

Do R packages for MAIC match each other? Insights into consistency and usability

Taylor K¹ & Hatswell AJ^{1,2}

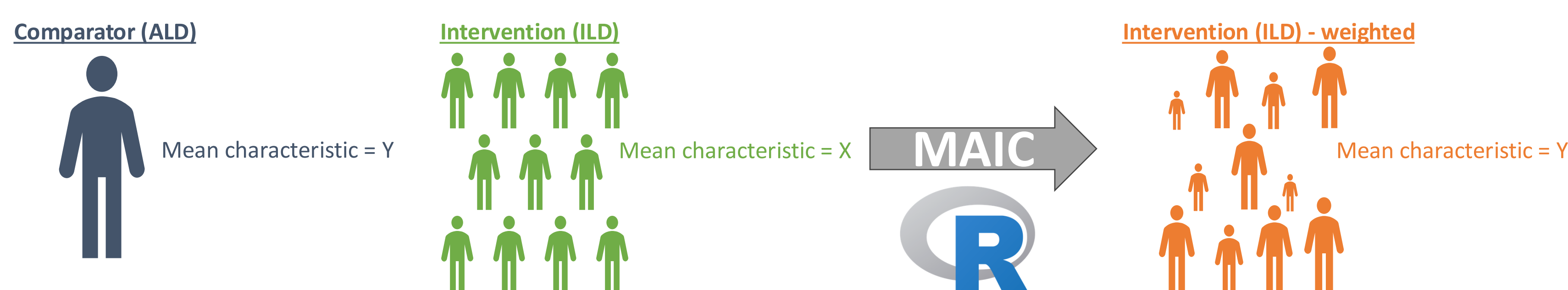
ktaylor@deltahat.com

¹Delta Hat Limited, Nottingham, UK; ²University College London, London, UK

Background and Aim

- Match-adjusted indirect comparison (MAIC) is a method commonly used to perform indirect treatment comparisons.
- MAIC involves reweighting the Individual Level Data (ILD) from one study to align to the aggregate characteristics of a study with only Aggregate Level Data (ALD) available, as illustrated in **Figure 1**.
- Several packages are available to implement MAICs using R; “*maic*”¹, “*MAIC*”², “*Maicplus*”³ and “*maicChecks*”⁴.
- This work compares the results and usability of the different packages, and the consistency of outputs across packages.

Figure 1. Illustration of the match-adjusted indirect comparison method whereby a characteristic from aggregate level data (ALD) is used to re-weight the individual level data (ILD) to match.



Methods

- Unanchored MAICs were performed matching ILD (N = 1,000 patients) to ALD (based on N = 1,000), using data from a published simulation study⁵.
- The simulated example includes a control arm (ALD) with worse baseline characteristics for 4 variables that influence time to event, and a treatment arm for which ILD is assumed to be available (“Treatment X”), which has better outcomes. These better outcomes are then exaggerated in a naïve comparison due to the more favourable patient characteristics.
- Each R package was used to perform MAICs on the same dataset with different combinations of matching variables. The variables included means, medians, and proportions/percentages; all of which are typically seen in MAICs.
- Comparisons of effective sample sizes (ESS), and weighted survival outcomes were made.

