Published in final edited form as:

Med Care. 2009 April ; 47(4): 375–377. doi:10.1097/MLR.0b013e3181a14b65.

Improving the Quality of Quality Measurement: The Tinkerer, The Tailor and The Candlestick Maker

Monika M. Safford, MD

University of Alabama at Birmingham

Editorial

The US healthcare industry has been assessing the quality of healthcare using public reporting for nearly 15 years now. As far back as 2003, reports began to indicate measurable improvements in quality indicators.¹ Yet, these early signs of success were not enough, as many individuals continued to receive suboptimal care. The pay-for-performance experiment ensued, with reimbursements tied to quality in hopes of stimulating more improvements.

In the midst of this tour de force of commitment to improving quality of care in the United States, several dissenting voices were heard. Early critics pointed to shortcomings of quality measures that assess the proportion of individuals above a certain risk factor level, such as blood pressure >140/90 mm Hg. Thresholds ignore the very real benefits derived from lowering a risk factor from a high to a moderate level, which has greater health impact than bringing someone from just above goal to just below it. Yet, quality indicators based on thresholds equate these two very different clinical and quality scenarios. Additional troublesome signs began to emerge in reports confirming that some indicators based on risk factor achievement were influenced by patient characteristics, like the age of the patients or the severity of their illness.^{2,3} These characteristics do not reflect quality of care, and the possibility of unfairness in quality comparisons began to be discussed. Moreover, other reports began to raise the possibility that some quality indicators related to testing risk factors may not be associated with improved health outcomes.⁴ The report by Mangione et al found associations between health plan clinical care management and testing rates, but not risk factor levels.⁴ Added to these issues is a growing concern that an overall picture of quality of care is not reflected well in individual quality measures.

Onto this stage enter 3 important new actors, each with an appearance in this issue of *Medical Care*. The *tinkerer* takes current quality measures and improves on the way they are used. The report by Kaplan et al⁵ is an example of this approach. No novel measures are proposed, rather, existing measures are aggregated in a mathematical way, improving reliability and providing the sought-after overall picture of the quality of care. The authors make a compelling case that their approach may be less prone to the influences of patient characteristics on quality comparisons. The article takes an increasingly popular approach in studies of quality comparisons, using multilevel models and focusing on intraclass correlation coefficients (ICC) to detect physician “thumbprints.” Several cautions may be warranted with this approach. First, although variation in care not explained by patient factors is certainly a cause for concern, this approach will not work if all physicians render roughly consistent levels of inadequate care. Racial and ethnic biases come to mind, especially those that are not overt.⁶ Consider a disturbing recent report of hysterectomy

among middle-aged women, demonstrating an over 3-fold higher rate for African Americans compared with European American women, after accounting for a host of physiologic and socioeconomic factors.⁷ Further, this novel kind of use of ICC in a quality and accountability context raises the ante for certainty. We are not accustomed to examining confidence bands around an ICC, but in this context we should consider doing so. How certain are Kaplan et al that a given ICC is 0.3, the threshold for a sufficiently large “thumbprint” and inclusion in their summary measure, and not 0.1, which would not be included? The authors are to be commended for the mathematical ingenuity that will permit a different view of albeit the same familiar measures, polished to a new gleam, but still only representing the same small sliver of the whole quality pie that we have assessed in the accountability context to date.

A slightly different approach is taken by the *tailor*. Persell et al⁸ demonstrate the limitations of the blood pressure control quality indicator by basting on a few more clinical scenarios, such as including individuals who at the previous visit were controlled, or whose mean blood pressure over the past year was controlled, or those on aggressive treatment. These whip stitches added to the original measure improve the proportion assessed as receiving appropriate care by nearly 50%. Although the precise tailoring of the measure can be argued, a new concept emerges from this work: measuring the proportion of patients with appropriate care, rather than the proportion above some fixed threshold. This is an important step in a new direction, beginning to cut into another slice of the quality pie. As we recognize that current quality indicators and public reporting have not led to sufficient improvements, it may be wise to pause to consider why not. The measures that assess processes like A1c testing tell us about a proportion of patients who received “good care,” but that care is not necessarily followed downstream by better outcomes. The measures that assess the proportion of patients that achieved a desired outcome do not tell us anything about how they got there. As we begin to develop and propose new quality indicators, it may be wise to consider whether they not only assess care but also provide a roadmap toward improvement. By focusing on care processes that are more directly under the provider’s control, such as the aggressive treatment component of Persell’s approach, the pathway toward achieving better control may be made clearer.

This concept is taken a considerable step further by the *candlestick maker*, here embodied in the paper by Selby et al.⁹ This approach takes a firm step into new territory, slicing off an entirely new dimension of the quality of care pie. The targeted process of care is one that few physicians would argue is directly under the doctor’s control: medication management. This group takes their previous work in this area an important step further¹⁰: they validate that their measure of medication intensification is in fact associated with the desired outcome, and also that variations at the facility level are associated. The pathway to improvement is clearly laid out. Systems at the higher performing facilities could be examined to identify how to structure care so that appropriate medication management is facilitated. Although their focus in this report is on the system level, one could envision this quality indicator being quite useful at the physician level as well, possibly in audit and feedback programs.

