Received: 17 July 2023 | Revised: 11 January 2024 | Accepted: 19 January 2024

DOI: 10.1002/jrsm.1709

SOFTWARE FOCUS

Research Synthesis Methods WILEY

A comprehensive review and shiny application on the matching-adjusted indirect comparison

Ziren Jiang[1] | **Joseph C. Cappelleri**[2] | **Margaret Gamalo**[3] | **Yong Chen**[4] | **Neal Thomas**[2] | **Haitao Chu**[1,2]

[1]Division of Biostatistics and Health Data Science, University of Minnesota School of Public Health, Minneapolis, Minnesota, USA

[2]Statistical Research and Data Science Center, Pfizer Inc., New York, New York, USA

[3]Inflammation & Immunology Statistics, Pfizer Inc., New York, New York, USA

[4]Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania, Philadelphia, Pennsylvania, USA

Correspondence
Ziren Jiang and Haitao Chu, Division of Biostatistics and Health Data Science, University of Minnesota School of Public Health, Minneapolis, MN 55455, USA.
Email: jian0746@umn.edu and chux0051@umn.edu

Abstract

Population-adjusted indirect comparison (PAIC) is an increasingly used technique for estimating the comparative effectiveness of different treatments for the health technology assessments when head-to-head trials are unavailable. Three commonly used PAIC methods include matching-adjusted indirect comparison (MAIC), simulated treatment comparison (STC), and multilevel network meta-regression (ML-NMR). MAIC enables researchers to achieve balanced covariate distribution across two independent trials when individual participant data are only available in one trial. In this article, we provide a comprehensive review of the MAIC methods, including their theoretical derivation, implicit assumptions, and connection to calibration estimation in survey sampling. We discuss the nuances between anchored and unanchored MAIC, as well as their required assumptions. Furthermore, we implement various MAIC methods in a user-friendly R Shiny application Shiny-MAIC. To our knowledge, it is the first Shiny application that implements various MAIC methods. The Shiny-MAIC application offers choice between anchored or unanchored MAIC, choice among different types of covariates and outcomes, and two variance estimators including bootstrap and robust standard errors. An example with simulated data is provided to demonstrate the utility of the Shiny-MAIC application, enabling a user-friendly approach conducting MAIC for healthcare decision-making. The Shiny-MAIC is freely available through the link: https://ziren.shinyapps.io/Shiny_MAIC/.

KEYWORDS

indirect treatment comparison, matching-adjusted indirect comparison, population-adjusted indirect comparison, R shiny, software

Highlights

What is already known

- Matching-adjusted indirect comparison (MAIC) methods are increasingly used in the health technology assessment submissions where there exist imbalance of effect modifiers and restricted access to the individual participant data (IPD).
- Researchers proposed various MAIC methods including the MAIC with method of moment, MAIC with largest effective sample size, and the two-stage MAIC method. These methods estimate a set of weights for the IPD to balance the distribution of selected variables between IPD and aggregate-level data (AgD).

What is new

- We propose a web-based R Shiny tool, "Shiny-MAIC," which estimates and visualizes the MAIC results for various MAIC methods. We also provide step-by-step instructions with an illustrative example using simulated data.
- We perform a comprehensive review of the anchored and unanchored MAIC methods, including different weights estimation, the corresponding required assumptions, and their connections to the entropy balancing, a technique used in survey sampling.
- Shiny-MAIC is freely available at: https://zireen.shinyapps.io/Shiny_MAIC/.

Potential impact for *Research Synthesis Methods* readers

- Shiny-MAIC provides a way for all users to perform various MAIC methods regardless of their experience in R programming.
- The review of MAIC methods offers an introduction to the concept, providing readers with a deeper understanding of the MAIC methods and their connections with entropy balancing.

1 | INTRODUCTION

In medical product development, it is very common that there are multiple treatments available for the same disease condition. To receive regulatory approval, a medicine must demonstrate efficacy compared with placebo or the current standard of care (SoC), commonly evaluated through a randomized control trial (RCT). In order to be reimbursed by some governments and/or private sectors, the medicine at least needs to show noninferiority and cost-effectiveness over the competitor medicine (usually the current SoC) through health technology assessment (HTA) submissions to HTA agencies such as the National Institute of Health and Care Excellence (NICE),[1] Healthcare Improvement Scotland, Australian Government Department of Health and Aged Care, and Canadian Agency for Drugs and Technologies in Health.

While the RCTs are considered the gold standard for comparing different treatments, it is often impractical to conduct a head-to-head randomized trial that compares all available drugs for a particular disease condition. When there is no head-to-head comparison trial, the indirect treatment comparisons (ITCs) provide the highest level of evidence available to guide treatment and reimbursement decisions.[2,3] It is widely accepted and becomes increasingly common in the HTAs.[4] The ITC evaluates comparative effectiveness of two or more treatments by synthesizing the indirect evidence from two or more RCTs, commonly with a common comparator (e.g., the SoC or placebo).

However, the standard ITC such as Bucher's method[5] and network meta-analysis[6–9] assumes that the study populations are similar and the distributions of the effect modifiers are balanced across different RCTs. Thus, there could be substantial bias when the study populations are heterogeneous and/or the distribution of the effect modifiers are not balanced across different trials. Additionally, researchers often only have access to the individual participant data (IPD) of their own trials and the aggregate-level data (AgD) of published trials conducted by others.

The AgD only provides summary statistics (e.g., mean and standard deviation) of patient characteristics included in the trial as well as the population-averaged (or marginal) treatment effect.

To deal with the problem, researchers had developed a set of tools known as population-adjusted indirect comparison (PAIC), which allows for the estimation of comparative effectiveness without access to the IPD of competitors' trials while accounting for between-trial imbalance of the covariates.[4,10] PAIC has three main approaches: matching-adjusted indirect comparison (MAIC), STC, and ML-NMR. The MAIC involves assigning weights to the IPD to match with the summary statistics, such as mean and standard deviation of each covariate, in the published AgD. Examples of these approaches include the traditional MAIC method via methods of moment[11–13] and the calibration estimation method from the survey literature.[14,15] The STC approach[13,16] involves fitting parametric or semi-parametric models to the IPD and extrapolating the fitted model to the competitors' trials with AgD. ML-NMR[17,18] models IPD and AgD in a joint likelihood and integrate evidence over the evidence network.

As the distribution of covariates varies across different populations, the estimation of comparative effectiveness corresponds to a specific population. The MAIC and STC methods are only able to estimate the comparative effectiveness for the population in the competitor's trial. Furthermore, the STC estimates the expected treatment effect of the drug of IPD conditional on the mean covariates value of the aggregate data (so-called conditional effect) which makes it incompatible to the marginal treatment effect of the competitor's drug reported in the competitor's trial.[19] The ML-NMR can also estimate the comparative effectiveness in arbitrary population with additional information of the covariate distribution. Nevertheless, in the particular scenario with two trials with IPD available only in one trial, estimating the comparative effectiveness for a different population requires an untestable assumption that the individual-level relative effects are transportable or transitive across populations.[18]

In this article, we mainly focused on different weighting approaches to balance the covariate distribution in the scenario of two trials with IPD of one trial and only AgD of the other trial. Various statistical methods have been proposed to estimate the weights in recent years. Signorovitch et al.,[11,12] introduced the original MAIC method with weights estimated through method of moments. Petto et al.,[20] proposed a modification to anchored MAIC which balances the covariates separately for the active treatment and common comparator arms in the two studies. Jackson et al.,[21] proposed to estimate the weights that match the moments of covariates and have the largest possible effective sample size (ESS) under the restriction of all weights being non-negative. Remiro-Azócar[22] introduced a two-stage MAIC method that takes into account the treatment assignment mechanism within each study as the first stage, followed by the trial assignment mechanism (which is identical to the MAIC with method of moment) as the second stage. The weights are computed by multiplying the inverse probability weight of the two stages.

Calibration estimation is a technique commonly employed in survey sampling, where the sampling weights are adjusted to ensure that the weighted sample matches some targeted population.[15] Additionally, calibration estimation seeks to identify weights that minimize the distance to a constant weight, thereby decreasing the variability of weights and weighted estimators.[23] Recent literature has demonstrated that the MAIC using the method of moments and the MAIC with the largest ESS are special cases of calibration estimation[23,24] and additional methods can be constructed based on different distance measures implemented in the calibration estimation. Cassidy et al.[25] performed an empirical comparison of the MAIC with method of moment and the standard indirect comparison approach (without adjustment). Several simulation studies have been conducted to evaluate the relative performance of different MAIC methods under various scenarios.[10,20,23,26–28] Yet, no user-friendly tools are available to implement those approaches without some statistical and programming knowledge.

