Editorial

Regarding “Implicit Review Instrument to Evaluate Quality of Care Delivered by Physicians to Children in Emergency Departments”

There is an ongoing interest in the development and implementation of quality measures for pediatric care, highlighted by the initial funding and renewal of the Pediatric Quality Measurement Program. This initial legislation charged the Agency for Healthcare Research and Quality and the Centers for Medicare and Medicaid Services to identify “an initial set of pediatric core measures to assess the state of children’s health care quality across and within State Medicaid and CHIP programs and then for establishment of the PQMP to improve and strengthen the initial core set of measures and develop new measures as needed” (Mistry et al. 2014, https://www.ahrq.gov/sites/default/files/publications/files/pqmpfact_0.pdf [accessed on August 20, 2017]). As part of this initiative, numerous groups developed pediatric quality measures for both general pediatric care and specific subspecialties such as sickle cell disease, asthma, and children with complex medical conditions. These measures primarily took the form of a specific outcome or process measure that could be assessed using administrative datasets such as insurance data from Medicaid or private insurance claims; hospital administrative data usually taken from state claims-based data; or hospital-complied data using an explicit quality instrument for a specific condition.

Within this framework, other groups have developed additional quality measures, such as the instrument developed by Marcin et al. (2018) and the PECARN research network. Unlike most other quality measures, this study provides a thorough overview of the consistency, reliability, and validity of an *implicit* measure of overall emergency department care. This instrument asks

reviewers to assess care among four measures of care processes (data gathering, integration of information and the development of appropriate diagnoses, initial treatment plan and orders, and disposition and follow-up) and a fifth measure of the overall quality of care. The measure demonstrated strong internal consistency with an overall Cronbach's alpha of 0.9, moderate inter-rater reliability that is consistent or better than previously published implicit measures (Goodman 1992; Smith et al. 1997), and modest correlation with four disease-specific explicit review instruments. However, given the plethora of pediatric quality measures already developed, we may ask the question, "Do we need another measure of care quality?," in this case for pediatric emergency medical care, but that could easily be asked for any medical specialty or condition.

To answer this question, it is important to place this measure of pediatric emergency care within the scope of what the public, policy makers, insurers, and medical care professionals would ideally like to have in a quality measure. There are two key aspects to this question. First, we have to assess the characteristics of the specific quality measure. An ideal quality measure would have the following characteristics, as adapted from Lorch (2017a):

- Easy to classify
- Easy to measure
- High reliability and reproducibility
- Face and construct validity
- Present in sufficient numbers to minimize loss of statistical power to detect significant differences between health care providers and facilities
- Adequate risk adjustment if necessary

As stated above, most quality measures focus on outcome measures such as medical complications, infection rates, or readmission rates, which are easy to measure within an administrative dataset and thus available for health care organizations within that dataset. Compared with these measures, the implicit review tool developed by Marcin et al. (2018) is much more complicated. For this project, the authors employed eight pediatric emergency medicine

Address correspondence to Scott A. Lorch, M.D., M.S.C.E., Roberts Center for Pediatric Research, 2716 South Street, Floor 10, Room 10215, Philadelphia, PA 19146; e-mail: lorch@email.chop.edu. Scott A. Lorch, M.D., M.S.C.E., is with the Department of Pediatrics, Division of Neonatology, The Children's Hospital of Philadelphia, Philadelphia, PA; Center for Pediatric and Perinatal Health Disparities Research and PolicyLab, The Children's Hospital of Philadelphia, Philadelphia, PA; Leonard Davis Institute, University of Pennsylvania, Philadelphia, PA, and the Center for Clinical Epidemiology and Biostatistics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA.

physicians to examine over 300 medical records each, each of whom spent 1 day in in-person training and additional electronic meetings and chart review. While the internal consistency of their tool was high, the Intraclass Correlation Coefficient of 0.49 for their overall summary score of the quality of care was only moderate. Finally, prior studies consistently raise the concern about poor inter-rater reliability of implicit measures and potential biases that may result in variation between reviewers (Hayward, McMahon, and Bernard 1993). Such variation was seen in the Marcin measure. The amount of time needed to implement such a tool on a population level may explain why few organizations or policy makers use implicit or even explicit chart-based tools to assess the quality of care.

