Masthead Logo

DePaul Law Review

Volume 68
Issue 2 *Winter 2019*

Article 13

When Public Reporting Misleads the Public: The Case of Medicare's Hospital Compare Mortality Model

Jeffrey H. Silber

WHEN PUBLIC REPORTING MISLEADS THE PUBLIC: THE CASE OF MEDICARE'S HOSPITAL COMPARE MORTALITY MODEL

*Jeffrey H. Silber, M.D., Ph.D.**

INTRODUCTION

In 2007, the new Medicare mortality model was introduced with great fanfare. It was a web-based application that allowed the public to click on hospitals of their choosing and compare, among other things, the thirty-day mortality rate for patients admitted for various conditions, including "heart attack" or acute myocardial infarction. This was a special moment for Medicare, as a previous mortality model utilized for hospital comparisons was introduced over twenty years earlier, only to be discredited and retired from use over a decade later because of demonstrably biased estimates.[1] The new Medicare Hospital Compare mortality model was fundamentally different from previous public reporting models. It was thought to introduce advances from the field of Bayesian statistics to aid in making better predictions for patients and to solve a particularly difficult problem concerning how to report hospital mortality rates when hospitals are small and mortality rates are unstable.

Unfortunately, the Hospital Compare model made (and continues to make) assumptions that produce large underestimates of mortality rates in small hospitals and overestimates of mortality rates in large hospitals. Over the past decade the model has not been appreciably changed, despite a growing realization that problems exist with the model[2] and consequently, the advice given by the Medicare model to

* Professor, Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania. Professor, Department of Health Care Management, The Wharton School, University of Pennsylvania. Silber presented this topic at the Clifford Symposium held at DePaul University College of Law on April 19, 2018.

1. Jesse Green, *The Importance of Severity of Illness in Assessing Hospital Mortality*, 263 JAMA 241, 243 (1990).

2. Jeffrey H. Silber et al., *The Hospital Compare Mortality Model and the Volume–Outcome Relationship*, 45 HEALTH SERVS. RES. 1148, 1161 (2010) [hereinafter Silber et al., *The Hospital Compare Mortality Model*]; *see also* ARLENE S. ASH ET AL., STATISTICAL ISSUES IN ASSESSING HOSPITAL PERFORMANCE 1 (2012); Edward I. George et al., *Mortality Rate Estimation and Standardization for Public Reporting: Medicare's Hospital Compare*, 112 JAMA 933, 933 (2017); Dana B. Mukamel et al., *Measuring Quality for Public Reporting of Health Provider Quality:*

the public remains misleading. This Article will present a brief history of public reporting of hospital outcomes, place the present Medicare Hospital Compare model in context regarding this history, describe the flaw in the present Hospital Compare model, suggest reasonable ways to correct the problem, and examine why Medicare has not acted to change the model.

I. A Brief History of the Public Reporting of Hospital Outcomes

While many may say that the reporting of hospital mortality rates dates back to Florence Nightingale's examination of infection rates during the Crimean War,[3] or to Ernest Codman's examination of mortality rates at Massachusetts General Hospital in 1918,[4] I would contend that the use of formal statistical analysis to explore variation in hospital quality through examining hospital mortality rates really dates back to the time of Moses—that is Lincoln Moses and Frederick Mosteller's classic 1968 paper in the Journal of the American Medical Association titled *Institutional Differences in Postoperative Death Rates: Commentary on Some of the Findings of the National Halothane Study*.[5] Their paper was a secondary analysis of the National Halothane Study, performing formal statistical adjustments on the thirty-four participating hospitals in the study to determine whether there were excess deaths at some hospitals.[6] Before adjusting for patient characteristics, they found a twenty-four-fold difference between the highest and lowest hospital mortality rates.[7] After statistical adjustment they found a three-fold difference, but this difference was statistically significant.[8] Moses and Mosteller concluded that such differences may relate to quality, but feared a "circus" of analyses examining hospital quality through such adjustments.[9]

Making It Meaningful to Patients, 100 Am. J. Pub. Health 264, 266–67 (2010); Jeffrey H. Silber et al., *Improving Medicare's Hospital Compare Mortality Model*, 51 Health Servs. Res. 1229, 1230 (2016) [hereinafter Silber et al., *Improving Medicare's Hospital Compare Mortality Model*].