Another important feature of Selby’s approach that points the way to improvement is the focus of their measure on a subset of all patients with hypertension. Rather than providing an opaque proportion controlled, the proportion intensified out of a denominator of good candidates for intensification is highly transparent. Although the reasons why physicians do not always intensify medications at encounters with patients with uncontrolled high blood pressure are poorly understood, there is ample evidence that risks rise along with the blood pressure level. By targeting a group that has strong indications for treatment, the influence of other patient characteristics is lessened, and the potential reasons for not intensifying may

be fewer. The variation on intensification rates described in this report is certainly concerning, and provides a firm foundation for ongoing assessment of this quality indicator as a new dimension of quality of care.

The authors sound a caution that is worth pausing on: what about patients who are having trouble taking medications? This is a common problem, and providers are not very good at detecting it.^{11,12} Few would consider it good clinical care to intensify medication regimens for people who are not even taking the medicines in the first place. Indeed, it may be wise to construct 2 quality indicators for this slice of the quality pie: one focusing on appropriate intensification among candidates who are both poorly controlled and demonstrate good medication adherence, guided by Selby's approach, and a second assessing appropriate clinical actions for patients having trouble taking medications as prescribed. Exactly what this second measure should look like is worth thoughtful consideration.

The fans of a summary measure of quality of care may not be enthusiastic about honing in on a targeted subgroup, albeit one at high risk. What about the quality of care for everyone else? Persell's report points the way to constructing a picture of the overall "appropriateness of care" of patients with high blood pressure. One could amalgamate the proportion in control, the proportion of adherent patients with poor control who were intensified, the proportion of all who were already aggressively treated, and the proportion of the nonadherent patients who had adherence addressed. Different clinical actions may be appropriate in each of these subgroups—intensifying medications, not intensifying medications, or reinstating the regimen. A key element emerging from this work is the concept of "smart" quality indicators, each assessing quality while identifying which clinical process needs work, and together providing an overview of the quality of care being provided.

The tinkerer, the tailor and the candlestick maker open exciting new doors into the future of performance measurement. Whether demonstrating how to improve the way we use existing measures or proposing new measures, these 3 excellent reports should move quality assessment forward in the continual quest for improved healthcare quality and outcomes in the United States.

REFERENCES

1. Jencks SF, Huff ED, Cuerdon T. Change in the quality of care delivered to Medicare beneficiaries, 1998–1999 to 2000–2001. *JAMA*. 2003; 289:305–312. [PubMed: 12525231]
2. Zaslavsky AM, Hochheimer JN, Schneider EC, et al. Impact of sociodemographic case mix on the HEDIS measures of health plan quality. *Med Care*. 2000; 38:981–992. [PubMed: 11021671]
3. Zhang Q, Safford M, Ottenweller J, et al. Performance status of health care facilities changes with risk adjustment of HbA1c [comment]. *Diabetes Care*. 2000; 23:919–927. [PubMed: 10895841]
4. Mangione CM, Gerzoff RB, Williamson DF, et al. The association between quality of care and the intensity of diabetes disease management programs. *Ann Intern Med*. 2006; 145:107–116. [PubMed: 16847293]
5. Kaplan SH, Griffith JL, Price LL, et al. Improving the reliability of physician performance assessment: identifying the "physician effect" on quality and creating composite measures. *Med Care*. 2009; 47:378–387. [PubMed: 19279511]
6. Schulman KA, Berlin JA, Harless W, et al. The effect of race and sex on physicians' recommendations for cardiac catheterization. *N Engl J Med*. 1999; 340:618–626. [PubMed: 10029647]
7. Bower JK, Schreiner PJ, Sternfeld B, et al. Black-white differences in hysterectomy prevalence: the CARDIA study. *Am J Public Health*. 2009; 99:300–307. [PubMed: 19059854]
8. Persell SD, Kho AN, Thompson JA, et al. Improving hypertension quality measurement using electronic health records. *Med Care*. 2009; 47:388–394. [PubMed: 19330887]

9. Selby JV, Uratsu CS, Fireman B, et al. Treatment intensification and risk factor control: toward more clinically relevant quality measures. *Med Care*. 2009; 47:395–402. [PubMed: 19330888]
10. Rodondi N, Peng T, Karter AJ, et al. Therapy modifications in response to poorly controlled hypertension, dyslipidemia, and diabetes mellitus. *Ann Intern Med*. 2006; 144:475–484. [PubMed: 16585661]
11. Heisler M, Hogan MM, Hofer TP, et al. When more is not better: treatment intensification among hypertensive patients with poor medication adherence. *Circulation*. 2008; 117:2884–2892. [PubMed: 18506011]
12. Schmittiel JA, Uratsu CS, Karter AJ, et al. Why don't diabetes patients achieve recommended risk factor targets? Poor adherence versus lack of treatment intensification. *J Gen Intern Med*. 2008; 23:588–594. [PubMed: 18317847]