PharmaSUG China 2023 - Paper SA-182

# Introduction of Matching Adjusted Indirect Comparison (MAIC)

Cheng Zhang, BeiGene (Beijing) Co.,Ltd.

## ABSTRACT

Head-to-head trials refer to clinical trials conducted under the same trial conditions using clinically used therapeutic drugs or methods as direct comparisons, but such trials are often difficult, expensive and risky. In the absence of direct comparison trials, we can perform indirect comparisons. The most common scenario is that the comparator' data can be from publish summary data or real-world database which won't provide individual patient data due to patient privacy protection. While researchers can obtain individual data for their own study drugs.

Several different statistical frameworks for adjusted indirect comparisons are available, widely used and generally accepted by health technology assessment (HTA) bodies for elicitation of comparative evidence. This article will briefly introduce one of the commonly used methods, the MAIC method.

MAIC is a non-parametric likelihood reweighting method of comparing treatment effects, while minimizing bias that results from prognostic or effect-modifying baseline characteristics that are imbalanced across trial populations. Where the evidence is connected by a common comparator (e.g., trial AB vs trial AC, where common treatment A acts as the common comparator), anchored approaches are preferred because they respect randomization within studies.

Statistical analysis was performed using the IPD on a wide set of baseline characteristics variables to identified potential prognostic variables and/or effect modifiers for study. Then the IPD will be assigned statistical weights in order to adjust comparisons with existing sources of comparative evidence, and mean baseline characteristics (mean and variance or proportion of patients within the category) between individual patients and the comparison population will be balanced. After the matching procedure is performed and weights are derived, analysis containing derived weights are used to compare efficacy and safety outcomes between balanced treatment groups.

## INTRODUCTION

Indirect comparisons and network meta-analysis (NMA) based on aggregate data assume that the effect-modifying variables are similar across studies. However, this may not be true, and population-adjusted indirect comparisons are often used for reimbursement submissions, such as to the National Institute for Health and Care Excellence (NICE). The best way to adjust for patient differences is to use individual patient data (IPD) from all studies and do network meta-regression, as aggregate data network meta-regression has low power and is subject to ecological bias. However, full IPD is rarely available. A common situation is when a company has IPD from its own trial but only published aggregate data from its competitor's trial, usually with average treatment effects and summary patient characteristics. Population adjustment methods use the IPD to balance the differences in the observed covariates between trials. These methods cannot account for other factors, such as treatment delivery, co-treatments, or treatment switching, as these are mixed with treatment. We focus on a new method: Matching-Adjusted Indirect Comparison (MAIC).

## MATCHING-ADJUSTED INDIRECT COMPARISON

### MAIC OVERVIEW

MAIC is a method that allows researchers to compare the effects of different treatments, while reducing the bias that may arise from differences in the baseline characteristics of patients in different trials. These characteristics may influence the outcome of interest or the response to treatment, and they may not be distributed equally across trial populations. To address this issue, MAIC uses a non-parametric likelihood reweighting technique, which assigns weights to individual patients in a trial for which individual patient data (IPD) are available. The weights are calculated using a logistic regression model that matches the weighted summary statistics of the IPD trial to those of a comparator trial population for which only

aggregate data are reported. By doing so, MAIC creates a pseudo-population that is more similar to the comparator population in terms of the baseline characteristics. The treatment effects can then be estimated using the weighted outcomes of the IPD trial and the aggregate outcomes of the comparator trial. The methodology of MAIC, which is illustrated in Figure 1, is well-established and suggested in the NICE guidelines for submissions involving indirect comparisons.