3. Lisa I. Iezzoni, *100 Apples Divided by 15 Red Herrings: A Cautionary Tale from the Mid-19th Century on Comparing Hospital Mortality Rates*, 124 Annals Internal Med. 1079, 1079 (1996).

4. Lisa I. Iezzoni, Risk Adjustment for Measuring Health Care Outcomes 11 (4th ed. 2013) [hereinafter Iezzoni, Risk Adjustment for Measuring Health Care Outcomes].

5. *See generally* Lincoln E. Moses & Fredrick Mosteller, *Institutional Differences in Postoperative Death Rates: Commentary on Some of the Findings of the National Halothane Study*, 203 JAMA 492, 492 (1968).

6. *Id.*

7. *Id.*

8. *Id.*

9. *Id.* at 494.

After the National Halothane Study was published, there were numerous studies examining hospital quality of care. Of course, since patients are not randomized to hospitals, these studies are observational in nature and hence, always must account for selection bias. Some hospitals treat sicker patients and therefore these variations in severity of illness must be accounted for in any analysis.

"Indirect standardization" is the typical method by which hospitals are compared.[10] Indirect standardization can be thought of as a comparison of the observed number of deaths (O) at a hospital to the expected number of deaths (E), and this is often referred to as the "O/E ratio."[11] If one hospital has more deaths than expected, its quality of care is considered worse than another hospital with a lower O/E. Hospitals want to have O/E ratios below 1, and it is even better if these hospitals can be shown to have a ratio below 1 with statistical confidence. When examining the many hospital report card systems currently in use, most use indirect standardization of one form or another and most report O/E for specific hospitals. One of the earliest and best-known hospital and physician report cards was the New York State Coronary Artery Bypass Surgery report, which began to be produced in 1986.[12] While these reports have gained in sophistication over the years, the report uses a risk adjustment model to estimate the probability of dying for each patient at a hospital undergoing coronary artery bypass grafting (CABG) surgery, and sums these probabilities to get an expected death rate.[13] It then provides the actual death rate and allows the reader to compare the O/E ratio for each hospital.[14]

In 1983, Medicare changed the way it paid hospitals, moving from a retrospective payment system (i.e., after a hospitalization, hospitals sent bills to Medicare and were paid some fraction of usual and customary charges), to a prospective payment system (i.e., hospitals were paid a pre-specified amount based on the diagnosis related group

10. Jeffery H. Silber et al., *Template Matching for Auditing Hospital Cost and Quality*, 49 HEALTH SERVS. RES. 1446, 1447 (2014).

11. *Id.*; *see also* JOSEPH L. FLEISS ET AL., STATISTICAL METHODS FOR RATES AND PROPORTIONS 631 (2003).

12. *See* Edward L. Hannan et al., *The New York State Cardiac Registries: History, Contributions, Limitations, and Lessons for Future Efforts to Assess and Publicly Report Healthcare Outcomes*, 59 J. AM. COLL. CARDIOLOGY 2309, 2309 (2012) [hereinafter Hannan et al., *The New York State Cardiac Registries*].

13. *Id.* at 2310–11; *see also* Edward L. Hannan et al., *The Decline in Coronary Artery Bypass Graft Surgery Mortality in New York State: The Role of Surgeon Volume*, 273 JAMA 209, 209 (1995).

