Title

Going down the right path: Using causal diagrams and multiple comparison procedures to steer clear of the Table 2 Fallacy and multiplicity

Authors

McIntyre, Kevin J[1¶] and Wiener, Joshua C[1¶] and Davies Smith, Emma[1,2¶]

Affiliation

[1] Department of Epidemiology and Biostatistics, Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada

[2] Department of Biostatistics, T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts, USA

¶ authors contributed equally to this work.

Correspondence

Kevin J McIntyre
Centre for Public Health and Family Medicine
1465 Richmond Street
London, Ontario, Canada, N6A 2M1
kmcint67@uwo.ca

Abstract

The Table 2 Fallacy is an interpretation error commonly encountered in medical research. This error occurs when coefficient estimates in multivariable regression models, apart from that of the primary exposure, are interpreted as valid total effects on the outcome. Causal diagrams, notably directed acyclic graphics (DAGs), can be used to avoid this error. DAGs can be used to identify sets of covariates that, when adjusted for, allow for unbiased estimation and correct interpretation of the total effect for each primary exposure. However, the risk of a false positive finding, or the type I error rate, increases when inference is conducted for multiple associations of interest. While multiple comparison procedures are recognized as a critical feature of randomized controlled trial (RCT) design and analysis, opinions remain divided on their use within observational studies. Target trial emulation is an increasingly used framework that aligns the design and analysis of causal observational studies with that of RCTs. This article demonstrates that when multiple causal relationships are of primary interest to an observational study, causal diagrams can be used in conjunction with multiple comparison procedures to simultaneously avoid the Table 2 Fallacy and control the risk of spurious findings. The proposed framework should further align these two study types and assist in ensuring that results are reproducible. Methods are demonstrated by simulating a hypothetical causal mechanism using the dagR library.

Keywords: Causal inference, Multivariable regression, Directed acyclic graphs, Multiple comparison procedures, Multiplicity, Table 2 Fallacy, Target trials

Main Text

The Foundation

Observational Studies: There Comes a Time

Randomized controlled trials (RCTs) are considered the ‘gold standard’ study design in medical research. Randomly allocating participants to different intervention groups inhibits associations between the intervention and other covariates, including (un)measured confounders. However, due to feasibility issues or ethical concerns, many research questions cannot be answered using RCTs. Separately, there is also an increasing wealth of electronic health record (EHR) data that, when used in a principled manner, can provide a means of answering important questions about health today.

In these scenarios, epidemiologists often turn to observational study designs, using various techniques to address confounding during the design and analytic stages of the study. It has been proposed that the design of observational studies can be guided by conceptualizing and emulating an ideal RCT, or target trial, that could answer the research question if a trial were possible (Hernán et al., 2008). In addition to encouraging the unification of RCTs and observational studies in terms of study design, the target trial framework also adopts features of RCT analyses, such as the identification of intention-to-treat and per-protocol causal contrasts.

The target trial framework represents a step in the right direction by better aligning results from causal observational studies with that of their RCT counterparts. If RCTs are considered the standard to which other designs are compared, we argue that observational studies should also employ crucial principles of RCT analysis. In particular, the control of type I error rates is increasingly recognized as critical to RCTs with multiple treatment arms, outcomes, or subgroups. However, attitudes remain divided on the importance of this error within observational studies (Althouse, 2016; Patel & Ioannidis, 2014; Rothman, 1990; U.S. Food and Drug Administration, 2022). Here, we focus on the identification of multiple distinct exposures of primary interest via

causal directed acyclic graphs (DAGs) and the use of multiple comparison procedures (MCPs) to control the resulting inflation of type I errors.

The Table 2 Fallacy: Don't Let Me Be Misunderstood

Modern statistical analyses of observational studies commonly rely upon multivariable regression methods. Under necessary causal assumptions, multivariable regression can provide unbiased estimates of the total effect of the primary exposure on the outcome by adjusting for the influence of covariates, notably confounders. However, coefficient estimates for additional covariates within the same model are *not* interpretable as total effects, as the model structure is determined specifically by the primary exposure and the outcome. Incorrect interpretation of secondary effect estimates as valid total effects was dubbed the "Table 2 Fallacy" by Westreich & Greenland (2013), and this phenomenon is fairly common in the medical literature (Akinkugbe et al., 2021; Ponkilainen et al., 2021; Westreich & Greenland, 2013).