Shiny is an R package that helps users to build a web page implementing R functions. The web-based R Shiny tool does not require the installation of R or any background knowledge of statistical programming. It is particularly helpful for users (regardless of their proficiency in R) who need to run R functions that may not be available in other programming languages. In this article, we develop and present a user-friendly R Shiny application called Shiny-MAIC that enables researchers to implement and compare various MAIC methods.

Compared with the R package "maic," which enables the method of moment for MAIC, the proposed Shiny application incorporates multiple MAIC methods with various visualization approaches evaluating the weights. Specifically, the application implements four different MAIC methods: (1) MAIC with method of moment (equivalent to calibration estimation with entropy distance, see details in Section 2); (2) MAIC with largest ESS (equivalent to calibration estimation with quadratic distance, see details in Section 2); (3) two-stage MAIC; and (4) the Bucher method, which uses equal weights for all subjects.

Shiny-MAIC also offers flexible options for researchers to perform anchored or unanchored MAIC, select different matching types for covariates, and use different variance

estimators including robust and bootstrap standard errors. In addition, it implements flexible weight truncation to remove extreme large weights and mitigate their impact on the treatment comparison. To perform MAIC, researchers need to upload two CSV files, one for the IPD and the other for the AgD, and select their desired options. The application can summarize results from different methods in a table and then create various visualization plots assessing the distribution of the estimated weights.

This article comprehensively reviews different MAIC methods, which reinforces key facets and assumptions of these methods, and develops an R Shiny app. Section 2 presents an overview of MAIC including the notation and key assumptions of both unanchored and anchored MAIC. Section 3 focuses on the four MAIC methods included in our Shiny application. Section 4 provides an overview on how to use the Shiny application, while an example with simulated data is presented in Section 5 to further demonstrate the implementation. Section 6 concludes the article with a brief discussion.

2 | AN OVERVIEW OF MAIC

2.1 | Unanchored MAIC

2.1.1 | Notation and setup of unanchored MAIC

Assume we would like to compare two interventions A and B in two separate single-arm trials A and B where we only have the IPD from trial A. For trial A, denote the K covariates for subject $i = 1,...,n$ as $\boldsymbol{X_i^A} = \left(X_{i1}^A,...,X_{iK}^A\right)$ and the corresponding outcome as Y_i^A. Then the effect of intervention A in the population of trial A can be calculated as the average of the outcome $\overline{Y}_{A(\mathrm{A})} = \frac{1}{n}\sum_i Y_i^A$, where the notation $A(\mathrm{A})$ denotes treatment A's effect in trial A's population. For trial B, since we only have the AgD, denote the sample mean for the K covariates as $\widehat{\mu}_1^B,...,\widehat{\mu}_K^B$ and average outcome for intervention B in the population of trial B as $\overline{Y}_{B(\mathrm{B})}$ with estimated standard error $\widehat{\sigma}_Y^B$. Additionally, denote the corresponding sample standard deviations of the covariates in trial B's population as $\widehat{\sigma}_1^B,...,\widehat{\sigma}_K^B$. It is important to note that MAIC can only adjust covariates that overlap between trials A and B. The effectiveness of MAIC methods may be limited when some important prognostic variables or effect modifiers are not included in the model.

Because of the potential difference between the population for trials A and B, a naïve comparison of two average outcomes $\overline{Y}_{A(\mathrm{A})}$ and $\overline{Y}_{B(\mathrm{B})}$ may produce substantial bias. Instead, the MAIC assigned each subject in the IPD trial (trial A) with a weight w_i $\left(i = 1,...,n \text{ where } \sum_i w_i = 1\right)$ in order to ensure that the weighted population of trial A aligns with the population mean (and standard deviation) of trial B. Then the effect of intervention A in the population of trial B can be estimated as the weighted outcome $\widehat{Y}_{A(\mathrm{B})} = \sum_i w_i Y_i^A$. For matching the population mean, the weights need to satisfy the condition $\sum_i w_i X_{ik}^A = \widehat{\mu}_k^B$, for any k ranging from 1 to K. For the standard deviation of a continuous covariate X_k^A, note that $Var(X_k) = E\left(X_k^2\right) - E(X_k)^2$ and then we use $\sum_i w_i \left(X_{ik}^A\right)^2 = \left(\widehat{\mu}_k^B\right)^2 + \left(\widehat{\sigma}_k^B\right)^2$ to match the standard deviation $\widehat{\sigma}_k^B$ observed in trial B. This is equivalent to adding a new covariate $\left(\boldsymbol{X_k^A}\right)'$ with value equaling $\left(X_{ik}^A\right)' = \left(X_{ik}^A\right)^2$ and matching the target value of $\left(\widehat{\mu}_k^B\right)^2 + \left(\widehat{\sigma}_k^B\right)^2$. The comparative effectiveness of the intervention A and B in the population of trial B is then generated as $g\left(\widehat{Y}_{A(B)}\right) - g\left(\overline{Y}_{B(B)}\right)$, where $g(.)$ is the link function corresponding to the specific type of outcome (e.g., the logit link for the binary outcome).

2.1.2 | Assumptions of the unanchored MAIC

Phillippo et al.[1,4] provided a detailed review of the assumptions required for the anchored and unanchored MAIC in their NICE submission guidelines. In this context, we briefly discuss the assumptions necessary for the unanchored MAIC methods. Since there is no common comparator in the unanchored MAIC, we must extrapolate the absolute outcome of intervention A in the population of trial A to the population of trial B. In order to achieve this, the MAIC methods need to balance all the covariates that can influence the absolute outcomes for intervention A (i.e., the prognostic variables and effect modifiers). Note that effect modifiers are those that influence the relative outcome under treatment (i.e., outcome under treatment subtract by the outcome under control), whereas prognostic variables are those that influence the absolute outcome. Unanchored MAIC requires the so-called conditional constancy of absolute effects assumption, which postulates that the difference in absolute outcomes in different populations can be entirely explained by the imbalances of the included covariates in the corresponding scales.[1,4] It is important to note that this requires the MAIC method to include all the prognostic variables and effect modifiers of the outcome in the weighting model which can be challenging when not all prognostic variables and effects modifiers are known and collected.

As a result, this is considered impractical because one of the main goals of conducting a RCT is to reduce the impact of unknown confounders on the outcome. Therefore, researchers are always recommended to perform

anchored indirect comparison whenever it possible.[4] From the NICE submission guidelines,[1] unanchored indirect comparison should only be considered when the anchored indirect comparison cannot be performed when there is no common comparator or when only single-arm trials are available.

In addition, MAIC requires additional assumptions on the correlations between covariates. Since we typically do not know the correlation matrix between the covariates from the summary level statistics of trial B, the validity of MAIC depends on the untestable assumptions that the true outcome model is additive of the covariates (outcome does not depend on the correlation between covariates) or the correlation of covariates is similar between trials A and B.[4]

2.2 | Anchored MAIC

2.2.1 | Notation and setup of anchored MAIC

As stated above, researchers are recommended to perform the anchored indirect comparison (compared with the unanchored MAIC) when there is a common comparator in the two trials. The anchored indirect comparison has weaker assumptions which are more plausible in practice. Assume that we would like to compare intervention A and B studied in two separate randomized trials with a common comparator C. Let's name the first trial as the AC trial which we have the IPD and the second as the BC trial that we only have AgD. The anchored indirect comparison does not directly compare the average outcomes of intervention A and B; it instead compares the relative treatment effects of A and B as contrast to the common intervention C (i.e., the anchor). For the AC trial, denote the random variable of K covariates for subject $i = 1,...,n_t$ as $\boldsymbol{X_i^t} = (X_{i1}^t,...,X_{iK}^t)$ where $t = A,C$ indicates the treatment allocation, and the corresponding outcome as Y_i^t. Then the relative treatment effect for intervention A in contrast to intervention C in the population of AC trial is

$$\overline{Y}_{AC(\mathrm{AC})} = g\left(\frac{\sum_{i=1}^{n_A} Y_i^A}{n_A}\right) - g\left(\frac{\sum_{i=1}^{n_C} Y_i^C}{n_C}\right). \quad (1)$$

Here, the $g(.)$ is the link function. For the BC trial, since we only have access to the aggregate data, denote the population mean and standard deviation for the K covariates as $\widehat{\mu}_1^{BC},...,\widehat{\mu}_K^{BC}$ and $\widehat{\sigma}_1^{BC},...,\widehat{\sigma}_K^{BC}$. If we want to match the treatment groups and control groups in the two trials separately, we will need the arm level summary statistics $\widehat{\mu}_1^B,...,\widehat{\mu}_K^B$, $\widehat{\sigma}_1^B,...,\widehat{\sigma}_K^B$ and $\widehat{\mu}_1^C,...,\widehat{\mu}_K^C$, $\widehat{\sigma}_1^C,...,\widehat{\sigma}_K^C$ (see next paragraph for more details). The relative treatment effect for intervention B as contrast to intervention C in the BC trial population is estimated by $\widehat{Y}_{BC(\mathrm{BC})}$ with standard error $\widehat{\sigma}_Y^{BC}$. Here, we focus on the marginal (or unadjusted) mean differences in the BC trial. Because even if trial BC reports some adjusted mean differences (e.g., stratified risk difference), it is likely adjusted by a set of covariates that is different from the set of covariates included in the estimation of weights. In return, it can unnecessarily complex the comparison.

