

## VIEWPOINT

## Are Patient Satisfaction Instruments Harming Both Patients and Physicians?

**Barak D. Richman, JD, PhD**

Duke University School of Law, Durham, North Carolina; and Clinical Excellence Research Center, School of Medicine, Stanford University, Palo Alto, California.

**Kevin A. Schulman, MD**

Clinical Excellence Research Center, School of Medicine, Graduate School of Business, Stanford University, Palo Alto, California.

**Patient satisfaction surveys** have become a mainstay of the health care experience. Patients have become accustomed to responding to surveys, physicians attentively monitor their aggregate satisfaction scores, and public payers reward physicians and institutions with strong results. This reliance comes from good intentions. Patient satisfaction is an integral element of care, and scholars have argued that positive patient experience represents an important quality dimension not captured in other metrics.<sup>1</sup>

However, there is an important difference between patient experience as a component of quality and the measurement tools used to assess patient experience in clinical care. Measures can lose value as they age, and just like the Google search algorithm, patient satisfaction measurement strategies need to be updated to remain useful. For patient experience measures, the issue is whether the measures being used in clinical care accurately reflect clinical performance and support efforts to improve the patient experience. We suggest that prevailing patient satisfaction surveys fail these critical questions and are not “fit for purpose”<sup>2</sup> in evaluating physician performance.

### Patient Satisfaction as a Value and as a Measure

Patient satisfaction is an essential component of quality. Studies have found that higher patient satisfaction scores are correlated with desirable health outcomes, such as lower rates of 30-day readmission<sup>3</sup> and mortality.<sup>4</sup> These studies evaluate ratings of patient experience across large populations, not surveys used as a measure of individual physician performance.

However, more recent studies have found that efforts to increase mean patient satisfaction scores are either ineffectual or counterproductive when evaluated as a measure of physician performance.<sup>5</sup> One famous study, cleverly named “The Cost of Satisfaction,” found that the patients giving the highest satisfaction ratings exhibited higher costs of care and had higher mortality than those offering lower satisfaction ratings.<sup>6</sup>

Most instruments were developed more than 20 years ago and were based on survey design constructs that originated in a world of paper-based, mailed surveys. These tools were optimized to evaluate patient experience as a broad component of care quality. But when Medicare and private plans started using these same scales to measure and evaluate individual physicians and reward those with superior scores, the underlying limitations of these instruments became more apparent. Although they may measure some elements of the patient experience, they are not responsive to the performance of individual physicians or care teams. These instruments have become unable to reflect or inform meaningful differences in care, and physicians struggle to respond to this method of evaluation.

As an illustration, we gathered raw data from the Clinician and Group Consumer Assessment of Healthcare Provider and Systems (CG-CAHPS) survey for 2015 to 2019. This instrument helps to determine bonus payments under Medicare's Merit-based Incentive Payment System (MIPS) program. Patients asked to rate their outpatient physician overwhelmingly respond with the highest possible rating, with nearly 60% of all respondents giving their physician a perfect 10 out of 10, with a mean rating of 9.16 (SD, 1.5) (Figure). Given such skewed results, small perturbations in scores can result in significant changes in the ranking of physicians within a practice or nationally, causing percentile rankings effectively to become random. Nonetheless, MIPS bases some bonus payments on marginal relative differences in CG-CAHPS scores. If the CG-CAHPS scores approximately represent the distribution of scores across clinicians, a ranking system could penalize physicians who earn a perfect score from a majority of their patients but nonetheless find themselves in the bottom 50% of providers.<sup>7</sup> The Medicare Access and CHIP Reauthorization Act of 2015 states that “...such performance threshold for a year shall be the mean or median (as selected by the Secretary) of the composite performance scores for all MIPS-eligible professionals with respect to a prior period specified by the Secretary” (§1395w-4[q][6]).

