Causal mediation analysis: How to avoid fooling yourself that X causes Y

Stanley E. Lazic[1,*]

1. Prioris.ai Inc., 459-207 Bank St., Ottawa ON K2P 2N2, Canada

*Corresponding author: stan.lazic@cantab.net

Abstract

The purpose of many preclinical studies is to determine if an experimental intervention affects an outcome through a particular mechanism, but the analytical methods and inferential logic typically used cannot answer this question, leading to erroneous conclusions about causal relationships, which can be highly reproducible. A causal mediation analysis can directly test if a hypothesised mechanism is partly or completely responsible for a treatment's effect on an outcome. Such an analysis can be easily implemented with modern statistical software. We show how a mediation analysis can distinguish between three different causal relationships that are indistinguishable when using a standard analysis.

Keywords

Causal, Mechanism, Mediation, Statistics

Introduction

Once researchers establish that an intervention or treatment (T) affects an outcome (O), they often try to determine the mechanism (M) by which the treatment works. For example, a common problem in neuroscience is understanding how a treatment (T) changes some aspect of brain function (M), which in turn leads to a change in behaviour (O). Consider an example where animals are randomly assigned to a drug or control group. Compared to controls, animals in the drug group show less depressive-like behaviour and have higher levels of serotonin

in the brain. Did the drug affect behaviour because it increased serotonin? The inferential logic typically follows this pattern: we observe a statistically significant difference between treatment groups on (1) the outcome (the drug affects behaviour), and (2) on the mechanism or mediating variable (the drug affects serotonin levels), and then (3) conclude that the change in behaviour was likely caused by a change in serotonin levels. Unfortunately, causal conclusions cannot be drawn with this approach, but it is standard in many fields[1]. In toxicological research, a key question is if a drug (T) directly affects an organ's weight (O), or if changes in organ weight reflect changes in the size and weight of the entire body (M)[2]. Similar problems arise whenever a biomarker is used as a surrogate or proxy for a clinical outcome[3].

Methods

To demonstrate the inability of the standard approach to find the true causal relationship, data from three models with known causal relationships were simulated (Fig. 1, top row). Arrows in these diagrams represent the true causal structure from which the data were generated, and missing arrows between two variables indicate there is no causal relationship between them. Results are shown beneath each model (units are arbitrary, but are comparable across rows in the figure). Data were simulated for two groups of $N = 12$ animals per group and the effect of the treatment on the mediator is the same for all three models. The data were then analysed using standard methods and a Bayesian mediation analysis. The simulations and analyses were carried out using Julia, and the code is available on Github (https://github.com/stanlazic/lab-animals-mediation).

Results

For the left causal model, the treatment affects the hypothesised mechanism/mediating variable (serotonin), which in turn affects the outcome (behaviour). The drug has no direct effect on behaviour; 100% of the effect acts through changes in serotonin. The two bar graphs below the diagram show the data, with significant differences ($p < 0.05$) between groups on both the outcome and mediator, thus meeting the typical criteria for concluding that the treatment affected behaviour via changes in serotonin. In the middle causal model (Fig. 1) the treatment still affects the mediator, but now it also directly affects the outcome (through some other unmeasured mechanism) and the mediator has absolutely no effect on behaviour. The data, however, look nearly identical to the previous model and also show significant group differences. In the right causal model (Fig. 1) the treatment affects both the mediator and the outcome, and the mediator also affects the outcome. Therefore, there are two paths by which treatment affects the outcome: the direct $T \rightarrow O$ path and the indirect $T \rightarrow M \rightarrow O$ path. Once again, the data appear identical to those from the previous models.