As is the case with an unanchored indirect comparison, we cannot simply compare the two relative treatment effects $\overline{Y}_{AC(\mathrm{AC})}$ and $\widehat{Y}_{BC(\mathrm{BC})}$ due to the potential difference between the population of AC trial and BC trial. For anchored MAIC, there are generally two matching strategies to balance the covariates. The first approach is to match all the subjects in the AC trial (as a single population) to the summary statistics of the BC trial, without considering group-specific information. The second approach is to match the treatment groups (intervention A or B) and control groups (intervention C) in the two trials separately. That is, we match the population characteristics of the intervention A group in the AC trial to the summary statistics of subjects from the intervention B group in the BC trial and, in addition, match the population characteristics of the intervention C group in the AC trial to the summary statistics of subjects from the intervention C group in the BC trial, separately. Based on a simulation study comparing different weighting approaches for anchored MAIC, Petto et al.[19] concluded that matching arms separately leads to a more precise estimates when all effect modifiers are captured. Nonetheless, Remiro-Azócar et al.[10] argued that matching arms separately may distort the balance (due to randomization) between treatment arms A and C on covariates that are not accounted for.

With the estimated balancing weights w_i^t for $i = 1,...,n_t$ and $t = A,C$, we can estimate the weighted relative treatment effect of intervention A in contrast to intervention C in the BC trial as

$$\widehat{Y}_{AC(\mathrm{BC})} = g\left(\frac{\sum_{i=1}^{n_A} w_i^A Y_i^A}{\sum_{i=1}^{n_A} w_i^A}\right) - g\left(\frac{\sum_{i=1}^{n_C} w_i^C Y_i^C}{\sum_{i=1}^{n_C} w_i^C}\right). \quad (2)$$

The indirect comparison between A and B is then computed as $\widehat{Y}_{AC(\mathrm{BC})} - \overline{Y}_{BC(\mathrm{BC})}$.

2.2.2 | Assumptions of the anchored MAIC

Because there exists a common comparator group C in both trials, the anchored MAIC requires a weaker assumption compared with the unanchored MAIC.

Specifically, the anchored MAIC extrapolates the relative treatment effect (i.e., the effect of intervention A in contrast to intervention C) instead of the absolute treatment effect (i.e., the outcome of intervention A). Therefore, we only need to balance the covariates that can modify the relative treatment effect for intervention A compared with intervention C, the so-called effect modifiers (or the baseline covariates that can influence relative treatment effect). Thus, the anchored MAIC requires all effect modifiers to be matched by weighting. As same as the unanchored MAIC, the anchored MAIC also requires that the correlations between the covariates in the BC trial either has no impact on the outcome or is similar to the corresponding correlations in the AC trial.

3 | METHODS ESTIMATING THE MAIC WEIGHTS

3.1 | Methods applicable for both unanchored and anchored MAIC

In this subsection, we introduce three MAIC methods that are applicable to both unanchored and anchored indirect comparison. For each method, we present the derivation for the unanchored indirect comparison. For anchored indirect comparison with separate matching approach, researchers can apply the MAIC methods for the intervention groups (A and B) and control groups (C) in the two trials separately to estimate the balancing weights.

3.1.1 | MAIC weights via the method of moment

Signorovitch et al.[12] proposed the original MAIC method that estimates the weights via the method of moments. Specifically, they consider modeling the propensity score of trial allocation (i.e., the probability of a subject being allocated to trial A as contrast to trial B) as follows:

$$\text{logit}(P(T_i = \text{A}|\boldsymbol{X_i})) = \alpha_0 + \boldsymbol{X_i}\boldsymbol{\alpha}, \tag{3}$$

where T_i is the treatment indicator variable of which a participant i has been assigned. The weights for participants in trial A can be calculated as proportional to the inverse of the propensity score ratio:

$$\begin{aligned} w_i &\propto \frac{P(T_i = \text{B}|\boldsymbol{X_i})}{P(T_i = \text{A}|\boldsymbol{X_i})} \\ &= \left(\frac{P(T_i = \text{A}|\boldsymbol{X_i})}{1 - P(T_i = \text{A}|\boldsymbol{X_i})}\right)^{-1} \text{subject to} \sum\nolimits_i w_i = 1, \end{aligned} \tag{4}$$

which is the inverse of treatment or trial assignment odds weight. Because we do not have IPD for participants in trial B, the maximum likelihood method cannot be directly applied to estimate the parameters α_0 and $\boldsymbol{\alpha}$ in the propensity score model. Signorovitch[11,12] estimated the weights via the method of moment. Since we only have the IPD for participants in trial A, based on Equations (3) and (4), we have $w_i \propto \exp\left(-\left(\alpha_0 + \boldsymbol{X_i^A}\boldsymbol{\alpha}\right)\right) \propto \exp\left(-\boldsymbol{X_i^A}\boldsymbol{\alpha}\right)$ subject to $\sum_i w_i = 1$, thus:

$$w_i = \frac{\exp\left(-\boldsymbol{X_i^A}\boldsymbol{\alpha}\right)}{\sum_i \exp\left(-\boldsymbol{X_i^A}\boldsymbol{\alpha}\right)} = \frac{\exp\left(\boldsymbol{X_i^A}(-\boldsymbol{\alpha})\right)}{\sum_i \exp\left(\boldsymbol{X_i^A}(-\boldsymbol{\alpha})\right)} = \frac{\exp\left(\boldsymbol{X_i^A}\boldsymbol{\alpha'}\right)}{\sum_i \exp\left(\boldsymbol{X_i^A}\boldsymbol{\alpha'}\right)} \tag{5}$$

where $\boldsymbol{\alpha'} = -\boldsymbol{\alpha}$. To estimate $\boldsymbol{\alpha'}$, we only need to solve the equation:

$$\sum\nolimits_i w_i x_{ik}^A = \frac{\sum_{\boldsymbol{i}} \exp\left(\boldsymbol{x_i^A}\boldsymbol{\alpha'}\right) x_{ik}^A}{\sum_i \exp\left(\boldsymbol{x_i^A}\boldsymbol{\alpha'}\right)} = \widehat{\mu}_k^B$$

for $k = 1, \ldots, K$ and where $\boldsymbol{x_i^A} = \left(x_{i1}^A, \ldots, x_{iK}^A\right)$ be the observed covariates for subject i in trial A. Note that, here we have K equations corresponding to K parameters $\boldsymbol{\alpha'} = \left(\alpha_1', \ldots, \alpha_K'\right)$. Therefore, there should exist a unique solution of parameters $\boldsymbol{\alpha'}$, unless the covariate matrix $\boldsymbol{X^A} = \left(\boldsymbol{x_1^A}, \ldots, \boldsymbol{x_n^A}\right)$ with multicollinearity such that some equations are redundant. To simplify the computation, define the centered covariate $\tilde{\boldsymbol{X}}_{\boldsymbol{i}}^{\boldsymbol{A}} = \boldsymbol{X_i^A} - \widehat{\boldsymbol{\mu}}^{\boldsymbol{B}}$, where $\widehat{\boldsymbol{\mu}}^{\boldsymbol{B}}$ is the mean value of the covariates in the AgD study (i.e., trial B). Then the right-hand side of the equation becomes zero:

$$\frac{\sum_{\boldsymbol{i}} \exp\left(\tilde{\boldsymbol{X}}_{\boldsymbol{i}}^{\boldsymbol{A}}\boldsymbol{\alpha'}\right)\tilde{\boldsymbol{X}}_{\boldsymbol{i}}^{\boldsymbol{A}}}{\sum_i \exp\left(\tilde{\boldsymbol{X}}_{\boldsymbol{i}}^{\boldsymbol{A}}\boldsymbol{\alpha'}\right)} = \frac{\sum_{\boldsymbol{i}} \exp\left(\boldsymbol{X_i^A}\boldsymbol{\alpha'}\right)\left(\boldsymbol{X_i^A} - \widehat{\boldsymbol{\mu}}^{\boldsymbol{B}}\right)}{\sum_i \exp\left(\boldsymbol{X_i^A}\boldsymbol{\alpha'}\right)} = \boldsymbol{0}.$$

Because the denominator is always positive, the above equation is equivalent to:

$$\sum\nolimits_{\boldsymbol{i}} \exp\left(\tilde{\boldsymbol{X}}_{\boldsymbol{i}}^{\boldsymbol{A}}\boldsymbol{\alpha'}\right)\tilde{\boldsymbol{X}}_{\boldsymbol{i}}^{\boldsymbol{A}} = \boldsymbol{0}.$$

Note that the left-hand side is the derivative of the function:

$$g(\boldsymbol{\alpha'}) = \sum\nolimits_{\boldsymbol{i}} \exp\left(\tilde{\boldsymbol{X}}_{\boldsymbol{i}}^{\boldsymbol{A}}\boldsymbol{\alpha'}\right).$$

Then, solving this equation is equivalent to minimizing the object function $g(\boldsymbol{\alpha'})$, which can be easily implemented through the optimization functions in R[21].

