populations from diverse settings. When we choose to employ statistical techniques involving propensity scores, as well as evaluate studies that use these methods, we must evaluate both internal and external validity to ensure our conclusions drawn from observational data are robust and reliable.

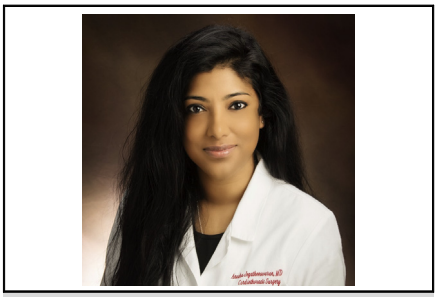
References

1. Baiocchi M, Woo YJ, Chiu P, Goldstone AB. The role and significance of sensitivity analyses in enhancing the statistical validity of clinical studies. *J Thorac Cardiovasc Surg*. 2022;163:749-53.
2. Freemantle N, Marston L, Walters K, Wood J, Reynolds MR, Petersen I. Making inferences on treatment effects from real world data: propensity scores, confounding by indication, and other perils for the unwary in observational research. *BMJ*. 2013;347:f6409.

See Article page 749.

Commentary: Sensitivity analyses: Mitigating the problem of garbage in, equals garbage out?

Anusha Jegatheeswaran, MD, PhD, FRCSC



Anusha Jegatheeswaran, MD, PhD, FRCSC

CENTRAL MESSAGE

Those both consuming and producing research in the domain of cardiac surgery should be aware of the utility of sensitivity analyses to mitigate the problem of garbage in, equals garbage out.

In this issue of the *Journal*, Baiocchi and colleagues[1] present the importance of sensitivity analyses in improving the validity and diminishing the bias inherent in observational studies. Of course, the gold standard in outcome evaluation, due to its ability to eliminate bias in studies, is the implementation of randomized controlled trials to answer questions where the determination of causality is the primary goal. However, in cardiac surgery, there are many questions for which a randomized controlled trial is simply not possible, practical, or in some cases ethical. This may occur due to a lack of equipoise, the inability of different sites to carry out techniques with which they are unfamiliar, or secondary to exceedingly small patient numbers, which is often the case in the domain of pediatrics. As such, we are often left to rely on observational data and quasi-experimental designs to provide us with the best possible results to inform important decisions regarding patient care.

Those unfamiliar with performing statistical analyses themselves may not truly appreciate the number of "decisions" that must be made to get from a raw data set to a final output with conclusions. These decisions can have a wide spectrum of influence on the results of any study and may range from something as simple as how to split patients into comparison groups, to more complex decisions involving the specific nature of how study variables are created, and determination of the variables used when matching patients. While it may seem like these are tiny decisions of little consequence, ultimately it is this multitude of decisions which, when compounded, can cause the results to lie farther and farther from the truth. This is in addition to issues that may arise from small datasets or poor data quality.

Similarly, those without a statistical background may not be aware that all statistical models are based on assumptions. As such, the validity of the conclusions that can be drawn from these models is highly dependent on the degree to which the assumptions are met. The primary drawback of

From the Division of Cardiovascular Surgery, Department of Surgery, Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada.

Disclosures: The author reported no conflicts of interest.

The *Journal* policy requires editors and reviewers to disclose conflicts of interest and to decline handling or reviewing manuscripts for which they may have a conflict of interest. The editors and reviewers of this article have no conflicts of interest.

Received for publication Sept 20, 2020; revisions received Sept 20, 2020; accepted for publication Sept 23, 2020; available ahead of print Sept 28, 2020.

Address for reprints: Anusha Jegatheeswaran MD, PhD, FRCSC, Division of Cardiovascular Surgery, Hospital for Sick Children, 555 University Ave, Toronto, Ontario M5G 1X8, Canada (E-mail: anusha.jegatheeswaran@utoronto.ca).

J Thorac Cardiovasc Surg 2022;163:755-6
0022-5223/$36.00

https://doi.org/10.1016/j.jtcvs.2020.09.086

observational research is that we often must assume that there are no unmeasured confounders. However, violation of this assumption can invalidate our results. This is where sensitivity analyses become important and, in particular, gamma sensitivity for studies using propensity matching, which is the focus of this manuscript.

I congratulate the authors for this contribution that provides an entrée for surgeons to achieve a greater understanding of sensitivity analyses including their interpretation, utility, and limitations, with a focus on the methods used for studies involving propensity score–based matching. My hope is that surgeons will come away understanding that the data, statistical decisions, and assumptions made in observational studies can have a significant impact on the outcomes that may result and the robustness of study conclusions. Most importantly, my hope is that they will take away an appreciation that quantitative methods exist to assess the degree of robustness of results, not only in studies they are reading, but how these may apply to their own work to mitigate issues that may result from garbage in equaling garbage out.

Reference

1. Baiocchi M, Woo YJ, Chiu P, Goldstone AB. The role and significance of sensitivity analyses in enhancing the statistical validity of clinical studies. *J Thorac Cardiovasc Surg*. 2022;163:749-53.