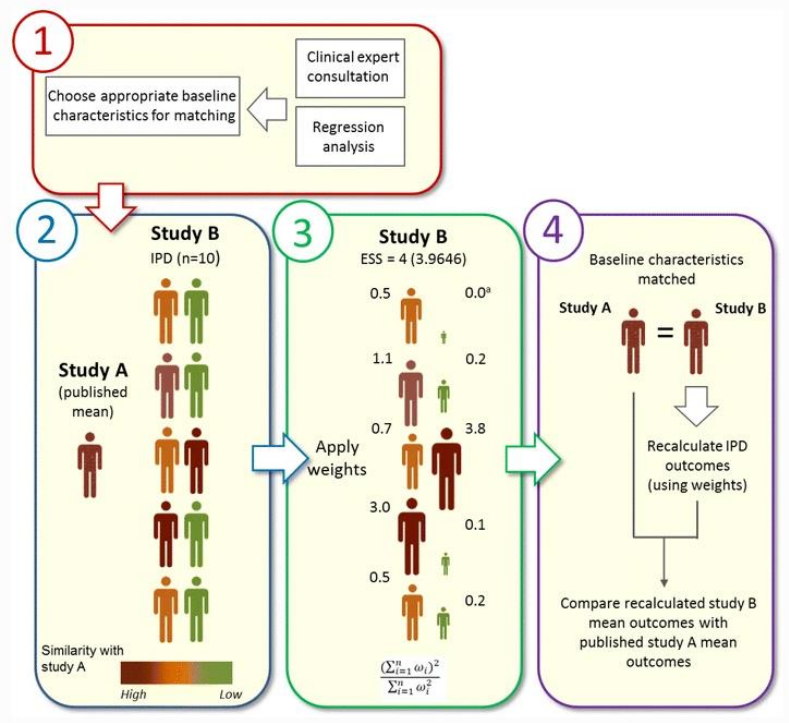


**Figure 1: Overview of the MAIC procedure. Graph adopted from Nash et al., 2018**

MAICs are methods that allow researchers to compare the effects of different treatments based on different trials, when individual patient data (IPD) are available for some but not all trials. MAICs can be classified into two types: 'anchored' or 'unanchored' indirect comparison, depending on whether the trials share a common treatment comparator arm or not. For example, if there are two trials, one comparing treatment A with treatment B (trial AB), and another comparing treatment A with treatment C (trial AC), then treatment A is the common comparator that connects the evidence from both trials. In this case, an anchored approach is preferred, because it preserves the randomisation within each study and avoids confounding by other factors. However, if there is no common comparator between the trials (e.g., trial BD vs trial CE), or if there are single-arm studies that do not compare any treatments (e.g., trial F), then the treatment network is disconnected and an unanchored approach is needed. The Figure 2 illustrates the difference between anchored and unanchored indirect comparisons.

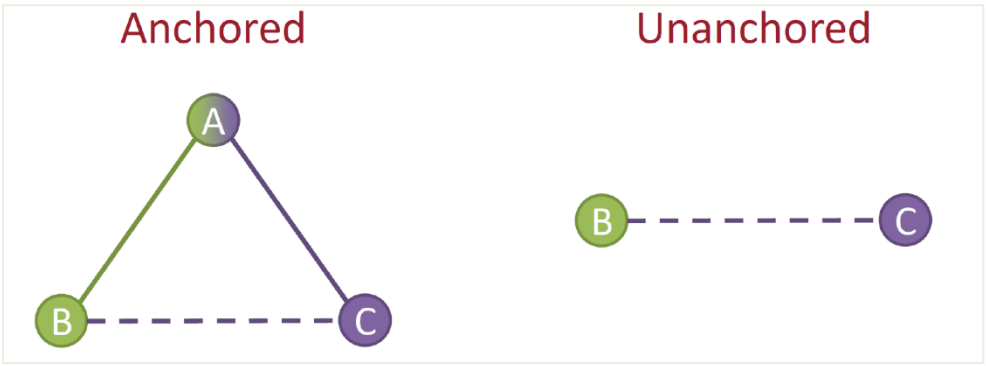


**Figure 2: An anchored versus and unanchored indirect comparison of treatments B and C**

## MATCHING VARIABLES

Effect modifiers (EMs) are variables that affect the magnitude or direction of the treatment effect, meaning that the difference between the treatment and control groups may vary depending on the value of the EM. For example, age may be an EM if the treatment works better for younger than older patients. Prognostic variables (or prognostic factors) are variables that affect the outcome of interest, regardless of the treatment group. For example, gender may be a prognostic variable if males and females have different outcomes on average, but the treatment effect is the same for both genders. EMs may also be prognostic variables, if they influence both the outcome and the treatment effect. For example, smoking status may be both an EM and a prognostic variable if smokers have worse outcomes than non-smokers, and also respond differently to the treatment than non-smokers. The Figure 3 below illustrates the difference between EMs and prognostic variables.

