

## Do Clinicians Understand Quality Metric Data? An Evaluation in a Twitter-Derived Sample

Sushant Govindan, MD<sup>1\*</sup>, Vineet Chopra, MD, MSc<sup>1-3</sup>, Theodore J. Iwashyna, MD, PhD<sup>1,2</sup>

<sup>1</sup>Department of Medicine, University of Michigan Health System, Ann Arbor, Michigan; <sup>2</sup>Center for Clinical Management Research, Ann Arbor VA Healthcare System, Ann Arbor, Michigan; <sup>3</sup>Patient Safety Enhancement Program, Ann Arbor VA Medical Center, Ann Arbor, Michigan.

**OBJECTIVE:** Despite significant efforts and cost, quality metrics do not consistently influence practice. While research has focused on improving data through statistical risk-adjustment, whether clinicians understand these data is unknown. Therefore, we assessed clinician comprehension of central line-associated blood stream infection (CLABSI) quality metric data.

**DESIGN:** Cross-sectional survey with an 11-item test of CLABSI data comprehension. Each question assessed 1 of 3 concepts concerning CLABSI understanding: basic numeracy, risk-adjustment numeracy, and risk-adjustment interpretation. Hypothetical data were used and presented in a validated format.

**PARTICIPANTS:** Clinicians were recruited from 6 nations via Twitter to take an online survey. Clinician eligibility was confirmed by assessing responses to a question regarding CLABSI.

**MAIN MEASURES:** The primary outcome was percent correct of attempted questions pertaining to the presented CLABSI data.

**RESULTS:** Ninety-seven clinicians answered at least 1 item, providing 939 responses; 72 answered all 11 items. The mean percentage of correct answers was 61% (95% confidence interval [CI], 57%-65%). Overall, doctor performance was better than performance by nurses and other respondents (68% [95% CI, 63%-73%] vs. 57% [95% CI, 52%-62%],  $P = 0.003$ ). In basic numeracy, mean percent correct was 82% (95% CI, 77%-87%). For risk-adjustment numeracy, the mean percent correct was 70% (95% CI, 64%-76%). Risk-adjustment interpretation had the lowest average percent correct, 43% (95% CI, 37%-49%). All pairwise differences between concepts were statistically significant at  $P < 0.05$ .

**CONCLUSIONS:** CLABSI quality metric comprehension appears low and varies substantially among clinicians. These findings may contribute to the limited impact of quality metric reporting programs, and further research is needed. *Journal of Hospital Medicine* 2017;12:18-22. © 2017 Society of Hospital Medicine

Central line-associated bloodstream infections (CLABSIs) are common and serious occurrences across healthcare systems, with an attributable mortality of 12% to 25%.<sup>1,2</sup> Given this burden,<sup>3-5</sup> CLABSI is a focus for both high-profile public reporting and quality improvement interventions. An integral component of such interventions is audit and feedback via quality metrics. These measures are intended to allow decision makers to assess their own performance and appropriately allocate resources. Quality metrics present a substantial cost to health systems, with an estimated \$15.4 billion dollars spent annually simply for reporting.<sup>6</sup> Despite this toll, “audit and feedback” interventions have proven to be variably successful.<sup>7-9</sup> The mechanisms that limit the effectiveness of these interventions remain poorly understood.

One plausible explanation for limited efficacy of quality metrics is inadequate clinician numeracy—that is, “the ability to understand the quantitative aspects of clinical medi-

cine, original research, quality improvement, and financial matters.”<sup>10</sup> Indeed, clinicians are not consistently able to interpret probabilities and or clinical test characteristics. For example, Wegwarth et al. identified shortcomings in physician application of lead-time bias toward cancer screening.<sup>11</sup> Additionally, studies have demonstrated systematic misinterpretations of probabilistic information in clinical settings, along with misconceptions regarding the impact of prevalence on post-test probabilities.<sup>12,13</sup> Effective interpretation of rates may be a key—if unstated—requirement of many CLABSI quality improvement efforts.<sup>14-19</sup> Our broader hypothesis is that clinicians who can more accurately interpret quality data, even if only from their own institution, are more likely to act on it appropriately and persistently than those who feel they must depend on a preprocessed interpretation of that same data by some other expert.