If the total effect of a covariate is also of interest, an additional model should be specified by considering the distinct biasing pathways of this new relationship. Explicitly outlining all relationships of interest and their respective causal pathways *a priori* can help ensure that regression analyses are not biased by model misspecification. Misspecification may occur, for example, when colliders or mediators are included as covariates or confounders are omitted (Schisterman et al., 2009).

The Tools

Creating Causal Diagrams: Think (About It)

Causal diagrams, namely DAGs, can be used as a tool to avoid the Table 2 Fallacy. DAGs provide investigators with a framework to explicitly state their assumptions regarding the hypothesized causal mechanism (Pearl, 1995). Within a DAG, exposures or covariates are represented as nodes within a graph, and causal paths as arcs between nodes. By constructing a DAG, investigators can visually identify confounders, which can help formalize the selection of covariates for adjustment in their statistical model prior to analysis. Gaskell and Sleigh (2020) and Lipsky and Greenland (2022)

provide accessible tutorials on how to use DAGs in medical research (Gaskell & Sleigh, 2020; Lipsky & Greenland, 2022). Unfortunately, despite DAGs being broadly accepted within the epidemiological community, they have been underutilized in published medical research (Tennant et al., 2021).

Multiple Comparison Procedures: Don't Think Twice, It's All Right

The type I error rate or false positive rate, denoted commonly by α, refers to the risk of falsely concluding that an effect estimate is statistically significant. The type I error rate is often inflated when significance testing is performed for each of multiple exposures or outcomes. For example, if hypothesis tests are performed separately for each of five exposures, permitting an $\alpha = 0.05$ error rate for each test, the probability of falsely rejecting at least one hypothesis is $1 - (1 - 0.05)^5 = 0.23$ or 23%. As the number of tests grows, the probability of at least one false positive result approaches one. This phenomenon is often referred to as multiplicity or the multiple comparisons (testing) problem.

Multiple comparison procedures (MCPs) are statistical methods that adjust the level of significance or p-value for each association to control or prevent inflation of the overall type I error rate. The simplest of the MCPs, the Bonferroni adjustment, assigns α/K rather than alpha to each of the K associations of interest, thereby ensuring that the overall type I error rate is, at most, the desired α. There are a variety of MCPs, each presenting their own strengths and weaknesses, and investigators should ensure that the assumptions of a chosen MCP suit their research question or scenario (Vickerstaff et al., 2019). In addition to using DAGs to properly identify models for each exposure of interest, we propose that investigators pre-specify the use of MCPs to minimize the risk of false positives and increase the likelihood of reproducibility.

Statistical Significance and Interval Estimation: Think It Over

Over the past several decades, the use of hypothesis testing and p-values specifically, and the concept of statistical significance more broadly, have received criticism for oversimplifying effect interpretation (Amrhein et al., 2019; Evans et al., 1988; McShane et al., 2017; Rothman, 1978, 1986; Stang et al., 2010; Wasserstein et al., 2019). Trial

reporting guidelines including CONSORT (Schulz et al., 2010), as well as leaders in trial methodology, (Gardner & Altman, 1986) now recommend that confidence intervals be reported alongside point estimates and p-values, permitting assessment of both statistical and clinical significance.

For example, the common $(1-\alpha)\times 100\%$ Wald confidence interval takes the form:

(Lower Bound, Upper Bound) = Point Estimate -/+ $Z_{\alpha/2}$ ** Standard Error*

where $(1-\alpha)$ is the desired confidence level or coverage and $Z_{\alpha/2}$ is the corresponding critical value. The lower and upper bounds of the interval provide a range of effects compatible with the observed data, and assist in understanding the error or uncertainty associated with point estimates. Confidence intervals do, however, share a direct relationship to hypothesis testing. Instead of judging statistical significance using p-value < α, it can be equivalently assessed by determining whether the $(1-\alpha)\times 100\%$ confidence interval includes the null effect (commonly 0 or 1 depending on whether the effect measure is absolute or relative, respectively). That is, if a null effect is compatible with the data, it suggests that the effect estimate is statistically non-significant.