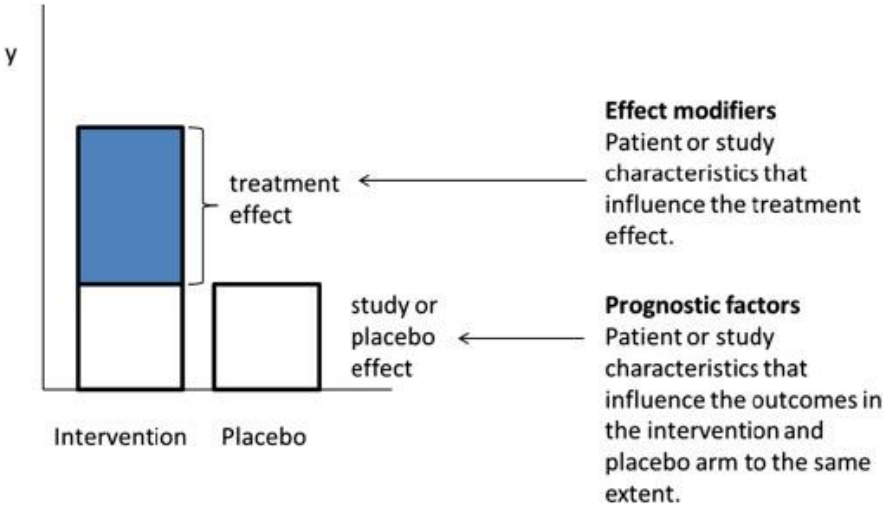


**Figure 2: Prognostic variables/factors and EMs. Graph from Abramson et al., 2020**

## IDENTIFICATION OF EFFECT MODIFIERS AND PROGNOSTIC VARIABLES

The final list of treatment effect modifiers and prognostic factors to be adjusted for in the MAIC analysis will be determined through a deliberative process that considers Key Opinion Leader (KOL) input and statistical analysis conducted using the IPD trials. This process will involve identifying the variables that affect the outcome of interest or the response to treatment, and assessing their distribution and balance across the study populations. The KOL input will provide clinical expertise and insights on the relevance and importance of the variables, and the statistical analysis will provide empirical evidence and estimates of the variables. The aim of this process is to select the most appropriate and informative variables to adjust for in the MAIC analysis, in order to reduce bias and increase precision.

### Quantitative analysis

Two statistical methods are employed for the assessment of EMs:

1. **Likelihood ratio tests (LRT)** between a model that includes the interaction of treatment and covariate of interest, and a model without the interaction.

E.g, the following two models are fitted, where X is the covariate of interest, Y is the outcome.

$$\text{model0: } Y \sim X + treatment$$

$$\text{model1: } Y \sim X + treatment + X * treatment$$

If the covariate is not an effect modifier, the interaction term is not statistically significant given the significance level (e.g., 0.05, 0.1, etc). The p-value of the interaction term can be obtained by LRT where the log transformed likelihood ratio follows chi-square distribution.

R function *anova(model0, model1)* could be used to obtain the p-value from LRT. The potential effect

modifiers will be assessed one by one by repeating the above LRT for respective outcomes in each study.

R code for Likelihood ratio tests:

- Survival outcome:

```
fit.0 <- survival::coxph(Surv(Time,Event)~ Trt + EM, ties = 'efron',data =
sdata)

fit.1 <- survival::coxph(Surv(Time,Event)~ Trt * EM, ties = 'efron',data =
sdata)

p.value <- anova(fit.0,fit.1,test = "Chisq") $`P(>|Chi|)`[2]
```

- Rate outcome:

```
fit.0 <- glm(outcome ~ Trt + EM, data = sdata, family=poisson(link="log"))

fit.1 <- glm(outcome ~ Trt * EM, data = sdata, family=poisson(link="log"))

p.value <- anova(fit.0,fit.1, test = "Chisq")$`Pr(>Chi)`[2]
```

2. **Stepwise selection process** based on Akaike Information Criterion (AIC) using backward elimination and starting with the full model including all covariates of interest, along with their interaction with treatment.

Suppose we have full models:

$$Y \sim trt + X_1 + X_2 + \cdots + X_k + X_1 * trt + X_2 * trt + \cdots + X_k * trt$$

*R* function *stepAIC* can be used to facilitate the backward selection process by the following call:

*step.model <- stepAIC(full.model, direction = "backward", trace = FALSE)*

After selection, the reduced model with the smallest AIC can be obtained. Covariates whose interactions with treatment are remained in the final model are considered as effect modifiers. By calling *step.model$anova*, the final model can be exhibited and the potential effect modifier will be assessed simultaneously for respective outcome in each study.

