Inferring causality

Daniel J. Wong, MD, MHS, *and* Elliot L. Chaikof, MD, PhD, *Boston, Mass*

Whether a ruptured abdominal aortic aneurysm (rAAA) is best treated via open surgical repair (rOSR) or endovascular repair (rEVAR) remains a topic of debate. Despite the preponderance of observational data demonstrating that rEVAR is superior, the handful of prospective, randomized trials comparing rEVAR with rOSR have not demonstrated a significant short-term survival benefit. The largest of these trials, the IMPROVE trial, did demonstrate an advantage with rEVAR in midterm survival and cost effectiveness.[1] However, significant concerns persist about the generalizability of the randomized, controlled trials in rAAA. Thus, many consider the real-world observational studies favoring rEVAR to better reflect reality and, indeed, society guidelines have evolved to advocate for an rEVAR strategy first, if anatomically feasible.[2]

Articles by Wang et al,[3] as well as Lundgren and Troëng[4] in this edition of the *Journal of Vascular Surgery*, provide further insight regarding outcomes after rAAA with notably discordant results. Wang et al evaluate cases of rAAA from the Vascular Quality Initiative (VQI) database between 2003 and 2018 to construct a large propensity-matched cohort demonstrating significantly decreased mortality at 30 days for rEVAR vs rOSR (18% vs 32%) and improved all-cause survival at 1 year for rEVAR vs rOSR (73% vs 59%). In contrast, Lundgen and Troëng present the results of all operative rAAA cases in Sweden between 2012 and 2015 with the adjusted midterm results (3 months to 3 years) favoring rOSR (hazard ratio, 0.63; 95% confidence interval, 0.43-0.93). Short-term mortality rates do favor rEVAR, although this survival advantage is lost after 90 days.

Traditionally, randomized clinical trials provide the most direct approach to determine causality between treatment and outcome. However, as is the case for rAAA, the difficulty of conducting randomized trials means that comparative effectiveness studies increasingly come from observational data. There are several theories on how to infer causality from observational data.[5,6] A central principle of these theories is the necessity to control for confounding factors. This goal can be accomplished with methods such as multivariable regression or propensity score matching as Lundgen and Troëng and Wang et al have conducted in these articles. Regardless of the analytic method, a critical step in these analyses is the selection and inclusion of potential confounders. Inadequate covariate inclusion allows residual confounding or bias, whereas inappropriate covariate designation, for instance of an instrumental variable (a variable strongly associated with the treatment rather than outcome), may lead to bias amplification rather than reduction.[7] There is no single correct method for covariate selection or variable inclusion in a model.[8] In these reports, the authors use different methods for variable selection; however, they ultimately controlled for similar factors in their analyses and yet achieved opposing results with respect to midterm survival.

Unmeasured, unknown, or unavailable variables present serious challenges to determining causality from observational studies. For instance, in these studies neck anatomy is known for only one-third of rEVAR patients in the VQI and none of the patients who underwent rOSR or were included in the Swedish cohort. Thus, one could imagine that an uncontrolled bias may exist against patients who had a hostile neck or difficult anatomy and more frequently underwent rOSR in the VQI but not in the Swedish cohort. Unfortunately, no registry is comprehensive for all factors that potentially influence treatment selection or outcomes. Investigators hope that, by the inclusion of appropriate variables, unknown but important variables become balanced. Although there are limitations to all registry data, it is important to acknowledge the strength of the VQI and the Swedvasc cohort in that they capture critical details such as hemodynamic instability, which was not available in previous studies with administrative data.

Both multivariable analysis and propensity score matching attempt to arrive at causal conclusions using analytic methods for controlling confounders in observational data.[9] Another observational method for causal inference is to construct an instrument-based study or quasi-experimental study. In their analysis, Lundgen and Troëng group hospitals into those that routinely perform both EVAR and OSR and those that almost exclusively perform open repairs (OAR). In this case, the instrument is the hospital OAR or EVAR/OSR practice and the results are surprising in that short-term mortality is decreased with an OAR-only practice. An important consideration with instrument-based studies is to be

From the Department of Surgery, Beth Israel Deaconess Medical Center, Harvard Medical School.

Author conflict of interest: none.

