

PERSPECTIVE

Directly Observed Care: Crossing the Chasm of Quality Measurement



A. Taylor Kelley, MD, MPH, MSc^{1,2,3} , Saul J. Weiner, MD^{4,5}, and Joseph Francis, MD, MPH⁶

¹Vulnerable Veteran Innovative Patient Aligned Care Team (VIP) Initiative, Informatics, Decision-Enhancement, and Analytic Sciences (IDEAS) Center, VA Salt Lake City Health Care System, Salt Lake City, UT, USA; ²Division of General Internal Medicine, Department of Internal Medicine, University of Utah School of Medicine, 30 North 1900 East, Room 5R218, Salt Lake City, UT, USA; ³Program for Addiction Research, Clinical Care, Knowledge and Advocacy (PARCKA), Division of Epidemiology, Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, UT, USA; ⁴Center of Innovation for Complex Chronic Healthcare, Jesse Brown VA Medical Center, Chicago, IL, USA; ⁵Division of Academic Internal Medicine and Geriatrics, Department of Medicine, University of Illinois-Chicago, Chicago, IL, USA; ⁶Office of Analytics and Performance Integration, Veterans Health Administration, Washington, DC, USA.

After more than two decades of national attention to quality improvement in US healthcare, significant gaps in quality remain. A fundamental problem is that current approaches to measure quality are indirect and therefore imprecise, focusing on clinical documentation of care rather than the actual delivery of care. The National Academy of Medicine (NAM) has identified six domains of quality that are essential to address to improve quality: patient-centeredness, equity, timeliness, efficiency, effectiveness, and safety. In this perspective, we describe how directly observed care—a recorded audit of clinical care delivery—may address problems with current quality measurement, providing a more holistic assessment of healthcare delivery. We further show how directly observed care has the potential to improve each NAM domain of quality.

KEY WORDS: directly observed care; healthcare quality; unannounced standardized patients; patient-centered care.

J Gen Intern Med 38(1):203–7

DOI: 10.1007/s11606-022-07781-1

© This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2022. This article is an open access publication

Two decades ago, the National Academy of Medicine (NAM) released a landmark report that provocatively described a gap between “the care we have” and “the care we could have” as a quality “chasm.”¹ Quality measurement organizations then in their infancy, including the National Quality Forum (NQF) and National Committee for Quality Assurance (NCQA), responded swiftly, ushering in a rapid proliferation of quality measures to capture the structures,

processes, and outcomes of clinical care. Today, the National Quality Forum (NQF) endorses more than 440 quality measures for use in healthcare settings, while similar measures in the NCQA’s Healthcare Effectiveness and Information Set (HEDIS) are used in health plans covering more than 191 million individuals.²

Surprisingly, none of these measures attempts direct observation of clinical care delivery. Instead, each relies on surrogates—provider self-report, electronic medical records, billing claims, or patient survey—that are steps removed from the actual clinical encounter. This limited understanding about the “care we have” curtails our ability to cross the quality chasm.

The problem is two-fold: first, surrogate measures are inaccurate and incomplete. When compared to direct audiovisual assessment of patient care, provider documentation in the medical record reveals substantial inaccuracies, both of commission (services charted but not delivered) and omission (services delivered but not charted).^{3,4} Claims data are similarly lacking (e.g., measuring whether blood pressure monitoring occurred but *not* whether it was taken correctly). Further, while patient surveys may capture perceptions of care, such as how well a provider listens, they cannot discern whether the provider acts appropriately on the information presented.

Second, clinicians are rewarded more for how well they document than how well they care for patients. In current practice, quality measures based on documentation alone generally fail to capture provider behaviors that are important to quality, such as appropriately utilizing motivational interviewing to facilitate lifestyle change or taking time to find out why their patients are not taking their medication as directed. By contrast, providers may document actions to reflect high-quality care they are expected to deliver but in fact either deliver them inadequately or not at all. Unfortunately, research using observers or incognito standardized patients demonstrates that the latter scenario is all too common.^{3,4} Such lapses in quality cannot be addressed using current measures of quality assessment.

Prior Presentations A portion of this work was previously presented at an internal seminar for substance use care within the Department of Veterans Affairs.