For the method of moment, it estimates the coefficient $\boldsymbol{\alpha}'$ instead of directly estimating the weights w_i, but still makes an implicit assumption of propensity score model. Because the length of vector $\boldsymbol{\alpha}'$ equals the number of equations $\sum_i \exp\left(\tilde{\boldsymbol{X}}_i^{\boldsymbol{A}}\boldsymbol{\alpha}'\right)\tilde{\boldsymbol{X}}_i^{\boldsymbol{A}} = \boldsymbol{0}$, the solution of $\boldsymbol{\alpha}'$ will be unique if it exists. Also note that the solution may not exist when all the covariates $\boldsymbol{X}_i^{\boldsymbol{A}} < \widehat{\boldsymbol{\mu}}^{\boldsymbol{B}}$ or $\boldsymbol{X}_i^{\boldsymbol{A}} > \widehat{\boldsymbol{\mu}}^{\boldsymbol{B}}$ which indicates that the overlap of the two populations is extremely poor, and the population weighting is not useful in this situation.

3.1.2 | MAIC weights by maximizing the ESS

Jackson et al.[21] proposed an alternative weighting approach which, unlike the method of moment, directly estimates the weights w_i for each participant in trial A such that: (1) the weights $\boldsymbol{w} = \{w_i\}_{i=1}^n$ balance the covariate distribution for the two trials in the sense that $\sum_i w_i X_{ik}^A = \widehat{\mu}_k^B$ for $k = 1, \ldots, K$; (2) the same weights $\boldsymbol{w}$ are non-negative and sum to 1; and (3) the same weights $\left[w = \{w_i\}_{i=1}^n\right]$ give the largest ESS defined as $ESS(\boldsymbol{w}) = \left(\sum_i w_i^2\right)^{-1}$. Therefore, the weight $\boldsymbol{w}$ can be solved by the following optimization problem:

$$\min_{\boldsymbol{w}} \left\{\sum_i w_i^2\right\} \text{ subject to } \sum_i w_i \boldsymbol{X}_{\boldsymbol{i}}^{\boldsymbol{A}} = \widehat{\boldsymbol{\mu}}^{\boldsymbol{B}}, \sum_i w_i = 1 \text{ and } w_i \geq 0 \text{ for all } i.$$

Since assigning patients with negative weights in the IPD trial is not permitted in the standard method, the last constraint requires that all weights being larger than or equal to 0. Note that here we directly solve the weights $\boldsymbol{w} = \{w_i\}_{i=1}^n$ from the optimization problem constrained with the specified conditions. There are n unknown variables $\{w_i\}_{i=1}^n$ with only $K+1$ restricted equations leaving $n-K-1$ degrees of freedom for the feasible weights. This method selects the weight that have the largest ESS $\left(\sum_i w_i^2\right)^{-1}$ among all feasible weights (i.e., non-negative weights that satisfy the $K+1$ equations).

The above optimization problem can be transformed to a standard optimization problem proposed in Zubizarreta,[29] which can be solved by an R function; see the derivation in Jackson et al.[21] for more details on R programming. Jackson and colleagues also demonstrate that, through the analytical result for the scenario with only one covariate, the weights are a linear combination of the covariate which is substantially different from the MAIC with method of moment where the weights are the exponentiated linear combination of covariates. These researchers also indicated that the MAIC via the method of moment is close to the MAIC with largest ESS when the population of the two trials is similar (such that the weights are closer to the constant weights).

3.1.3 | MAIC weights via calibration estimation

Calibration estimation is a widely used method in survey sampling that aims to adjust the sampled variables to match some target population characteristics. This method was first introduced by Deville and Sarndal[15] which defined calibration estimation as a procedure for minimizing the distance between the estimated weights and some basic survey weights (e.g., the uniform weights). In other words, calibration estimation selects weights that can match the target mean (and standard deviation) while minimizing its variability. This method is often used when researchers have a desired population mean (and standard deviation) for some covariates, which is exactly the same situation as in an PAIC. More specifically, calibration estimation determines the weights by solving the following optimization problem:

$$\min_{\boldsymbol{w}} \sum_i D\left(w_i, \frac{1}{n}\right) \text{ subject to } \sum_i w_i \boldsymbol{X}_{\boldsymbol{i}}^{\boldsymbol{A}} = \widehat{\boldsymbol{\mu}}^{B}, \sum_i w_i = 1, \text{ and } w_i \geq 0 \text{ for all } i.$$

where $D\left(w_i, \frac{1}{n}\right)$ is some measurement of the distance between the weight w_i and the uniform weight $\frac{1}{n}$. The smaller the distance, the less variability of the weight, making the weights more efficient. There are different ways to define measurements for the variability of the weights, leading to different types of calibration estimation weights. Among the several typical choices of the distance measure D are entropy distance $D\left(w, \frac{1}{n}\right) = w \log w$, the quadratic distance $D\left(w, \frac{1}{n}\right) = \left(w - \frac{1}{n}\right)^2$, and the absolute distance $D\left(w, \frac{1}{n}\right) = \left| w - \frac{1}{n} \right|$.

3.1.4 | MAIC weights via the method of moment and maximizing ESS are special cases of the calibration estimation

Recent researchers have pointed out the connection between the calibration estimation to the original MAIC via the method of moment and the MAIC method with the largest ESS.[23,24] In fact, the MAIC via the method of moment is equivalent to the calibration estimation with the entropy distance. The MAIC with the largest ESS is equivalent to the calibration estimation with quadratic distance. We briefly elaborate the derivation by Philippo

et al.[24] for the MAIC via the method of moment and Wang[23] for the MAIC with largest ESS. For calibration estimation with the entropy distance, the optimization problem is $\min_{\boldsymbol{w}} \sum_i w_i \log w_i$ subject to $\sum_i w_i \boldsymbol{X_i^A} = \widehat{\boldsymbol{\mu}}^{\boldsymbol{B}}$, $\sum_i w_i = 1$ and $w_i \geq 0$ for all i. Hainmueller[30] proposed a dual optimization without any constraint using the Lagrange multipliers by $\min_{\boldsymbol{\alpha}} \log\left(\sum_{i=1}^{n} \exp\left(\tilde{\boldsymbol{X}}_{\boldsymbol{i}}^{\boldsymbol{A}}(-\boldsymbol{\alpha})\right)\right)$.

Since the logarithm is a monotonic function, Philippo et al.[24] pointed out that minimizing this objective function is equivalent to minimizing $g(\boldsymbol{\alpha}')$ as in the MAIC by method of moment. For the calibration estimation with quadratic distance, note that the quadratic distance can be decomposed as $\sum_{i=1}^{n}\left(w_i - \frac{1}{n}\right)^2 = \sum_{i=1}^{n} w_i^2 - \frac{2}{n}\sum_{i=1}^{n} w_i + \frac{1}{n} = \sum_{i=1}^{n} w_i^2 - \frac{1}{n}$. Because it is a monotonically increasing function of $\sum_{i=1}^{n} w_i^2$, the calibration estimation with quadratic distance finds the weights $\min_{\boldsymbol{w}} \sum_{i=1}^{n} w_i^2$, subject to $\sum_{i=1}^{n} w_i \boldsymbol{X_i^A} = \widehat{\boldsymbol{\mu}}^{\boldsymbol{B}}$ and $\sum_i w_i = 1$, which is equivalent to the MAIC method with largest ESS as $ESS(\boldsymbol{w}) = \left(\sum_i w_i^2\right)^{-1}$.

3.2 | Method only for anchored MAIC

3.2.1 | The two-stage MAIC method

For the anchored indirect comparison, Remiro-Azócar et al.[22] proposed a two-stage MAIC method which is an extension to the MAIC method of moment. The two-stage MAIC method employs two parametric models: one estimates the treatment assignment mechanism in the study with individual participant data (IPD), whereas the other estimates the trial assignment mechanism among the two studies. These models produce inverse probability of treatment weights within IPD and trial assignment odds weights across trials, respectively. The two weights are combined to generate the final weights that balance covariates between treatment arms and across trials. Specifically, for the first stage, it fits the propensity score model for the treatment assignment mechanism in the IPD study (AC trial):

$$\text{logit}(e_i) = \beta_{0i} + \boldsymbol{x_i}\boldsymbol{\beta_1},$$

where e_i be the probability of subject i being assigned to intervention A as contrast to the common comparator C. The second stage is to calculate the weights that balance the covariates between the AC trial and the BC trial, which can be either the MAIC via the method of moment, via maximizing the ESS, or via calibration estimation. Then, the final weights $\widehat{\omega}_i$ can be computed by:

$$\widehat{\omega}_i = \frac{t_i \widehat{w}_i}{\widehat{e}_i} + \frac{(1-t_i)\widehat{w}_i}{(1-\widehat{e}_i)}.$$

Here, $\widehat{w}_i$ is the estimated weights for subject i in the second stage, t_i is the indicator of treatment allocation for subject i with $t_i = 1$ if subject i is in the group A and $t_i = 0$ if subject i is in the group C, $\widehat{e}_i$ is the predicted propensity score for subject i in the first stage. With the final weights $\widehat{\omega}_i$, the indirect comparison can be implemented directly following Section 3.1.

