10 Cipolletta S, Ravasio G, Bussotti M. Sexual and reproductive health in women with pulmonary hypertension: a qualitative study. *Arch Sex Behav* 2022;51:1647–1657.

11 Pulmonary Hypertension Association. Living with PH: a guide to intimacy. Silver Spring, MD: Pulmonary Hypertension Association; 2019. Available from: https://phassociation.org/a-guide-to-intimacy/.

12 Greenwood BN, Carnahan S, Huang L. Patient-physician gender concordance and increased mortality among female heart attack patients. *Proc Natl Acad Sci USA* 2018;115:8569–8574.

13 Greenwood BN, Hardeman RR, Huang L, Sojourner A. Physician-patient racial concordance and disparities in birthing mortality for newborns. *Proc Natl Acad Sci USA* 2020;117:21194–21200.

14 Galiè N, Channick RN, Frantz RP, Grünig E, Jing ZC, Moiseeva O, *et al.* Risk stratification and medical therapy of pulmonary arterial hypertension. *Eur Respir J* 2019;53:1801889.

15 Vizza CD, Lang IM, Badagliacca R, Benza RL, Rosenkranz S, White RJ, *et al.* Aggressive afterload lowering to improve the right ventricle: a new target for medical therapy in pulmonary arterial hypertension? *Am J Respir Crit Care Med* 2022;205:751–760.

16 Rubin EB, Buehler AE, Halpern SD. States worse than death among hospitalized patients with serious illnesses. *JAMA Intern Med* 2016;176:1557–1559.

Check for updates

Knowledge from the Noise: A Regression Discontinuity Design to Inform Optimal Transfusion Thresholds for Critically Ill Patients

Stephanie P. Taylor, M.D., M.S.[1], and Andrew J. Admon, M.D., M.P.H., M.Sc.[2,3]

[1]Department of Internal Medicine, Wake Forest School of Medicine Atrium Health Enterprise, Critical Illness, Injury, and Recovery Research Center, Wake Forest School of Medicine, Winston-Salem, North Carolina; [2]Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan; and [3]University of Michigan Institute for Healthcare Policy and Innovation, Ann Arbor, Michigan

ORCID ID: 0000-0002-7432-3764 (A.J.A.).

The landmark TRICC (Transfusion Requirements in Critical Care) trial found no difference in 30-day mortality for critically ill patients treated using a restrictive (hemoglobin <7 g/dl) compared with liberal (hemoglobin <10 g/dl) transfusion threshold (1). After this, the optimal threshold for administering packed red blood cells (pRBCs) has been a topic of great interest for clinicians and researchers. Physiologically, pRBC transfusions improve tissue oxygenation by increasing oxygen-carrying capacity. However, the theoretical benefits of increased oxygen delivery are balanced by potential harms of pRBC transfusion, such as volume overload, transfusion reactions, and infectious disease transmission (2–5). In addition, blood transfusion for those unlikely to benefit is costly, inconsistent with high-value health care, and diminishes the availability of blood products for others (6).

The results of the TRICC trial and other evidence have been broadly incorporated by practice guidelines to support pRBC at a hemoglobin level of <7 g/dl for critically ill but clinically stable intensive care unit (ICU) patients over more liberal transfusion practices (7). Notably, the threshold of 7 g/dl represents the upper bound of conservative strategies used in trials but has itself never been shown to be superior to lower thresholds (or to physiologic triggers without a laboratory-based threshold). There has been little enthusiasm to date to allocate patients to a transfusion threshold <7 g/dl, so randomized controlled trial (RCT) data to guide optimal transfusion thresholds are lacking.

Fortunately, advances in healthcare data richness and analytic techniques for causal estimation have enabled more reliable comparative effectiveness estimates outside of RCTs. A major challenge in generating unbiased effect estimates from such observational data is confounding by indication—when factors influencing a clinician's treatment decision (e.g., severity of illness or likelihood to benefit) also influence outcomes. Common techniques to adjust for these factors may fail to fully account for the differences between patients that clinicians use to make treatment decisions, or data sources may not contain sufficient data to accurately model these factors. The unmeasured confounding that remains violates the assumptions necessary for causal conclusions from many observational study designs. This can mean false or distorted results regarding a treatment's benefits and harms.