Received January 28, 2022

Accepted August 31, 2022

Published online September 20, 2022

This chasm between care as currently measured and the excellence we aspire to achieve calls for a transformative strategy that encompasses holistic measurement *throughout* the care cycle: before, during, and after a clinical encounter. Directly observed care—a recorded audit of care provided to real or standardized patients—enables such a quality measurement strategy (see Fig. 1).⁵⁻⁸ As Fig. 1 illustrates, directly observed care complements asynchronous measures of quality that occur before and after care delivery by providing synchronous data of actual care the patient receives. Although directly observed care draws upon evaluation techniques such as “mystery shoppers” initially found in hospitality industries and the social sciences, it is increasingly common in healthcare evaluation and research.

The potential of directly observed care to transform healthcare quality has implications for each quality domain identified in the NAM report—patient-centeredness, equity, timeliness, efficiency, effectiveness, and safety. We describe how each may benefit from this approach.

PROMOTING PATIENT-CENTEREDNESS

Current quality measures can promote a “one size fits all” approach, inattentive to evidence that many aspects of care delivery require appropriate tailoring to each patient’s values and preferences.⁹ Knowing whether such attention was given requires observing the interaction between the clinician and patient. Did the physician ask the patient which treatment

option they prefer after describing the risks and benefits? Did they attempt to understand why the patient is not taking a critical medication as directed and attempt to intervene where feasible? Directly observed care instead evaluates patient-centeredness by objectively capturing whether relevant patient life circumstances are elicited, recognized, and subsequently addressed within the clinical encounter—a predictor of healthcare outcomes.¹⁰ Professionalism, communication skills, and the use of stigmatizing language can likewise be evaluated.¹¹ As shown in Fig. 1, these data complement patient experience data, medical records, and claims to create a 360-degree assessment of the entire care cycle.

Directly observed care can also reveal relative effectiveness of various practice patterns in delivering patient-centered care. For example, a meta-analysis of directly observed care across three datasets showed that patient concerns identified by providers are more likely to be incorporated into care plans when they are actively elicited by a provider rather than spontaneously stated by the patient.¹² In addition, directly observed care allows for intrinsic risk adjustment by incorporating individual patient characteristics into quality measurement—a major challenge using claims data alone.¹³

ENSURING EQUITY

Secret shoppers have played a pivotal role in studies of racism, including disparities in hiring practices, lending, and property