Based on a simulation study, Remiro-Azócar et al.[10] reported that the two-stage MAIC method improved precision and efficiency relative to the standard MAIC, particularly in situations with low IPD trial sample sizes. Nevertheless, its effectiveness decreases when the overlap between trial target populations is poor and the extremity of weights is high.

3.3 | Statistical inference

Once we estimate the comparative effectiveness $\overline{Y}_w$, we need to estimate the uncertainty of $\overline{Y}_w$ in order to perform the statistical inference. There are generally two approaches for estimating the uncertainty of $\overline{Y}_w$: the nonparametric bootstrap method and the robust sandwich estimator. Signorovitch et al.[12] proposed to use the robust sandwich estimator to estimate the uncertainty that incorporates the fact that the weights are estimated rather than fixed.[4] Yet, the sandwich estimator relies on the large-sample properties and can potentially underestimate the variability of MAIC estimator when the ESS is small.[10] On the other hand, the nonparametric bootstrap method involves sampling the IPD in trial A with replacement and calculating the treatment effect $\overline{Y}_w^b$ for the bth bootstrap iteration, where $b = 1, ..., B$. The uncertainty of $\overline{Y}_w$ can then be estimated by the standard deviation of the bootstrap sample or through the percentile of the bootstrap sample (e.g., 2.5, 97.5 percentiles). Nevertheless, using the bootstrap method also presents some challenges. If there is poor overlap between the two populations, resampling the IPD could worsen this issue leading to extremely high weights for some subjects, inflating the estimated variance. As a consequence, we do not recommend using the nonparametric bootstrap method as the default method, and users should be more cautious as the results may be unstable.

4 | THE R SHINY APPLICATION: SHINY-MAIC

In this article, we develop an R Shiny application Shiny-MAIC that implements various methods for anchored and unanchored MAIC. We offer some guidance in this

FIGURE 1 Roadmap for using Shiny-matching-adjusted indirect comparison (MAIC). Summary of procedures presented in Sections 4.1–4.3. [Colour figure can be viewed at wileyonlinelibrary.com]

section, which outlines the application's functions and features. Users can refer to Section 5 for an illustrative example using a simulated data from the example 2 in Jackson et al.[21] The Shiny-MAIC consists of three pages: (1) "Home" page provides basic information about the app; (2) "Load Data" page enables users to upload IPD and AgD for indirect comparison and choose different matching options (such as matching only the mean value or both mean and standard deviation); and (3) "Result" page displays the outcomes of the indirect comparison. Figure 1 is a flow diagram that summarizes the steps for using Shiny-MAIC as a roadmap.

4.1 | Load data

In order to get results from Shiny-MAIC, users should have the IPD for the first trial (the user's own trial) and

the AgD for the second trial (the competitor's trial). For the IPD, users can directly upload the corresponding CSV file through the "Load Data" page. The CSV file for IPD should contain the covariates, the outcome variable, and the treatment indicator variable (only for anchored MAIC). However, there is no restriction on the variable names, and the CSV file can have variables that are not included in the MAIC model (i.e., users can decide which variables to balance).

Once the IPD is uploaded, the application will automatically display a summary table containing each variable from the IPD and the corresponding summary statistics (mean and standard deviation) for the IPD. Then, users can directly input the corresponding target mean and standard deviation (if applicable) in the second trial (i.e., the AgD trial) by double-clicking the table cells. In order to perform the indirect comparison, users should let Shiny-MAIC know (by selecting the corresponding type in the summary table): (1) the outcome variable (i.e., Y variable), (2) the corresponding type of the outcome variable (i.e., continuous, binary, or survival outcome), (3) the treatment indicator variable (only for anchored MAIC), (4) the variables that only match the mean value, (5) the variables that match both the mean and standard deviation, and (6) the variables that is not used in the application. If users select to separately match the treatment group and control group, they should provide additional information of the mean and standard deviation values for each variable in the control group of the second trial. Instead of directly inputting the aggregate data for indirect comparison, users can also upload this table in a CSV file with the corresponding column names and variable names.

Users should note that Shiny-MAIC has several restrictions on the input value of the IPD. First, Shiny-MAIC only allows for the dichotomized categorical variable with the value either 0 or 1. For the categorical variable with multiple categories, users should first convert it into several dichotomized variables (with 0 and 1 value) before uploading to Shiny-MAIC. Second, users should label the treatment indicator with either 0 (representing the shared common intervention C among the two trials) or 1 (representing the treatment group A which is intended to be compared with the competitor's trial treatment B).

4.2 | Flexibility of Shiny-MAIC

After uploading the IPD and AgD for MAIC, users can navigate to the "Result" page for additional customization of the indirect comparison. Users can first choose between anchored and unanchored MAIC. For unanchored MAIC, available methods include the Bucher method, the MAIC with the method of moments, and the MAIC method with the largest ESS. For anchored MAIC, the two-stage MAIC method is also selectable. Regarding statistical inference, users can opt for either the nonparametric bootstrap method or the robust sandwich estimator to estimate the standard error of the respective MAIC estimators.

Shiny-MAIC also allows users to truncate weights. Truncation is a technique that replaces extreme weights with a specified percentile of the estimated weights aiming to reduce the influence of extreme weights (outliers).[31] Extreme weights can occur when there is poor overlap between the target populations in the studies being compared, potentially increasing the variance. By truncating or capping these weights at a predetermined threshold, the influence of extreme weights on the analysis can be limited.

Nonetheless, it is essential to note that truncation may introduce some bias into the estimation, as it can lead to some loss of information and may not perfectly balance the covariates between the treatment arms; see Remiro-Azócar et al.[22] for a simulation study involve truncation. Therefore, careful consideration should be given to the choice of truncation threshold and its potential impact on the overall analysis. In a specific context, the optimal approach for implementing truncation in MAIC should be determined by evaluating the trade-off between bias reduction and precision improvement.

4.3 | Results from Shiny-MAIC

The results for the indirect comparison are displayed at the right panel of the "Result" page with several separate windows. The first window displays the summary table of the indirect comparison results for different MAIC methods, including the estimated comparative effects along with their estimated standard errors and 95% confidence intervals. We also display the estimated ESS for the corresponding weights. The second window shows a histogram (along with the smoothed empirical probability density curve in the fourth window) of the estimated weights for each method.

The third window displays a plot for visualizing estimated weights with the y-axis representing the order of the weights and the x-axis reflecting the weight value. The red-dashed line represents the average weights that give all subjects the same preference (same as the weights in the Bucher method). Note that, to clearly visualize the weights, we adjust them to have an average value of 1 in these plots. Therefore, this plot can be seen as an empirical CDF of the estimated weights with y-axis goes to the sample size in IPD. The fourth window displays a plot for the piece-wise comparison of each estimated weights.

Users can examine the relationship between the weights estimated with different methods by using scatter plots and the calculated correlation value. The final window displays the mean and standard deviation of the covariates after weighting for each method. Shiny-MAIC also allows users to download all the tables and plots in a .zip file.

Although Shiny-MAIC provides a convenient tool to conduct MAIC in a web environment, it faces the challenges of the data privacy policy. In practice, uploading the private IPD online may be prohibited as it may violate the data privacy policy. Thus, we provide the code for conducting the MAIC methods in the "MAIC Code" window. The code is generated automatically according to the settings in the left panel. Users need to substitute only the data.AC for the IPD and data.BC for the AgD to get the MAIC results.

5 | AN ILLUSTRATIVE EXAMPLE

We now illustrate the use of our Shiny-MAIC through an example from Jackson et al.[21] This example uses

Method	Indirect Comparison Estimate	SE	Inference method	95% CI Lower	95% CI Upper	Effective sample size
Bucher	−0.324	0.413	Robust sandwich estimator	−1.135	0.486	500.000
MAIC	0.032	0.518	Robust sandwich estimator	−0.983	1.046	185.602
MAIC_ESS	−0.038	0.517	Robust sandwich estimator	−1.052	0.976	190.789
MAIC_2S	0.031	0.518	Robust sandwich estimator	−0.983	1.046	185.598

FIGURE 2 Shiny-MAIC output for the result of the illustrative example. The second and third columns are the estimated log-odds ratio between intervention *A* and intervention *C* and its robust standard error. The fourth and fifth columns are the lower and upper bounds of the estimated comparative effectiveness. The last column is the effective sample size (ESS) of the weights.

