Dear Editor:

We read with great interest 'The Oral Health Statistics Guidelines for Reporting Observational Studies and Clinical Trials in Oral Health Research: Explanation and Elaboration'.[1] We are members of the United Kingdom National Institute for Health and Care Research Statistics Group, which aims to translate best methodological practice into applied research. We are therefore delighted to see collaborative efforts to improve reporting, and hope that other health areas might follow, but believe the guidelines are unclear on some important points.

Multivariable modelling is used for many different purposes throughout health research; despite both involving modeling, predicting an outcome versus understanding what causes it (and how to change it) are different tasks requiring tailored analytical approaches rather than a one-size-fits-all solution.[2] The artificial intelligence explosion within health requires researchers to understand this distinction because here inappropriate methodology can have huge impact.

The guidance does not encourage clarity regarding predictive vs causal aims. Specifically, items 15 (variables), 29 (multivariable modeling) and 40 (modeling) do not highlight that variable roles, variable selection and model interpretation differ between the two. Therefore, we suggest that a future revision of the guidelines should go beyond simply classifying items as predictors and response in the context of multivariable modelling.

It is well established, but not widely understood, that you cannot interpret the effects of multiple variables from the same multivariable model. This is deemed the Table 2 Fallacy, because multivariable models are often presented in Table 2 of a publication.[3] To understand whether one particular variable (exposure) is causing an outcome, we need to estimate the strength and direction of that specific relationship, adjusting for confounders, i.e. variables that cause both the exposure and the outcome, but should not include mediators, which sit between them on the causal pathway. This a) adjusts away some of the total causal effect of the exposure, which operates through the mediator (Figure 1) and b) introduces unpredictable

biasing effects that can even reverse the apparent direction of effect.[3] Therefore, if we switch focus to a different exposure variable, we will need a different analysis model, because variables that are confounders of one exposure-outcome relationship will be mediators of another.

The elaboration paper refers to the strengthening the reporting of observational studies in epidemiology (STROBE) guidelines, which pay greater attention in this regard. Section 16a states: *Decisions about excluding or including variables should be guided by knowledge, or explicit assumptions, on causal relations. Inappropriate decisions may introduce bias, for example by including variables that are in the causal pathway between exposure and disease (unless the aim is to assess how much of the effect is carried by the intermediary variable)*.[4] We believe that item 29 might ask whether directed acyclic graphs were used to guide model development and covariate selection[5] and that item 40 should be reworded to discourage interpretation of multiple 'primary factors of interest' from the same model.

We appreciate the huge effort that went into the creation of these excellent guidelines but believe they would be further enhanced were these distinctions to be drawn.

Elizabeth MA Hensor, PhD; Leeds, UK

Catey Bunce, DSc; London, UK

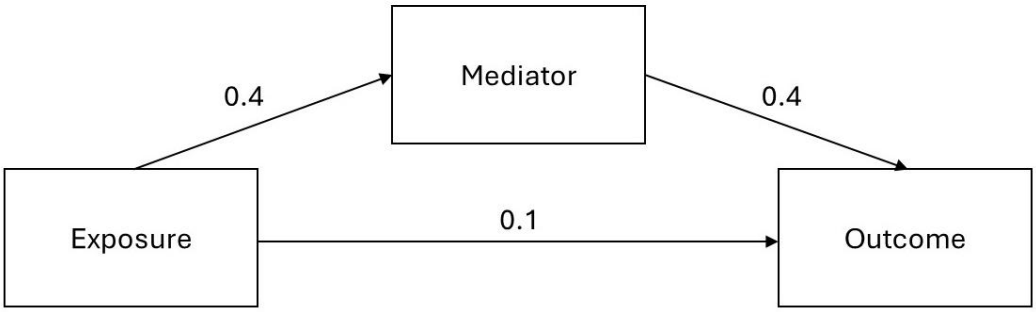
Sam D Leary, PhD; Bristol, UK

References

1. Best AM, Lang TA, Greenberg BL, Gunsolley JC, Ioannidou E; Task Force on Design and Analysis in Oral Health Research. The Oral Health Statistics Guidelines for Reporting Observational Studies and Clinical Trials in Oral Health Research: Explanation and Elaboration. **J Oral Maxillofac Surg**. 2024 Jul 12:S0278-2391(24)00584-6. doi: 10.1016/j.joms.2024.06.174. Epub ahead of print.
2. Hernán MA, Hsu J, Healy B. A second chance to get causal inference right: a classification of data science tasks. **Chance** 2019;32:42–9.
3. Westreich D, Greenland S. The table 2 fallacy: presenting and interpreting confounder and modifier coefficients. **Am J Epidemiol**. 2013 Feb 15;177(4):292-8.
4. Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, Poole C, Schlesselman JJ, Egger M; STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. **Int J Surg**. 2014 Dec;12(12):1500-24.
5. Tennant PWG, Murray EJ, Arnold KF, Berrie L, Fox MP, Gadd SC, Harrison WJ, Keeble C, Ranker LR, Textor J, Tomova GD, Gilthorpe MS, Ellison GTH. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. **Int J Epidemiol**. 2021 May 17;50(2):620-632.

Figure legend

Figure 1: Basic directed acyclic graph highlighting that including a mediator adjusts away some of the total causal effect of an exposure on an outcome



In this very simplified example, the total effect of the exposure is 0.1 + (0.4 x 0.4) = 0.26.

However, if we included the mediator in the analysis model, the coefficient for the exposure would be just 0.1. In the absence of other bias, this is the direct effect, not the total effect. We might decide the exposure did not have a strong effect on the outcome, as the coefficient from the model was small; however, intervening on the exposure would have a larger total effect.