The Journal of Open Source Software

outstandR: An R Package for Indirect Treatment Comparison with Limited Subject-Level Data

Nathan Green [1*], **Antonio Remiro-Acócar**[2*], **and Gianluca Baio**[1¶]

1 University College London (UCL), UK **2** Novo Nordisk ¶ Corresponding author * These authors contributed equally.

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor:

Submitted: 29 July 2025
Published: unpublished

Summary

The goal of `outstandR` is to provide functionality to perform population adjustment methods are increasingly used to compare marginal treatment effects when there are cross-trial differences in effect modifiers and limited patient-level data (Remiro-Azócar et al., 2022). This presents a significant challenge in evidence synthesis, particularly in health technology assessment. Existing methods often face limitations such as sensitivity to poor covariate overlap or inability to extrapolate beyond observed covariate spaces. The `outstandR` package addresses these challenges by offering a robust framework for out- come regression standardisation, focusing on G-estimation. It enables researchers to perform model-based standardization with two additional crucial steps: covariate simulation (to over- come limited subject-level data for one of the studies) and indirect comparison across studies (Remiro-Azócar et al., 2022). The target audience of `outstandR` is those who want to perform model-based standardization in the specific context of two-study indirect treatment comparisons with limited subject-level data.

Statement of need

Indirect treatment comparisons (ITCs) are a cornerstone of modern evidence synthesis, especially in health technology assessment (HTA) where decision-makers must compare novel treatments against a range of competitors. A common and challenging scenario arises when individual patient data (IPD) is available for one trial, but only aggregate-level data (ALD) is available for the comparator trial. Naively comparing these studies can introduce significant bias because of differences in patient populations. There is a clear need for a unified and robust software tool that implements a range of modern population adjustment techniques to address these challenges systematically.

Method

We developed the `outstandR` R package to provide a comprehensive framework for performing anchored ITCs using a suite of population adjustment methods, with a focus on robust G-computation techniques. `outstandR` streamlines the entire analytical workflow: from fitting outcome models on IPD and standardizing them to the ALD population, to performing the final indirect comparison. By implementing multiple methods — including Matching-Adjusted Indirect Comparison (MAIC), Simulated Treatment Comparison (STC) (Phillippo et al., 2016), parametric G-computation (both frequentist and Bayesian) (Remiro-Azócar et al., 2022), and the Multiple Imputation Marginalization (MIM) method (Remiro-Azócar et al., 2024) — within a single interface, our package empowers researchers to conduct sensitivity analyses and select the most appropriate approach for their data. `outstandR` lowers the technical barrier to entry

for these complex analyses, promoting more reliable and transparent evidence synthesis for healthcare decision-making.

Related R packages we are aware of are more general-purpose tools implementing specific methods. The `marginaleffects` (Arel-Bundock, 2024) package is not designed for population adjustment between studies. `stdReg2` focuses on standardising outcomes with a single data set (G. Sofer et al., 2023). For G-formula implementations, `gfoRmula` can estimate effects in the presence of time-varying treatments and confounders (Sjölander & Dahlqwist, 2023). It is designed for estimating causal effects from longitudinal data with one study. `gFormulaMI` employs multiple imputation using the `mice` package (Sterne et al., 2023). Finally, `maicplus` is a specialist ITC package but focused only on the MAIC approach (maicplus, 2024).

Our analysis performs an anchored indirect treatment comparison (ITC) to estimate the relative effect of two treatments, A and B. This comparison uses individual patient data (IPD) from a trial comparing treatments A and C (the AC trial) and aggregate level data (ALD) from a trial comparing treatments B and C (the BC trial). To account for potential differences in the patient populations between the two trials, which can bias the ITC, we use population adjustment methods. These methods standardize the results from the AC trial to the baseline characteristics of the BC trial population.

The general procedure involves two main stages. First, we fit an outcome regression model using the IPD from the AC trial. This model describes the relationship between the outcome, treatment assignment, and a set of baseline covariates, including both prognostic factors and treatment effect modifiers. Second, we use this fitted model to predict the outcomes for treatment A and C in a target population that reflects the aggregate baseline characteristics of the BC trial. For G-computation, this step can involve simulating a large synthetic dataset that mirrors the covariate distributions of the BC population. The resulting adjusted treatment effect, ΔAC(BC), is then indirectly compared against the observed effect from the ALD, ΔBC(BC), to yield the final adjusted estimate for A versus B, ΔAB(BC). Uncertainty is quantified using non-parametric bootstrapping for frequentist methods or by propagating parameter uncertainty from the full posterior distribution in Bayesian implementations.

References

Arel-Bundock, V. (2024). marginaleffects: An R Package to Estimate and Plot Predictions, Marginal Means, and Marginal Effects. *Journal of Open Source Software*, *9*(97), 6224. https://doi.org/10.21105/joss.06224

G. Sofer, T., L. D., R., & Richardson. (2023). *tsa-lab/stdReg2: stdReg2* (Version v1.1.0). Zenodo. https://doi.org/10.5281/zenodo.10022204

maicplus, M. (2024). *maicplus: Matching-Adjusted Indirect Comparison Plus* (Version 0.1.0). https://cran.r-project.org/package=maicplus

Phillippo, D. M., Dias, S., Ades, A. E., Belger, M., Briscoe, T., & Abrams, K. R. (2016). *Methods for population-adjusted indirect comparisons in submissions to NICE* (NICE DSU Technical Support Document No. 18). NICE Decision Support Unit, School of Health; Related Research (ScHARR), University of Sheffield.

Remiro-Azócar, A., Heath, A., & Baio, G. (2022). Parametric g-computation for compatible indirect treatment comparisons with limited individual patient data. *Research Synthesis Methods*, *11*(6), 716–744. https://doi.org/10.1002/jrsm.1565

Remiro-Azócar, A., Heath, A., & Baio, G. (2024). Model-based standardization using multiple imputation. *BMC Medical Research Methodology*, *24*(32), 1–15. https://doi.org/10.1186/s12874-024-02157-x

Sjölander, A., & Dahlqwist, E. (2023). gfoRmula: An R package for implementing the g-formula. *Observational Studies*, *9*(1), 1–19.

Sterne, J. A. C., Daniel, R. M., Ferguson, A. M. B., Sari, S., Moreno-Betancur, M., Wallace, J. W. R., White, I. R., Wood, A. M., & Tilling, K. (2023). A unified framework for multiple imputation and missing data in the parametric g-formula. *Statistics in Medicine*, *42*(24), 4455–4470. https://doi.org/10.1002/sim.9863