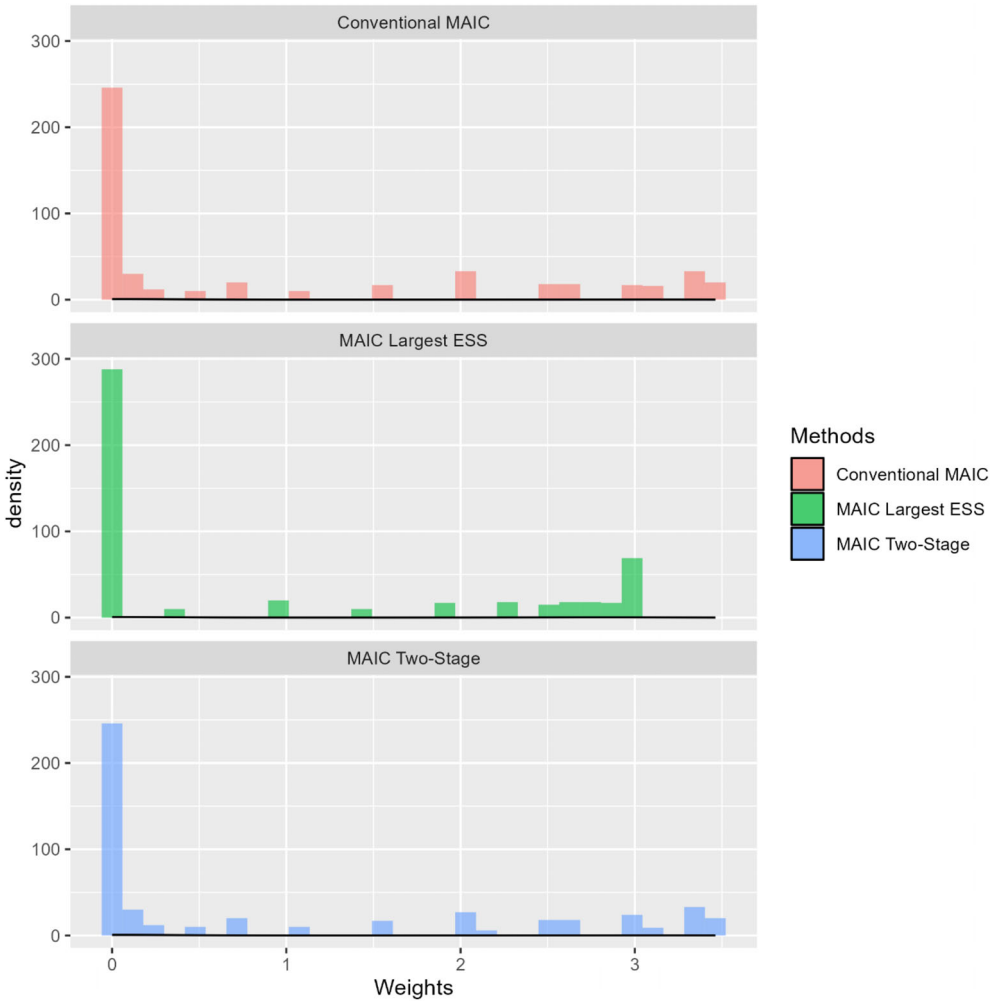

FIGURE 3 Histogram of the estimated weights. Histogram for the weights are estimated by matching-adjusted indirect comparison (MAIC) with method of moment, MAIC with largest effective sample size (ESS), and the two-stage MAIC. The weights are adjusted to have an average value of 1. [Colour figure can be viewed at wileyonlinelibrary.com]

simulated data with a realistic setting which are originally proposed by the NICE DSU Technical Support Document 18: Methods for PAICs in submissions to NICE.[1] We assume the AC trial has 500 patients and the BC trial has 300 patients. The age of the patients is uniformly distributed from 45 to 75 years old in the AC trial while being uniformly distributed from 45 to 55 years old in the BC trial. The proportion of female is 0.64 in the AC trial and is 0.8 in the BC trial. Both trials assume half of their patients are randomized in the intervention groups (A and B) and the shared control groups (C). The outcome of interest is a binary outcome with probability derived from a multivariate logistic regression equation:

$$\begin{aligned}\text{logit}(p_{it}) = {} & 0.85 + 0.12\text{male}_{it} + 0.05(\text{age}_{it} - 40) \\ & + (\beta_t - 0.08(\text{age}_{it} - 40))I(t \neq C),\end{aligned}$$

where $t = A$, B or C indicating the treatment, p_{it} is the probability of the recurrence of a disease for patient i who is receiving intervention t, and $I(t \neq C)$ is the indicator function for $t \neq C$, $\beta_A = -2.1$ and $\beta_B = -2.5$. From the model, we can see that the age is an effect modifier while the gender is a prognostic variable.

Assuming that with the clinical information, data analyst correctly assumes that gender is not an effect modifier and adheres to the NICE guidelines,[1] choosing to adjust the age variable only for its mean and standard deviation. Once the data are uploaded, we set the "match_type" for the age variable to "mean & sd" and for the gender variable to "none." We then proceed to the "Result" page for further MAIC method settings. As we have a common comparator group C in this example, we opt for the anchored MAIC with a robust sandwich estimator. We first examine the results of all four MAIC methods without weight truncation to prevent potential bias, with the results displayed in the right panel.

In line with the example 2 from Jackson et al.,[21] the three MAIC methods yield similar results (Figure 2), while the unadjusted method (Bucher method) appears to introduce bias in the indirect comparison of log odds ratio. The MAIC with the method of moments and the two-stage MAIC have nearly identical results up to the second decimal point, likely due to the random

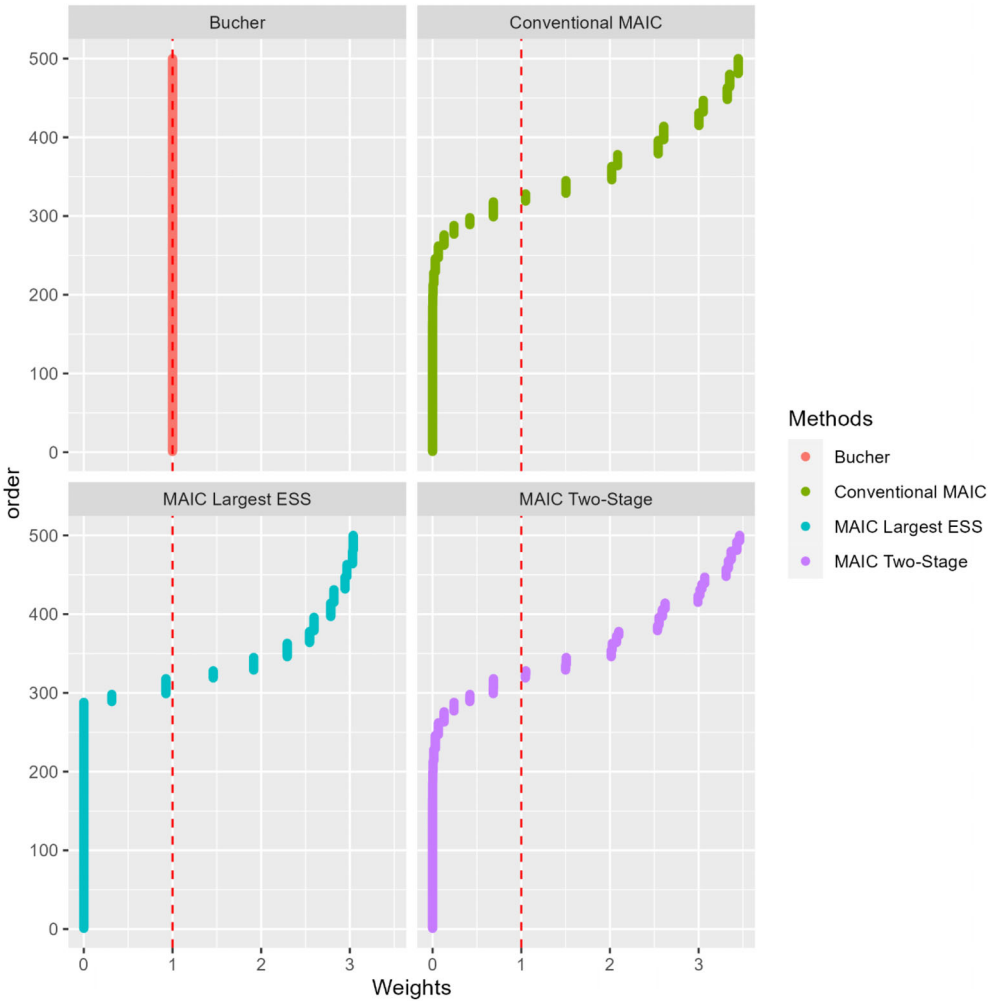

FIGURE 4 Empirical cumulative distribution function (CDF) of the fitted weights. The ranks of the weights are plotted against the weights value. Note that the plot is equivalent to the empirical CDF curve if we draw a curve between the points and divide the *y*-axis by the sample size. The weights are adjusted to have an average value of 1. [Colour figure can be viewed at wileyonlinelibrary.com]

allocation and perfect balance of treatment across different ages. The MAIC method with the largest ESS has a slightly larger ESS relative to the other MAIC methods, although the three methods have similar results in this example. We proceed to examine the estimated weights in the other windows.