Similar to how the significance level may be adjusted for multiple hypothesis tests, the confidence level for multiple intervals may also be adjusted to ensure simultaneous confidence or coverage for the true effect (*E 9 Statistical Principles for Clinical Trials*, 2006; Schulz et al., 2010). For example, borrowing conceptually from the Bonferroni adjustment, one may construct $(1-\alpha/K)\times 100\%$ confidence intervals for K effects. These intervals provide the same information on statistical significance as K tests at the adjusted α/K significance level but provide additional insight into effect magnitude and direction(Vittinghoff et al., 2012, pp. 40–41). However, for more complex MCPs, the form of corresponding confidence intervals is not always straight-forward. Further guidance on this issue can be found in Vickerstaff, Omar & Ambler

(2019). Lastly, it should be noted that statistical significance is only one component of interpreting effect estimates and researchers should ensure that estimates are also assessed through the lens of clinical significance.

The Example

All Together Now

In summary, RCTs cannot always be conducted for ethical or logistical reasons. The target trial framework can be used to emulate a hypothetical RCT using observational data to assist in the inference of causal effects. When multiple exposures are of interest, the "Table 2 Fallacy" can occur when all coefficient estimates within a regression model are interpreted as valid total effects. This interpretation error can be avoided by using DAGs to explicitly specify causal pathways and construct multiple models for each relationship of interest. This solution, however, introduces the secondary problem of multiplicity due to the assessment of multiple effects using the same data. This challenge can be addressed by applying MCPs, as is common practice in the RCT literature.

An Illustration

Figure 1A presents the DAG for a hypothetical causal mechanism between a series of eight variables; details of DAG visualization and analysis using R packages dagitty (Textor et al., 2016) and ggdagR (Barrett, 2018) are available in the Online Appendix. The DAG features a primary exposure (E) and an outcome (O), as well as six covariates that are causally related to the exposure, outcome, and/or each other (A, B, C, D, F, G). In this DAG, there are three different combinations of covariates that, when adjusted for, produce an unbiased estimate of the total effect of E on O. These “sufficient sets” of covariates are illustrated in Figure 1B: (1) A, C, G; (2) C, D, G; and (3) F, G.

Suppose an investigator conducts an initial analysis conditioning on the first sufficient set (A, C, G), which yields an unbiased estimate of the total effect of E on O. If the investigator was also interested in estimating the total effect of C on O, interpreting the estimated coefficient for C within the initial model would be erroneous. The research

question has changed: C is the primary exposure, and thus the covariates needed to achieve an implied sufficient set have also changed (see Figure 2A). To estimate the total unbiased effect of C on O, it would only be necessary to adjust for A, which is the minimally sufficient set (see Figure 2B). The effect estimate of C on O presented within a hypothetical Table 2 would not estimate the total effect, as the mediated effects of C on O through the variables G and E are “blocked” in this model. As such, interpretation of the effect estimate of C in the initial model is guaranteed to be a biased estimate of the total effect of C on O.

A Simulation

We used data simulated according to the DAG for our causal mechanism to demonstrate the Table 2 Fallacy; details of DAG simulation using the dagR library (Breitling, 2010) are available in the Online Appendix. The true direct effect of the primary binary exposure E on the continuous outcome O was specified as 0.50.

In Table 1, we present the results of a multivariable linear regression of E on O, adjusting for the first sufficient set (A, C, G). As expected, adjustment for this sufficient set yielded an unbiased estimate of the total effect of E on O, indicated by a coefficient estimate of 0.51. If this effect were the only effect of interest, the corresponding 95% CI is (0.22, 0.80) and the p-value is 0.001 < $\alpha = 0.05$, suggesting statistically significant evidence of a total effect of E on O.

