The Journal of Open Source Software

medRCT: Causal mediation analysis estimating interventional effects mapped to a target trial in R

Tong Chen [1,2¶], **S. Ghazaleh Dashti**[1,3], **and Margarita Moreno-Betancur**[1,3]

1 Clinical Epidemiology and Biostatistics Unit, Murdoch Children's Research Institute, Australia **2** Melbourne Dental School, University of Melbourne, Australia **3** Department of Paediatrics, University of Melbourne, Australia ¶ Corresponding author

DOI: 10.21105/joss.08063

Software

- Review
- Repository
- Archive

Editor: Fabian Scheipl

Reviewers:

- @ronikobrosly
- @nhejazi

Submitted: 25 February 2025
Published: 21 June 2025

Summary

Causal mediation analysis is often used to explore the causal pathways linking exposures to outcomes across various research domains, including epidemiology, public health, and social sciences. It quantifies the extent to which the causal effect of an exposure on an outcome is mediated through one or more intermediate variables, known as mediators. The goal of studies conducting mediation analyses, particularly in health research, is often to understand how hypothetical interventions targeting mediators might counter exposure effects in order to guide future treatment and policy decisions as well as intervention development.

The modern causal inference literature initially defined the natural (in)direct effects (Pearl, 2001; Robins & Greenland, 1992) as the estimands of interest in causal mediation analysis. Grounded in the potential outcomes framework, these effects are defined based on cross-world counterfactuals (Robins & Richardson, 2011), and their identifiability relies on a cross-world independence assumption. This assumption, however, can never be guaranteed to hold, not even with an experiment (Didelez et al., 2006; Robins & Richardson, 2011). Further, this assumption is not met in the presence of a mediator-outcome confounder that is itself affected by the exposure (Avin et al., 2005; VanderWeele et al., 2014; Vansteelandt & VanderWeele, 2012), and more generally, in the presence of multiple mediators of interest (VanderWeele & Vansteelandt, 2014).

Interventional effects (Geneletti, 2007; VanderWeele et al., 2014) have been proposed as an alternative to address these limitations. In particular, these effects have been shown to implicitly emulate target trials (Hernán & Robins, 2016) that assess the impact of hypothetical interventions shifting the distribution of the mediators (Moreno-Betancur & Carlin, 2018). It has thus been proposed that interventional effects be explicitly defined by mapping to a hypothetical randomised trial (a target trial) that assesses the hypothetical interventions of interest (Moreno-Betancur et al., 2021). Specifying the target trial clarifies the causal question and makes the interventional effects directly policy-relevant and practically meaningful.

Moreno-Betancur et al. (2021) proposed three interventional effects, with each one examining a distinct policy-relevant question about how hypothetical interventions that would shift mediator distributions individually, together, or in sequence might impact the outcome. For example, they examined the extent to which the increased risk of financial hardship in mid-adulthood (outcome) among adolescent self-harmers (exposed group) could be countered by a hypothetical intervention that would shift the distribution of weekly cannabis use in young adulthood (mediator) among adolescent self-harmers to the levels of those who did not self-harm (unexposed group), either treating other mediators as independent or allowing for flow-on effects on its causal descendants, such as education and employment. They also considered an interventional effect capturing the effect of an intervention that shifts the joint distribution of all mediators (in their example, depression or anxiety, cannabis use in young

adulthood, education, and employment status) among adolescent self-harmers to the levels in those who did not self-harm.

The `medRCT` package provides a user-friendly interface for estimating policy-relevant interventional effects as defined in Moreno-Betancur et al. (2021) using a Monte Carlo simulation-based g-computation approach. Key features of `medRCT` include support for multiple mediators, intermediate confounders (exposure-induced mediator-outcome confounders), and a Shiny application (Chang et al., 2024) for comprehensive model assessment. Detailed definitions of the mediation estimands that can be estimated using the `medRCT` package, as well as the target trial to which they map, their identifiability and estimation procedures are given in Moreno-Betancur et al. (2021).

Statement of need

Several R (R Core Team, 2025) software packages are available for causal mediation analysis. The `mediation` package (Tingley et al., 2014) implements the estimation of natural effects using a generic approach based on Monte Carlo integration methods as described by Imai et al. (2010). The package can also estimate path-specific effects when there are multiple mediators, and allows researchers to conduct sensitivity analyses to evaluate the robustness of their results to potential unmeasured confounding. The `medflex` package (Steen et al., 2017) implements the methods proposed by Vansteelandt et al. (2012), which directly model the natural effects based on a class of mean models for nested counterfactuals. The `medoutcon` package (Hejazi et al., 2022) implements asymptotically efficient causal machine learning-based estimators of both the natural and interventional direct and indirect effects (Díaz et al., 2020). However, the interventional effects considered are not defined explicitly in terms of a target trial examining policy-relevant interventions, which may limit their direct practical relevance in informing decision-making.