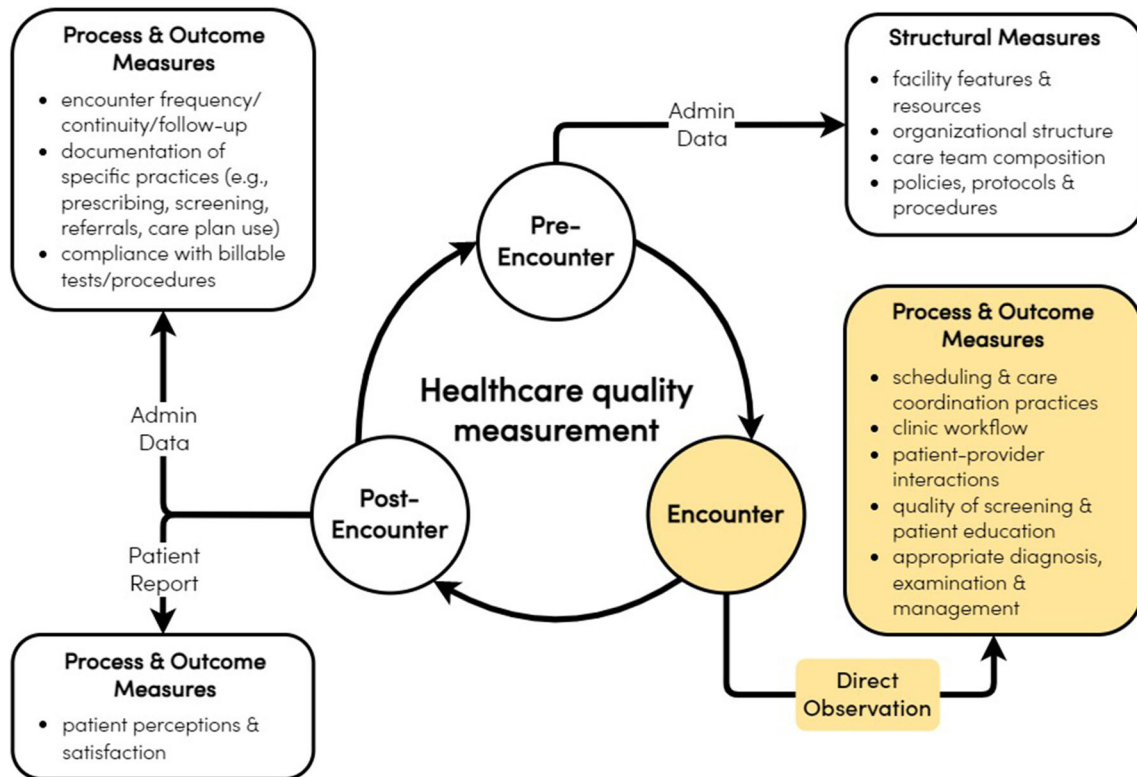


Fig. 1 Holistic quality measurement throughout the cycle of healthcare delivery.

rentals and sales.¹⁴ Unannounced standardized patients (USPs) may likewise serve an equally powerful role in identifying disparities in care at the level of the individual provider or health system. The advantage of evaluating healthcare equity through direct observation with USPs is that it avoids selection bias and confounding present in retrospective data while isolating one or more independent variables of interest (typically patient characteristics or concerns). Unbiased data are particularly helpful for provider feedback and subsequent evaluations that follow quality improvement interventions. For example, after a USP study of US Veterans experiencing homelessness showed that skin color predicted barriers to homeless services at community-based organizations, data feedback to these organizations resulted in resolution of the disparity within two years.¹⁵

USPs can likewise promote equitable comparisons of care quality across heterogeneous practice settings through a single standardized patient profile, which effectively serves as a reference standard. Recent USP evaluations using this approach have exposed wide practice variation in the provision of treatment for opioid use disorder, including residential treatment and pregnancy.^{16,17} Data such as these are compelling because they represent “ground truth” about the patient experience, inform targeted quality improvement interventions, and establish a baseline for subsequent assessments.

IMPROVING TIMELINESS AND EFFICIENCY

A major drawback of current quality measures is the time lag—typically 12 months or longer—that separates care delivery and quality reporting.¹⁸ These delays in feedback undermine principles of continuous quality improvement. In contrast, data collected from directly observed care and coded by trained staff can be shared with providers as feedback soon after an encounter. In one study, USPs showed that providers receiving training to engage patients in shared decision-making for prostate cancer screening decreased utilization of screening tests among patients with uncertain clinical benefit—all within a three-month evaluation period.¹⁹

While current quality measurement strategies engender inefficiencies and an increasing burden on healthcare providers,²⁰ there is growing evidence that observation data may be captured efficiently and sustainably over time when initial content coding validation by USPs is followed by direct observation of real patients (e.g., patient-collected audio) in regular clinical encounters.^{21,22} These two methods of directly observed care are complementary: USPs can portray specific scripts (e.g., high-stakes clinical scenarios where good communication is vital to treatment effectiveness) to assess how clinicians perform; patient-collected audio, in which patients volunteer to record their visits for quality improvement purposes, can assess how providers perform across a range of typical clinical situations. The former provides an

experimental method of quality assessment, while the latter is observational.

As additional directly observed care data from both USPs and real patients become available, “big data” approaches, such as automating content analysis using machine learning algorithms, may improve efficiency. Such efforts are already underway.

INCREASING EFFECTIVENESS AND SAFETY

The NAM report emphasizes that quality measurement must ultimately lead to quality improvement without resulting in excessive costs or self-defeating inefficiencies in care delivery. Unfortunately, quality measurement development since the report’s release has been subject to administrative burden, limited capability to account for risk, and other implementation challenges.²⁰ Currently available data are not only inadequate to assess many aspects of quality^{3,4} but are also often problematic for developing accurate, sustainable, and comprehensive quality measures that appeal across stakeholders, improve outcomes, and reduce costs. Directly observed care shows promise in generating effective performance data. For example, in the US Veterans Health Administration, a provider feedback intervention based on directly observed care delivery improved providers’ ability to identify and address contextualized patient risks. The result of the intervention was a significant reduction in avoidable hospitalizations, signaling safer delivery of care and saving millions of dollars in acute care expenses—more than a 70-fold return on investment.²¹

By its nature, directly observed care may be used in different environments, which can increase the effectiveness of quality measurement. For example, measuring care quality for substance use treatment within a rural American Indian community, while normally difficult for conventional quality measurement approaches, can be adapted to assess verbal cues, treatment plans, and cultural sensitivity pertinent to the environmental context.²³ Such adaptations are enhanced when they rely on stakeholder input to ensure representativeness of quality measures and on continuous audit and feedback to assess and improve care.