Therefore, we designed a survey to assess the numeracy of clinicians on CLABSI data presented in a prototypical feedback report. We studied 3 domains of comprehension: (1) basic numeracy: numerical tasks related to simple data; (2) risk-adjustment numeracy: numerical tasks related to risk-adjusted data; and (3) risk-adjustment interpretation: inferential tasks concerning risk-adjusted data. We hypothesized that clinician performance would vary substantially across domains, with the poorest performance in risk-adjusted data.

\*Address for correspondence and reprint requests: Sushant Govindan, MD, Taubman Center, Floor 3 Room 3920, 1500 East Medical Center Drive, SPC 5360, Ann Arbor, MI 48109; Telephone: 734-763-9077; Fax: 734-764-4556; E-mail: sushantg@med.umich.edu

Additional Supporting Information may be found in the online version of this article.

Received: May 11, 2016; Revised: July 20, 2016; Accepted: August 4, 2016

2017 Society of Hospital Medicine DIO 10.1002/jhm.2680

## METHODS

We conducted a cross-sectional survey of clinician numeracy regarding CLABSI feedback data. Respondents were also asked to provide demographic information and opinions regarding the reliability of quality metric data. Survey recruitment occurred on Twitter, a novel approach that leveraged social media to facilitate rapid recruitment of participants. The study instrument was administered using a web survey with randomized question order to preclude any possibility of order effects between questions. The study was deemed Institutional Review Board exempt by the University of Michigan: protocol HUM00106696.

### Data Presentation Method

To determine the optimal mode of presenting data, we reviewed the literature on quality metric numeracy and presentation methods. Additionally, we evaluated quality metric presentation methods used by the Centers for Disease Control and Prevention (CDC), Centers for Medicare & Medicaid Services (CMS), and a tertiary academic medical center. After assessing the available literature and options, we adapted a CLABSI data presentation array from a study that had qualitatively validated the format using physician feedback (Appendix).<sup>20</sup> We used hypothetical CLABSI data for our survey.

### Survey Development

We developed a survey that included an 11-item test regarding CLABSI numeracy and data interpretation. Additional questions related to quality metric reliability and demographic information were included. No preexisting assessment tools existed for our areas of interest. Therefore, we developed a novel instrument using a broad, exploratory approach as others have employed.<sup>21</sup>

First, we defined 3 conceptual categories related to CLABSI data. Within this conceptual framework, an iterative process of development and revision was used to assemble a question bank from which the survey would be constructed. A series of think-aloud sessions were held to evaluate each prompt for precision, clarity, and accuracy in assessing the conceptual categories. Correct and incorrect answers were defined based on literature review in conjunction with input from methodological and content experts (TJI and VC) (see Appendix for answer explanations).

Within the conceptual categories related to CLABSI risk-adjustment, a key measure is the standardized infection ratio (SIR). This value is defined as the ratio of observed number of CLABSI over the expected number of CLABSIs.<sup>22</sup> This is the primary measure to stratify hospital performance, and it was used in our assessment of risk-adjustment comprehension. In total, 54 question prompts were developed and subsequently narrowed to 11 study questions for the initial survey.

The instrument was then pretested in a cohort of 8 hospitalists and intensivists to ensure appropriate comprehension, retrieval, and judgment processes.<sup>23</sup> Questions were revised

based on feedback from this cognitive testing to constitute the final instrument. During the survey, the data table was reshown on each page directly above each question and so was always on the same screen for the respondents.

### Survey Sample

We innovated by using Twitter as an online platform for recruiting participants; we used Survey Monkey to host the electronic instrument. Two authors (TJI, VC) systematically sent out solicitation tweets to their followers. These tweets clearly indicated that the recruitment was for the purpose of a research study, and participants would receive no financial reward/incentive (Appendix). A link to the survey was provided in each tweet, and the period of recruitment was 30 days. To ensure respondents were clinicians, they needed to first answer a screening question recognizing that central lines were placed in the subclavian site but not the aorta, iliac, or radial sites.

To prevent systematic or anchoring biases, the order of questions was electronically randomized for each respondent. The primary outcome was the percentage correct of attempted questions.

