



HHS Public Access

Author manuscript

Circulation. Author manuscript; available in PMC 2018 January 24.

Published in final edited form as:

Circulation. 2017 January 24; 135(4): 317–319. doi:10.1161/CIRCULATIONAHA.116.025653.

Challenges in Risk Adjustment for Hospital and Provider Outcomes Assessment

William S. Weintraub, MD and **Kirk N. Garratt, MD**

Christiana Care Health System, Newark, Delaware

How are we to evaluate outcomes for the care that we provide, and how are we to benchmark ourselves compared to colleagues at other institutions? It is critical that we do this to improve our own performance, and it is also being demanded of us by other stakeholders in society. Indeed, regular public reporting of outcomes is now effectively mandated of providers in many medical disciplines, and standards for statistical adjustment have been developed.¹ As a result, performance reports based on data from both administrative and clinical databases are commonplace, with a few elite databases serving as *de facto* sources of truth for medical performance in America. However, important questions must be asked about the registries that are being used to set performance standards and inform medical regulatory policies.

Within cardiology, one of the most commonly reported metrics is mortality after revascularization, either in-hospital or at some time thereafter, often 30 days. Such reporting may be at the level of the individual operator, the hospital, or the health care system. Other metrics, including non-fatal events and cost, are also reported commonly.² Initial public reports several decades ago offering raw outcome statistics on these endpoints resulted in a firestorm of criticism that the published data did not appropriately account for patient differences. In recent years, efforts have been made to account for variation in severity of illness, acuteness of presentation, and comorbidities using statistical methods that risk adjust in order to create fairer comparisons.¹ Concerns now focus on metrics such as risk adjusted mortality or the ratio of observed to expected (O/E) event rates.

Risk adjusted data are certainly more informative than raw data, but are adjusted data sufficiently reliable for the purpose of measuring quality of care? A major determinant of mortality in the setting of acute myocardial infarction (AMI) is the presence of shock. Providers, being aware of this, may respond appropriately, by ensuring that shock is carefully documented or less appropriately by avoiding care of these sicker patients where the expected adverse event rate is high. Evidence exists that such avoidance behavior has, in fact, occurred in parts of the country with public reporting, and may lead to deterioration, not improvement, in patient outcomes.³

Correspondence to: William S. Weintraub, MD, Suite 2E99, Christiana Care Health System, 4755 Ogletown-Stanton Road, Newark, DE 19718, wweintraub@christianacare.org, Phone: 302-733-1200, Fax: 302-733-4998.

Disclosures

None.

Other endpoints besides mortality have some degree of subjectivity. For example, the definition of myocardial infarction after a procedure is fraught with challenges, and troponin is not routinely measured in all hospitals after revascularization procedures, and when measured can be difficult to interpret. Covariates used for risk adjustment may be subjective too, even for common cardiovascular conditions. For example, a clinical diagnosis of heart failure captured by a data abstractor from an electronic health record may appear important in a registry but will not reliably identify patients with or without meaningful left ventricular dysfunction. Covariates may not be consistently collected, and data which are missing cannot be assumed to be missing at random. And as with mortality, important variables known to influence other outcomes are almost certain to go uncollected. In 2011, McNulty et al⁴ found that a series of important clinical factors, such as frailty measures, which are not routinely collected do influence both decisions about elective surgical versus catheter-based revascularization and outcomes of treatment in patients with left main coronary artery disease. Thus, problems with covariates limit the ability to risk adjust accurately, and may result in error throughout the spectrum of risk.

Clinicians are generally most concerned about the issues noted above, but the statistical issues are just as critical. If a hospital does 400 PCI procedures annually, and the expected mortality is about 1.5% (6 deaths), how can we evaluate if the hospital is truly an outlier if it reports, say, 9 deaths? This becomes much worse for the individual: if an operator performs 80 PCIs annually and has that same expected annual mortality rate of 1.5%, she is expected to lose 1.2 patients each year of practice. Does she become an outlier if she experiences 2 deaths (66% over target)? Three deaths (150% over target)? Is she truly providing superior care if she experiences no deaths? This problem is exacerbated by having a large number of hospitals, and thousands of operators, wherein some hospitals will be seen as being better or worse than their peers simply by the play of chance. This has been dealt with statistically by using Bayesian hierarchical modeling.⁵ This statistical approach will pull hospitals toward the average, i.e., an O/E of 1, and shrink the confidence interval. This reduces the issue of multiplicity, but this approach tends to make all providers look the same, and makes it quite difficult to identify a true outlier among low volume providers. This may be very important if low provider and hospital volumes really do contribute to increased risk. Conversely, larger providers can be penalized or rewarded in this scenario as their numbers permit them to stand out even after shrinkage of the confidence interval.