POTENTIAL CHALLENGES

As with any quality measurement or quality improvement strategy, directly observed care is not without challenges. Unannounced standardized patients are resource-intensive. However, as the studies we have described here demonstrate, very few clinical encounters are needed to adequately assess care quality using USPs. Further, with development of patient-collected audio, USPs may not be necessary for ongoing quality assessment and improvement, adding to cost savings and improving return on investment.²¹ There are also cultural concerns about the privacy of a provider-patient interaction

that could lead to consternation about directly observing care. Fortunately, early evidence suggests that when audio-recordings of visits are used for coaching and quality improvement rather than punitive or remedial purposes, provider satisfaction is high and *increases* over time.²⁴ Broad stakeholder input has also been shown to improve implementation and sustainability.²² Finally, there is a potential challenge that the Hawthorne effect could bias results when providers are aware they are being observed. While the Hawthorne effect has been demonstrated in many research settings, early results of directly observed care find similar patterns of care quality between studies when providers are aware they are being observed and when they are not aware.^{7,21} While more data is needed, these early results are reassuring that biases related to the Hawthorne effect are minimal in directly observed care.

CROSSING THE CHASM

Two decades ago, the NAM envisioned a future that would carry healthcare safely across the quality chasm. But before we can effectively improve quality, we must first cross the chasm of quality measurement. Those who deliver healthcare—as dedicated and well-intentioned as they may be—are unlikely to perform at their best if evaluated primarily by what they document rather than by what they do. While logistical and cultural challenges remain in implementing directly observed care into current quality measurement, it should be noted that direct observation is widely employed in other industries and has been implemented within healthcare at substantial scale in piloted quality improvement programs (full disclosure: including by the Institute for Practice and Provider Performance Improvement, of which one of the authors, SJW, is a principal). Although beyond the scope of this piece, the barriers are surmountable, and the costs of maintaining the status quo are likely greater.

As quality measurement organizations, policy makers, and other stakeholders consider the future of measuring and improving the care we provide, they should leverage data throughout the care cycle, including directly observed care. Only then can we cross the quality chasm in decades to come.

Acknowledgements: *The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs.*

Corresponding Author: A. Taylor Kelley, MD, MPH, MSc; Division of General Internal Medicine, Department of Internal Medicine, University of Utah School of Medicine, 30 North 1900 East, Room 5R218, Salt Lake City, UT 84132, USA (e-mail: Alan.Kelley@hsc.utah.edu).

Authors Contribution ATK completed the initial draft of the manuscript. All authors reviewed, edited, and contributed content to subsequent manuscript revisions.

Funding ATK is supported by funding from the Department of Veterans Affairs (the Vulnerable Veteran Innovative Patient-aligned Care Team (VIP) Initiative). SJW is supported by funding from Department of Veterans Affairs Health Services Research &

Development (HSR&D IIR-19-068 and IRP 20-001) and the Agency for Healthcare Research and Quality (R01 HS25374-04).

Declarations:

Conflict of Interest: SJW is a Principal of the Institute for Practice and Provider Performance Improvement (I3PI), founded to employ unannounced standardized patient assessments as a quality improvement service. There are no other relationships or activities that could appear to have influenced the submitted work.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