The `medRCT` package addresses these gaps by estimating policy-relevant interventional effects, explicitly mapped to a target trial. However, as `medRCT` uses a Monte Carlo simulation-based g-computation approach, it requires a large number of replications to produce stable and reliable inference, which can be computationally intensive. Additionally, the approach is sensitive to model misspecification, as all nuisance parameters can only be estimated via restrictive parametric modelling. Furthermore, the method is not suited for settings with high-dimensional mediators or intermediate confounders, where causal machine learning-based estimators are required (Chen et al., 2025; Díaz et al., 2020; Liu et al., 2024; Rudolph et al., 2024).

There are several published applications of this methodology in real-world studies, including Dashti et al. (2022), Goldfeld et al. (2023), and Afshar et al. (2024), demonstrating its utility across diverse research contexts. For instance, Dashti et al. (2022) studied the mediating roles of C-reactive protein, leptin, fasting insulin, and estradiol in the effect of adiposity on cancer risk among postmenopausal women. Goldfeld et al. (2023) assessed the extent to which inequities could be mitigated by improving disadvantaged children's parental mental health and preschool attendance. Before the development of `medRCT`, these studies had to rely on user-written code, which is prone to errors and limits both accessibility and reproducibility. Several ongoing studies at the Murdoch Children's Research Institute are now using `medRCT`. For example, one study aims to investigate the impact on cardiovascular outcomes of a hypothetical intervention that shifts the distribution of inflammatory markers in adolescents from high-income households to the levels in those from low-income households, using data from multiple longitudinal cohort studies.

`medRCT` has also been used in education and training. It has been used in workshops and training sessions, such as the ViCBiostat Summer School and workshop at the Society for Epidemiologic Research 2024 meeting.

The package is extensively documented at its website. Future developments for `medRCT` include

extending its functionality to handle survey design weights to broaden its applicability to real-world studies involving complex survey designs.

Acknowledgements

This work was supported by an Investigator Grant fellowships to MMB [grant ID: 2009572] and SGD [grant ID: 2027171] from the National Health and Medical Research Council. The Murdoch Children's Research Institute is supported by the Victorian Government's Operational Infrastructure Support Program.

References

Afshar, N., Dashti, S. G., Mar, V., Marvelde, L. te, Evans, S., Milne, R. L., & English, D. R. (2024). Do age at diagnosis, tumour thickness and tumour site explain sex differences in melanoma survival? A causal mediation analysis using cancer registry data. *International Journal of Cancer*, *154*(5), 793–800. https://doi.org/10.1002/ijc.34752

Avin, C., Shpitser, I., & Pearl, J. (2005). Identifiability of path-specific effects. *Proceedings of the 19th International Joint Conference on Artificial Intelligence*, 357–363. https://dl.acm.org/doi/10.5555/1642293.1642350

Chang, W., Cheng, J., Allaire, J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson, J., Dipert, A., & Borges, B. (2024). *Shiny: Web application framework for R*. https://doi.org/10.32614/CRAN.package.shiny

Chen, T., Vansteelandt, S., Burgner, D., Mansell, T., & Moreno-Betancur, M. (2025). Causal machine learning for high-dimensional mediation analysis using interventional effects mapped to a target trial. *arXiv Preprint arXiv:2504.15834*. https://doi.org/10.48550/arXiv.2504.15834

Dashti, S. G., Simpson, J. A., Viallon, V., Karahalios, A., Moreno-Betancur, M., Brasky, T., Pan, K., Rohan, T. E., Shadyab, A. H., Thomson, C. A., Wild, R. A., Wassertheil-Smoller, S., Ho, G. Y. F., Strickler, H. D., English, D. R., & Gunter, M. J. (2022). Adiposity and breast, endometrial, and colorectal cancer risk in postmenopausal women: Quantification of the mediating effects of leptin, c-reactive protein, fasting insulin, and estradiol. *Cancer Medicine*, *11*(4), 1145–1159. https://doi.org/10.1002/cam4.4434

Didelez, V., Dawid, P., & Geneletti, S. (2006). Direct and indirect effects of sequential treatments. *Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence*, 138–146. https://dl.acm.org/doi/10.5555/3020419.3020437

Díaz, I., Hejazi, N. S., Rudolph, K. E., & Der Laan, M. J. van. (2020). Nonparametric efficient causal mediation with intermediate confounders. *Biometrika*, *108*(3), 627–641. https://doi.org/10.1093/biomet/asaa085