14. Hannan et al., *The New York State Cardiac Registries*, *supra* note 12, at 2311.

(DRG) that the patient fell into).[15] This new payment system (which is still used today) raised concerns that hospitals would become incentivized to provide less-costly care, since hospitals' profits or losses were determined by whether they spent more or less on a patient compared to the DRG rate. In response to these concerns in the new payment environment, the Center for Medicare and Medicaid Services (CMS), previously known as the Health Care Financing Administration (HCFA), became very keen on examining hospital quality of care, so they developed the HCFA Mortality model. The model estimated O/E for every hospital in the country and was published in state-specific books.[16] The HCFA model results appeared on each page of the report and a letter from the Hospital's CEO appeared on the following page. When the O/E numbers were not good (i.e., above 1) the excuses from the CEOs generally stated that the patients at their hospital were sicker than what the HCFA model estimated. So, while reports correctly stated the O (just a count of deaths at that hospital), Hospital CEOs argued the E was too low (underestimating the sickness of their patients), resulting in an O/E ratio above 1.[17] The reports continued to be produced, but began to meet opposition from those concerned that the E was not being correctly estimated. Finally, in 1996, Jesse Green demonstrated convincingly that the HCFA model was profoundly biased because it "underestimate[d] the death rates for high-risk patients and overestimate[d] them for low-risk patients."[18] By 1997, the model stopped being provided to the public. There would not be a new mortality model for public consumption from Medicare until 2007.

II. The Small Numbers Problem

For many years, healthcare analysts have been concerned about public reporting when there are small numbers of patients at a hospital. The problem with small numbers is that reported outcome rates may be unstable. Mark Chassin and his co-authors published a paper in 1988 that clearly described the problem of small numbers when an-

15. *See generally* John K. Iglehart, *Medicare Begins Prospective Payment of Hospitals*, 308 New Eng. J. Med. 1428 (1983).

16. *See, e.g.*, U.S. Dep't of Health & Human Servs., Health Care Fin. Admin., Medicare Hospital Mortality Information: Illinois, Indiana, Michigan Region V (1987), https://babel.hathitrust.org/cgi/pt?id=mdp.39015012739465;view=1up;seq=3.

17. Iezzoni, Risk Adjustment for Measuring Health Care Outcomes, *supra* note 4, at 345–48.

18. Green, *supra* note 1, at 243.

alyzing hospital mortality in heart attack admissions.[19] The paper described the problem by plotting the mortality rate for each hospital on the vertical axis (y-axis) and hospital volume on the horizontal axis (x-axis).[20] Hospital mortality rates went from 0% to 100%, and hospital volume of heart attack patients went from 1 to 300 per year.[21] While the central tendency was approximately a 20% mortality rate, there was tremendous variability at low-volume hospitals and far less variation at high-volume hospitals. This funnel shape is well-known in statistics and reflects a binomial process. The authors were concerned that calling a low-volume hospital a high-mortality outlier may just be due to statistical variation, i.e., if a hospital only sees one heart attack patient and the patient dies, the hospital suddenly has a 100% death rate even though their average might approach 20% if they saw more patients. So these authors decided to rank hospitals not by the O/E ratio, but instead by the *p*-value associated with how different O was from E.[22] Since the *p*-values only indicate the magnitude of statistical significance, the ranking also had to take into consideration the direction of the difference—a very good or high rank of a hospital (associated with a statistically significant O below E), and a poor or low rank for a hospital (associated with a statistically significant O above E)—with all rankings performed on these constructed *p*-values. While it was admirable for Hospital Compare's developers to be concerned about these small-volume hospitals and their unstable mortality rates, the problem with ranking by *p*-value is that small hospitals will always look average, since small numbers will prevent any difference of O and E from being statistically significant. Hence, small hospitals look average, and only big hospitals have the potential to look better or worse than average. Not surprisingly, ranking by *p*-value did not catch on for hospital public reporting.

III. The Hospital Compare Mortality Model

In response to Medicare's need for monitoring hospital quality, a new model was developed and introduced in 2007. HCFA had changed its name to the Center for Medicare and Medicaid Services (CMS) by 2007, and it created a CMS website where patients could examine individual hospitals for their adjusted mortality, among other

19. Mark R. Chassin, *Differences Among Hospitals in Medicare Patient Mortality*, 24 Health Servs. Res. 1, 29 (1989). Mark Chassin became the Public Health Commissioner of New York and helped introduce the New York State CABG report cards.

20. *Id.* at 7.