In Table 2, we present the results of a multivariable linear regression for our second exposure of interest, C. The model is specified as a regression of O on C adjusting for A, after identification of its minimally sufficient set using the DAG. Within this new model, the coefficient estimate of C is 0.36 (95% CI: 0.05, 0.67; p-value = 0.033 < $\alpha = 0.05$), suggesting a positive, statistically significant total effect of C on O. In contrast, the coefficient estimate of C in the previous model (Table 1) was -0.18 (95% CI: -0.51, 0.15; p-value = 0.296 > $\alpha = 0.05$). With proper covariate adjustment, the coefficient estimate has completely changed direction, exemplifying the risk of the Table

2 Fallacy and the importance of principled model selection, construction, and interpretation.

Since multiple associations were of interest (K=2), we next used the Bonferroni adjustment to correct confidence intervals for multiplicity. That is, wider $(1-\alpha/K)\times 100\%$ confidence intervals (i.e., 97.5%) were constructed for each association rather than the typical $(1-\alpha)\times 100\%$ confidence intervals (i.e., 95%). Prior to adjustment, the 95% CI for the total effect of C on O in Model 2 was (0.05, 0.67) which excludes the null effect of 0 and the p-value was 0.033 $< \alpha = 0.05$. In comparison, the multiplicity adjusted 97.5% CI of (-0.01, 0.72) includes the null value and the p-value 0.033 > $\alpha/2$ = 0.025, no longer suggesting statistically significance. In comparison, the multiplicity adjusted confidence interval for the total effect of E on O is (0.17, 0.85) and the p-value is 0.001 $<$ 0.025, suggesting the effect remains statistically significant.

The Finale

The Table 2 Fallacy is a major interpretation error that can occur in observational studies when coefficients, other than that of the primary exposure, are interpreted as total effects despite having a different causal relationship with the outcome variable. To avoid this fallacy, researchers can use causal diagrams to specify multiple models for each exposure of interest. Assessing multiple models fitted using the same data can, however, lead to an inflation of the type I error rate or multiplicity. While multiplicity is well acknowledged within the RCT literature, corresponding multiple comparison procedures (MCPs) are not commonly used in observational studies, despite growing adoption of target trial emulation. Here, we have proposed and demonstrated how the use of MCPs and DAGs can be used in conjunction to simultaneously avoid the Table 2 Fallacy and multiplicity within causal observational analyses. We feel this strategy better aligns the analysis of observational data within the target trial framework to current best practices in the RCT literature and has the potential to create more rigorous and reproducible medical research.

Acknowledgements

Thank you to Stephanie Armbruster, Georg Hahn, and Emma Crenshaw for their thoughtful comments which improved the quality of this manuscript (statistically significant).

Declarations

Competing Interests

The authors declare that no conflicts of interest exist with respect to the research, authorship, and/or publication of this article.

Financial Support

The authors declare that no funds, grants, and/or other support were received during the preparation of this manuscript.

References

Akinkugbe, A. A., Simon, A. M., & Brody, E. R. (2021). A scoping review of Table 2 fallacy in the oral health literature. *Community Dentistry and Oral Epidemiology*, *49*(2), 103–109. https://doi.org/10.1111/cdoe.12617

Althouse, A. D. (2016). Adjust for Multiple Comparisons? It's Not That Simple. *The Annals of Thoracic Surgery*, *101*(5), 1644–1645. https://doi.org/10.1016/j.athoracsur.2015.11.024

Amrhein, V., Greenland, S., & McShane, B. (2019). *Retire statistical significance*. *567*, 305–307.

Barrett, M. (2018). *ggdap: Analyze and create elegant directed acyclic graphs in R* [Computer software]. https://github.com/r-causal/ggdag

Breitling, L. P. (2010). dagR: A Suite of R Functions for Directed Acyclic Graphs. *Epidemiology*, *21*(4), 586–587. https://doi.org/10.1097/EDE.0b013e3181e09112

E 9 Statistical Principles for Clinical Trials. (2006).