R code for Stepwise selection process:

- Survival outcome:

```
full.model <- survival::coxph(Surv(Time,Event)~ Trt + EM1 + EM2 +...+ EMk +
Trt:EM1 + Trt:EM2 + ... + Trt:EMk , ties = 'efron',data = sdata)

step.model <- MASS::stepAIC(full.model, direction = "backward", trace =
FALSE)

step.model$anova
```

- Rate outcome:

```
full.model <- glm(outcome ~ Trt + EM1 + EM2 +...+ EMk + Trt:EM1 + Trt:EM2
+ ... + Trt:EMk , data = sdata, family=poisson(link="log"))

step.model <- MASS::stepAIC(full.model, direction = "backward", trace =
FALSE)

step.model$anova
```

The type of actual model employed depends on the type of outcome of interest. In case of survival outcomes, such as OS and PFS, Cox proportional hazards models are used. In case of binary outcomes, binomial generalized linear models (GLM) are employed and in case of rate outcomes, Poisson GLMs.

## ESTIMATION OF MAIC WEIGHTS

To enable an adjusted comparison between the treatments of interest and the available comparative evidence sources, individual patients in the IPD trials will be given statistical weights that adjust for their over- or under-representation relative to the average effect modifying/prognostic variables observed in the comparative evidence source. As a result, after weighting, the average baseline characteristics (mean and variance or proportion of patients within a category) will be matched between the patients in the IPD trials and the comparator populations.

The balancing weights will be estimated using a propensity score-type logistic regression predicting whether a given type of patient originated from the index or the comparator population as a function of population characteristics. More specifically, weight for the *i*th patient in the index population will be estimated by the odds calculated as

$$w_i = exp(\alpha + x_i'\beta)$$

, where $x_i'$ is the vector of population characteristics included in the matching, and $\alpha$ and $\beta$ are coefficients to be estimated from the data.

The $\beta$ coefficients will be determined by the method of moments rather than the maximum likelihood (as is usually the case), because only aggregate data for the $x$'s are typically available for the comparator population [5]. Robust estimators of variance will be used. Once the coefficients are estimated, the equation will be applied to each patient in the index population.

In order to do this, we define the objective function to minimise (as above), and the gradient function (its derivative) which will be used by the minimisation algorithm.

```
objfn <- function(a1, X){
  sum(exp(X %*% a1))
}
gradfn <- function(a1, X){
  colSums(sweep(X, 1, exp(X %*% a1), "*"))
}
```

Here we make use of the sweep function to simultaneously centre the vectors by subtracting matching variables which in the comparative evidence sources.

```
X.EM.0 <- sweep(as.matrix(ipd[, match.var]), 2, agg[match.var], '-')
```

To estimate $\beta$, we use the function optim to minimise the function objfn. The method we tell optim to use is BFGS (after Broyden, Fletcher, Goldfarb and Shanno), which makes use of the gradient function gradfn that we specified to aid minimisation. We have to specify an initial value in the par argument (The zero vector that matches the number of elements), and X = X.EM.0 is passed to objfn and gradfn as an additional argument.

```
optim(par = rep(0,length(match.var)), fn = objfn, gr = gradfn, X = X.EM.0,
method = "BFGS")
```

R code for Estimation weights:

```
N_ipd <- nrow(ab_ipd_data)
N_agg <- nrow(ac_ipd_data)

## select variable for baseline matching
match_var <- c("AGEGE40", "ECOGLE2", "BMKMUT")
COVBL_ipd <- ab_ipd_data %>% select(all_of(match_var))
```

```
COVBL_agg <- map_dbl(ac_ipd_data %>% select(all_of(match_var)),
sum)/nrow(ac_ipd_data)

## MAIC procedure
objfn <- function(a1, X){
  sum(exp(X %*% a1))
}
gradfn <- function(a1, X){
  colSums(sweep(X, 1, exp(X %*% a1), "*"))
}

X.EM.0 <- sweep(as.matrix(COVBL_ipd), 2, COVBL_agg, "-")
opt1 <- optim(par = rep(0, length(match_var)), fn = objfn, gr = gradfn, X =
X.EM.0, method = "BFGS")

## Effective Sample Size
a1 <- opt1$par
wts <- exp(X.EM.0 %*% a1)
```

Following the estimation of the weights, the distribution of the re-scaled weights (calculated as in Equation 1) will be visually inspected to determine whether specific patient(s) or groups of patients (based on covariate values) are over- or under-represented in the analysis. The use of re-scaled weights aids interpretation; a re-scaled weight > 1 means that an individual carries more weight in the re-weighted sample than in the original sample, and a re-scaled weight < 1 means that an individual carries less weight.