Quasi-experimental study designs aim to address this limitation by identifying naturally occurring variation in receipt of a treatment or intervention. Practically, naturally occurring variation is induced by measurement error in in data such as
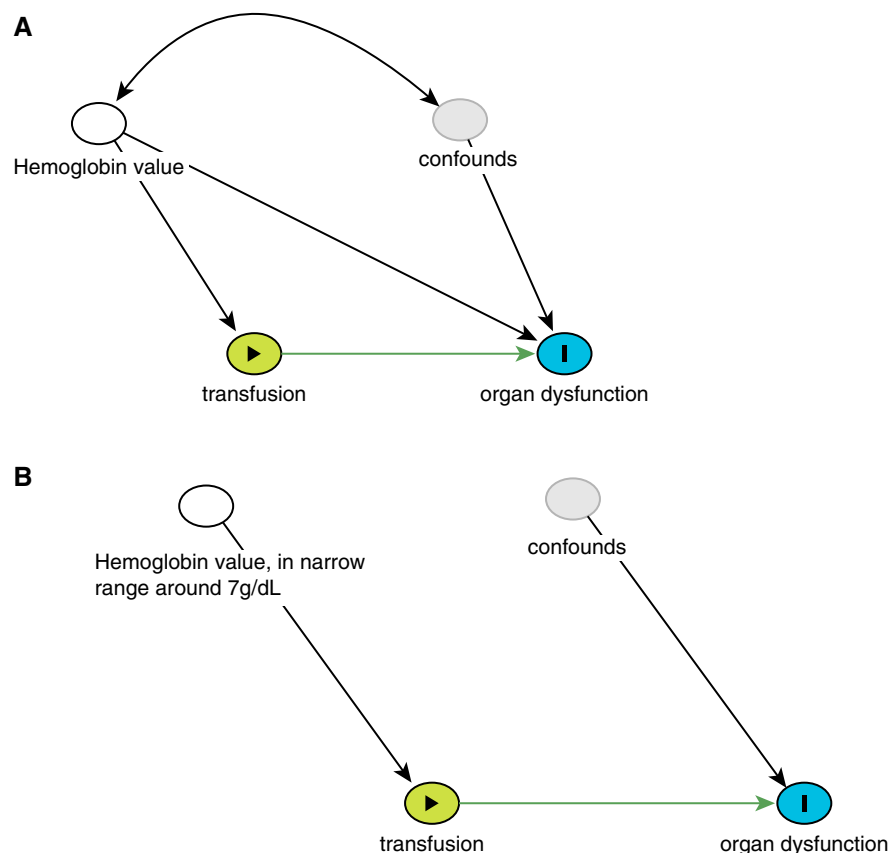
Figure 1. Causal graphs representing regression discontinuity design. (*A*) Data generating graph showing hemoglobin threshold of 7 g/dl as the assignment variable resulting in transfusion. (*B*) Data limited graph in which hemoglobin values are restricted to a narrow range around 7 g/dl. The central assumption of regression discontinuity designs is that severity of illness (and other confounders) are no longer related to hemoglobin levels in such a narrow band, so there is no confounding.

laboratory results or standardized test scores. This natural variation is highly unlikely to be related to an outcome of interest, One technique for leveraging such variation for causal inference is a regression discontinuity design (RDD). RDD is a powerful quasi-experimental design for use when interventions are assigned on the basis of a cutoff value on a scale measured with noise (e.g., a hemoglobin of 7 g/dl or a test score of 80%). The underlying assumption of the RDD is that patients with scores just above versus just below the threshold value differ in their likelihood of receiving an intervention but are otherwise very similar. In effect, this noise functions similarly to a coin flip in an RCT, determining receipt of treatment independent of other patient characteristics. RDD analyses can be "sharp" when the relationship between the treatment and the running variable is absolute or "fuzzy"

when the threshold does not deterministically move all observations from unexposed to exposed, but rather increases the probability of receiving treatment.

In this issue of *AnnalsATS*, Bosch and colleagues (pp. 1177–1184) leveraged the noise in measurement of hemoglobin to set up a "fuzzy" RDD that estimates the effect of transfusing pRBCs at a threshold of hemoglobin 7 g/dl by comparing outcomes in patients just above with those just below the threshold (8). Selecting from three large, feature-rich datasets, Bosch and colleagues identified 191,987 patients who were admitted to an ICU with at least one hemoglobin measurement on Days 2–28 of ICU admission and who did not have active bleeding or myocardial infarction (conditions that may warrant higher hemoglobin targets

and that likely would have resulted in exclusion from a comparable RCT). The assumptions necessary for fuzzy RDD were met, including general balance of other patient characteristics for patients just above and just below the hemoglobin threshold. Overall, the authors found an abrupt increase (discontinuity) in transfusion at the hemoglobin threshold but no improvement in organ dysfunction as measured by the maximum Sequential Organ Failure Assessment score within 72 hours after transfusion. In an exploratory analysis among patients with sepsis, pRBC transfusion was associated with increased organ dysfunction. Extensive sensitivity and secondary analyses largely confirmed the primary findings.

There are considerable strengths to this study. The combined dataset is large, generalizable, and granular with information to permit detailed analysis and check of necessary assumptions. The use of RDD mitigates the risk of residual confounding that has not been possible in prior studies. Indeed, several studies have shown that RDDs are remarkably effective at replicating results from RCTs, whereas other observational designs often are not (9–11). The fewer assumptions required by RDD to identify a causal effect compared with model-based designs (e.g., regression adjustment) are clear when representing the RDD with a causal graph (12, 13). Figure 1 shows data-generating and data-limited directed acyclic graphs representing Bosch and colleagues' RDD. In the full dataset (Figure 1A), transfusion decisions are driven by multiple factors that may also affect the outcome. This confounding leads to biased estimates unless all these factors are measured and adjusted for. In the limited dataset of patients with hemoglobin values very close to the threshold (Figure 1B), transfusion decisions are only caused by the random noise in hemoglobin measurement and not by severity of illness and other confounder variables. In other words, measurement error in the hemoglobin value serves a role similar to randomization in an RCT.