### Statistical Analysis

Descriptive statistics were calculated for all demographic variables. The primary outcome was evaluated as a dichotomous variable for each question (correct vs. incorrect response), and as a continuous variable when assessing mean percent correct on the overall survey. Demographic and conceptual associations were assessed via t-tests, chi-square, or Fisher exact tests. Point biserial correlations were calculated to assess for associations between response to a single question and overall performance on the survey.

**TABLE 1. Respondent Demographics**

	No., n = 97 <sup>a</sup> (%)
Country	United States: 68 (85) Other: 12 (15)
Profession	Doctor: 39 (48) Nurse: 31 (39) Other: 11 (13)
Specialty (Doctor)	Internal medicine, pulmonary and critical care: 23 (59) Internal medicine, general medicine/hospitalist medicine: 8 (20) Other <sup>b</sup> : 8 (21)
Seen hospitalized patients in past 12 months	Yes: 77 (96) No: 3 (4)
Years of experience	In training: 7 (9) 1-5: 15 (19) 6-10: 24 (30) 11-20: 20 (25) 21-30: 10 (13) 31-40: 3 (4)
Member of a hospital quality committee	Yes: 31 (39) No: 49 (61)

<sup>a</sup>Not all demographic questions were answered by every respondent.

<sup>b</sup>Other professions (n): physiotherapist (3), occupational therapist (1), nurse practitioner (2), physician assistant (1), infection preventionist (1), researcher (1), quality improvement professional (1), medical technologist (1).

To evaluate the association between various respondent characteristics and responses, logistic regression analyses were performed. An ANOVA was performed to assess the association between self-reported reliability of quality metric data and the overall performance on attempted items. Analyses were conducted using STATA MP 14.0 (College Station, TX);  $P < 0.05$  was considered statistically significant.

## RESULTS

A total of 97 respondents attempted at least 1 question on the survey, and 72 respondents attempted all 11 questions,

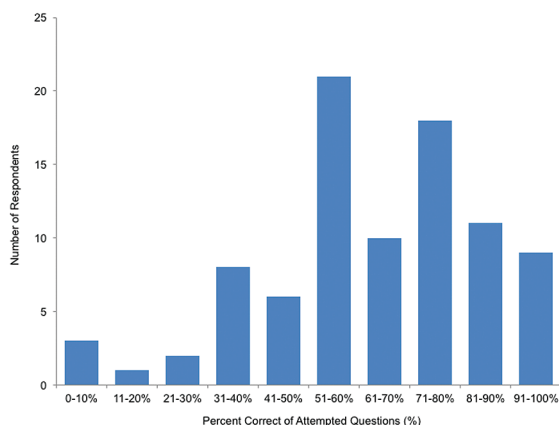


FIG. 1. Percent correct of attempted questions

TABLE 2. CLABSI Numeracy and Interpretation Assessment

Question	Cohort Percent Correct
<b>Basic Numeracy</b>	
Which hospital uses the most central lines?	90%
Which hospital has the lowest CLABSI rate?	80%
If hospital A doubled its central-line use but other practice patterns remained the same, how many actual infections would hospital A expect to have?	79%
If hospital G's number of actual infections doubled, what would its CLABSI rate be?	77%
<b>Risk-Adjustment Numeracy</b>	
Which is better: a higher or lower SIR?	95%
If hospital B had its number of projected infections halved, what is its SIR?	46%
<b>Risk-Adjustment Interpretation</b>	
The presence of a gastrostomy (g) tube is a risk factor for CLABSI. If this variable is not accounted for in CLABSI reporting, how would this impact the interpretation of the number of infections projected by national experience?	75%
Which hospital is most effective at preventing CLABSI?	51%
Suppose hospitals A and H have the exact same CLABSI prevention practices. Which hospital will have the higher number of CLABSI?	34%
Which hospital's patients are the most predisposed to developing CLABSI?	32%
Suppose hospital A begins using a central line with an antibiotic coating that halves infections. What would hospital A's number of projected infections be?	17%

NOTE: All surveys had a randomized order of questions, and the data table was shown directly above each question. Abbreviations: CLABSI, central line-associated blood stream infection; SIR, standardized infection ratio.

yielding 939 unique responses for analysis. Seventy respondents (87%) identified as doctors or nurses, and 44 (55%) reported having 6 to 20 years of experience; the survey cohort also came from 6 nations (Table 1). All respondents answered the CLABSI knowledge filter question correctly.

