Assessing Hospital Performance for Emerging Technologies
Essential but Elusive

Nihar R. Desai, MD, MPH; Karen E. Joynt, MD, MPH

We are in the midst of a technological revolution in cardiology. Over the past decade, new devices, tools, and techniques have proliferated, changing the face of cardiovascular care. The ability to relieve the hemodynamic burdens of a severely stenotic aortic valve without the need for cardiopulmonary bypass or a sternotomy with transcatheter aortic valve implantation has been heralded as an example of the best of modern medical science. Several large, randomized trials have established the efficacy and safety of transcatheter aortic valve replacement (TAVR) relative to medical therapy or surgical aortic valve replacement.\(^1\)\(^2\) This procedure has impacted hundreds of thousands of people with previously inoperable disease or with increased operative risk, offering both symptom relief and additional years of life. TAVR has spread rapidly, now widely available across the United States, and the platform is iteratively improving to become safer and better. Although the experience in the United States is in its relative infancy, experience in Europe foreshadows continued expansion of TAVR to intermediate and low-risk patients with several randomized trials reported\(^3\) or currently underway.\(^4\)

Concurrently, we are in the midst of an information revolution with growing demands for quality reporting, transparency, and performance assessment. The development and implementation of public reporting on Hospital Compare, value-based payment programs such as the Hospital Value Based Purchasing program, and alternative payment models such as accountable care organizations have grown out of Medicare and other payers’ desire to more closely align reimbursement with the quality of care delivered and the outcomes achieved.\(^5\) At the same time, healthcare consumers are increasingly willing and able to seek and interpret quality information, with Hospital Compare being joined by US News and World Report, Leapfrog, and Consumer Reports to deliver quality information to the public.

The worlds of technological advancement and information advancement collide in the article by O’Brien et al\(^6\) and highlight both the importance and limitations of creating robust methods by which to judge hospital performance on emerging technologies. Postmarket surveillance of the real-world utilization, efficacy, and safety of novel technologies is essential. Hospitals performing TAVR are subject to a mandate to collect data on all patients who receive a Food and Drug Administration–approved transcatheter valve as a condition for Medicare reimbursement. The Society of Thoracic Surgeons—American College of Cardiology Transcatheter Valve Therapy Registry, which was developed in 2011 at the time of the first approval of TAVR for commercial use, not only satisfies this regulatory requirement but also provides participating hospitals benchmarked reports that can be used to support performance improvement initiatives. Consequently, developing risk models appropriate for judging hospital performance on TAVR have become a priority. Typically, hierarchical modeling that includes both patient-level covariates and a hospital-specific intercept are used for such purposes; these models can adjust for differences in case-mix and account for the clustering of patients within a site, and in doing so, provide an estimate of the hospital’s quality that can then be used for comparative performance assessment.

The devil, however, is in the details. Both the opportunities and challenges of hierarchical modeling and hospital-level performance measurement are highlighted by O’Brien et al.\(^6\) To build a model suitable for assessing hospital performance, investigators included 22,248 TAVR procedures across 318 sites. The outcome of interest was in-hospital mortality, and the model included 40 covariates based on clinical relevance, data availability and quality, and variation across sites. Notably, frailty and patient-reported functional status, two powerful predictors of mortality after TAVR, were missing too frequently to be included.\(^7\)\(^8\) The mortality model demonstrated modest predictive capacity with a C statistic of 0.71 for the overall sample. The authors found >2-fold variation in the range of hospital risk-adjusted mortality rates (3.4%–7.7%) with an interquartile range of 4.8% to 5.4%. Still further, the authors report that a patient’s predicted odds of dying was 80% higher if they were treated by a hospital 1 SD above the mean when compared with a hospital 1 SD below the mean (odds ratio, 1.8; 95% confidence interval, 1.4–2.2). Such significant differences in outcomes, even after risk adjustment, illustrate the importance of such hospital comparisons, not only for performance improvement but also ultimately for public reporting, pay for performance, and consumer choice.

The opinions expressed in this article are not necessarily those of the editors or of the American Heart Association.

From the Section of Cardiovascular Medicine, Department of Internal Medicine, Yale University School of Medicine, New Haven, CT (N.R.D.); Center for Outcomes Research and Evaluation, New Haven, CT (N.R.D.); Division of Cardiovascular Medicine, Department of Medicine, Brigham and Women’s Hospital, Boston, MA (K.E.J.); and Department of Health Policy and Management, Harvard T.H. Chan School of Public Health, Boston, MA (K.E.J.).