In the "Weights Distribution" window (see Figure 3), most MAIC weights are close to zero, whereas other weights are uniformly distributed between 0 and 0.007. The "Weights Check" window (Figure 4) displays the empirical CDF of the different weights. Note that, both Figure 3 and Figure 4 adjust weights to have an average value of 1 for visualization purpose. Since there are no apparent extreme weights or outliers, we do not truncate the weights. Users can perform a sensitivity analysis to truncate weights if they identify any outliers. The

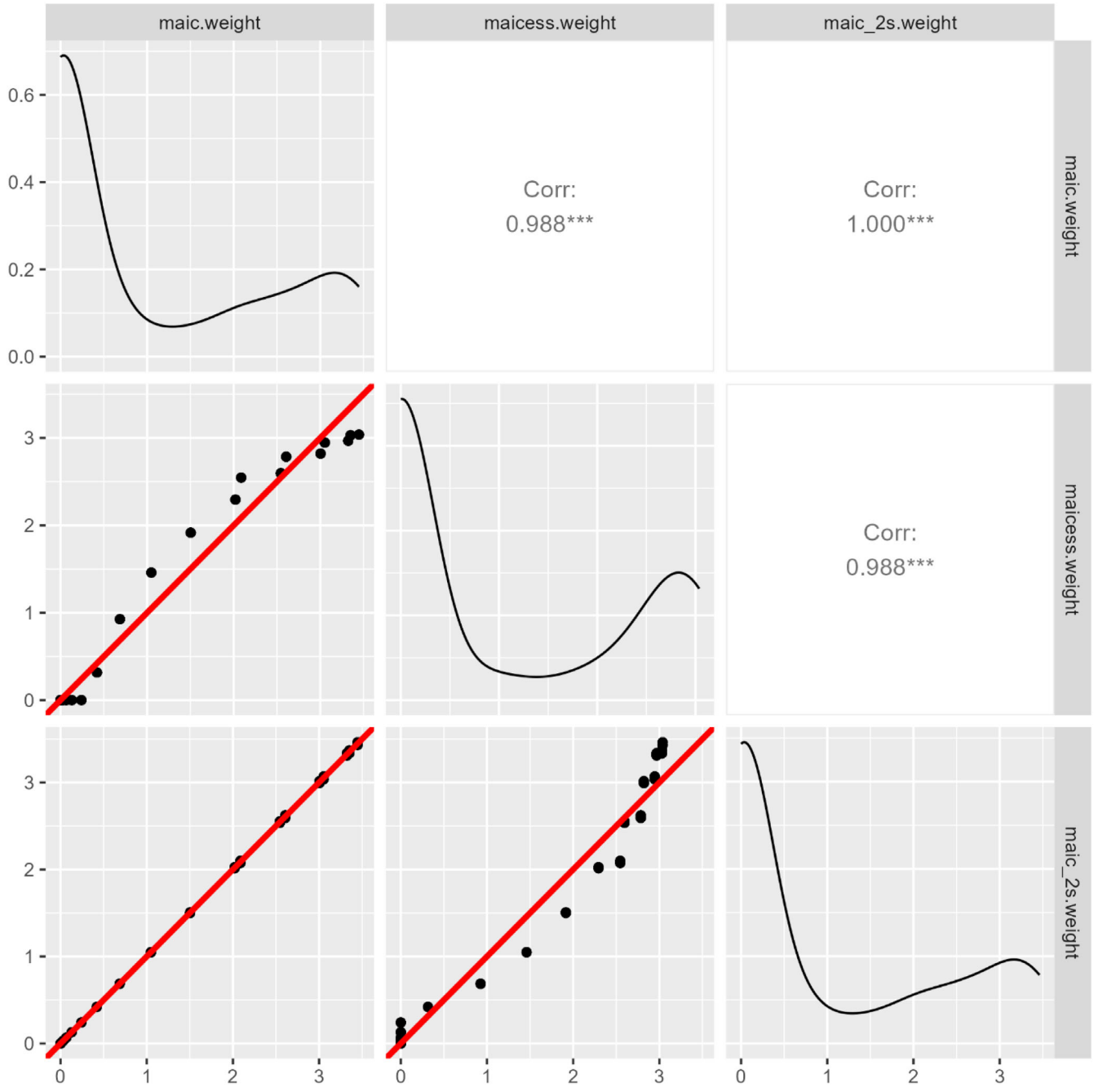

FIGURE 5 Pair-wise comparison of the weights estimated with different matching-adjusted indirect comparison (MAIC) methods. The pair-wise comparison scatter plots for the weights are estimated by MAIC with method of moment, MAIC with largest effective sample size, and the two-stage MAIC. Burcher method is not included in this plot as it uses constant weights. ***The pairwise correlation suggests the weights estimated by different methods are highly correlated. [Colour figure can be viewed at wileyonlinelibrary.com]

Summary of the weighted mean

Variable	IPD.mean	AgD.mean	Match.type	Bucher.mean	MAIC_MM.mean	MAIC_ESS.mean	TSMAIC.mean
age	60.05	49.81	sd	60.05	49.81	49.81	49.81

Summary of the weighted standard deviation

Variable	IPD.sd	AgD.sd	Match.type	Bucher.sd	MAIC_MM.sd	MAIC_ESS.sd	TSMAIC.sd
age	9.17	3.08	sd	9.16	3.08	3.08	3.08

FIGURE 6 Shiny-MAIC output for the summary tables of the balancing performance. The MAIC methods aim to match the mean and standard deviation for the age variable. IPD.mean and IPD.sd are the sample mean and standard deviation of the IPD study. AgD.mean and AgD.sd are the targeted mean and standard deviation that we would like to be adjusted for after weighting. The Match.type are the corresponding matching type for the variable age, with "sd" indicates to match both the mean and standard deviation. The rest of the columns are the weighted sample means and standard deviations correspond to different methods, including Bucher method, MAIC with method of moment, MAIC with largest effective sample size, and the two-stage MAIC.

"Weights Comparison" window (Figure 5) presents a pairwise comparison of the different weights, resembling Figure 2 in Jackson et al.[21] Finally, users can check the summary statistics for the weighted IPD sample in the "Appendix" panel (Figure 6). While in most cases the weighted means (and standard deviations if applied) align with the target AgD means (and standard deviations if applied), imbalances may occasionally occur, particularly in the presence of multicollinearity.

6 | CONCLUSION

In this article, we review different MAIC methods for PAIC when we only have access to IPD in one trial. We developed a user-friendly R Shiny application called Shiny-MAIC that enables researchers to perform different MAIC methods without the need for R programming expertise. We also provide guidance for users to use Shiny-MAIC with an illustrative example. The application offers several MAIC methods, anchored and unanchored options, flexibility in matching types for covariates, and a choice of variance estimators. By providing an accessible tool for conducting MAICs, researchers can easily make more informed decisions in the context of multiple competing drugs for HTAs.

There are other MAIC methods that are not comprehensively reviewed in this article and the Shiny application. Malangone and Sherman[32] proposed to draw repeated random samples of the IPD without replacement that mimic the distribution of the comparator study for pre-selected categorical matching variables, and illustrated its implementation for survival data using the SAS® 9.2 procedure. Alsop et al.[33] proposed to use a fourth-order polynomials within an optimization problem to identify weights that has the largest ESS while the unbalance of the covariates is controlled in a tolerable range. Han[34] directly estimated the covariate distributions in the two trials with the assumption that the two distributions share the same parametric form. The weight for each individual in the IPD sample can then be calculated as the density ratio of the two distributions.

While Shiny-MAIC makes the statistical analysis easier to perform, there are at least two aspects that need further consideration. First, although Shiny-MAIC provides various plots to check the estimated weights, it cannot reveal some key aspects for the validation of the MAIC methods including the overlap of the population, and the validity of no unmeasured effect modifiers. Therefore, users should check the validity of the implicit assumptions before interpreting the results of MAIC.

