

Reliability of Reliability Adjustment for Quality Improvement and Value-based Payment

Elliot Wakeam, M.D., M.P.H., Joseph A. Hyder, M.D., Ph.D.

RECENT work has demonstrated the importance of “reliability adjustment” when comparing surgical performance across hospitals.^{1,2} Reliability adjustment is a technique that most anesthesiologists have never heard of, but nearly all anesthesiologists are ranked using it because of its widespread adoption by nearly all programs that measure surgical outcomes including the Society of Thoracic Surgeons, the American College of Surgeons National Surgical Quality Improvement Program, the Scientific Registry of Transplant Recipients, and the Center for Medicare and Medicaid Services. Here, we describe the technique, why it is useful and why it merits scrutiny by anesthesiologists engaging in quality improvement and public reporting programs.

What Is Reliability Adjustment and Why Is It Needed?

Reliability adjustment is most commonly applied in conjunction with risk adjustment to improve the certainty of comparisons across hospitals—the goal being an “apples-to-apples” comparison of outcomes across hospitals. Although risk adjustment takes into account differences in patient disease severity and case mix, reliability adjustment takes into account repeatability of estimates related to the relative number of cases and outcomes that are being used to calculate the indicator of interest. This kind of adjustment, alternatively referred to as “empirical Bayes estimation” or “shrinkage adjustment,” intentionally “shifts” the observed-to-expected (O:E) ratio of a given hospital toward the average O:E ratio for all hospitals based that hospital’s number of patients and events. The new, shifted, O:E is referred to as the P:E, as in predicted-to-expected ratio. Hospitals that remain as outliers after this type of adjustment can thus be more “reliably” categorized as high or low performers.



“Reliability adjustment is an important statistical technique that will take on greater importance in policy and hospital payment.”

The result of “reliability adjustment” is improved “rankability” of hospitals including improved prediction of future performance.^{1,3} A simple example would be to compare mortality between two hospitals, each with a 50% mortality rate. One hospital contributes two cases (one death) and the other 1,000 cases (500 deaths)—clearly the latter hospital’s rate is a more stable, reliable estimate of that institution’s performance, whereas the first hospital’s rate would fluctuate drastically with the addition of even a single additional case. Thus, with reliability adjustment, hospitals identified as outliers are more reliably considered true performance outliers rather than simply hospitals that had a stretch of “bad luck” with a few patients. This technique not only provides more “reliable” estimates of performance but also permits small-volume hos-

pitals to be included in performance comparisons rather than excluded, as had been done previously. By improving certainty and adding more hospitals to performance benchmarking, reliability adjustment is widely considered an important advancement in the science of performance measurement.

Reliability Adjustment Merits Scrutiny

Despite its important advantages, reliability adjustment may have important unintended consequences. The first limitation of reliability adjustment is that its results may challenge intuition. Some anesthesiologists and surgeons may find it difficult to understand why their relative performance is “altered” by empirical Bayes estimation. Suppose the average mortality rate for all comers is 5%, and two hospitals, A and B, have different event rates during a year. Because case volume (and not just mortality rate) is a determinant of rank, hospital A with 3 of 33 (9%) deaths will be ranked as “better” than hospital B with 18 of 200 (9%) deaths. If, the following year, hospital A increases its

Image: A. Johnson.

Accepted for publication July 2, 2015. From the Department of Surgery, University of Toronto, Toronto, Ontario, Canada (E.W.); and Department of Anesthesiology, Division of Critical Care Medicine, and Center for the Science of Health Care Delivery, Mayo Clinic, Rochester, Minnesota (J.A.H.).

Copyright © 2015, the American Society of Anesthesiologists, Inc. Wolters Kluwer Health, Inc. All Rights Reserved. Anesthesiology 2016; 124:16-8

volume and has 18 of 200 (9%) deaths whereas hospital B loses volume and has 3 of 33 deaths (9%), then hospital A will be ranked as “worse” than hospital B. In real terms, these counterintuitive changes may be more reliable as estimates, particularly for identifying outliers, but surgeons and anesthesiologists may be confounded by their changes in rank with seemingly no change in relative or absolute performance. Fluctuating changes in rates in the absence of changes in perceived performance could frustrate and undermine local quality improvement efforts.

