

# Observing Blood Management Programs through the Retrospectroscope

Richard J. Cook, B.Sc., M.Math, Ph.D., Richard B. Weiskopf, M.D.

**I**N this issue of ANESTHESIOLOGY, Gupta *et al.*<sup>1</sup> report on the results of a retrospective cohort study comparing transfusion practices and clinical outcomes before and after the implementation of a blood management program in orthopedic surgery at a single center. The motivation stems in part from interest in reducing the number of transfusions by the use of more restrictive hemoglobin triggers for red blood cell transfusions in orthopedic surgery patients. The authors observe that both lower utilization and comparable or improved patient outcomes followed implementation of the blood management program and conclude that a “hemoglobin threshold of 7 g/dl appears to be safe for many orthopedic patients.” There is a clear need to understand the relationship between transfusion triggers and outcomes to ensure that limited resources are used judiciously, to minimize exposures of patients, and to optimize patient outcomes.

The study by Gupta *et al.*<sup>1</sup> offers a good illustration of the kinds of retrospective analyses often conducted based on data from large registries or administrative databases. There have been a number of studies in transfusion medicine which involved retrospective database analyses yielding findings that, when tested in prospective randomized trials, were not validated. For example, in a large retrospective analysis of 4,470 intensive care unit patients, Hébert *et al.*<sup>2</sup> observed an association between lower hemoglobin concentrations and death. However, when they tested this hypothesis in a subsequent randomized trial in 838 similar intensive care unit patients, there was no evidence of differences in mortality between the liberally and restrictively transfused groups that were transfused to hemoglobin concentrations of 10.7 and 8.5 g/dl.<sup>3</sup> In a retrospective analysis directed at the effect of red blood cell storage duration, Koch *et al.*<sup>4</sup> suggested that cardiac surgery patients transfused with red cells



***“Large databases are appealing to exploit for investigating focused scientific questions, but the data necessary for a rigorous analysis are often lacking ...”***

of anemia-induced neurocognitive function in a wide range of populations: cardiac surgery, critically ill adults, children with severe anemia, low-weight premature infants, all hospitalized patients, and healthy volunteers.<sup>8–14</sup>

The current publication affords an opportunity to discuss some challenges arising in retrospective analyses, which are highlighted below. The themes include the *post hoc* definition exposure variables and the interpretation of their effects, the challenge of dealing completely and rigorously with the effect of confounding variables, incomplete data, and the use of composite outcomes. These, and other issues, are important to bear in mind when trying to explain conflicting findings between publications on different database analyses, and the results of randomized trials.

## **Post Hoc Definition of Exposure**

Although a central theme of this work is examination of the effect of a new blood management program on red blood cell

stored for longer periods of time experienced a higher mortality rate than did patients transfused with red cells having shorter storage durations; a subsequent larger retrospective analysis of all transfusions in Denmark and Sweden by Edgren *et al.*<sup>5</sup> showed different results. Exposure of patients with cardiovascular disease to red cells of particularly long storage duration was also associated with increased in-hospital mortality in a retrospective registry analysis by Eikelboom *et al.*<sup>6</sup> but these findings were not validated in a subsequent analysis based on an expanded dataset.<sup>7</sup> Seven prospective randomized trials addressing this question did not substantiate the findings of Koch *et al.*, finding no difference between red cells storage duration and mortality, change in multiple organ dysfunction scores, composite morbidity, pulmonary and immune function, lactate clearance, and rever-

Image: ©ThinkStock.

Corresponding article on page 1082.

Accepted for publication September 10, 2018. From the Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada (R.J.C.); and the Department of Anesthesia and Perioperative Care, University of California, San Francisco, California (R.B.W.).