Correspondence: Elliot L. Chaikof, MD, PhD, Department of Surgery, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA 02215 (e-mail: echaikof@bidmc.harvard.edu).

The editors and reviewers of this article have no relevant financial relationships to disclose per the JVS policy that requires reviewers to decline review of any manuscript for which they may have a conflict of interest.

J Vasc Surg 2020;72:387-8

0741-5214

Copyright © 2020 by the Society for Vascular Surgery. Published by Elsevier Inc.

https://doi.org/10.1016/j.jvs.2020.01.036

conscious of the research question posed by the instrument. Thus, it is not rOSR vs rEVAR, but rather OAR-only practice vs EVAR/OSR mixed practice. Although not answering the question of efficacy of rEVAR vs rOSR, this type of analysis is very helpful to examine the effectiveness of an OAR-only practices with significant policy implications given the costs to enable a hospital to perform rEVAR.

When confronted with discordant results, we must resist the temptation to choose the study that supports a personal preference. Instead, we must decide if the inferred causality between treatment and outcome is valid both internally and externally. Wang et al and Lundgen and Troëng present contemporary, relevant, and seemingly internally valid studies comparing rEVAR with rOSR, with opposing results. There are clear systemic differences between hospitals contributing to the VQI and the Swedish healthcare system that limit their generalizability to one another. Nevertheless, it would be an error to attribute these conflicting outcomes simply to context. Choices in methodology, analysis, or presentation often determine a study's findings. The authors of these two articles are to be commended on their disclosure of methodology and analysis. However, in the era of mandated data sharing in clinical trials, we should require more from authors of observational studies than the inclusion of a comprehensive methods section alone. When retrospective studies are of sufficient impact to change or challenge practice, sharing of patient-level data is necessary. Initial studies from the International Consortium of Vascular Registries demonstrate not only the power and complexity of data sharing, but also its feasibility.[10] In the absence of clear and convincing evidence from randomized trials, we have an obligation to maximize the quality, reproducibility, and transparency of our observational findings. It is only through this difficult and collaborative dissection of the rapidly evolving real-world practice of modern vascular surgery that we can ensure optimal patient care.

REFERENCES

1. IMPROVE Trial Investigators. Comparative clinical effectiveness and cost effectiveness of endovascular strategy v open repair for ruptured abdominal aortic aneurysm: Three year results of the IMPROVE randomised trial. BMJ 2017;359:j4859.
2. Chaikof EL, Brewster DC, Dalman RL, Makaroun MS, Illig KA, Sicard GA, et al. SVS practice guidelines for the care of patients with an abdominal aortic aneurysm: executive summary. J Vasc Surg 2009;50:880-96.
3. Wang LJ, Locham S, Al-Nouri O, Eagleton MJ, Clouse WD, Malas MB. Endovascular repair of ruptured abdominal aortic aneurysm is superior to open repair: propensity-matched analysis in the Vascular Quality Initiative. J Vasc Surg 2020;72:498-507.
4. Lundgren F, Troeng T. Treatment choice and survival after ruptured abdominal aortic aneurysm: a population-based study. J Vasc Surg 2020;72:508-17.
5. Little RJ, Rubin DB. Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. Annu Rev Public Health 2000;21:121-45.
6. Pearl J. An introduction to causal inference. Int J Biostat 2010;6.
7. Pearl J. Invited commentary: understanding bias amplification. Am J Epidemiol 2011;174:1223-7.
8. Kahlert J, Gribsholt SB, Gammelager H, Dekkers OM, Luta G. Control of confounding in the analysis phase — an overview for clinicians. Clin Epidemiol 2017;9:195-204.
9. Matthay EC, Hagan E, Gottlieb LM, Tan ML, Hlahov D, Adler NE, et al. Alternative causal inference methods in population health research: Evaluating tradeoffs and triangulating evidence. SSM Popul Heal 2020;10:100526.
10. Beck AW, Sedrakyan A, Mao J, Venermo M, Faizer R, Debus S, et al. Variations in abdominal aortic aneurysm care: a report from the international consortium of vascular registries. Circulation 2016;134:1948-58.

Submitted Jan 11, 2020; accepted Jan 17, 2020.