This shifting in position due to reliability adjustment is a statistical challenge without a single answer. One proposed alternative is “targeted shrinkage” where hospital performance is shrunk not to the global average but to a targeted average, such as the average performance of demographically similar hospitals according to procedure volume. As with risk adjustment, there is no definitive standard to determine whether one approach is superior to another. This may complicate decision-making about whether and how much reliability adjustment or targeted shrinkage is useful. No one is certain how much is correct, and this is an area of controversy.⁴

Shortcomings for Local Performance Measurement

Because reliability adjustment alters the performance of a hospital in relation to that of other hospitals, the published P:E value for a given hospital may confuse local quality improvement efforts. Published changes in P:E performance may be inconsistent with local performance, limiting their interpretability and risking misdirection of institutional priorities for value-improvement efforts. For instance, if an integrated care delivery system or Perioperative Surgical Home is attempting to establish which cardiac patients are best managed at a community hospital and which should be referred to a more expensive referral center, the two published P:E values, skewed by the case volume adjustments, may not provide the needed information for comparison among centers. Such “in-network” comparisons will become more important for the cost-control efforts that Accountable Care Organizations are designed to execute. Furthermore, integrated perioperative quality improvement and benchmarking of surgical outcomes are increasingly important to anesthesiologists as they engage with surgeons and with their own hospitals through the mechanisms such as the Perioperative Surgical Home. In these instances, it may be important *not* to include reliability adjustment in the calculations so they are more comparable and useful for local improvement efforts.

Why Does This Matter?

Reliability adjustments are intended to improve comparisons of hospital quality but do so by altering which hospitals are among performance outliers. In modern health care,

these redefinitions of high and low performers have financial consequences and may have implications for patients as they choose hospitals. Two key issues emerge. First, reliability adjustment, which accounts for the signal:noise ratio in performance, typically reduces the number of hospitals that can be meaningfully identified as performance outliers. In some instances, no hospitals are reliably identified as performance outliers.⁵ The second important issue is that the kinds of hospitals that are preferentially removed from outlier positions are most likely smaller hospitals with lower case volumes. These hospitals will thus be underrepresented among both high and low performers. If a hospital cannot attain outlier status by virtue of its size, it may not be motivated (either by positive or negative incentives) to improve, which may have significant impact on pay-for-performance schemes. Finally, the impact on safety net hospitals (hospitals serving the poor and uninsured) is unknown. Risk adjustment for these hospitals is controversial,^{6,7} and changes in ranking methodology need to be evaluated to ensure that care for the poor is not adversely affected.

Conclusion

Reliability adjustment is an important statistical technique that will take on greater importance in policy and hospital payment. Although internal quality improvement efforts may not benefit from this adjustment, public reporting of statistical outliers or payment-for-performance decisions may require it. Until the science and aims of reliability adjustment are clarified through research and policy evolution, anesthesiologists and surgeons must know whether their results were adjusted for “reliability” and to understand the limitations of this technique as they impact local decision-making for quality improvement, publicly reported quality rankings, and value-based payments.

Acknowledgments

Dr. Wakeam received salary support from the Ontario Ministry of Health and Long-Term Care (Concord, Ontario, Canada) and Dr. Hyder received institutional support from the Kern Center for the Science of Health Care Delivery (Rochester, Minnesota).

Competing Interests

The authors are not supported by, nor maintain any financial interest in, any commercial activity that may be associated with the topic of this article.