Second, although Shiny-MAIC provides a convenient tool to conduct MAIC, it faces the challenges of the data privacy policy. Users may not be allowed to upload their private IPD online due to the consideration of the data security (although Shiny-MAIC does not keep the input of the IPD file). Shiny-MAIC can automatically generate the code for the selected MAIC method according to the settings in the application. Users can directly copy the code and replace the data with their own IPD and AgD to get the MAIC results. Alternatively, with the access of the source code of Shiny-MAIC, users can run this Shiny application on local network which avoids the risk of data leakage. However, this will negate the benefits of not requiring R installation for the Shiny application.

Future extensions of Shiny-MAIC will support the following: (1) the Poisson distribution outcome with count data, (2) the survival outcome with potential truncation and censoring, and (3) the digitalization of the Kaplan–Meier (KM) survival curve. As survival data contain additional information of censoring and truncation, enabling survival data may make application more complex. Since most of the trials will publish the KM survival curve, obtaining the individual data through digitalizing the plot can be valuable for the indirect comparison.

We expect to further implement some packages for digitalizing the KM curve (such as IPDfromKM[35]) in our Shiny application.

AUTHOR CONTRIBUTIONS

Ziren Jiang: Methodology; software; visualization; writing—original draft; writing—review and editing. **Joseph Cappelleri:** Writing—review and editing; validation. **Margaret Gamalo:** Writing—review and editing; validation. **Yong Chen:** Writing—review and editing; validation. **Neal Thomas:** Writing—review and editing; validation. **Haitao Chu:** Writing—review and editing; conceptualization; supervision; resources; formal analysis; validation.

ACKNOWLEDGMENTS

We thank the editor, associate editor, and two anonymous referees for their contributions to the peer review process.

CONFLICT OF INTEREST STATEMENT

Joseph C. Cappelleri, Margaret Gamalo, Neal Thomas, and Haitao Chu are employed by Pfizer and may own stocks of their company. However, all of the contents in this article are strictly educational, instructive, and methodological, not involving any real medicinal intervention.

DATA AVAILABILITY STATEMENT

The data that used in the illustrative example (Section 5) are simulated according to the example from Jackson et al.[21] The AgD is automatically loaded in the default table in the application and the IPD is available upon request to the authors.

ORCID

Ziren Jiang https://orcid.org/0000-0002-5830-327X
Haitao Chu https://orcid.org/0000-0003-0932-598X

REFERENCES

1. Phillippo D, Ades T, Dias S, Palmer S, Abrams KR, Welton N. NICE DSU technical support document 18: methods for population-adjusted indirect comparisons in submissions to NICE. Published Online 2016.
2. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med Decis Making*. 2013;33(5):607-617.
3. Dias S, Ades AE, Welton NJ, Jansen JP, Sutton AJ. *Network Meta-Analysis for Decision-Making*. John Wiley & Sons; 2018.
4. Phillippo DM, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. Methods for population-adjusted indirect comparisons in health technology appraisal. *Med Decis Making*. 2018;38(2):200-211. doi:10.1177/0272989X17725740
5. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol*. 1997;50(6):683-691.
6. Zhang J, Carlin BP, Neaton JD, et al. Network meta-analysis of randomized clinical trials: reporting the proper summaries. *Clin Trials*. 2014;11(2):246-262.
7. Lin L, Zhang J, Hodges JS, Chu H. Performing arm-based network meta-analysis in R with the pcnetmeta package. *J Stat Softw*. 2017;80(5).
8. Lu G, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. *J Am Stat Assoc*. 2006;101(474):447-459.
9. White IR, Turner RM, Karahalios A, Salanti G. A comparison of arm-based and contrast-based models for network meta-analysis. *Stat Med*. 2019;38(27):5197-5213.
10. Remiro-Azócar A, Heath A, Baio G. Methods for population adjustment with limited access to individual patient data: a review and simulation study. *Res Synth Methods*. 2021;12(6):750-775. doi:10.1002/jrsm.1511
11. Signorovitch JE, Wu EQ, Yu AP, et al. Comparative effectiveness without head-to-head trials: a method for matching-adjusted indirect comparisons applied to psoriasis treatment with adalimumab or etanercept. *Pharmacoeconomics*. 2010;28:935-945.
12. Signorovitch JE, Sikirica V, Erder MH, et al. Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. *Value Health*. 2012;15(6):940-947. doi:10.1016/j.jval.2012.05.004
13. Ishak KJ, Proskorovsky I, Benedict A. Simulation and matching-based approaches for indirect comparison of treatments. *Pharmacoeconomics*. 2015;33(6):537-549. doi:10.1007/s40273-015-0271-1
14. Kim JK, Park M. Calibration estimation in survey sampling. *Int Stat Rev*. 2010;78(1):21-39.
15. Deville JC, Särndal CE. Calibration estimators in survey sampling. *J Am Stat Assoc*. 1992;87(418):376-382.
16. Caro JJ, Ishak KJ. No head-to-head trial? Simulate the missing arms. *PharmacoEconomics*. 2010;28(10):957-967. doi:10.2165/11537420-000000000-00000
17. Phillippo DM. *Calibration of Treatment Effects in Network Meta-Analysis Using Individual Patient Data*. University of Bristol; 2019.
18. Phillippo D, Dias S, Ades T, Welton N. Multilevel network meta-regression for population adjustment based on individual and aggregate level data.
19. Remiro-Azócar A, Heath A, Baio G. Conflating marginal and conditional treatment effects: comments on "assessing the performance of population adjustment methods for anchored indirect comparisons: a simulation study". *Stat Med*. 2021;40(11):2753-2758.
20. Petto H, Kadziola Z, Brnabic A, Saure D, Belger M. Alternative weighting approaches for anchored matching-adjusted indirect comparisons via a common comparator. *Value Health*. 2019;22(1):85-91.
21. Jackson D, Rhodes K, Ouwens M. Alternative weighting schemes when performing matching-adjusted indirect comparisons. *Res Synth Methods*. 2021;12(3):333-346.

22. Remiro-Azócar A. Two-stage matching-adjusted indirect comparison. *BMC Med Res Methodol*. 2022;22(1):1-16.
23. Wang J. On matching-adjusted indirect comparison and calibration estimation. *ArXiv Prepr ArXiv210711687*. Published online 2021.
24. Phillippo DM, Dias S, Ades AE, Welton NJ. Equivalence of entropy balancing and the method of moments for matching-adjusted indirect comparison. *Res Synth Methods*. 2020;11(4):568-572.
25. Cassidy O, Harte M, Trela-Larsen L, et al. A comparison of relative-efficacy estimate (s) derived from both Matching-Adjusted Indirect Comparisons and Standard Anchored Indirect Treatment Comparisons. Value Health. Published online 2023.
26. Phillippo DM, Dias S, Ades AE, Welton NJ. Assessing the performance of population adjustment methods for anchored indirect comparisons: a simulation study. *Stat Med*. 2020;39(30):4885-4911. doi:10.1002/sim.8759
27. Jiang Y, Ni W. Performance of unanchored matching-adjusted indirect comparison (MAIC) for the evidence synthesis of single-arm trials with time-to-event outcomes. *BMC Med Res Methodol*. 2020;20:1-9.
28. Cheng D, Ayyagari R, Signorovitch J. The statistical performance of matching-adjusted indirect comparisons. Published Online April 8, 2020. Accessed May 20, 2023. http://arxiv.org/abs/1910.06449
29. Zubizarreta JR. Stable weights that balance covariates for estimation with incomplete outcome data. *J Am Stat Assoc*. 2015;110(511):910-922.
30. Hainmueller J. Entropy balancing for causal effects: a multivariate reweighting method to produce balanced samples in observational studies. *Polit Anal*. 2012;20(1):25-46. doi:10.1093/pan/mpr025
31. Xiao Y, Moodie EE, Abrahamowicz M. Comparison of approaches to weight truncation for marginal structural cox models. *Epidemiol Methods*. 2013;2(1):1-20.
32. Malangone E, Sherman S. Matching-adjusted indirect comparison analysis using common SAS® 9.2: procedures. In: *SAS Global Forum 2011*. 2016.
33. Alsop JC, Pont LO. Matching-adjusted indirect comparison via a polynomial-based non-linear optimization method. *J Comp Eff Res*. 2022;11(8):551-561.
34. Han S. Statistical Methods for Aggregation of Indirect Information. (Doctoral dissertation). 2014. https://dash.harvard.edu/handle/1/12274529
35. Liu N, Zhou Y, Lee JJ. IPDfromKM: reconstruct individual patient data from published Kaplan-Meier survival curves. *BMC Med Res Methodol*. 2021;21(1):111.

How to cite this article: Jiang Z, Cappelleri JC, Gamalo M, Chen Y, Thomas N, Chu H. A comprehensive review and shiny application on the matching-adjusted indirect comparison. *Res Syn Meth*. 2024;1-16. doi:10.1002/jrsm.1709