However, this implicit review tool truly shines when we examine the second aspect of the ideal quality measure: What does it truly measure? In truth, stakeholders would like a quality measure that measures the nebulous construct, “quality of care.” The quality “black box,” though, resists easy measurement, primarily because there are so many different dimensions and aspects of clinical care that require measurement. As a result, quality measurements can be divided into two groups, depending on how they assess this construct (Lorch 2013, 2017a):

- Measures that broadly assess the quality of an organization, but provide limited data about a specific dimension of care or topic area. Observed improvements in these “broad” measures should impact numerous dimensions of clinical practice and outcomes, but the specific size of this improvement for a specific dimension may be difficult to assess *a priori*.
- Measures that assess a specific dimension of care, which may or may not be correlated with another “narrow” measure depending on the specific structures and processes of care that determine the outcome of that dimension. Observed improvements in one “narrow” measure may not impact another measure; for example, measures to improve asthma and pulmonary care may not result in changes to an organization’s central-line associated blood stream infection rate. The challenge for any of these narrowly defined measures is to include enough information to fully capture the care delivered by a health care profession (Lorch 2017b).

As we might imagine, the implicit review tool provides an insight into the overall care provided by a provider that may not be assessed with a single outcome or process measure, or even by a composite measure of these outcomes. For example, one study of complications as a potential quality metric found that characteristics such as patient illness, the complexity of the specific case, and potential technical difficulties especially for surgical cases may have affected physician’s judgement concerning the role of quality in the development of a specific complication (Weingart et al. 2000). Given this framework, then, it should come as little surprise that the correlation between this implicit

review measure and four disease-specific explicit care tools only ranged from 0.24 to 0.38.

So, do we really need to develop additional quality measures, or this one in particular? The answer to this question depends on the user and the capabilities of current quality measures for the condition measured. In most cases, different quality measures will assess a different part of the quality box with different accuracy and validity. For pediatric emergency medicine, and pediatric care in general, there is no one measure that truly captures the care delivered by hospitals in total. Broad-based measures, such as the one proposed by Marcin et al. (2018), likely better summarize the overall care of health care provider without providing the details needed to implement quality improvement projects or identify dimensions of care for further evaluation and root-cause analyses. Narrow-based measures, while generally easier to implement and monitor, may only truly assess care delivery for the dimension that the measure assesses. To provide a snapshot of the care delivered by a health care organization, we should be considering what dimensions of quality are assessed by a specific measure, the ease to implement the measure, and the needs of the various stakeholders. A careful development of a combination of validated broad and narrow measures may be more effective than adding hundreds of additional measures—usually narrow—of the same dimension of care, which challenge providers and other stakeholders to collect and interpret these data. Ignoring broad measures because of their limitations secondary to challenges to data collection and time may limit our ability to identify truly low-performing providers across all dimensions of care.

ACKNOWLEDGMENT

Joint Acknowledgment/Disclosure Statement: None.

Disclosures: None.

Disclosure: None.

Scott A. Lorch

REFERENCES

- Goodman, R. L. 1992. "The Reliability of Peer Assessments of Quality of Care." *Journal of the American Medical Association* 267 (7): 958–60.

- Hayward, R. A., L. F. McMahon Jr, and A. M. Bernard. 1993. "Evaluating the Care of General Medicine Inpatients: How Good Is Implicit Review?" *Annals of Internal Medicine* 118 (7): 550–6.
- Lorch, S. A. 2013. "Quality Measurements in Pediatrics: What Do They Assess?" *JAMA Pediatrics* 167 (1): 89–90.
- . 2017a. "National Quality Measures in Perinatal Medicine." *Clinics in Perinatology* 44 (3): 485–509.
- . 2017b. "Challenges to Measuring the Quality of Low-Risk Newborns." *Pediatrics* 139 (3): e20164025.
- Marcin, J. P., P. S. Romano, M. Dharmar, J. M. Chamberlain, N. Dudley, C. G. Macias, L. E. Nigrovic, E. C. Powell, A. J. Rogers, M. Sonnett, L. Tzimenatos, E. R. Alpern, R. Andrew-Dickert, D. A. Borgialli, E. Sidney, C. Casper, J. M. Dean, and N. Kuppermann. 2018. "Implicit Review Instrument to Evaluate Quality of Care Delivered by Physicians to Children in Emergency Departments." *Health Services Research* 53 (3): 1316–34. <https://doi.org/10.1111/1475-6773.12800>
- Mistry, K. B., F. Chesley, K. Llanos, and D. Dougherty. 2014. "Advancing Children's Health Care and Outcomes through the Pediatric Quality Measures Program." *Academic Pediatrics* 14 (5 Supp): S19–26.
- Smith, M. A., A. J. Atherly, R. L. Kane, and J. T. Pacala. 1997. "Peer Review of the Quality of Care. Reliability and Sources of Variability for Outcome and Processes Assessments." *Journal of the American Medical Association* 278 (19): 1573–8.
- Weingart, S. N., L. I. Iezzoni, R. B. Davis, R. H. Palmer, M. Cahalane, M. B. Hamel, K. Mukamal, R. S. Phillips, D. T. Davies Jr, and N. J. Banks. 2000. "Use of Administrative Data to Find Substandard Care: Validation of the Complications Screening Program." *Medical Care* 38 (8): 796–806.