There is a problem of another sort with this type of risk adjustment. The idea is to find differences between providers. This is dependent on risk adjustment accounting for non-quality of care variables. However, for acutely ill patients, mortality and morbidity are largely predicted by patient variables rather than provider variables. Thus, the smaller influence of the provider must be seen through the thicket of patient variables, which can actually form a type of competing risk – that is, deaths due to patient variables vs provider variables. On the other hand, for non-acute patients, the mortality risk may be so small (a fraction of a percent) that finding evidence of increased risk confidently attributable to a provider or hospital may be impossible, and perhaps not a clinically meaningful measure.

The best way to make constructive use of benchmarked provider level outcome data from large national databases is for performance improvement. In the world of public reporting,

there is a much greater challenge. Professional societies have been well aware of the difficulties and have urged those who favor unfettered public reporting to proceed cautiously. Other stakeholders may not be responsive to the challenges, and the public may perceive calls to constrain public reporting as a sign that doctors and healthcare systems are hiding something. Indeed, there may be a perception that providers are using statistical arguments to avoid scrutiny. There is no perfect answer to this, and providers will have to accept that their outcomes will become increasingly available to the public. Professional societies must be responsive to the public and sanction, in some manner, practitioners or hospitals with convincing evidence of consistently poor outcomes, while also defending practitioners who find themselves inappropriately in the public eye. Furthermore, agencies advocating for full disclosure and transparency also have a responsibility to inform the public of the limitations of information we have to share with them, and of the dangers of misinterpreting numbers taken out of context. Importantly, formal collaboration between patient advocacy groups and medical societies is sorely needed to develop education plans that can help patients and their families understand the information contained in public performance reports, lest patients make poor judgements about physician quality or physicians conclude that taking care of the sickest patients is simply too risky.

Acknowledgments

Sources of Funding

Funded in part by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number U54-GM104941 (PI: Binder-Macleod).

References

1. Krumholz HM, Brindis RG, Brush JE, Cohen DJ, Epstein AJ, Furie K, Howard G, Peterson ED, Rathore SS, Smith SC Jr, Spertus JA, Wang Y, Normand SL, American Heart A, Quality of C, Outcomes Research Interdisciplinary Writing G, Council on E, Prevention, Stroke C and American College of Cardiology F. Standards for statistical models used for public reporting of health outcomes: an American Heart Association Scientific Statement from the Quality of Care and Outcomes Research Interdisciplinary Writing Group: cosponsored by the Council on Epidemiology and Prevention and the Stroke Council. Endorsed by the American College of Cardiology Foundation. *Circulation*. 2006; 113:456–62. [PubMed: 16365198]
2. Chassin MR, Loeb JM, Schmaltz SP, Wachter RM. Accountability measures—using measurement to promote quality improvement. *N Engl J Med*. 2010; 363:683–8. [PubMed: 20573915]
3. Joynt KE, Blumenthal DM, Orav EJ, Resnic FS, Jha AK. Association of public reporting for percutaneous coronary intervention with utilization and outcomes among Medicare beneficiaries with acute myocardial infarction. *JAMA*. 2012; 308:1460–8. [PubMed: 23047360]
4. McNulty EJ, Ng W, Spertus JA, Zaroff JG, Yeh RW, Ren XM, Lundstrom RJ. Surgical candidacy and selection biases in nonemergent left main stenting: implications for observational studies. *JACC Cardiovasc Interv*. 2011; 4:1020–7. [PubMed: 21939943]
5. Christiansen CL, Morris CN. Improving the statistical approach to health care provider profiling. *Ann Intern Med*. 1997; 127:764–8. [PubMed: 9382395]