However, even though the data were gathered from rich clinical databases, they remain secondary data with inherent limitations such as missingness and accuracy. Furthermore, contextual knowledge of hospital or unit transfusion

protocols or knowledge of clinician-level decision making is not available in these types of studies. For example, it is possible that hospitals or physicians with greater propensity to transfuse above the threshold are also more likely to provide other unproven therapies that may inadvertently cause harm. These possibilities could not be explored thoroughly in the study datasets. Finally, clinicians wishing to apply these findings broadly to critically ill patients must keep in mind that the results of the fuzzy RDD are only applicable to the subpopulation of patients whose hemoglobin values are in a narrow range around the threshold.

Study limitations notwithstanding, the lack of demonstrable benefit of a transfusion threshold of 7 g/dl challenges conventional wisdom about the usefulness of an intervention that can cause harm and incurs substantial cost. The results of this study may provide adequate equipoise for a future RCT that better establishes optimal transfusion thresholds. Alternatively, advancements in physiologic measurement and "big data" methodology may soon permit individualized transfusion recommendations based on time-varying clinical data and real-time predictions regarding the probability of benefit. In the meantime, the present study is an elegant demonstration of the role that quasi-experimental methods can play in providing real-world evidence when RCT data are lacking. This approach is already supported by the 21st Century Cures Act requiring the U.S. Food and Drug Administration to consider real-world evidence during approval decisions (14). ■

References

1 Hébert PC, Wells G, Blajchman MA, Marshall J, Martin C, Pagliarello G, *et al.* A multicenter, randomized, controlled clinical trial of transfusion requirements in critical care. Transfusion Requirements in Critical Care Investigators, Canadian Critical Care Trials Group. *N Engl J Med* 1999;340:409–417.

2 Vincent JL, Sakr Y, Sprung C, Harboe S, Damas P; Sepsis Occurrence in Acutely Ill Patients (SOAP) Investigators. Are blood transfusions associated with greater mortality rates? Results of the Sepsis Occurrence in Acutely Ill Patients study. *Anesthesiology* 2008;108:31–39.

3 Toy P, Popovsky MA, Abraham E, Ambruso DR, Holness LG, Kopko PM, *et al.*; National Heart, Lung and Blood Institute Working Group on TRALI. Transfusion-related acute lung injury: definition and review. *Crit Care Med* 2005;33:721–726.

4 Zallen G, Offner PJ, Moore EE, Blackwell J, Ciesla DJ, Gabriel J, *et al.* Age of transfused blood is an independent risk factor for postinjury multiple organ failure. *Am J Surg* 1999;178:570–572.

5 Vincent JL, Baron JF, Reinhart K, Gattinoni L, Thijs L, Webb A, *et al.*; ABC (Anemia and Blood Transfusion in Critical Care) Investigators. Anemia and blood transfusion in critically ill patients. *JAMA* 2002;288:1499–1507.

6 Trentino KM, Farmer SL, Swain SG, Burrows SA, Hofmann A, Ienco R, *et al.* Increased hospital costs associated with red blood cell transfusion. *Transfusion* 2015;55:1082–1089.

7 Vlaar AP, Oczkowski S, de Bruin S, Wijnberge M, Antonelli M, Aubron C, *et al.* Transfusion strategies in non-bleeding critically ill adults: a clinical practice guideline from the European Society of Intensive Care Medicine. *Intensive Care Med* 2020;46:673–696.

8 Bosch NA, Law AC, Bor J, Myers LC, Roubinian NH, Liu VX, *et al.* Red blood cell transfusion at a hemoglobin threshold of 7 g/dl in critically ill patients: a regression discontinuity study. *Ann Am Thorac Soc* 2022;19:1177–1184.

9 Buddelmeyer H, Skoufias E. An Evaluation of the Performance of Regression Discontinuity Design on PROGRESA. World Bank Policy Research Working Paper No. 3386; IZA Discussion Paper No. 827. Washington, DC: World Bank; August 25, 2004.

10 Berk R, Barnes G, Ahlman L, *et al.* When second best is good enough: a comparison between a true experiment and a regression discontinuity quasi-experiment. *J Exp Criminol* 2010;6:191–208.

11 Shadish WR, Galindo R, Wong VC, Steiner PM, Cook TD. A randomized experiment comparing random and cutoff-based assignment. *Psychol Methods* 2011;16:179–191.

12 Steiner PM, Kim Y, Hall CE, Su D. Graphical models for quasi-experimental designs. *Sociol Methods Res* 2017;46:155–188.

13 Lederer DJ, Bell SC, Branson RD, Chalmers JD, Marshall R, Maslove DM, *et al.* Control of confounding and reporting of results in causal inference studies. Guidance for authors from editors of respiratory, sleep, and critical care journals. *Ann Am Thorac Soc* 2019;16:22–28.

14 Kesselheim AS, Avorn J. New "21st Century Cures" legislation: speed and ease vs science. *JAMA* 2017;317:581–582.