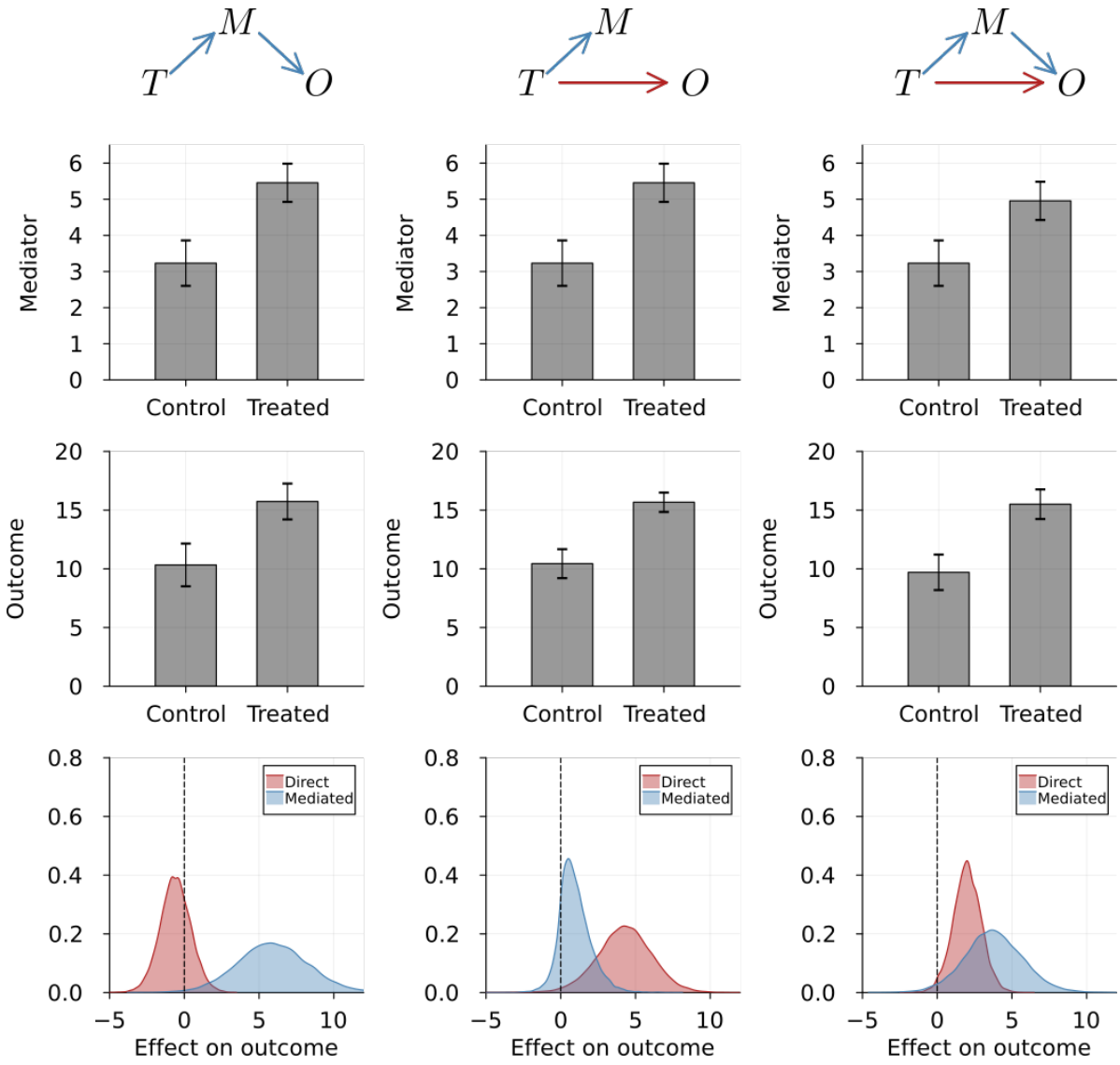

Figure 1: Three causal models for how a treatment (T) can affect an outcome (O) through a hypothesised mechanism (M). All six bar graphs show significant differences between groups ($p < 0.05$), which makes it impossible to distinguish between causal models. The mediation analysis (bottom row) clearly identifies the correct model. Error bars are 1 standard deviation.

Clearly, the standard analysis cannot distinguish between these three causal models. A Bayesian mediation analysis was then conducted, which estimates the effect of the treatment separately through the direct and indirect paths, and the results are presented as posterior distributions for the treatment effect (Fig. 1, bottom row). When 100% of the drug effect is mediated through serotonin levels (Fig. 1, left), the direct effect (red) is correctly estimated to be close to zero and the indirect mediated effect (blue) is far from zero. When the drug acts completely through the direct effect (Fig. 1, middle), the analysis correctly estimates this – the red distribution is far from zero, and the blue distribution is close to zero. Finally, when the drug acts both directly on behaviour and through serotonin levels, the model correctly recovers these estimates and both distributions are far from zero. Increasing the sample size will narrow these distributions and make the results more conclusive.

Discussion

Mediation analysis follows a straightforward logic: define a causal model that reflects the hypothesised causal relationships, fit the model to the data, and interpret the results, which directly address the causal questions of interest. Researchers are then able to eliminate competing explanations that follow the direct $T \rightarrow O$ path, providing support for the hypothesised mechanism. Instead, if the intervention's effect is primarily or entirely mediated by the direct pathway, then the mediation hypothesis can be discarded. The hunt can then begin for other mechanisms, thereby advancing the field.

Some may argue that mediation analyses are unnecessary for controlled experiments and that causal conclusions can be drawn from a standard analysis if the mediating variable is controlled. For example, if serotonin receptors can be blocked with another drug, the increase in serotonin due to the first drug should have no effect on the outcome, which provides evidence for a causal $T \rightarrow M \rightarrow O$ path. Experimentally controlling the mediating variable is certainly beneficial, but it does not guarantee that causal relationships will be identified, and researchers can still be misled (see supplementary material in reference[4]). A more complex causal model can be used to represent such an experiment, and the inferential logic remains the same.