**Equation 1: Calculation of re-scaled weights [1]:**

$$Rescaled\ weight_i = \frac{weight_i}{\sum_{i=1}^{n} weight_i} \times N$$

R code:

```
wts_rescaled <- (wts / sum(wts)) * nrow(COVBL_ipd)
```

The robustness of the analysis will also be assessed by estimating the effective sample size (ESS). For a weighted estimate, the ESS is the number of independent non-weighted individuals that would be needed to give an estimate with the same precision as the weighted sample estimate. The calculation of the ESS estimate is specified in Equation 2. A low ESS, relative to the original sample size, is a sign that the weights are highly variable due to a lack of population overlap, and that the resulting relative effect estimate may be unreliable.

**Equation 2: Calculation of the effective sample size [1]:**

$$ESS = \frac{\left(\sum_{i=1}^{n} weight_i\right)^2}{\sum_{i=1}^{n} weight_i^2}$$

R code:

```
ess <- sum(wts)^2 / sum(wts^2)
```

## STATISTICAL ANALYSIS INCORPORATING MAIC WEIGHTS

After the matching procedure is conducted and the weights are derived, efficacy and safety outcomes will be compared between balanced treatment groups using analysis that incorporate the derived weights. These analysis aim to estimate the average outcomes of each treatment group in the target population, after adjusting for the differences in the covariate distributions between the IPD and aggregate populations. For anchored MAICs, the reweighted relative effect (in the form of HR for survival outcomes

and RR for safety outcomes) obtained for the IPD study will be compared with the relative effect reported in the comparator study in order to indirectly obtain the relative effect of the treatment of interest against the relevant comparator treatment.

The balancing weights will be applied to patients in the IPD study to derive adjusted outcomes. For the survival outcomes, the adjusted KM curves will be estimated by a weighted KM analysis and plotted alongside the KM curves of the unadjusted population and the corresponding population in the comparator study to illustrate the direction and the magnitude of the shift due to the adjustment. Similarly, binary outcomes will be tabulated before and after population adjustment along with their 95% CIs.

Hazard ratios along with 95% CI will be reported for the unweighted and weighted Cox proportional regression models to provide naïve and MAIC-adjusted estimate of the relative efficacy.

## DISCUSSION

Treatment effects may differ by population and affect decisions. MAIC compares aggregate data, often from a competitor's study, which may not suit the target population. If the competitor used its IPD and reversed the analysis, results could be opposite. But the real conflict is not the MAIC results, but which population is more representative. Ironically, each company thinks their competitor's trial is more representative. The shared effect modifier assumption can move indirect comparisons to the target population without representativeness. Methods to relax or check this assumption are further research areas.

## CONCLUSION

This paper has introduced Matching Adjusted Indirect Comparison (MAIC) and its role in clinical trials. MAIC is a valuable option when direct head-to-head trials are difficult or impossible. By matching for baseline characteristics and using a common comparator, MAIC reduces bias and enables the comparison of treatment effects. We have explained the main elements of MAIC, such as the matching variables and weighting methods that balance the differences between treatment groups. MAIC is a useful tool for researchers and clinicians who want to make informed treatment choices. However, MAIC is not a replacement for head-to-head trials, and should be combined with other methods to ensure the best and most trustworthy results. We hope this paper has given a helpful introduction to the MAIC method and its uses in clinical research. As comparative effectiveness research advances, we expect that MAIC will have a growing role in the assessment of new treatments and therapies.

## REFERENCES

[1] Phillippo, D, Ades, T, Dias, S, Palmer, S, Abrams, KR & Welton, N 2016, "NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submissions to NICE." Technical Support Documents, vol. 18.

[2] Nash, P, McInnes, IB, Mease, PJ, Thom, H, Hunger, M, Karabis, A, Gandhi, K, Mpofu, S & Jugl, SM 2018, "Secukinumab Versus Adalimumab for Psoriatic Arthritis: Comparative Effectiveness up to 48 Weeks Using a Matching-Adjusted Indirect Comparison." Rheumatology and Therapy, vol. 5, no. 1, pp. 99-122.

[3] Phillippo, D, Ades, T, Dias, S, Palmer, S, Abrams, K & Welton, N 2018, "Methods for Population-Adjusted Indirect Comparisons in Health Technology Appraisal." *Medical Decision Making*, vol. 38, no. 2, pp. 200-211.

[4] Abramson, J.S., et al., 2020, "Lisocabtagene maraleucel for patients with relapsed or refractory large B-cell lymphomas (TRANSCEND NHL 001): a multicentre seamless design study." VOLUME 396, ISSUE 10254, P839-852

[5] Signorovitch, J.E., et al., 2012, "Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research." Volume 15, Issue 6, Pages 940-947.

## CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Cheng Zhang
BeiGene (Beijing) Co.,Ltd.
18120776522
cheng2.zhang@beigene.com