User Guide: Conducting Mediation Analysis at Fixed Time

Contents

Introduction

This RShiny application enables users to conduct **mediation analysis with single mediator at a fixed time point** using the `CMAverse` R package. The app provides an intuitive interface for selecting your exposure, mediator, and outcome variables, and computing causal effects including total effect, direct effect, indirect effect, and proportion mediated. See Valeri and VanderWeele (2013) and VanderWeele (2015) for details about these effects.

This guide walks you through each component of the app.

Uploading Data

Click `Browse` to upload your **fully cleaned CSV file**. This guide uses a sample dataset based on the SUMMIT trial. Detailed trial results can be found in Packer et al. (2025).

Your dataset must contain:

- A **binary exposure variable**
- A **continuous or binary mediator**
- A **continuous or binary outcome**
- *(Optional)* **Baseline covariates/confounders**
- *(Optional)* **Post exposure confounders**

Model Setup

Choose a Mediation Analysis Approach

- Select `"rb"` (regression-based) for standard fixed-time single mediator analysis **WITHOUT** confounders affected by the exposure (post-exposure confounders).
- Select `"gformula"` (g-formula approach) for fixed-time single mediator analysis **WITH** confounders affected by the exposure (post-exposure confounders).
- Other options are in CMAverse, but this guide assumes the use of `"rb"`.

Exposure Variable

- Must be **coded as 0 (control) and 1 (treatment)**.
- Example: `TRTP_n=0` for placebo group, `TRTP_n=1` for tirzepatide MTD group.

Mediator

- Variable hypothesized to lie on the causal path between exposure and outcome.
- Example: `PCHG_1052.wt`, percent change from baseline in body weight to week 52.

Outcome

- Final outcome of interest.
- Example:
  - `CHG_1052.kccq`, change from baseline in KCCQ-CSS to week 52.
  - `LCHG_1052.hscrp`, change from baseline in hsCRP to week 52 on log scale.
  - `bin_outcome`, a binary variable, =1 if change from baseline in 6-minute walking distance $>$ 10; otherwise, =0.

Mediator/Outcome Model Model Types

- Choose `linear` if the mediator or outcome is continuous.
- Choose `logistic` if the variable is binary.

Baseline Covariates / Confounders

- Enter covariates that confound the relationship between exposure–mediator or mediator–outcome.
- Can be continuous or categorical, usually include baseline measurements of the mediator and outcome variables, and stratification factors.
- Example:
  - `BASE.kccq`, baseline KCCQ-CSS. For different outcomes, we recommend including the corresponding baseline measurement.

  - `BASE.wt`, baseline body weight. For different mediators, we recommend including the corresponding baseline measurement.
  - `HFDCOMFL`, HF decompensation within 12 months of screening (Y/N), stratification factor.
  - `T2DMFL`, diagnosed T2DM (Y/N), stratification factor.
  - `BMIGR2`, baseline BMI (<35, 35 kg/m^2), stratification factor.

Post-Exposure Confounders

- Enter confounders affected by the exposure
- Example: when the outcome is `CHG_1052.kccq`, change from baseline in 6-min walking distance, `CHG_1052.smwt`, may be a potential confounder.
- Do **NOT** include post-exposure confounders unless using the appropriate model (consult experts and select the g-formula approach for model fitting).

Exposure–Mediator Interactions

- Set to `TRUE` to include interaction terms between exposure and mediator in the outcome model (recommended).

Multiple Imputation

The App is designed to automatically check if the input data has missing values for any of the above selected variables. If missing value is detected, the following options will be available:

- Use `completed data only`: the App will removing all missing data.
- Use `multiple imputation`: the App will impute missing data and perform the analysis on the imputed completed data. **NOTE**: the multiple imputation is different from the `imputation` estimation method, which imputes counterfactual values.

Estimation Method

Two options available:

- `imputation`: Recommended for general use and compatible with various outcome types. Set the number of `bootstrap replications` (e.g., 200) to obtain robust confidence intervals.
- `paramfunc`: Use only if single mediator analysis and you want a closed-form solution. Delta method is conducted at back-end to estimate standard error and make inference.

Model output

DAG Visualization

The right side of the app displays a **Directed Acyclic Graph (DAG)** showing:

- Exposure
- Mediator
- Outcome
- Baseline Covariates

- Effect paths and estimates:
  - **Direct Effect** (X $\rightarrow$ Y)
  - **Indirect Effect** (X $\rightarrow$ M $\rightarrow$ Y)
  - **Proportion mediated** that quantifies how much of the effect of an exposure on an outcome operates through mediators, rather than directly.

Results Output

The app displays the following estimates:

Continous outcome

Example output:

Path	Estimate	SE	CI (Lower)	CI (Upper)	P-value
Direct Effect	1.611	2.159	-2.621	5.843	0.456
Indirect Effect	8.409	1.706	5.066	11.752	<0.001
Total Effect	10.02	1.443	7.256	12.784	<0.001
Proportion Mediated	0.839	0.202	0.443	1.236	<0.001

Path	Description
Direct Effect	Path from Exposure $\rightarrow$ Outcome
Indirect Effect	Path via Mediator
Total Effect	**Direct + Indirect**
Proportion Mediated	**Indirect / Total**

Closed-form estimation: P-value is calculated by $2 \cdot \Phi\left(Z > \left|\frac{\text{Estimate}}{\text{SE}}\right|\right)$.

Binary outcome

Example output:

Path	Estimate	SE	CI (Lower)	CI (Upper)	P-value
Direct Effect	2.282	0.641	1.316	3.957	0.003
Indirect Effect	1.687	0.433	1.025	2.793	0.040
Total Effect	3.861	0.805	2.566	5.809	<0.001
Proportion Mediated	0.552	0.197	0.167	0.937	0.005

Path	Description
Direct Effect	Path from Exposure $\rightarrow$ Outcome
Indirect Effect	Path via Mediator
Total Effect	**Direct * Indirect**
Proportion Mediated	**Direct * (Indirect - 1) / (Total - 1)**

Closed-form estimation: P-value is calculated by $2 \cdot \Phi\left(Z > \left|\frac{\log(\text{Estimate})}{\text{SE/Estimate}}\right|\right)$.

Interpretation Tips

- A significant indirect effect ($P < 0.05$) suggests the mediator carries part of the exposure effect.
- Proportion mediated > 0.5 indicates most of the effect is through the mediator.
- If direct effect is not significant but indirect is, this may suggest *full mediation*.
- Use confidence intervals to assess inference if `imputation` and `bootstrap` are selected.

Notes and Recommendations

- Ensure variables are **appropriately coded** (e.g., exposure as binary 0/1).
- Double check model assumptions (e.g., linearity).
- Interpret results in the **causal inference** context, assuming no unmeasured confounding.
- This tool is best suited for **observational or interventional studies** with clear temporal ordering.

References

- Shi, Q., et al. (2021). https://journals.lww.com/epidem/fulltext/2021/09000/cmaverse___a_suite_of_functions_for_reproducible.23.aspx
- CMAverse CRAN: https://cran.r-project.org/web/packages/CMAverse/
- CMAverse GitHub: https://bs1125.github.io/CMAverse/index.html
- Valeri, L and VanderWeele TJ. (2013). https://pmc.ncbi.nlm.nih.gov/articles/PMC3659198/
- VanderWeele TJ. (2014). https://pmc.ncbi.nlm.nih.gov/articles/PMC4220271/
- Packer et al. (2025). https://www.nejm.org/doi/full/10.1056/NEJMoa2410027/

Prepared for internal use and education. Please consult a statistician for advanced mediation design or interpretation.