Evans, S. J., Mills, P., & Dawson, J. (1988). The end of the p value? *Heart*, *60*(3), 177–180. https://doi.org/10.1136/hrt.60.3.177

Gardner, M. J., & Altman, D. G. (1986). Confidence intervals rather than P values: Estimation rather than hypothesis testing. *BMJ*, *292*(6522), 746–750. https://doi.org/10.1136/bmj.292.6522.746

Gaskell, A. L., & Sleigh, J. W. (2020). An Introduction to Causal Diagrams for Anesthesiology Research. *Anesthesiology*, *132*(5), 951–967. https://doi.org/10.1097/ALN.0000000000003193

Lipsky, A. M., & Greenland, S. (2022). Causal Directed Acyclic Graphs. *JAMA*, *327*(11), 1083–1084. https://doi.org/10.1001/jama.2022.1816

McShane, B. B., Gal, D., Gelman, A., Robert, C., & Tackett, J. L. (2017). Five ways to fix statistics. *Nature*, *551*, 557–559. https://doi.org/10.1038/d41586-017-07522-z

Patel, C. J., & Ioannidis, J. P. A. (2014). Placing epidemiological results in the context of multiplicity and typical correlations of exposures. *Journal of Epidemiology and Community Health*, *68*(11), 1096–1100. https://doi.org/10.1136/jech-2014-204195

Pearl, J. (1995). Causal Diagrams for Empirical Research. *Biometrika*, *82*(4), 669–688.

Ponkilainen, V. T., Uimonen, M., Raittio, L., Kuitunen, I., Eskelinen, A., & Reito, A. (2021). Multivariable models in orthopaedic research: A methodological review of

covariate selection and causal relationships. *Osteoarthritis and Cartilage*, *29*(7), 939–945. https://doi.org/10.1016/j.joca.2021.03.020

Rothman, K. J. (1978). A Show of Confidence. *New England Journal of Medicine*, *299*(24), 1362–1363. https://doi.org/10.1056/NEJM197812142992410

Rothman, K. J. (1986). Significance Questing. *Annals of Internal Medicine*, *105*(3), 445–447. https://doi.org/10.7326/0003-4819-105-3-445

Rothman, K. J. (1990). No Adjustments Are Needed for Multiple Comparisons: *Epidemiology*, *1*(1), 43–46. https://doi.org/10.1097/00001648-199001000-00010

Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomised trials. *Journal of Clinical Epidemiology*, *63*(8), 834–840. https://doi.org/10.1016/j.jclinepi.2010.02.005

Stang, A., Poole, C., & Kuss, O. (2010). The ongoing tyranny of statistical significance testing in biomedical research. *European Journal of Epidemiology*, *25*(4), 225–230. https://doi.org/10.1007/s10654-010-9440-x

Tennant, P. W. G., Murray, E. J., Arnold, K. F., Berrie, L., Fox, M. P., Gadd, S. C., Harrison, W. J., Keeble, C., Ranker, L. R., Textor, J., Tomova, G. D., Gilthorpe, M. S., & Ellison, G. T. H. (2021). Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: Review and recommendations. *International Journal of Epidemiology*, *50*(2), 620–632. https://doi.org/10.1093/ije/dyaa213

Textor, J., Van Der Zander, B., Gilthorpe, M. S., Liśkiewicz, M., & Ellison, G. T. H. (2016). Robust causal inference using directed acyclic graphs: The R package

‘dagitty.’ *International Journal of Epidemiology*, *45*(6), 1887–1894. https://doi.org/10.1093/ije/dyw341

U.S. Food and Drug Administration. (2022). Multiple Endpoints in Clinical Trials—Guidance for Industry. *Clinical Trials*. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/multiple-endpoints-clinical-trials-guidance-industry

Vittinghoff, E., Glidden, D. V., Shiboski, S. C., & McCulloch, C. E. (2012). *Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models* (2nd ed.). Springer.