21. *Id.*

22. *Id.* at 11–13.

items. The new model was developed by Krumholz and his co-authors[23] based on what is called a Hierarchical Random Effects Logistic Regression model.[24] The model was purported to be a good estimator of mortality, so that it could be used for the denominator in indirect standardization (O/E), and the model also purported to address the "small numbers" problem in the numerator by "borrowing strength"[25] from the overall mortality rate observed across the entire distribution of hospitals. The entire distribution of hospitals refers to the bell-shaped curve of all the hospital-specific heart attack mortality rates at all hospitals that see heart attack patients in the United States. Instead of having CMS report O/E rates to the public, the new CMS approach was to report Predicted to Expected rates (P/E), substituting O (i.e., a simple count of a hospital's deaths) for P (i.e., a prediction of these deaths which is stabilized for the small numbers problem).[26] Small hospital death rates were considered unstable, so these reported numerators were "shrunken" or moved in the direction of the mean mortality value over all hospitals. Here the term "shrunken" refers to a revised mortality estimate that is closer to the national mortality estimate for heart attack patients. Hospital observed death rates would be shrunken (moved toward the mean death rates across all hospitals) based on how unstable the model believed an individual hospital death rate was. The P in the P/E that is reported to the public for any one hospital can be approximately described as a weighted mixture of the observed hospital death rate and the national death rate. For very small hospitals, the P provided to the public is often, almost exactly just the national death rate because the model down-weights the O derived from the small hospital with small numbers of deaths. For large hospitals, P also generally reflects a number somewhere between the national rate and the large hospital's O. The public is no longer provided with the actual mortality rates because they are considered too unstable to report.

One way to consider how the shrinkage works in a hierarchical random effects model like Hospital Compare is to consider the estimate

23. Harlan M. Krumholz et al., *An Administrative Claims Model Suitable for Profiling Hospital Performance Based on 30-Day Mortality Rates Among Patients with an Acute Myocardial Infarction*, 113 Circulation 1683, 1684 (2006) [hereinafter Krumholz et al., *An Administrative Claims Model*]; Harlan M. Krumholz et al., *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*, 113 Circulation 456, 460 (2006).

24. Silber et al., *Improving Medicare's Hospital Compare Mortality Model*, *supra* note 2, at 1230.

25. Ash et al., *supra* note 2, at 15–16.

26. Krumholz et al., *An Administrative Claims Model*, *supra* note 23, at 1648.

of P as a linear combination of O and the overall estimated national mean mortality rate for the actual patients seen at the hospital of interest. Reference rate E is the expected rate of death for the patients at the hospital of interest had they been treated at the typical hospital.[27] The linear combination can be written as $P = \lambda O + (1 - \lambda) E$. One can solve for the weight (λ) on any given hospital's observed mortality (O) rate using the raw data used in the Hospital Compare model and obtaining the hospital's P and E. After doing this computation, we found that for small hospitals in the lowest twentieth percentile, λ is near 0.05, suggesting that P is almost entirely reporting the national mortality rate (which was weighted by 0.95). Even for very large hospitals, the Hospital Compare model assigns a λ weight generally below 0.5. Therefore, the public is being provided stabilized mortality rates that are made up anywhere from 50% to 95% from the national mortality rate. It is not hard to understand, therefore, that small hospitals look average when 95% of the estimate for small hospitals is simply the national average. Not surprisingly, it follows that Hospital Compare does not identify many outlier hospitals because the problematic small hospitals are made to look average according to the model.

It is important to realize that the Hospital Compare model utilized to produce their P/E estimate predicts mortality only though patient characteristics and the certainty with which the prediction is made. But there are many hospital features that are predictive of differential outcomes at different hospitals. So, hospital size, technology, nurse-to-bed ratio, teaching status, or even the ability of the hospital to perform percutaneous cardiac interventions or CABG surgery does not go into the model predicting the hospital-specific heart attack mortality rate. By not placing any hospital characteristics in the Hospital Compare predictive model, the model shrinks small hospital mortality rates to the overall typical hospital.[28] This is precisely because it takes no hospital characteristics into consideration when addressing the small numbers problem, and it borrows strength only from the overall distribution of all hospital mortality rates. Hospital Compare assumes that small hospitals, like the typical hospital, would have a central tendency toward 20% mortality if only there were enough data to show this. This is the crux of the problem with Hospital Compare's approach. In reality, small hospitals likely have a different central tendency from the typical hospital precisely because they see fewer

27. Silber et al., *The Hospital Compare Mortality Model*, *supra* note 2, at 1159.

28. *See generally* Mukamel et al., *supra* note 2; Silber et al., *The Hospital Compare Mortality Model*, *supra* note 2.

patients and have less experience treating them. Small hospitals have low patient volume, and because there are worse outcomes associated with a low volume (assuming a volume-outcomes relationship), small hospitals tend to have higher death rates than the typical hospital. Because Hospital Compare refuses to place hospital characteristics like hospital volume or available technology in their model to estimate P (in order to stabilize O), poor-performing small hospitals will be reported to the public as having average mortality rates when they may be performing far worse than average.

