] trial."*

**#5 IMMATURITY OF DATA:**
The immaturity of data frequently drew scrutiny regarding the choice between time-varying and constant-hazard models, as well as the extrapolations applied in the cost-effectiveness model.

> *"the assumption of proportional hazards add another level of uncertainty […]; the company tried to mitigate this by carrying out a time-varying HR analysis, but the results were such that the degree of uncertainty around the HR render such results unreliable. […] Issues around the extrapolation of immature data and violation of proportionality assumption of the survival analyses further increase the uncertainty around the estimates"*

**Abbreviations**: EAG, Evidence Assessment group; ESS, Effective sample size; HR, Hazard ratio; MAIC, Matching-adjusted indirect comparison; OS, Overall survival; UK, United Kingdom

- We considered whether using an ML-NMR approach could address some of these criticisms or if they would still apply.

### Criticisms that could be addressed by the ML-NMR framework

**#1 APPLICABILITY** **#2 EFFECTIVE SAMPLE SIZE (ESS)**

- **Criticism #1** can, with certain considerations, be addressed by an ML-NMR. Unlike an MAIC, which matches only to a specific trial population, an ML-NMR can produce treatment effect estimates adjusted to the population of interest, provided that there is evidence from all trials on the covariates that define the target population.
- ML-NMR avoids MAIC's pitfalls of discarding or heavily weighting certain observations thereby addressing **criticism #2**. Instead, ML-NMR adjusts for differences between populations in a more flexible way by allowing both individual patient data and aggregate data to be used together in a Bayesian hierarchical model.[2] However, estimates generated via ML-NMR can still be associated with considerable uncertainty if data are insufficient to estimate certain model parameters (*e.g.*, effect modifying covariates).

### Criticisms that would still apply within the ML-NMR framework

**#3 COMPARABILITY** **#4 MISSING COVARIATE DATA** **#5 IMMATURITY OF DATA**

- **Criticisms #3, #4 and #5** would still apply in the ML-NMR context.
- Anchor arms that are assumed to be the same but that may be associated with different efficacy have the potential to introduce bias to the generated results (**criticism #3**), irrespective of the PAIC method used.
- Missing covariate data (**criticism #4**) in any of the trials informing the ML-NMR limits the ability for the covariate(s) to be included for adjustment which is particularly concerning if the covariate conveys treatment effect modification.
- Data immaturity can lead to imprecise estimates of treatment effect, particularly in light of non-proportional hazards in the evidence base (**criticism #5**). While ML-NMR does not require the assumption of proportional hazards, guidance on the estimation of time-varying estimates using the ML-NMR framework has not yet been published.

## DISCUSSION

- As of May 2024, ML-NMR had not been used in NICE oncology submissions. Its recent development likely limited its inclusion in submissions completed by then.
- One further benefit of ML-NMR over MAIC not identified through this review in multiple MAICs. Each MAIC may adjust to a different set of covariates (depending on data availability) and population (that of the comparator trial). This piecemeal approach can introduce inconsistency in the population each estimate would generalize to. ML-NMR provides a unified model, eliminating these inconsistences and ensuring that all treatment effect estimates are generated for the target population of interest. However, as with other PAICs, reliability of the results depends on the correct specification of the regression model.[3]
- This research has several limitations. First, the search was restricted to NICE submissions; expanding the scope to include assessments from other bodies could have provided a broader perspective. Additionally, unadjusted indirect comparisons were excluded; however, these may have also been associated with criticisms by the EAG that could have been addressed with ML-NMR. Lastly, assessments were conducted by a single reviewer with a quality check, rather than being independently reviewed in duplicate, which may impact the rigor of the evaluation process.

## CONCLUSION

**To date, ML-NMR has not been used in NICE oncology submissions; however, given the advantages offered over existing PAIC methods, it is expected that to be increasingly used, particularly in more data rich disease areas with complex networks of clinical trial evidence.**

**REFERENCES**

[1] Phillippo, David M., et al. "Multilevel network meta-regression for population-adjusted treatment comparisons." Journal of the Royal Statistical Society Series A: Statistics in Society 183.3 (2020): 1189-1210. [2] Phillippo, David M., et al. "Population adjustment methods for indirect comparisons: a review of national institute for health and care excellence technology appraisals." International journal of technology assessment in health care 35.3 (2019): 221-228.[3] HTA CG Member State Coordination Group on Health Technology Assessment. " Methodological Guideline for Quantitative Evidence Synthesis: Direct and Indirect Comparisons." (2024).

**Contact: mbesada@broadstreetheor.com**

BROADSTREET HEALTH ECONOMICS & OUTCOMES RESEARCH