Geneletti, S. (2007). Identifying direct and indirect effects in a non-counterfactual framework. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *69*(2), 199–215. https://doi.org/10.1111/j.1467-9868.2007.00584.x

Goldfeld, S., Moreno-Betancur, M., Gray, S., Guo, S., Downes, M., O'Connor, E., Azpitarte, F., Badland, H., Redmond, G., Williams, K., Woolfenden, S., Mensah, F., & O'Connor, M. (2023). Addressing child mental health inequities through parental mental health and preschool attendance. *Pediatrics*, *151*(5), e2022057101. https://doi.org/10.1542/peds.2022-057101

Hejazi, N. S., Rudolph, K. E., & Díaz, I. (2022). medoutcon: Nonparametric efficient causal mediation analysis with machine learning in R. *Journal of Open Source Software*. https://doi.org/10.21105/joss.03979

Hernán, M. A., & Robins, J. M. (2016). Using big data to emulate a target trial when a randomized trial is not available. *American Journal of Epidemiology*, *183*(8), 758–764. https://doi.org/10.1093/aje/kwv254

Imai, K., Keele, L., & Yamamoto, T. (2010). Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science*, *25*(1), 51–71. https://doi.org/10.1214/10-STS321

Liu, R., Williams, N. T., Rudolph, K. E., & Díaz, I. (2024). General targeted machine learning for modern causal mediation analysis. *arXiv Preprint arXiv:2408.14620*. https://doi.org/10.48550/arXiv.2408.14620

Moreno-Betancur, M., & Carlin, J. B. (2018). Understanding interventional effects: A more natural approach to mediation analysis. *Epidemiology*, *29*(5), 614–617. https://doi.org/10.1097/EDE.0000000000000866

Moreno-Betancur, M., Moran, P., Becker, D., Patton, G. C., & Carlin, J. B. (2021). Mediation effects that emulate a target randomised trial: Simulation-based evaluation of ill-defined interventions on multiple mediators. *Statistical Methods in Medical Research*, *30*(6), 1395–1412. https://doi.org/10.1177/0962280221998409

Pearl, J. (2001). Direct and indirect effects. *Proceedings of the 17th Conference on Uncertainty in Artificial Intelligence*, 411–420. https://dl.acm.org/doi/10.5555/2074022.2074073

R Core Team. (2025). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://doi.org/10.32614/r.manuals

Robins, J. M., & Greenland, S. (1992). Identifiability and exchangeability for direct and indirect effects. *Epidemiology*, *3*(2), 143–155. https://doi.org/10.1097/00001648-199203000-00013

Robins, J. M., & Richardson, T. S. (2011). Alternative graphical causal models and the identification of direct effects. In *Causality and psychopathology: Finding the determinants of disorders and their cures*. Oxford University Press. https://doi.org/10.1093/oso/9780199754649.003.0011

Rudolph, K. E., Williams, N. T., & Díaz, I. (2024). Practical causal mediation analysis: Extending nonparametric estimators to accommodate multiple mediators and multiple intermediate confounders. *Biostatistics*, *25*(4), 997–1014. https://doi.org/10.1093/biostatistics/kxae012

Steen, J., Loeys, T., Moerkerke, B., & Vansteelandt, S. (2017). medflex: An R package for flexible mediation analysis using natural effect models. *Journal of Statistical Software*, *76*(11), 1–46. https://doi.org/10.18637/jss.v076.i11

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). mediation: R package for causal mediation analysis. *Journal of Statistical Software*, *59*(5), 1–38. https://doi.org/10.18637/jss.v059.i05

VanderWeele, T. J., & Vansteelandt, S. (2014). Mediation analysis with multiple mediators. *Epidemiologic Methods*, *2*(1), 95–115. https://doi.org/10.1515/em-2012-0010

VanderWeele, T. J., Vansteelandt, S., & Robins, J. M. (2014). Effect decomposition in the presence of an exposure-induced mediator-outcome confounder. *Epidemiology*, *25*(2), 300–306. https://doi.org/10.1097/EDE.0000000000000034

Vansteelandt, S., Bekaert, M., & Lange, T. (2012). Imputation strategies for the estimation of natural direct and indirect effects. *Epidemiologic Methods*, *1*(1), 131–158. https://doi.org/10.1515/2161-962X.1014

Vansteelandt, S., & VanderWeele, T. J. (2012). Natural direct and indirect effects on the exposed: Effect decomposition under weaker assumptions. *Biometrics*, *68*(4), 1019–1027.

https://doi.org/10.1111/j.1541-0420.2012.01777.x