Such a hierarchical random effects model is now the "gold standard" for many reporting systems. This Hospital Compare hierarchical random effects mortality model is used to rate hospitals by U.S. News Consumer Reports[29] and the Society for Thoracic Surgeons.[30] As the most widely-used hospital public reporting system in the country, one may ask if the reports are biased. As I will develop in this paper, the answer is, unfortunately, yes.

IV. Understanding the Problems in the Hospital Compare Mortality Model

In a series of studies at the University of Pennsylvania, my colleagues and I have been exploring if Hospital Compare may be providing misleading information to the public. In 2010, we published a paper titled *The Hospital Compare Mortality Model and the Volume-Outcome Relationship*.[31] We argued that, based on a vast literature,[32] there is a well-known volume-outcome relationship for most aspects of medicine, and that the more a provider does, the better the outcomes.[33] Accordingly, it would make sense that small hospitals have lower volume and may perform worse than their high-volume peers.

29. *See generally How We Rate Hospitals*, ConsumerReports (June 2018), http://article.images.consumerreports.org/prod/content/dam/cro/news_articles/health/PDFs/Hospital_Ratings_Technical_Report.pdf.

30. *See generally* Sean M. O'Brien et al., *An Empirically Based Tool for Analyzing Mortality Associated with Congenital Heart Surgery*, 138 J. Thoracic & Cardiovascular Surgery 1139 (2009); *STS CHSD Morality Risk Model*, STS Pub. Reporting Online, https://publicreporting.sts.org/chsd-risk-model (last visited Nov. 16, 2018).

31. Silber et al., *The Hospital Compare Mortality Model*, *supra* note 2, at 1162.

32. Afschin Gandjour et al., *Threshold Volumes Associated with Higher Survival in Health Care: A Systematic Review*, 41 Med. Care 1129, 1136 (2003); Ethan A. Halm et al., *Is Volume Related to Outcome in Health Care? A Systematic Review and Methodologic Critique of the Literature*, 137 Annals Internal Med. 511, 517 (2002); Harold S. Luft et al., *The Volume-Outcome Relationship: Practice-Makes-Perfect or Selective-Referral Patterns?*, 22 Health Servs. Res. 157, 158 (1987); David M. Shahian & Sharon-Lise T. Normand, *The Volume-Outcome Relationship: From Luft to Leapfrog*, 75 Annals Thoracic Surgery 1048, 1048–49 (2003).

33. Silber et al., *The Hospital Compare Mortality Model*, *supra* note 2, at 1149.

Because the Hospital Compare model shrinks small hospital death rates to the mean of all hospitals (precisely because the hospitals are small), we hypothesized that small hospitals would be reported as average in the Hospital Compare model. We also hypothesized that if each small hospital prediction was grouped over many small hospitals, those small hospitals would look average when using the Hospital Compare model. However, if we grouped all small hospital patients together (so that there were no small numbers problems), we suspected that because each hospital had small volume, the volume-outcome relationship should dictate that as a group, these hospitals' patients should display elevated mortality.[34] Indeed, that is what we reported. Hospital Compare was informing the public that thousands of small hospitals were all average in performance on the mortality model, suggesting that there was no difference across hospitals groups. However, we found that as a group, the relative risk of death was 50% higher at the smallest 20% of hospitals than the largest 20% of hospitals.[35] For example, consider a patient attempting to decide whether to stay affiliated with a small, local, suburban hospital or make the move to be affiliated with a larger, urban hospital. The mortality rates in the Hospital Compare model may reassure the suburban patient that the local hospital is at least average when compared to the data of the two options. However, had the patient considered a group of hospitals that look like the local, small hospital, the patient would have seen that the mortality rates were far higher at the small hospitals as a group. Thus, unbeknownst to the patient, the larger, urban hospital may be the superior option because it is likely to have a lower mortality rate. In subsequent work, we reconstructed the Hospital Compare model in a fully Bayesian framework to allow us to form an even better understanding of the strengths and weaknesses of the Hospital Compare model.[36]