Wasserstein, R. L., Schirm, A. L., & Lazar, N. A. (2019). Moving to a World Beyond “ $p < 0.05$.” *The American Statistician*, *73*(sup1), 1–19. https://doi.org/10.1080/00031305.2019.1583913

Westreich, D., & Greenland, S. (2013). The Table 2 Fallacy: Presenting and Interpreting Confounder and Modifier Coefficients. *American Journal of Epidemiology*, *177*(4), 292–298. https://doi.org/10.1093/aje/kws412

Figures

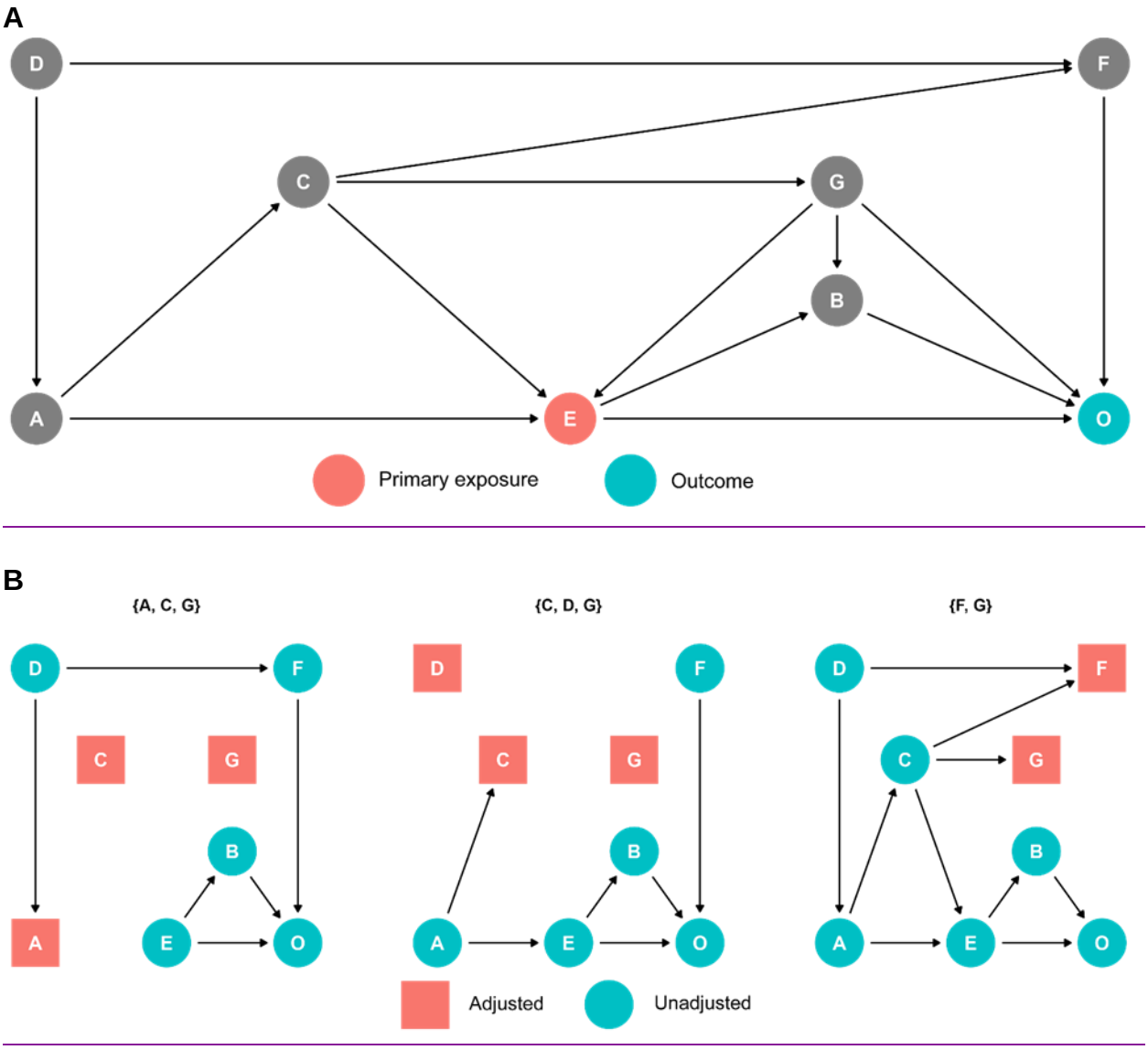
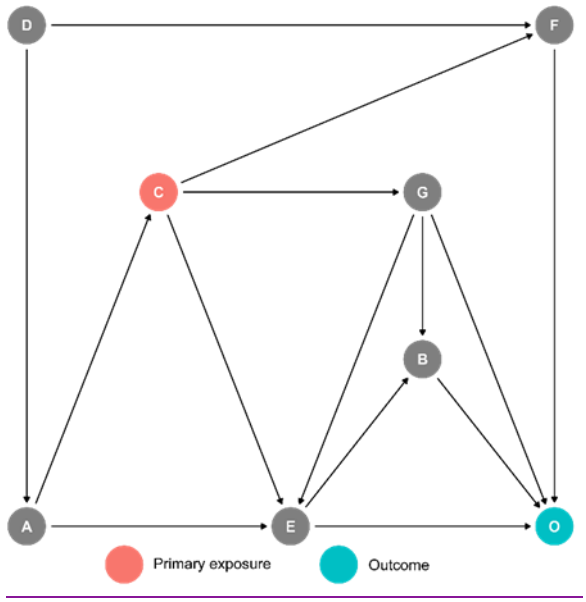