Correspondence

Address correspondence to Dr. Wakeam: elliott.wakeam@utoronto.ca

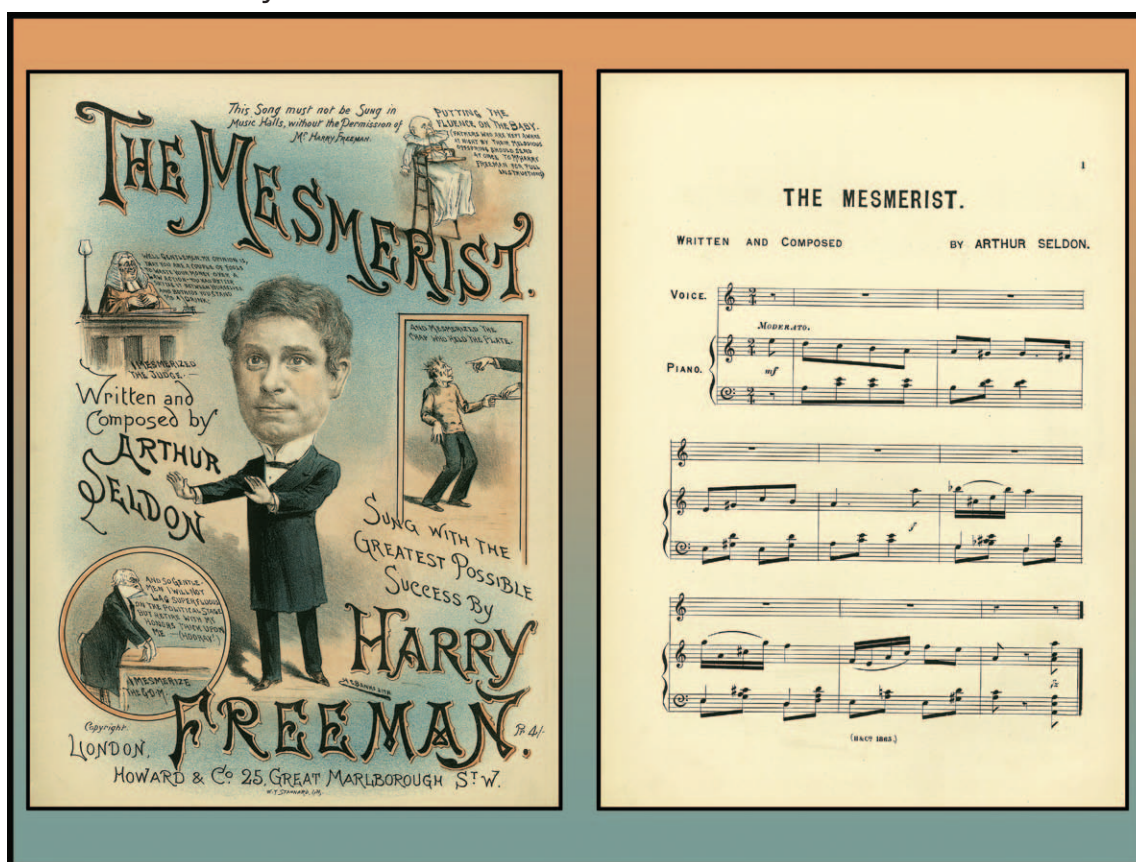
References

1. Dimick JB, Staiger DO, Birkmeyer JD: Ranking hospitals on surgical mortality: The importance of reliability adjustment. *Health Serv Res* 2010; 45(6 Pt 1):1614–29

- Dimick JB, Ghaferi AA, Osborne NH, Ko CY, Hall BL: Reliability adjustment for reporting hospital outcomes with surgery. *Ann Surg* 2012; 255:703–7
- Cohen ME, Ko CY, Bilimoria KY, Zhou L, Huffman K, Wang X, Liu Y, Kraemer K, Meng X, Merkow R, Chow W, Matel B, Richards K, Hart AJ, Dimick JB, Hall BL: Optimizing ACS NSQIP modeling for evaluation of surgical quality and risk: Patient risk adjustment, procedure mix adjustment, shrinkage adjustment, and surgical focus. *J Am Coll Surg* 2013; 217:336–46.e1
- Salkowski N, Snyder JJ, Zaun DA, Leighton T, Edwards EB, Israni AK, Kasiske BL: A scientific registry of transplant recipients Bayesian method for identifying underperforming transplant programs. *Am J Transplant* 2014; 14: 1310–7
- Lawson EH, Ko CY, Adams JL, Chow WB, Hall BL: Reliability of evaluating hospital quality by colorectal surgical site infection type. *Ann Surg* 2013; 258:994–1000
- Jha AK, Zaslavsky AM: Quality reporting that addresses disparities in health care. *JAMA* 2014; 312:225–6
- Fiscella K, Burstin HR, Nerenz DR: Quality measures and sociodemographic risk factors: To adjust or not to adjust. *JAMA* 2014; 312:2615–6

ANESTHESIOLOGY REFLECTIONS FROM THE WOOD LIBRARY-MUSEUM

The Mesmerist by Arthur Seldon



Whether called “animal magnetists,” hypnotists, or mesmerists, many entertainers introduced the public to the concept of preventing surgical pain well before Morton’s 1846 public demonstration of surgical etherization. And after that demonstration, some mesmerists offered hypnotic anesthesia as an alternative to the perceived perils of chemical anesthetics. “Sung with the greatest possible success” in the 1890s by London music hall performer Harry Freeman, “The Mesmerist” was a song “written and composed” by Arthur Seldon (right). This colorfully covered song sheet (left) was collected by the late Maurice Rickards, the founder of the Ephemera Society of Great Britain. Considered a throw-away item by some, such a piece of ephemera was regarded by Rickards as a “fragment of social history, a reflection of the spirit of its time.” In 2015, Drs. Ben Z. Swanson and George Bause donated this item to the Wood Library-Museum in memory of Swanson’s mentor, Maurice Rickards, whose not-so-ephemeral legacy included his masterwork “bible” for collectors of ephemera: *The Encyclopedia of Ephemera: A Guide to the Fragmentary Documents of Everyday Life for the Collector, Curator, and Historian*. (Copyright © the American Society of Anesthesiologists, Inc.)

George S. Bause, M.D., M.P.H., Honorary Curator, ASA’s Wood Library-Museum of Anesthesiology, Schaumburg, Illinois, and Clinical Associate Professor, Case Western Reserve University, Cleveland, Ohio. UJYC@aol.com.