V. Correcting the Assumptions of the Hospital Compare Model

In an attempt to better understand why the Hospital Compare model is providing such biased results, we have recently re-done the Hospital Compare model in a Bayesian framework to examine the cause and extent of the Medicare Hospital Compare model's inade-

34. *Id.* at 1158.

35. *Id.* at 1159–61.

36. *See generally* Silber et al., *Improving Medicare's Hospital Compare Mortality Model*, *supra* note 2.

quacies.[37] The underlying assumption in the Hospital Compare model is that all hospitals are identical and part of the same national hospital distribution, with a constant hospital heart attack mortality rate mean and constant variance. Note, the mean and variance define the shape of the hospital mortality rate distribution, assumed by the Hospital Compare model to be normally distributed (again, describing a bell-shaped distribution).[38] In a series of models, we gradually relaxed this assumption, suggesting instead that hospitals with specific characteristics may indeed have different mean mortality rates and different variances. Using a Bayesian framework, the data determined whether a model for different types of hospitals was best fit with an overall mean for all hospitals, or a separate mean for hospitals of different characteristics. We did this using split samples, so model development was tested for predictive accuracy on an outside sample to avoid a tautological improvement. We found dramatic improvements in the models when we "liberated" the mean and the variance, challenging the assumptions of the Hospital Compare model. We concluded that the Hospital Compare model's mistake was in shrinking all hospitals to the same mean and variance (or assuming all hospitals came from the same national distribution). We noted that, for example, shrinking the mortality rate estimates of small-volume hospitals to resemble the mean of the small-volume hospital death rates produced a far superior estimate of P/E—one that is less biased—than shrinking small-volume hospital mortality estimates to the mean of all hospital heart attack death rates for the entire country. This was also true for other hospital characteristics. Such accounting for hospital characteristics in the P of P/E is something that has not been done by the Hospital Compare model. Note, neither Hospital Compare nor our group would suggest including hospital characteristics in predicting E, as E (the standard against which hospital performance is judged) should only be based on patient characteristics.

The size of the bias in the present form of Hospital Compare is really quite remarkable. In a matched analysis recently published in the Journal of the American Statistical Association, we reported that hospitals in the lowest twentieth percentile by volume have an observed mortality rate of 28% for thirty-day mortality after admission, and hospitals in the highest twentieth percentile by volume, treating matched patients similar to the small-volume admissions, had an ob-

37. *See generally* George et al., *supra* note 2; Silber et al., *Improving Medicare's Hospital Compare Mortality Model*, *supra* note 2.

38. George et al., *supra* note 2, at 935; Silber et al., *Improving Medicare's Hospital Compare Mortality Model*, *supra* note 2, at 1232.

served mortality rate of only 20%.[39] It would appear that larger hospitals had far lower mortality rates for similar patients seen at the low-volume hospitals. However, using Hospital Compare's methods, hospitals in the lowest twentieth percentile by volume would have a publicly-reported mortality rate of 23%—5% lower than the 28% observed. Had the Hospital Compare model included the hospital characteristic of volume in the prediction, something we have suggested Hospital Compare could do, the model would have estimated the mortality rate at these smaller hospitals to be about 28% or 29% (depending on the change in the model, much closer to the observed 28%).