When using a mediation analysis, several factors should be considered. First, a larger sample size is needed compared with a standard analysis, and methods exist for mediation power calculations[5,6]. Second, the causal model must be correctly specified: the arrows must point in the correct direction. The simulation above assumed that serotonin affected behaviour ($M \rightarrow O$) but often animals' behaviour is assessed first, then their brain serotonin levels are measured. Thus, behavioural testing may affect serotonin levels ($M \leftarrow O$). This scenario may justify redesigning the experiment to measure serotonin levels prior to behavioural assessment (e.g., by imaging). Third, in addition to mediating, a variable may moderate the effect of one variable on another. This is a different type of relationship, and in our example could occur if the drug being tested does not affect serotonin levels, but it will only be effective if serotonin levels are high. Hence, a mediation analysis addresses how an intervention works while a moderation

analysis addresses when it works[7]. Fourth, there may be additional background variables to consider, such as sex, litter, cages, etc. Modern software can incorporate these variables into a more realistic and comprehensive model. Fifth, these methods can be applied to more complex experiments with multiple mediators, multilevel models, or high-dimensional data[8]. Finally, although a mediation analysis is straightforward to conduct, there are nuances related to the model specification, assumptions, and interpretation that should be considered[9,10].

Causal models are particularly useful in the following situations. First, when the treatment or intervention affects several variables, not just the mediator of interest. As a result, there are competing explanations for how the treatment affects the outcome. The more off-target effects there are, the more possible mechanisms that may be at work. Second, when it is difficult to experimentally control the mediator value, as in the example given above. Then, we can only attempt to identify causal relationships statistically. Third, when the mediator is under experimental control, but changing the mediator also changes other variables that might influence the outcome. In these cases, it may be difficult to "experiment one's way out of a situation" and more sophisticated mediation techniques may be necessary.

For researchers interested in learning more about these methods, many non-technical introductions are available[7,11,12]. Simple mediation models are freely available in the user-friendly JASP statistical package[13], as well as in the mediation[14], and brms[15] R packages. R packages by Yu and Li[8] contain the latest and more advanced methods.

In summary, combining high levels of experimental control with causal models will enable biologists to make stronger inferences about causal mechanisms by eliminating competing explanations.

Declaration of Conflicting Interests

No conflict of interest are declared.

Ethics Statement

This study did not require ethical board approval because it did not contain human or animal trials.

Data availability

The simulated data and Julia code are available on Github: https://github.com/stanlazic/lab-animals-mediation.

References

1. Lazic SE, Fuss J, Gass P. Quantifying the behavioural relevance of hippocampal neurogenesis. PLoS One 2014; 9: e113855.
2. Lazic SE, Semenova E, Williams DP. Determining organ weight toxicity with Bayesian causal models: Improving on the analysis of relative organ weights. Scientific Reports; 10. Epub ahead of print 2020. DOI: 10.1038/s41598-020-63465-y.
3. Elliott MR. Surrogate endpoints in clinical trials. Annual Review of Statistics and Its Application 2023; 10: 75–96.
4. Lazic SE. Using causal models to distinguish between neurogenesis-dependent and -independent effects on behaviour. J R Soc Interface 2012; 9: 907–917.
5. Fritz MS, Mackinnon DP. Required sample size to detect the mediated effect. Psychol Sci 2007; 18: 233–239.
6. Thoemmes F, Mackinnon DP, Reiser MR. Power analysis for complex mediational designs using Monte Carlo methods. Struct Equ Modeling 2010; 17: 510–534.
7. Hayes AF. Introduction to mediation, moderation, and conditional process analysis: A regression-based approach. New York, NY: Guilford Press, 2013.
8. Yu Q, Li B. Statistical methods for mediation confounding and moderation analysis using R and SAS. Chapman & Hall/CRC Press, 2022.
9. Fiedler K, Schott M, Meiser T. What mediation analysis can (not) do. Journal of Experimental Social Psychology 2011; 47: 1231–1236.
10. Fiedler K, Harris C, Schott M. Unwarranted inferences from statistical mediation tests – an analysis of articles published in 2015. Journal of Experimental Social Psychology 2018; 75: 95–102.
11. MacKinnon DP. Introduction to statistical mediation analysis. Erlbaum Psych Press, 2008.
12. Pearl J, Glymour M, Jewell NP. Causal inference in statistics: A primer. Chichester: Wiley, 2019.
13. JASP Team. JASP (Version 0.17.2)[Computer software].
14. Tingley D, Yamamoto T, Hirose K, et al. Mediation: R package for causal mediation analysis. Journal of Statistical Software 2014; 59: 1–38.
15. Burkner P-C. brms: An R package for Bayesian multilevel models using Stan. Journal of Statistical Software 2017; 80: 1–28.