Correspondence to Karen E. Joynt, MD, MPH, Brigham and Women’s Hospital, 75 Francis St, Boston, MA 02115. E-mail kjoynt@hsph.harvard.edu


Circ Cardiovasc Qual Outcomes is available at http://circoutcomes.ahajournals.org
DOI: 10.1161/CIRCOUTCOMES.116.003187
as well as for the prompt detection of any safety signals as this procedure continues to spread more broadly.

Or does it? O’Brien et al. also report that the model only identified 1 hospital, of the 316 examined, as having a mortality rate that were statistically significantly different from expected; certainly, this limits the model’s utility. If all hospitals are average, where are the opportunities for improvement? How can patients, providers, and payers make relevant decisions if the statistical methodology ascribes virtually all hospitals the same performance category?

This study illuminates a central challenge—although hospital quality assessment may be essential, in many ways it remains elusive, particularly for emerging technologies where sample sizes are small and adverse events are relatively rare. Hierarchical modeling approaches such as the one used by O’Brien et al. have many advantages and are used by The Centers for Medicare and Medicaid Services for public reporting on Hospital Compare for many conditions. Beyond the ability to adjust for differences in clinical characteristics of patients across hospitals and account for the clustered nature of the data within hospitals, this methodology can also account for differences in the number of observations across sites. Hospitals with a relatively small number of cases may have crude estimates of performance at the extremes (either worse or better performance). However, these may not represent the most accurate estimates of true performance. Although many statistical approaches simply exclude low-volume hospitals, this only creates a blind spot in the performance measurement system. Instead, hierarchical models can generate estimates of hospital variation in performance while also accounting for the uncertainty that comes with smaller volumes. Another important advantage from a public reporting standpoint is their conservatism—these models require a high degree of difference between observed and expected performance to achieve statistical significance and therefore are unlikely to accidentally penalize an average performer based on random variation in performance. This is perhaps wise for pay-for-performance applications, where financial stakes may be high.

On the other hand, because these models are highly sensitive to volume and tend to characterize low-volume centers as average performers, clinically meaningful variation in performance may be obscured. For the goal of supporting quality improvement efforts and fostering patient decision making, this represents a significant limitation, particularly relevant for emerging technologies where there is often a significant learning curve and where the lowest-volume centers may be poor performers. Indeed, previous literature supports a relationship between volume and outcomes for many invasive procedures; if such a relationship exists for TAVR, hierarchical models may not adequately stratify hospital performance by inappropriately crediting low-volume centers. This could hinder efforts at quality improvement and negatively impact transparency for consumer choice.

Hospital performance assessment is critical, but challenging. As payers’ and consumers’ desire for performance information continues to grow, and as new technologies continue to emerge, these issues will only become more salient. As ablation techniques for atrial fibrillation, left atrial appendage occlusion, peripheral arterial interventions, mitral valve clipping, and other procedures continue to expand in frequency and complexity, the unique needs of patients, providers, and payers for rapid and rigorous measurement of hospital performance need to be considered and addressed. Cardiology is, and ought to remain, at the leading edge, grappling with these difficult issues as we continue to find new ways to improve our patients’ health and outcomes.

Sources of Funding
Dr Desai was supported by grant K12 HS023000-01 from the Agency for Healthcare Research and Quality.

Disclosures
Dr Desai receives support from the Centers for Medicare and Medicaid Services to develop and maintain hospital performance measures that are used for public reporting and payment programs and research support from Johnson & Johnson, through Yale University, to develop methods of clinical trial data sharing. Dr Joynt currently serves as a senior advisor, US Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation.

References


Key Words: Editorials ☐ models, statistical ☐ outcome measures ☐ transcatheter aortic valve replacement
Assessing Hospital Performance for Emerging Technologies: Essential but Elusive
Nihar R. Desai and Karen E. Joynt

Circ Cardiovasc Qual Outcomes. 2016;9:498-500; originally published online September 13, 2016;
doi: 10.1161/CIRCOUTCOMES.116.003187
Circulation: Cardiovascular Quality and Outcomes is published by the American Heart Association, 7272
Greenville Avenue, Dallas, TX 75231
Copyright © 2016 American Heart Association, Inc. All rights reserved.
Print ISSN: 1941-7705. Online ISSN: 1941-7713

The online version of this article, along with updated information and services, is located on the
World Wide Web at:
http://circoutcomes.ahajournals.org/content/9/5/498

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published
in Circulation: Cardiovascular Quality and Outcomes can be obtained via RightsLink, a service of the
Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for
which permission is being requested is located, click Request Permissions in the middle column of the Web
page under Services. Further information about this process is available in the Permissions and Rights
Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation: Cardiovascular Quality and Outcomes is online
at:
http://circoutcomes.ahajournals.org//subscriptions/