To place that 5% underestimate of mortality at the smallest hospitals in context, such an error has very important consequences. If we defined outlier hospitals using the Hospital Compare model, we would report that out of 4,396 hospitals, only 30 hospitals were high-mortality outliers. However, if we improve the model to allow hospital characteristics to influence the prediction, we find that 1,028 hospitals would be considered as mortality outliers.[40] In yet another analysis, using Chicago as an example, we found that the Hospital Compare model portrays most hospitals as having very similar mortality rates. However, when allowing the model to shrink to a mean that is a function of hospital volume, differences in hospital mortality rates for patients admitted for heart attacks clearly emerged.[41]

To fix the Hospital Compare model, we have argued that hospital characteristics need to be placed in the predictive model.[42] One obvious characteristic is hospital volume, but many other variables need to be examined—such as available technology, nurse staffing, teaching hospital status, to name a few. By not including these variables in their model, Hospital Compare is assuming that shrinking mortality estimates to the mean of all hospitals produces better and unbiased estimates than shrinking to a target that accounts for potential hospital differences. The assumption to exclude hospital characteristics from the predictive model is clearly not consistent with the data.

39. George et al., *supra* note 2, at 941.

40. *Id.* at 945.

41. Silber et al., *Improving Medicare's Hospital Compare Mortality Model*, *supra* note 2, at 1236–37.

42. *See generally* George et al., *supra* note 2; Silber et al., *The Hospital Compare Mortality Model*, *supra* note 2; Silber et al., *Improving Medicare's Hospital Compare Mortality Model*, *supra* note 2.

VI. Why Has Medicare Not Yet Fixed the Hospital Compare Model?

Having repeatedly been shown to provide biased underestimates of mortality for small hospitals, and biased overestimates of mortality for larger hospitals, the question arises: Why has Medicare not yet fixed the model? Medicare could place hospital characteristics in their model to better estimate P, and therefore stabilize small-volume hospitals to the mortality rates of small-volume hospitals, as we have suggested above. The same would be true for any hospital characteristic that influences mortality rates. For patients using the Hospital Compare model, such a fix would mean that when inquiring about small hospitals, Hospital Compare could say something like "while this specific hospital is too small to provide a stable estimate, hospitals with very similar characteristics to this hospital performed poorly, with a P/E ratio of 1.3." Such a statement may not help the individual hospital know if it was performing better or worse than other, similar hospitals, but patients would be given far more useful information than simply a statement that suggests their small hospital performs exactly like the typical hospital (with a P/E near 1), a statement that is biased.

Medicare may object to this suggested fix, stating that such a policy of shrinking unstable estimates to peer group hospitals' rates will possibly punish those few, extraordinary, small hospitals with excellent outcomes. Our response would be that on average, the estimate that shrinks mortality rates to peer hospitals' rates is, by definition, demonstrably less biased than ignoring hospital characteristics. If hospital characteristics are ignored, then small hospitals, no matter how good or bad, will be reported as average. From the patient's perspective and across all hospitals, it is clear that by adding hospital characteristics to estimate P, the model will result in greater accuracy and the reduction of bias. From the perspective of many hospitals who now are reporting their outcomes as average, the fix to Hospital Compare will produce a potentially unflattering, though less biased, report to the public. At the very least, adding in hospital characteristics may produce more scrutiny of small-volume hospitals, or more scrutiny for any hospitals that have their results stabilized.

Deciding to continue to utilize a model that has been shown to produce misleading and biased information for a segment of hospitals seems to be an unwise choice. How to change the Hospital Compare mortality model remains a policy question. If patients understood that the mortality rates provided by Hospital Compare, and many other report card systems, do not reflect the actual mortality rates at their hospital, they may feel that change is needed. It is time that Medicare fix the Hospital Compare mortality model and improve the accuracy of public reporting.

Conclusion

It is hard to understand why Medicare has not fixed the Hospital Compare mortality model. One interesting question that may help explain this inaction is: Under what statutory authority are the rankings promulgated? It is possible that there is legal precedent for forcing Medicare to improve the model. On that same note, one could ask: Is there legal liability attached to misleading the public in the rankings, or could a large university hospital claim to be disparaged under any existing law? Clearly there are winners and losers associated with the decision by Medicare to not allow hospital characteristics in the predictive model. The result of this decision is that small hospitals tend to look average, when as a group they perform worse than average. Correspondingly, large hospitals that perform exceedingly well will have their outcomes shrunken (in the direction of reporting higher death rates) towards the national mean, making them look less outstanding. All this leads to a more confused and misinformed public when attempting to better understand the quality of care at their hospitals.