It is precisely because results are so skewed that, paradoxically, patient survey responses have such an outsized effect on behavior. Imagine a clinician who treats 120 patients each month and enjoys a mean patient rating of 9.5, well above the national mean. With a survey response rate of 20%, it takes only 1 disgruntled patient giving a score of zero to cause the clinician's mean score to plummet below 9.1. Perhaps it is this fear of the disgruntled patient that explains the systemic costs of patient satisfaction surveys. Patients who received discretionary services such as advanced imaging for acute back pain<sup>8</sup> and incurred higher costs of care<sup>5</sup> have been shown to provide higher mean satisfaction scores.

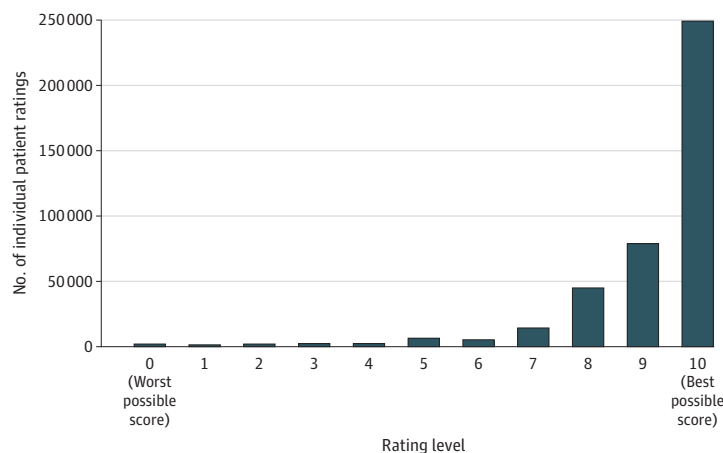
### Implications From Management Science

Forcing physicians to compete for bonus payments under a topped-out measure such as CG-CAHPS induces stress and burnout.<sup>9</sup> This is not a criticism of valuing patient satisfaction as a priority nor a criticism of using survey methods to assess patient satisfaction; it is a statement about the failure of a specific measurement system.

Patient satisfaction scores may have been well constructed psychometrically, but no static measure can retain utility for long in a setting such as a ranked pay-for-performance system. Professionals or their consultants learn from or engineer efforts to boost measured performance. This sets off a competition for the best-resourced individuals to drive metric-based

### Corresponding

**Author:** Kevin A. Schulman, MD, Clinical Excellence Research Center, Stanford University, CAM Bldg, 453 Quarry Rd, Ste 117B, Palo Alto, CA 94304 (kevin.schulman@stanford.edu).

**Figure. CG-CAHPS 2015-2019 Provider Ratings**

CG-CAHPS indicates Clinician and Group Consumer Assessment of Healthcare Provider and Systems. The survey states: "Using any number from 0 to 10, where 0 is the worst provider possible and 10 is the best provider possible, what number would you use to rate this provider?" Providers in this survey refer to clinicians, particularly physicians. Data from the Agency for Healthcare Research and Quality through a data use agreement with Duke University.

performance, often with exacting focus on factors wholly insignificant to physiologic health but potentially useful in climbing up the percentile ladder. In other words, scores can be improved with ever-greater investment of resources even if no meaningful clinical benefit is achieved. This counterintuitive result comes from a failure to retire metric-based assessments when they are no longer able to provide meaningful information for clinical care.

The management literature is familiar with the downside of high-stakes measurement systems when the underlying metrics offer no insights into collective performance improvement. One of the leading scholars in the field of performance management, W. Edwards Deming, warned that

[t]he merit rating nourishes short-term performance, annihilates long-term planning, builds fear, demolishes teamwork, [and] nourishes rivalry and politics. It leaves people bitter, crushed, bruised, battered, desolate, despondent, dejected, feeling inferior, some even depressed, unfit for work for weeks after receipt of rating, unable to comprehend why they are inferior. It is unfair, as it ascribes to the people in a group differences that may be caused totally by the system that they work in.<sup>10</sup>

This critique is as apt for the clinical setting as it is for the factory floor, and yet momentum for value-based payment and pay-for-performance schemes continues to proliferate across the health sector. Concurrent with the increasing use of these approaches, ever-higher rates of physician burnout and distress are occurring, exactly as anticipated by Deming.