Copyright © 2018, the American Society of Anesthesiologists, Inc. Wolters Kluwer Health, Inc. All Rights Reserved. Anesthesiology 2018; 129:1060-2

use and outcome, some statements made by the authors suggest a causal effect of hemoglobin threshold on clinical outcome. The *post hoc* definition of hemoglobin threshold used here is “the lowest (nadir) hemoglobin concentration during the hospital stay.” This crude summary of exposure over the course of a hospitalization may be reasonable for descriptive purposes, but there is danger in overinterpreting the relation between such an exposure summary and its association with the composite morbidity and mortality outcome. The principle reason is that this minimum hemoglobin concentration was simply an observed value over a period of time, with an unknown temporal relationship to any morbid event, rather than an actual predefined threshold as would be specified in a prospective randomized trial. The statement “[t]o our knowledge, ours is the first study in orthopedics to assess hemoglobin thresholds as low as 7 g/dl” and the concluding statement in the abstract are therefore inappropriate.

### Ecological Fallacy and Confounding

Large databases are appealing to exploit for investigating focused scientific questions, but the data necessary for a rigorous analysis are often lacking; the lack of preoperative hemoglobin values is one such example in this article, but there is a myriad of factors influencing patient care, some of which are dynamic and responsive to early treatment. When a relatively small number of factors are adjusted for such as age, sex, hip fracture status, surgical procedure, and a case-mix index are used, concerns arise about whether the data are sufficient and the adjustments are adequate. A prospective observational study would have enabled collection of a more comprehensive set of variables enabling a more complete adjustment for potential confounders and findings more consistent with the prospective interpretation of real clinical interest. The rationale for many interventions is not generally available in databases; information on the reason for the transfusion would also be useful for causal analysis and would be easily collected in a prospective study. We also note that the propensity score used by the authors appears to include the same factors adjusted for in the multivariate regression analysis, so it does not represent a causal sensitivity analysis in the usual sense. Some principle advantages of propensity score analyses include the ability to adjust for a larger number of confounders while maintaining a relatively simple model for the outcome. This can be achieved by matching, stratification, or regression on the propensity score, or using inverse probability of exposure weights. The latter would yield estimates of the effect of the blood management program on the response more in line with what would be estimated in a randomized clinical trial.<sup>15</sup>

The selection criteria for potentially confounding variables to adjust for raises challenging issues. Confounding variables have an association with the outcome and the exposure variable, and although this is a relatively simple concept when dealing with cross-sectional studies, when exposure variables change over time in complex feedback systems,

identifying, selecting, and modeling the effect of confounding variables is a daunting challenge. Moreover, selection of variables to adjust for in causal analyses should be based on scientific context rather than statistical significance. We refer readers to a timely, sobering, and stimulating recent paper by Hernán,<sup>16</sup> from which one can learn to calibrate expectations and interpretations from registry-based studies.

### Composite Outcomes

There is often compelling practical rationale for use of composite outcomes but these also introduce substantial challenges in interpretation of findings.<sup>17</sup> This challenge is particularly important when the components of a composite outcome are of unequal importance, when they represent quite different clinical outcomes, and when the relative weighting of the components is unclear in the final analysis. The magnitude of the effect reported on the composite outcome in Gupta *et al.* is striking, but for reasons stated earlier caution is warranted before attributing this large effect to a lower hemoglobin threshold. With such a large effect, however, it should be feasible to carry out a randomized clinical trial confirming this finding in the elderly population of orthopedic surgery patients. Some may consider that this had already been investigated in the prospective, randomized Transfusion Trigger Trial for Functional Outcomes in Cardiovascular Patients Undergoing Surgical Hip Fracture Repair (FOCUS) trial, in which patients age 65 yr or more with a history of, or risk factors for, cardiovascular disease and undergoing surgical repair of hip fracture were randomly allocated to a liberal or restrictive transfusion protocol.<sup>18</sup> There were no differences in functional recovery or mortality found in this trial. Even such a randomized, clinical trial can have limitations, however: there was a highly significant increase in the use of “rescue” transfusion for cardiovascular symptoms (*i.e.*, red blood cell transfusion at a trigger greater than that specified by the protocol) in the restrictive group compared with the liberal group. Moreover, the population in the FOCUS trial was quite different than that analyzed by Gupta *et al.* We note that the numbers of patients experiencing even the composite outcome are quite low. To gain better insight into the nature of any effects, larger samples would be useful so that the effects could be explored in the component outcomes of the composite outcome in a meaningful way.<sup>19</sup> Finally, we note that the FOCUS trial is yet another example of a randomized trial yielding different results from the database analyses from which it was spawned.<sup>20</sup>