Results

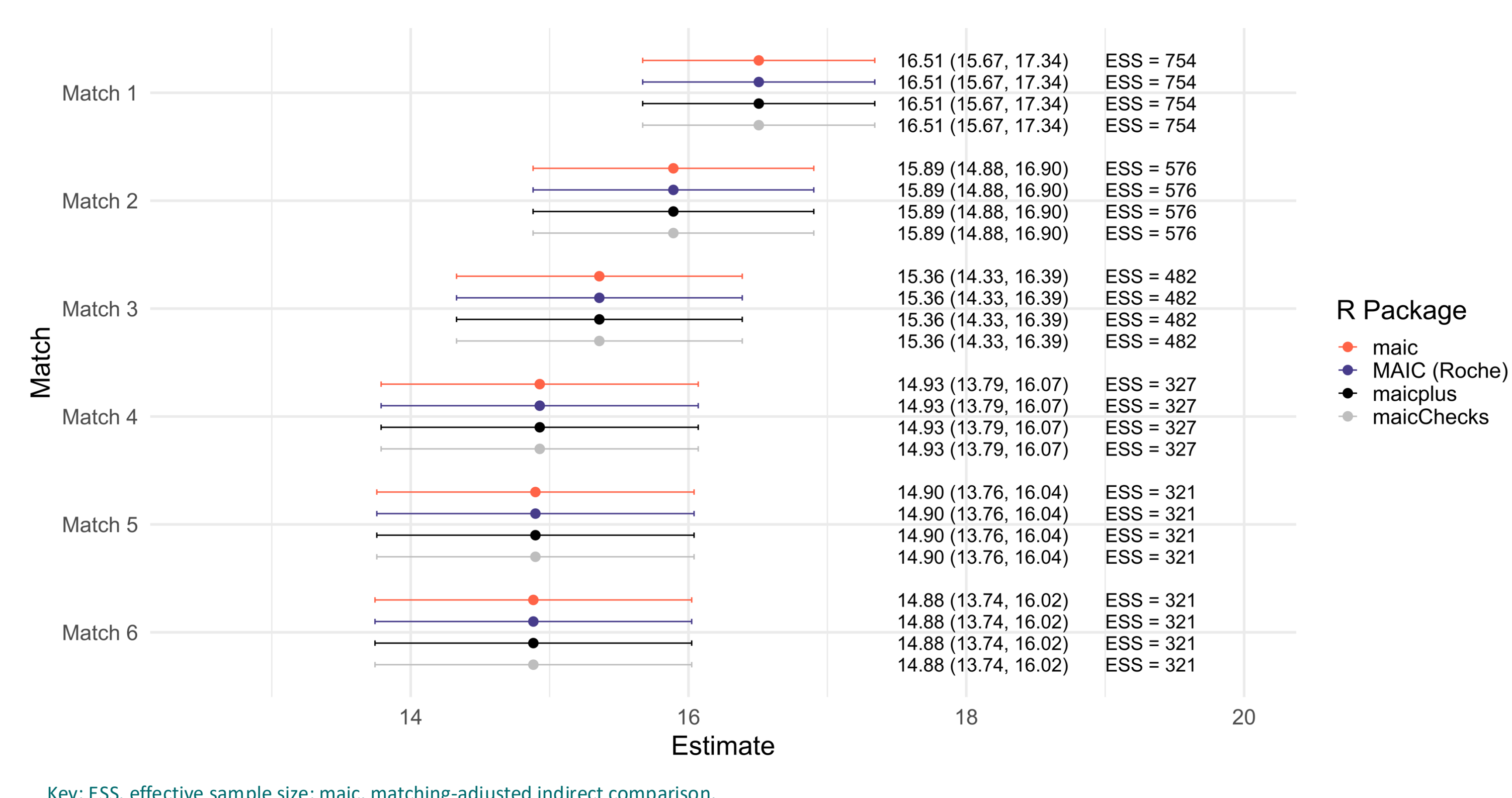
- Table 1** displays the characteristics for both the ALD group (labelled “Treatment X”) and the ILD group, before and after weighting for each match combination. The ESS decreased as more variables were matched (as would be expected). This shows how the ESS changes as variables are added, presenting the trade-offs between including a variable and its statistical impact on the results.
- All four packages generated identical ESS values and weighted outcomes (**Figure 2**).
- The only package that could natively handle ALD medians was the “*maic*” package. In other matching combinations [not shown], we found that ESS was higher when matching to medians than means, a methodological finding not previously observed in the literature.
- After observing the results, the source code for the various packages was investigated. To implement MAIC, all packages use the same underlying code to generate the weights, taken from the National Institute of Health and Care Excellence (NICE) Technical Support Document 18⁶.

Table 1. Illustrating characteristics for the ALD (Treatment X), the ILD unweighted and then how the ILD characteristics change following matching and re-weighting using a different combination of matching variables.

Characteristic	Hatswell et al - Treatment X	ILD unweighted	match-1	match-2	match-3	match-4	match-5	match-6
N Patients	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
ESS	-	-	754	576	482	327	321	321
ESS (%)	-	-	75	58	48	33	32	32
Mean Characteristic 1	0.244	0.296	0.244*	0.244*	0.244*	0.244*	0.244*	0.244*
Mean Characteristic 2	0.249	0.301	0.302	0.249*	0.249*	0.249*	0.249*	0.249*
Mean Characteristic 3	0.251	0.299	0.299	0.298	0.251*	0.251*	0.251*	0.251*
Mean Characteristic 4	0.243	0.307	0.306	0.305	0.305	0.243*	0.243*	0.243*
Median Characteristic 1	0.242	0.295	0.248*	0.251*	0.253*	0.252*	0.242*	0.242*
Percentage Characteristic 1 Yes	31	50	31*	31*	31*	32*	31*	31*
Median Characteristic 2	0.251	0.297	0.300	0.246*	0.247*	0.246*	0.244*	0.245*
Percentage Characteristic 2 Yes	33	50	52	30*	30	30*	30*	30*

ESS fill: Red = ESS is 50% or less than the original sample size. Yellow = ESS between 50% and 75%. Light green = ESS between 75% and 90%. Dark green = ESS is 90% or more of the original sample size.
Grey fill = The variable was not matched on.
* The value of a variable is within 10% of the comparators value for the same variable

Figure 2. Illustrating the weighted outcomes for each matching combination for all four R packages.



Conclusions

- The R packages currently available produce identical outcomes using the same source code, though do have differences in usability – most notably around the use of medians, which are frequently reported in clinical trial publications.
- Although demonstrated in unanchored MAICs, these results would also translate to anchored MAICs.
- Further research is needed to determine whether different analytical code for implementing MAICs produce variations in estimates, such as ESS or outcomes. Presently, there exists only one implementation in widespread use.
- Of the four packages reviewed, only “*maic*” and “*maicChecks*” were on the Comprehensive R Archive Network (CRAN).

References

- maic R package: <https://cran.r-project.org/web/packages/maic/index.html>
- MAIC R package: <https://zenodo.org/badge/latestdoi/419687389>
- Maicplus R package: <https://github.com/hta-pharma/maicplus>
- MaicChecks R package: <https://cran.r-project.org/web/packages/maicChecks/index.html>
- Hatswell AJ, Freemantle N, Baio G. The Effects of Model Misspecification in Unanchored Matching-Adjusted Indirect Comparison: Results of a Simulation Study. *Value Health J Int Soc Pharmacoeconomics Outcomes Res*. 2020;23:751–9.
- Phillippo D, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submissions to NICE (2016).