Figure 1. Panel A: Directed acyclic graph of hypothetical causal mechanism.
Panel B: Sufficient sets for unbiased total effect of exposure E on outcome O.

A



B

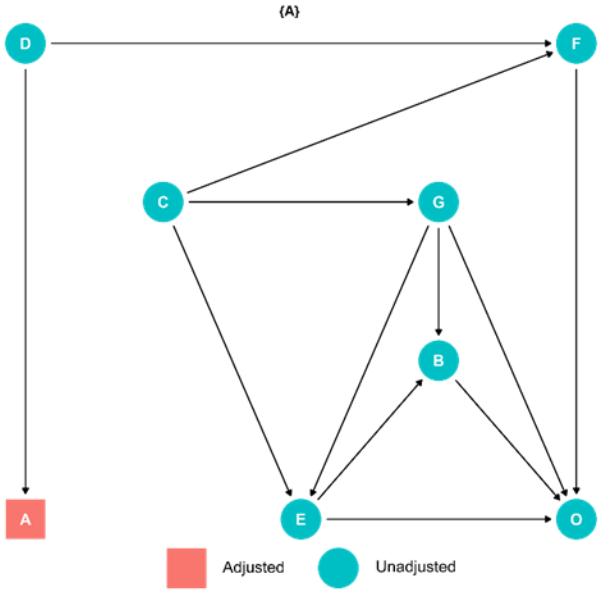


Figure 2. Panel A: Directed acyclic graph of hypothetical causal mechanism, with C as primary exposure. Panel B: Minimally sufficient set for unbiased total effect of C on O.

Tables

Table 1. Simulated multivariable linear regression of E on O, adjusted for A, C, G.

Variable	**Estimate**	**SE**	**95% CI**	**97.5% CI**	***p*-value**
E	0.51	0.15	(0.22, 0.80)	(0.17, 0.85)	0.001
A	0.15	0.18	(-0.20, 0.50)	(-0.26, 0.56)	0.402
C	-0.18	0.17	(-0.51, 0.15)	(-0.55, 0.20)	0.296
G	-0.12	0.20	(-0.51, 0.27)	(-0.57, 0.33)	0.533

CI = Confidence Interval, SE = Standard Error

Table 2. Simulated multivariable linear regression of C on O, adjusted for A.

Variable	**Estimate**	**SE**	**95% CI**	**97.5% CI**	***p*-value**
C	0.36	0.16	(0.05, 0.67)	(-0.01, 0.72)	0.033
A	0.31	0.19	(-0.06, 0.68)	(-0.12, 0.74)	0.109

CI = Confidence Interval, SE = Standard Error

Appendix

DAG Details

Initially, the DAG for the hypothetical causal mechanism in Figure 1A was constructed using the R package `dagitty`, with the following code:

```
dag {
A [pos="0.000,0.000"]
B [pos="1.500,0.250"]
C [pos="0.500,0.500"]
D [pos="0.000,0.750"]
E [exposure,pos="1.000,0.000"]
F [pos="2.000,0.750"]
G [pos="1.500,0.500"]
O [outcome,pos="2.000,0.000"]
A -> C
A -> E
B -> O
C -> E
C -> F
C -> G
D -> A
D -> F
E -> B
E -> O
F -> O
G -> B
G -> E
G -> O
}
```

Subsequently, the R package `ggdag` was used to create the DAGs in Figures 1 and 2.