### Primary Outcome

The mean percent correct of attempted questions was 61% (standard deviation 21%, interquartile range 50%-75%) (Figure 1). Of those who answered all 11 CLABSI questions, the mean percent correct was 63% (95% CI, 59%-67%). Some questions were answered correctly more often than others—ranging from 17% to 95% (Table 2). Doctors answered 68% of questions correctly (95% CI, 63%-73%), while nurses and other respondents answered 57% of questions correctly (95% CI, 52%-62%) ( $P = 0.003$ ). Other demographic variables—including self-reported involvement in a quality improvement committee and being from the United States versus elsewhere—were not associated with survey performance. The point biserial correlations for each individual question with overall performance were all more than 0.2 (range 0.24–0.62) and all statistically significant at  $P < 0.05$ .

### Concept-Specific Performance

Average percent correct declined across categories as numeracy requirements increased ( $P < 0.05$  for all pairwise comparisons). In the area of basic numeracy, respondents' mean percent correct was 82% (95% CI, 77%-87%) of attempted. This category had 4 questions, with a performance range of 77% to 90%. For example, on the question, "Which hospital has the lowest CLABSI rate?", 80% of respondents answered correctly. For risk-adjustment numeracy, the mean percent correct was 70% (95% CI, 64%-76%); 2 items assessed this category. For "Which is better: a higher or lower SIR?", 95% of the cohort answered correctly. However, on "If hospital B had its number of projected infection halved, what is its SIR?", only 46% of those who attempted the question answered correctly.

Questions featuring risk-adjustment interpretation had an average percent correct of 43% (95% CI, 37%-49%). Five questions made up this category, with a percent correct range of 17% to 75%. For example, on the question, "Which hospital's patients are the most predisposed to developing CLABSI?", only 32% of respondents answered this correctly. In contrast, for the question "Which hospital is most effective at preventing CLABSI?", 51% answered correctly. Figure 2 illustrates the cohort's performance on each conceptual category while Table 2 displays question-by-question results.

### Opinions Regarding CLABSI Data Reliability

Respondents were also asked about their opinion regarding the reliability of CLABSI quality metric data. Forty-three percent of respondents stated that such data were reliable at best 50% of the time. Notably, 10% of respondents indicated that CLABSI quality metric data were rarely or never reliable. There was no association between perceived reliability of quality metric data and survey performance ( $P = 0.87$ ).

## DISCUSSION

This Twitter-based study found wide variation in clinician interpretation of CLABSI quality data, with low overall performance. In particular, comprehension and interpretation of risk-adjusted data were substantially worse than unadjusted data. Although doctors performed somewhat better than nurses and other respondents, those involved in quality improvement initiatives performed no better than respondents who were not. Collectively, these findings suggest clinicians may not reliably comprehend quality metric data, potentially affecting their ability to utilize audit and feedback data. These results may have important implications for policy efforts that seek to leverage quality metric data to improve patient safety.

An integral component of many contemporary quality improvement initiatives is audit and feedback through metrics.<sup>6</sup> Unfortunately, formal audit and feedback, along with other similar methods that benchmark data, have not consistently improved outcomes.<sup>24-27</sup> A recent meta-analysis noted that audit and feedback interventions are not becoming more efficacious over time; the study further asserted that “new trials have provided little new knowledge regarding key effect modifiers.”<sup>9</sup> Our findings suggest that numeracy and comprehension of quality metrics may be important candidate effect modifiers not previously considered. Simply put: we hypothesize that without intrinsic comprehension of data, impetus or insight to change practice might be diminished. In other words, clinicians may be more apt to act on insights they themselves derive from the data than when they are simply told what the data “mean.”