In many circumstances it is not possible to conduct randomized clinical trials, and other types of data are the best that can be obtained. Randomized trials are possible to test the hypothesis that less frequent red cell transfusion does not increase risk, and the results have been mixed, attesting to the challenges in conducting experimental research in complex settings. Comroe<sup>21</sup> in his book, *Retrospectroscope*, described the origins of some great discoveries in medicine. He did not envision his imaginary instrument being used for retrospective database examinations; in such settings the instrument’s

lens can indeed be quite clouded. However, results generated from retrospective database analyses can be thought provoking, hypothesis generating, and help set the agenda for future investigations, as do the findings of Gupta *et al.*

## Competing Interests

Dr. Cook consults for the following entities that have an interest in red cell transfusion: Canadian Blood Services (Ottawa, Canada), U.S. Food and Drug Administration (Silver Spring, Maryland), TerumoBCT (Lakewood, Colorado), HbO2 Therapeutics (Souderton, Pennsylvania). Dr. Weiskopf consults for the following entities that have an interest in red cell transfusion: U.S. Department of Defense (Arlington, Virginia), U.S. Food and Drug Administration, TerumoBCT, HbO2 Therapeutics.

## Correspondence

Address correspondence to Dr. Weiskopf: rbw@theweiskopf-group.com

## References

- Gupta PB, DeMario VM, Amin RM, Gehrie EA, Goel R, Lee KHK, Yang WW, Khanuja HS, Sterling RS, Ness PM, Frank SM: Patient blood management program improves blood utilization and clinical outcomes in orthopedic surgery. *ANESTHESIOLOGY* 2018; 129:1082–91
- Hébert PC, Wells G, Tweeddale M, Martin C, Marshall J, Pham B, Blajchman M, Schweitzer I, Pagliarello G: Does transfusion practice affect mortality in critically ill patients? Transfusion Requirements in Critical Care (TRICC) Investigators and the Canadian Critical Care Trials Group. *Am J Respir Crit Care Med* 1997; 155:1618–23
- Hébert PC, Wells G, Blajchman MA, Marshall J, Martin C, Pagliarello G, Tweeddale M, Schweitzer I, Yetisir E: 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; 340:409–17
- Koch CG, Li L, Sessler DI, Figueroa P, Hoeltge GA, Mihaljevic T, Blackstone EH: Duration of red-cell storage and complications after cardiac surgery. *N Engl J Med* 2008; 358:1229–39
- Edgren G, Kamper-Jørgensen M, Eloranta S, Rostgaard K, Custer B, Ullum H, Murphy EL, Busch MP, Reilly M, Melbye M, Hjalgrim H, Nyrén O: Duration of red blood cell storage and survival of transfused patients (CME). *Transfusion* 2010; 50:1185–95
- Eikelboom JW, Cook RJ, Liu Y, Heddle NM: Duration of red cell storage before transfusion and in-hospital mortality. *Am Heart J* 2010; 159:737–43.e1
- Heddle NM, Eikelboom J, Liu Y, Barty R, Cook RJ: Exploratory studies on the age of transfused blood and in-hospital mortality in patients with cardiovascular diagnoses. *Transfusion* 2015; 55:364–72
- Steiner ME, Ness PM, Assmann SF, Triulzi DJ, Sloan SR, Delaney M, Granger S, Bennett-Guerrero E, Blajchman MA, Scavo V, Carson JL, Levy JH, Whitman G, D'Andrea P, Pulkrabek S, Ortel TL, Bornikova L, Raife T, Puca KE, Kaufman RM, Nuttall GA, Young PP, Youssef S, Engelman R, Greilich PE, Miles R, Josephson CD, Bracey A, Cooke R, McCullough J, Hunsaker R, Uhl L, McFarland JG, Park Y, Cushing MM, Klodell CT, Karanam R, Roberts PR, Dyke C, Hod EA, Stowell CP: Effects of red-cell storage duration on patients undergoing cardiac surgery. *N Engl J Med* 2015; 372:1419–29