Simulation Details

The DAG for the hypothetical causal mechanism was replicated using the `dag.init()` function. A total of *N=100* observations were generated from the DAG using the `dag.sim()` function. Observations for each node were assumed to follow a Normal distribution. The mean at each node was determined by the specified direct effects for each arc. In particular, the true direct effect of E on O was specified as 0.50. If a node had no ancestors, the Normal distribution from which observations were drawn was assumed to have a standard deviation of 1. If a node had ancestors, random noise was generated from a Normal distribution with a mean of 0 and a standard deviation of 1. The following summary of the data generation process was produced by `dag.sim()`:

```
Note: undirected arcs are discarded, an xy arc is appended
simulating node 5 from Normal, with mean SD 0 1
node 2 : calculated from nodes 5 via arcs 9 with b 0.1 , shifted to mean 0 (added
noise SD: 1 );
node 4 : calculated from nodes 2 via arcs 5 with b 0.5 , shifted to mean 0 (added
noise SD: 1 );
node 6 : calculated from nodes 4 5 via arcs 7 10 with b 0.35 0.5 , shifted to mean 0
(added noise SD: 1 );
node 7 : calculated from nodes 4 via arcs 8 with b -0.6 , shifted to mean 0 (added
noise SD: 1 );
```

```
node 1 : calculated from nodes 2 4 7 via arcs 4 6 12 with b 0.5 0.45 -0.9 , shifted
to mean 0 (added noise SD: 1 );
node 3 : calculated from nodes 1 7 via arcs 2 13 with b 0.45 -0.7 , shifted to mean 0
(added noise SD: 1 );
node 8 : calculated from nodes 1 3 6 7 1 via arcs 1 3 11 14 15 with b 0.3 0.75 -0.25
0.4 0 , shifted to mean 0 (added noise SD: 1 );
```

Further information on the simulation algorithm can be found within the `dagR` documentation.

The simulation code, for which results are presented in Tables 1 and 2, is provided below:

```
library(dagR)
# REPLICATE DAG IN FIG1A
# Specify exposure, outcome, and covariates as well as their arcs
dag_EO <-
       dag.init(outcome = NULL, exposure = NULL,
       covs = c(1, 1, 1, 1, 1, 1),
       arcs = c(0, -1,
                       0, 2,
                       2, -1,
                       1, 0,
                       1, 3,
                       3, 0,
                       3, 5,
                       3, 6,
                       4, 1,
                       4, 5,
                       5, -1,
                       6, 0,
                       6, 2,
                       6, -1),
             x.name = "Primary exposure",
             symbols   = c("E", "A", "B", "C", "D", "F", "G", "O"),
             y.name = "Outcome")
# Set (x, y) co-ords of nodes
dag_EO$x <- c(1, 0, 1.5, 0.5, 0, 2, 1.5, 2)
dag_EO$y <- c(0, 0, 0.5, 1.0, 2, 2, 1, 0)
# Draw DAG
       dag.draw(dag_EO, noxy = T, legend = F)
# Set seed for reproducibility
       set.seed(12345)
# Specify direct effects for each arc
betas <- c(EO =   0.30,  EB =   0.45,
           BO =   0.75,  AE =   0.50,
           AC =   0.50,  CE =   0.45,
           CF =   0.35,  CG =  -0.60,
           DA =   0.10,  DF =   0.50,
           FO =  -0.25,  GE =  -0.90,
           GB =  -0.70,  GO =   0.40)
# Simulate n = 100 observations based on DAG
sim.dat <- dag.sim(dag_EO,
              b=betas,
```

```
              stdev=rep(1, 14),
              seed=1,
              n=100,
              verbose=FALSE)
# Total effect of E on O
mod_EO <- lm(O ~ E + A + C + G, data=sim.dat)
summary(mod_EO)
# Total effect of C on O
mod_CO <- lm(O ~ A + C, data=sim.dat)
summary(mod_CO
```