Assessing Individual Hospitalist Performance: Domains and Attribution

John R Nelson, MD, MHM

Overlake Medical Center, Bellevue, Washington.

When asked by friend or family "Which hospital did you go to?" or "Which doctor did you see?" most are likely to answer with a single institution or clinician. Yet for hospital stays the patient's experience and outcomes are a product of many individuals and an entire system of care, so measuring performance at the group, or "team," level is appropriate.

Assessing and managing performance of *individuals* in healthcare is also important. In this regard, though, healthcare may be more like assessing individual baseball players prior to the widespread adoption of detailed statistics, a transition to what is often referred to as sabermetrics (and popularized by the 2004 book *Moneyball*).[1] An individual player's performance and future potential went from being assessed largely by the opinion of expert talent scouts to including, or even principally relying on, a wide array of measurements and statistics.

It sometimes seems healthcare has arrived at its "sabermetrics moment." There is a rapidly growing set of measures for individual clinicians, and nearly every week, hospitalists will open a new report of their performance sent by a payer, a government agency, their own hospitals, or other organizations. But most of these metrics suffer from problems with attributing performance to a single clinician; for example, many or most metrics attribute performance to the attending at the time of a patient's discharge according to the clinical record. Yet while clinical metrics (eg, administer beta-blocker when indicated, length of stay (LOS), readmissions), patient experience, financial metrics (eg, cost per case), and others are vital to understanding performance at an aggregate level such as a hospital or physician group, they are potentially confusing or even misleading when attributed entirely to the discharging provider. So healthcare leaders still tend to rely meaningfully on expert opinion—"talent scouts"—to identify high performers.

In this issue of the *Journal of Hospital Medicine*, Dow and colleagues have advanced our understanding of the current state of individual- rather than group-level hospitalist performance measurement.[2] This scoping review identified 43 studies published over the last 25 years reporting individual adult or pediatric hospitalist performance across one or more of the STEEEP framework domains of performance: Safe, Timely, Effective, Efficient, Equitable, Patient Centered.[3]

The most common domain assessed in the studies was Patient Centered (20 studies), and in descending order from there were Safe (16), Efficient (13), Timely (10), Effective (9). No studies reported individual hospitalist performance on Equitable care. This distribution of studied domains is likely a function of readily available data and processes for study more than level of interest or importance attached to each domain. Their research was not designed to assess the quality of each study, and some—or even many—might have weaknesses in both determining which clinicians met the definition of hospitalist and how performance was attributed to individuals. The authors appropriately conclude that "further defining and refining approaches to assess individual performance is necessary to ensure the highest quality."

Their findings should help guide research priorities regarding measurement of individual hospitalist performance. Yet each hospitalist group and individual hospitalist still faces decisions about managing their own group and personal performance and must navigate without the benefit of research providing clear direction. Many hospitalist metrics are tracked and reported to meet regulatory requirements such as those from Centers for Medicare & Medicaid Services, financial metrics for the local hospital and hospitalist group, and for use as components of hospitalist compensation. (The biennial State of Hospital Medicine Report captures extensive data regarding the latter.[4])

Many people and processes across an entire healthcare system influence performance on every metric, but it is useful and practical to attribute some metrics entirely to a single hospitalist provider, such as timely documentation and the time of day the discharge order is entered. And arguably, it is useful to attribute readmission rate entirely to the discharging provider—the last hospital provider who can influence readmission risk. But for most other metrics individual attribution is problematic or misleading and collective experience and expert opinion are helpful here. Two examples come to mind of relatively simple approaches that have gained some popularity in teasing out individual contribution to hospitalist performance.

One can estimate individual hospitalist contribution to patient LOS by calculating the ratio of current procedural terminology (CPT) codes for all follow-up services to all discharge codes. For each hospitalist in the group who cares for a similar population, those with the highest ratios likely manage patients in ways associated with longer LOS. It is relatively simple to use billing data to calculate the ratio, and some groups report it for all providers monthly.

Many metrics that aggregate performance across an entire hospital stay, such as patient experience surveys, can be ap-

Corresponding Author: John R Nelson, MD, MHM; Email: john.nelson@nelsonflores.com; Telephone: 425-467-3316.

portioned to each hospitalist who had a billed encounter with the patient. For example, if a hospitalist has 4 of a patient's 10 billed encounters within the same group, then 40% of the patient's survey score could be attributed to that hospitalist. It's still imperfect, but it's likely more meaningful than attributing the entire survey result to only the discharging provider.

These approaches have value but still leave us unsatisfied and unable to assess performance as effectively as we would like. Advancements in measurement have been slow and incremental, but they are likely to accelerate with maturation of electronic health records paired with machine learning or artificial intelligence, wearable devices, and sensors in patient rooms, which collectively may make capturing a robust set of metrics trivially easy (and raise questions regarding privacy and so forth). For example, it is already possible to capture via a smart speaker all conversations between patient, loved ones, and clinician.[5] Imagine you are presented with a word cloud summary of all conversations you had with all patients over a year. Did you use empathy words often enough? How reliably did you address all appropriate discharge-related topics?

As performance metrics become more numerous and ubiq-

uitous, the challenge will be to ensure they accurately capture what they appear to measure, are appropriately attributed to individuals or groups, and provide insights into important domains of performance. Significant opportunity for improvement remains.

References

1. Lewis M. *Moneyball: The Art of Winning an Unfair Game.* W.W. Norton & Company; 2004.
2. Dow AW, Chopski B, Cyrus JW, et al. A STEEEP hill to climb: a scoping review of assessments of individual hospitalist performance. *J Hosp Med.* 2020;15:599-605. https://doi.org/10.12788/jhm.3445
3. Institute of Medicine (US) Committee on Quality of Health Care in America. *Crossing the Quality Chasm: A New Health System for the 21st Century.* National Academy Press (US); 2001. https://doi.org/10.17226/10027
4. 2018 State of Hospital Medicine Report. Society of Hospital Medicine. Accessed May 19, 2020. https://www.hospitalmedicine.org/practice-management/shms-state-of-hospital-medicine/
5. Chiu CC, Tripathi A, Chou K, et al. Speech recognition for medical conversations. arXiv. Preprint posted online November 20, 2017. Revised June 20, 2018. https://arxiv.org/pdf/1711.07274.pdf