- Lacroix J, Hébert PC, Fergusson DA, Tinmouth A, Cook DJ, Marshall JC, Clayton L, McIntyre L, Callum J, Turgeon AF, Blajchman MA, Walsh TS, Stanworth SJ, Campbell H, Capellier G, Tiberghien P, Bardiaux L, van de Watering L, van der Meer NJ, Sabri E, Vo D; ABLE Investigators; Canadian Critical Care Trials Group: Age of transfused blood in critically ill adults. *N Engl J Med* 2015; 372:1410–8
- Fergusson DA, Hébert P, Hogan DL, LeBel L, Rouvinez-Bouali N, Smyth JA, Sankaran K, Tinmouth A, Blajchman MA, Kovacs L, Lachance C, Lee S, Walker CR, Hutton B, Ducharme R, Balchin K, Ramsay T, Ford JC, Kakadekar A, Ramesh K, Shapiro S: Effect of fresh red blood cell transfusions on clinical outcomes in premature, very low-birth-weight infants: The ARIPI randomized trial. *JAMA* 2012; 308:1443–51
- Heddle NM, Cook RJ, Arnold DM, Liu Y, Barty R, Crowther MA, Devereaux PJ, Hirsh J, Warkentin TE, Webert KE, Roxby D, Sobieraj-Teague M, Kurz A, Sessler DI, Figueroa P, Ellis M, Eikelboom JW: Effect of short-term vs. long-term blood storage on mortality after transfusion. *N Engl J Med* 2016; 375:1937–45
- Dhabangi A, Ainomugisha B, Cserti-Gazdewich C, Ddungu H, Kyeeyune D, Musisi E, Opoka R, Stowell CP, Dzik WH: Effect of transfusion of red blood cells with longer vs shorter storage duration on elevated blood lactate levels in children with severe anemia: The TOTAL randomized clinical trial. *JAMA* 2015; 314:2514–23
- Kor DJ, Kashyap R, Weiskopf RB, Wilson GA, van Buskirk CM, Winters JL, Malinchoc M, Hubmayr RD, Gajic O: Fresh red blood cell transfusion and short-term pulmonary, immunologic, and coagulation status: A randomized clinical trial. *Am J Respir Crit Care Med* 2012; 185:842–50
- Weiskopf RB, Feiner J, Hopf H, Lieberman J, Finlay HE, Quah C, Kramer JH, Bostrom A, Toy P: Fresh blood and aged stored blood are equally efficacious in immediately reversing anemia-induced brain oxygenation deficits in humans. *ANESTHESIOLOGY* 2006; 104:911–20
- Austin PC, Laupacis A: A tutorial on methods to estimating clinically and policy-meaningful measures of treatment effects in prospective observational studies: A review. *Int J Biostat* 2011; 7:6
- Hernán MA: The C-word: Scientific euphemisms do not improve causal inference from observational data. *Am J Public Health* 2018; 108:616–9
- Heddle NM, Cook RJ: Composite outcomes in clinical trials: What are they and when should they be used? *Transfusion* 2011; 51:11–3
- Carson JL, Terrin ML, Noveck H, Sanders DW, Chaitman BR, Rhoads GG, Nemo G, Dragert K, Beaupre L, Hildebrand K, Macaulay W, Lewis C, Cook DR, Dobbin G, Zakriya KJ, Apple FS, Horney RA, Magaziner J; FOCUS Investigators: Liberal or restrictive transfusion in high-risk patients after hip surgery. *N Engl J Med* 2011; 365:2453–62
- Freemantle N, Calvert M, Wood J, Eastaugh J, Griffin C: Composite outcomes in randomized trials: greater precision but with greater uncertainty? *JAMA* 2003; 289:2554–9
- Lawrence VA, Silverstein JH, Cornell JE, Pederson T, Noveck H, Carson JL: Higher Hb level is associated with better early functional recovery after hip fracture repair. *Transfusion* 2003; 43:1717–22
- Comroe JH: *Retrospectroscope. Insights into Medical Discovery.* Menlo Park, CA, Von Gehr Press, 1977