REFERENCES

1. Institute of Medicine (National Academy of Medicine) Committee on Quality of Health Care in America. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington (DC): National Academies Press (US). Copyright 2001 by the National Academy of Sciences. All rights reserved.; 2001.
2. HEDIS Measures and Technical Resources. Washington, D.C.: National Committee on Quality Assurance; 2021 [cited 2021 October 4]; Available from: <https://www.ncqa.org/hedis/measures/>.
3. Berdahl CT, Moran GJ, McBride O, et al. Concordance Between Electronic Clinical Documentation and Physicians' Observed Behavior. JAMA Netw Open. 2019;2(9):e1911390.
4. Weiner SJ, Wang S, Kelly B, et al. How accurate is the medical record? A comparison of the physician's note with a concealed audio recording in unannounced standardized patient encounters. J Am Med Inform Assoc. 2020;27(5):770-775.
5. Schwartz A, Peskin S, Spiro A, et al. Impact of Unannounced Standardized Patient Audit and Feedback on Care, Documentation, and Costs: an Experiment and Claims Analysis. J Gen Intern Med. 2021;36(1):27-34.
6. Schwartz A, Peskin S, Spiro A, et al. Direct observation of depression screening: identifying diagnostic error and improving accuracy through unannounced standardized patients. Diagnosis (Berl). 2020;7(3):251-6.
7. Weiner SJ, Schwartz A, Weaver F, et al. Contextual errors and failures in individualizing patient care: a multicenter study. Ann Intern Med. 2010;153(2):69-75.
8. Weiner SJ, Schwartz A. Directly observed care: can unannounced standardized patients address a gap in performance measurement? J Gen Intern Med. 2014;29(8):1183-7.
9. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. Lancet. 2017;390(10092):415-23.
10. Weiner SJ, Schwartz A, Sharma G, et al. Patient-centered decision making and health care outcomes: an observational study. Ann Intern Med. 2013;158(8):573-9.
11. Li L, Lin C, Guan J. Using standardized patients to evaluate hospital-based intervention outcomes. Int J Epidemiol. 2014;43(3):897-903.
12. Schwartz A, Weiner SJ, Binns-Calvey A, et al. Providers contextualise care more often when they discover patient context by asking: meta-analysis of three primary data sets. BMJ Qual Saf. 2016;25(3):159-63.
13. Risk Adjustment for Socioeconomic Status or Other Sociodemographic Factors. Accessed October 4, 2021, at https://www.qualityforum.org/Publications/2014/08/Risk_Adjustment_for_Socioeconomic_Status_or_Other_Sociodemographic_Factors.aspx: National Quality Forum, 2014 August 15. Report No.
14. Lederer A, Oros S, Bone S, et al. Lending Discrimination within the Paycheck Protection Program. Washington, D.C.: National Community Reinvestment Coalition, 2020 July 15. Report No.

15. **Weiner SJ, Schwartz A, Binns-Calvey A, et al.** Impact of an unannounced standardized veteran program on access to community-based services for veterans experiencing homelessness. *J Public Health (Oxf)*. 2021;44(1):207-213.
16. **Beetham T, Saloner B, Gaye M, et al.** Admission Practices And Cost Of Care For Opioid Use Disorder At Residential Addiction Treatment Programs In The US. *Health Aff (Millwood)*. 2021;40(2):317-25.
17. **Patrick SW, Richards MR, Dupont WD, et al.** Association of Pregnancy and Insurance Status With Treatment Access for Opioid Use Disorder. *JAMA Netw Open*. 2020;3(8):e2013456.
18. **Austin JM, Kachalia A.** The State of Health Care Quality Measurement in the Era of COVID-19: The Importance of Doing Better. *JAMA*. 2020;324(4):333-4.
19. **Feng B, Srinivasan M, Hoffman JR, et al.** Physician communication regarding prostate cancer screening: analysis of unannounced standardized patient visits. *Ann Fam Med*. 2013;11(4):315-23.
20. **Panzer RJ, Gitomer RS, Greene WH, et al.** Increasing demands for quality measurement. *JAMA*. 2013;310(18):1971-80.
21. **Weiner S, Schwartz A, Altman L, et al.** Evaluation of a Patient-Collected Audio Audit and Feedback Quality Improvement Program on Clinician Attention to Patient Life Context and Health Care Costs in the Veterans Affairs Health Care System. *JAMA Netw Open*. 2020;3(7):e209644.
22. **Weiner SJ, Schwartz A, Sharma G, et al.** Patient-collected audio for performance assessment of the clinical encounter. *Jt Comm J Qual Patient Saf*. 2015;41(6):273-8.
23. **Kelley AT, Smid MC, Baylis JD, et al.** Development of an unannounced standardized patient protocol to evaluate opioid use disorder treatment in pregnancy for American Indian and rural communities. *Addict Sci Clin Pract*. 2021;16(1):40.
24. **Ball SL, Weiner SJ, Schwartz A, et al.** Implementation of a patient-collected audio recording audit & feedback quality improvement program to prevent contextual error: stakeholder perspective. *BMC Health Serv Res*. 2021;21(1):891.

Publisher's Note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.