The present study further demonstrates that clinicians do not understand risk-adjusted data as well as raw data. Risk-adjustment has long been recognized as necessary to compare outcomes among hospitals.<sup>28,29</sup> However, risk-adjustment is complex and, by its nature, difficult to understand. Although efforts have focused on improving the statistical reliability of quality metrics, this may represent but one half of the equation. Numeracy and interpretation of the data by decision makers are potentially equally important to effecting change. Because clinicians seem to have

difficulty understanding risk-adjusted data, this deficit may be of growing importance as our risk-adjustment techniques become more sophisticated.

We note that clinicians expressed concerns regarding the reliability of quality metric feedback. These findings corroborate recent research that has reported reservations from hospital leaders concerning quality data.<sup>30,31</sup> However, as shown in the context of patients and healthcare decisions, the aversion associated with quality metrics may be related to incomplete understanding of the data.<sup>32</sup> Whether perceptions of unreliability drive lack of understanding or, conversely, whether lack of understanding fuels perceived unreliability is an important question that requires further study.

This study has several strengths. First, we used rigorous survey development techniques to evaluate the understudied issue of quality metric numeracy. Second, our sample size was sufficient to show statistically significant differences in numeracy and comprehension of CLABSI quality metric data. Third, we leveraged social media to rapidly acquire this sample. Finally, our results provided new insights that may have important implications in the area of quality metrics.

There were also limitations to our study. First, the Twitter-derived sample precludes the calculation of a response rate and may not be representative of individuals engaged in CLABSI prevention. However, respondents were solicited from the Twitter-followers of 2 health services researchers (TJI, VC) who are actively engaged in scholarly activities pertaining to critically ill patients and hospital-acquired complications. Thus, our sample likely represents a highly motivated subset that engages in these topics on a regular basis—potentially making them *more* numerate than average clinicians. Second, we did not ask whether the respondents had previously seen CLABSI data specifically, so we cannot stratify by exposure to such data. Third, this study assessed only CLABSI quality metric data; generalizations regarding numeracy with other metrics should be made with caution. However, as many such data are presented in similar formats, we suspect our findings are applicable to similar audit-and-feedback initiatives.

The findings of this study serve as a stimulus for further inquiry. Research of this nature needs to be carried out in samples drawn from specific, policy-relevant populations (eg, infection control practitioners, bedside nurses, intensive care unit directors). Such studies should include longitudinal assessments of numeracy that attempt to mechanistically examine its impact on CLABSI prevention efforts and outcomes. The latter is an important issue as the link between numeracy and behavioral response, while plausible, cannot be assumed, particularly given the complexity of issues related to behavioral modification.<sup>33</sup> Additionally, whether alternate presentations of quality data affect numeracy, interpretation, and performance is worthy of further testing; indeed, this has been shown to be the case in other forms of communication.<sup>34-37</sup> Until data from larger samples are available, it may be prudent for quality improvement leaders to assess the comprehension of local clinicians regarding

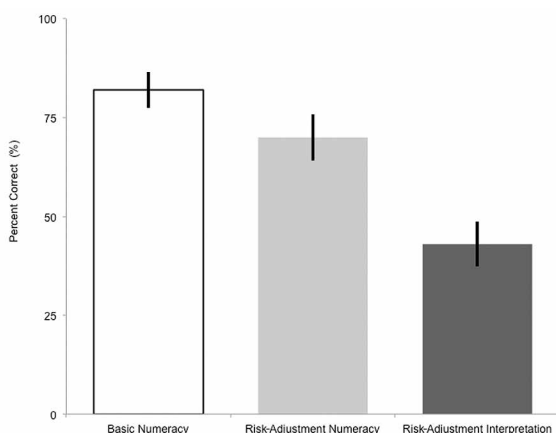


FIG. 2. Performance by conceptual category

feedback and whether lack of adequate comprehension is a barrier to deploying quality improvement interventions.

Quality measurement is a cornerstone of patient safety as it seeks to assess and improve the care delivered at the bedside. Rigorous metric development is important; however, ensuring that decision makers understand complex quality metrics may be equally fundamental. Given the cost of examining quality, elucidating the mechanisms of numeracy and interpretation as decision makers engage with quality metric data is necessary, along with whether improved comprehension leads to behavior change. Such inquiry may provide an evidence-base to shape alterations in quality metric deployment that will ensure maximal efficacy in driving practice change.