### Conclusions

Patient satisfaction remains an important component in the quality of care, and physicians widely agree that they should make a concerted effort to improve communication and care coordination. But current high-stakes use of static patient experience instruments renders them at best meaningless and at worst responsible for physician burnout, bad medical care, and the defrauding of health insurers by driving up use. Public and private payers should seriously reconsider relying on these surveys to influence quality assessment and payments. At a minimum, the Medicare Payment Advisory Commission should annually evaluate the performance of quality measures such as CG-CAHPS to ensure that they remain fit for purpose. Without reform, current survey practices might continue inflicting burnout on physicians and bad medical care on patients.

### ARTICLE INFORMATION

**Published Online:** November 17, 2022.

doi:10.1001/jama.2022.21677

**Conflict of Interest Disclosures:** Dr Schulman reported holding a patent (9305059; survey methods) assigned to the University of North Carolina (UNC), first filed in 2012, and assigned to UNC in 2016. No other disclosures were reported.

**Additional Contributions:** We thank Don Hopkins, PhD, data scientist for Empirical Research and Data Support Services (ERDS) at the Goodson Law Library, Duke Law School, for data collection and analysis, and Brian Smart, MD, Duly Health and Care, Lombard, Illinois, for insights into current patient survey uses. Neither received compensation for their input.

### REFERENCES

1. Manary MP, Boulding W, Stealing R, Glickman SW. The patient experience and health outcomes. *N Engl J Med*. 2013;368(3):201-203. doi:10.1056/NEJMp1211775

2. Edwards MC, Slagle A, Rubright JD, Wirth RJ. Fit for purpose and modern validity theory in clinical outcomes assessment. *Qual Life Res*. 2018;27(7):1711-1720. doi:10.1007/s11136-017-1644-z

3. Boulding W, Glickman SW, Manary MP, Schulman KA, Staelin R. Relationship between patient satisfaction with inpatient care and hospital readmission within 30 days. *Am J Manag Care*. 2011;17(1):41-48.

4. Jaipaul CK, Rosenthal GE. Do hospitals with lower mortality have higher patient satisfaction? *Am J Med Qual*. 2003;18(2):59-65. doi:10.1177/106286060301800203

5. Kohring JM, Pelt CE, Anderson MB, Peters CL, Gilliland JM. Press Ganey outpatient medical practice survey scores do not correlate with patient-reported outcomes after primary joint arthroplasty. *J Arthroplasty*. 2018;33(8):2417-2422. doi:10.1016/j.arth.2018.03.044

6. Fenton JJ, Jerant AF, Bertakis KD, Franks P. The cost of satisfaction: a national study of patient satisfaction, health care utilization, expenditures, and mortality. *Arch Intern Med*. 2012;172(5):405-411. doi:10.1001/archinternmed.2011.1662

7. Medicare Access and CHIP Reauthorization Act of 2015, 42 USC §1395w-4(q)(6).

8. Pham HH, Landon BE, Reschovsky JD, Wu B, Schrag D. Rapidity and modality of imaging for acute low back pain in elderly patients. *Arch Intern Med*. 2009;169(10):972-981. doi:10.1001/archinternmed.2009.78

9. Brilla RHC, Mostaghimi L. Patient satisfaction surveys and burnout. *Neurology*. May 12, 2017. Accessed November 14, 2022. <https://n.neurology.org/content/patient-satisfaction-surveys-and-burnout>

10. Deming WE. *The Essential Deming: Leadership Principles From the Father of Quality*. Nilsson Orsini J, ed. McGraw Hill; 2013:27-28.