Disclosures: This work was supported by VA HSR&D IIR-13-079 (TJI). Dr. Chopra is supported by a career development award from the Agency of Healthcare Research and Quality (1-K08-HS022835-01). The views expressed here are the authors' own and do not necessarily represent the view of the US Government or the Department of Veterans' Affairs. The authors report no conflicts of interest.

## References

1. Scott RD II. The direct medical costs of healthcare-associated infections in US hospitals and the benefits of prevention. Centers for Disease Control and Prevention. Available at: [http://www.cdc.gov/HAI/pdfs/hai/Scott\\_CostPaper.pdf](http://www.cdc.gov/HAI/pdfs/hai/Scott_CostPaper.pdf). Published March 2009. Accessed November 8, 2016.
2. O'Grady NP, Alexander M, Burns LA, et al. Guidelines for the prevention of intravascular catheter-related infections. *Am J Infect Control*. 2011;39(4 suppl 1):S1-S34.
3. Blot K, Bergs J, Vogelaers D, Blot S, Vandijck D. Prevention of central line-associated bloodstream infections through quality improvement interventions: a systematic review and meta-analysis. *Clin Infect Dis*. 2014;59(1):96-105.
4. Mermel LA. Prevention of intravascular catheter-related infections. *Ann Intern Med*. 2000;132(5):391-402.
5. Siempos II, Kopterides P, Tsangaris I, Dimopoulou I, Armaganidis AE. Impact of catheter-related bloodstream infections on the mortality of critically ill patients: a meta-analysis. *Crit Care Med*. 2009;37(7):2283-2289.
6. Casalino LP, Gans D, Weber R, et al. US physician practices spend more than \$15.4 billion annually to report quality measures. *Health Aff (Millwood)*. 2016;35(3):401-406.
7. Hysong SJ. Meta-analysis: audit and feedback features impact effectiveness on care quality. *Med Care*. 2009;47(3):356-363.
8. Ilgen DR, Fisher CD, Taylor MS. Consequences of individual feedback on behavior in organizations. *J Appl Psychol*. 1979;64:349-371.
9. Ivers NM, Grimshaw JM, Jamtvedt G, et al. Growing literature, stagnant science? Systematic review, meta-regression and cumulative analysis of audit and feedback interventions in health care. *J Gen Intern Med*. 2014;29(11):1534-1541.
10. Rao G. Physician numeracy: essential skills for practicing evidence-based medicine. *Fam Med*. 2008;40(5):354-358.
11. Wegwarth O, Schwartz LM, Woloshin S, Gaissmaier W, Gigerenzer G. Do physicians understand cancer screening statistics? A national survey of primary care physicians in the United States. *Ann Intern Med*. 2012;156(5):340-349.
12. Bramwell R, West H, Salmon P. Health professionals' and service users' interpretation of screening test results: experimental study. *BMJ*. 2006;333(7562):284.
13. Agoritsas T, Courvoisier DS, Combescur C, Deom M, Perneger TV. Does prevalence matter to physicians in estimating post-test probability of disease? A randomized trial. *J Gen Intern Med*. 2011;26(4):373-378.
14. Warren DK, Zack JE, Mayfield JL, et al. The effect of an education program on the incidence of central venous catheter-associated bloodstream infection in a medical ICU. *Chest*. 2004;126(5):1612-1618.
15. Rinke ML, Bundy DG, Chen AR, et al. Central line maintenance bundles and CLABSIs in ambulatory oncology patients. *Pediatrics*. 2013;132(5):e1403-e1412.
16. Pronovost P, Needham D, Berenholtz S, et al. An intervention to decrease catheter-related bloodstream infections in the ICU. *N Engl J Med*. 2006;355(26):2725-2732.
17. Rinke ML, Chen AR, Bundy DG, et al. Implementation of a central line maintenance care bundle in hospitalized pediatric oncology patients. *Pediatrics*. 2012;130(4):e996-e1004.
18. Sacks GD, Diggs BS, Hadjizacharia P, Green D, Salim A, Malinoski DJ. Reducing the rate of catheter-associated bloodstream infections in a surgical intensive care unit using the Institute for Healthcare Improvement Central Line Bundle. *Am J Surg*. 2014;207(6):817-823.
19. Berenholtz SM, Pronovost PJ, Lipsitt PA, et al. Eliminating catheter-related bloodstream infections in the intensive care unit. *Crit Care Med*. 2004;32(10):2014-2020.
20. Rajwan YG, Barclay PW, Lee T, Sun IF, Passaretti C, Lehmann H. Visualizing central line-associated blood stream infection (CLABSI) outcome data for decision making by health care consumers and practitioners—an evaluation study. *Online J Public Health Inform*. 2013;5(2):218.
21. Fagerlin A, Zikmund-Fisher BJ, Ubel PA, Jankovic A, Derry HA, Smith DM. Measuring numeracy without a math test: development of the Subjective Numeracy Scale. *Med Decis Making* 2007;27(5):672-680.
22. HAI progress report FAQ. 2016. Available at: <http://www.cdc.gov/hai/surveillance/progress-report/faq.html>. Last updated March 2, 2016. Accessed November 8, 2016.
23. Collins D. Pretesting survey instruments: an overview of cognitive methods. *Qual Life Res*. 2003;12(3):229-238.
24. Ivers N, Jamtvedt G, Flottorp S, et al. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev*. 2012;(6):CD000259.
25. Chatterjee P, Joynt KE. Do cardiology quality measures actually improve patient outcomes? *J Am Heart Assoc*. 2014;3(1):e000404.
26. Joynt KE, Blumenthal DM, Orav EJ, Resnic FS, Jha AK. Association of public reporting for percutaneous coronary intervention with utilization and outcomes among Medicare beneficiaries with acute myocardial infarction. *JAMA*. 2012;308(14):1460-1468.
27. Ryan AM, Nallamothu BK, Dimick JB. Medicare's public reporting initiative on hospital quality had modest or no impact on mortality from three key conditions. *Health Aff (Millwood)*. 2012;31(3):585-592.
28. Thomas JW. Risk adjustment for measuring health care outcomes, 3rd edition. *Int J Qual Health Care*. 2004;16(2):181-182.
29. Iezzoni LI. *Risk Adjustment for Measuring Health Care Outcomes*. Ann Arbor, Michigan: Health Administration Press; 1994.
30. Goff SL, Lagu T, Pekow PS, et al. A qualitative analysis of hospital leaders' opinions about publicly reported measures of health care quality. *Jt Comm J Qual Patient Saf*. 2015;41(4):169-176.
31. Lindenauer PK, Lagu T, Ross JS, et al. Attitudes of hospital leaders toward publicly reported measures of health care quality. *JAMA Intern Med*. 2014;174(12):1904-1911.
32. Peters E, Hibbard J, Slovic P, Dieckmann N. Numeracy skill and the communication, comprehension, and use of risk-benefit information. *Health Aff (Millwood)*. 2007;26(3):741-748.
33. Montano DE, Kasprzyk D. Theory of reasoned action, theory of planned behavior, and the integrated behavioral model. In: Glanz K, Rimer BK, Viswanath K, eds. *Health Behavior and Health Education: Theory, Research and Practice*. 5th ed. San Francisco, CA: Jossey-Bass; 2015:95-124.
34. Hamstra DA, Johnson SB, Daignault S, et al. The impact of numeracy on verbatim knowledge of the longitudinal risk for prostate cancer recurrence following radiation therapy. *Med Decis Making*. 2015;35(1):27-36.
35. Hawley ST, Zikmund-Fisher B, Ubel P, Jancovic A, Lucas T, Fagerlin A. The impact of the format of graphical presentation on health-related knowledge and treatment choices. *Patient Educ Couns*. 2008;73(3):448-455.
36. Zikmund-Fisher BJ, Witteman HO, Dickson M, et al. Blocks, ovals, or people? Icon type affects risk perceptions and recall of pictographs. *Med Decis Making*. 2014;34(4):443-453.
37. Korfae IJ, Fuhrer-Forbis A, Ubel PA, et al. Informed choice about breast cancer prevention: randomized controlled trial of an online decision aid intervention. *Breast Cancer Res*. 2013;15(5):R74.