ORIGINAL ARTICLE

Do Experts Understand Performance Measures? A Mixed-Methods Study of Infection Preventionists

Sushant Govindan, MD;[1,3] Beth Wallace, MD;[1] Theodore J. Iwashyna, MD, PhD;[1,2] Vineet Chopra, MD, MSc[1,2,3]

OBJECTIVE.   Central line-associated bloodstream infection (CLABSI) is associated with significant morbidity and mortality. Despite a nationwide decline in CLABSI rates, individual hospital success in preventing CLABSI is variable. Difficulty in interpreting and applying complex CLABSI metrics may explain this problem. Therefore, we assessed expert interpretation of CLABSI quality data. DESIGN. Cross-sectional survey PARTICIPANTS. Members of the Society for Healthcare Epidemiology of America (SHEA) Research Network (SRN) METHODS. We administered a 10-item test of CLABSI data comprehension. The primary outcome was percent correct of attempted questions pertaining to the CLABSI data. We also assessed expert perceptions of CLABSI reporting.

RESULTS.   The response rate was 51% (n = 67).Among experts, the average proportion of correct responses was 73% (95% confidence interval [CI], 69%–77%). Expert performance on unadjusted data was significantly better than risk-adjusted data (86% [95% CI, 81%–90%] vs 65% [95% CI, 60%–70%]; P < .001). Using a scale of 1 to 100 (0, never reliable; 100, always reliable), experts rated the reliability of CLABSI data as 61. Perceived reliability showed a significant inverse relationship with performance (r = −0.28; P = .03), and as interpretation of data improved, perceptions regarding reliability of those data decreased. Experts identified concerns regarding understanding and applying CLABSI definitions as barriers to care.

CONCLUSIONS.   Significant variability in the interpretation of CLABSI data exists among experts. This finding is likely related to data complexity, particularly with respect to risk-adjusted data. Improvements appear necessary in data sharing and public policy efforts to account for this complexity.

Central line-associated bloodstream infection (CLABSI) is a serious healthcare-associated infection (HAI) that carries an attributable mortality of 12%–25%.[1,2] Each CLABSI episode increases average hospital length of stay by an estimated 14 days and raises costs associated with hospitalization by US$32,000–$45,814.[3–5] Given the clinical and economic burden of CLABSI, it has been the focus of scientific enquiry, public reporting, and national performance measurement. CLABSI is designated a tier 1 preventable infection by the US Department of Health and Human Services, a category denoting the highest prevention priority among HAIs.[6] Recent data suggest that much progress has been made in reducing national rates of CLABSI as a result of myriad efforts.[7]

Despite this progress, success of an individual institution in reducing CLABSI remains highly variable.[8] One reason for this inconsistency may relate to challenges in measuring CLABSI. Multiple studies have reported variation in CLABSI rates based solely on the definition used and measurement practices.[9,10]

With recent definition changes to CLABSI measurement, this variability is likely to persist.[11] Importantly, inconsistency in performance measurement is not limited to CLABSI; rather, it is well described with a number of other quality metrics.[12–16] Unfortunately, the mechanisms for inconsistency in performance measurement, and how these can be improved, are not well understood.

A second explanation for variation in CLABSI reduction is limited comprehension of CLABSI-related data and reports. Indeed, methods of reporting quality metrics such as CLABSI have had little formal cognitive testing.[12] Thus, whether CLABSI data are interpreted correctly and are actionable by end users remains unknown. A recent Twitter-based study of clinicians found that understanding of CLABSI metrics was highly variable, with comprehension decreasing as risk-adjusted data and concepts related to implementing reports were introduced.[17] If CLABSI experts experience similar difficulties, a stronger case to revisit data presentation and reporting may exist.

Therefore, we performed a survey-based study to evaluate (1) how well experts interpret CLABSI performance data and (2) opinions regarding reliability and challenges related to these metrics.

METHODS

The development of a CLABSI comprehension assessment has previously been described in detail.[17] In brief, we utilized a broad, exploratory approach to develop a survey instrument to evaluate comprehension of CLABSI data. The methodology included literature review, survey development, and instrument refinement. Individual items within the assessment were constructed through an iterative process with pilot testing among clinicians. We divided questions into 2 categories: those pertaining to non–risk-adjusted data (eg, unadjusted metrics) and those pertaining to risk-adjusted outcomes (eg, adjusted metrics). For the survey, a contemporary, published CLABSI performance report was adapted from Rajwan et al,[18] who qualitatively validated the format using physician feedback (see Supplement online). The assessment underwent cognitive testing within a cohort of 8 hospitalists and intensivists outside the sampling frame to ensure appropriate internal validity (eg, comprehension, retrieval, and judgment).[19] Feedback from infection prevention experts regarding the content of the assessment was also solicited to ensure external validity. The instrument was refined based on feedback and psychometric analysis to a 10-item assessment of CLABSI data comprehension.

Assessment of CLABSI Interpretation in Experts

The survey was deployed to members of the Society for Healthcare Epidemiology of America (SHEA) Research Network (SRN), an international consortium of more than 100 hospitals dedicated to preventing HAIs. Members of the SRN include infection prevention experts, hospital epidemiologists, and infectious disease physicians. We sent 3 email solicitations to the members for recruitment, and no incentives were provided. SurveyMonkey served as the electronic platform for survey distribution and data collection, and question order was randomized. The primary outcome was the percentage of correctly answered questions. Secondary outcomes included percent correct on questions related to unadjusted data and percent correct on question related to risk-adjusted data. A priori, information was collected regarding respondent education, years in practice and in the SRN, country of residence, and number of hospital beds in the respondent's home institution.

Qualitative Assessment of Expert Opinions

In addition to quantitative assessment, we subjectively evaluated the perceptions of experts regarding CLABSI data. The first question asked, "In your opinion, how reliable are quality metric data about CLABSI?" The scale was 0 to 100 (0, never; 100, always). Experts were also asked, "In your opinion, what are the 3 biggest problems for reliability of quality metric data at your hospital?" Independently, 2 study authors (S.G. and B.W.) reviewed all responses and used inductive analysis to identify key themes gleaned from expert responses.[20] A code book was created to identify themes that were then reconciled and ratified in a series of meetings by both authors. The reviewers coded 10% of the responses and resolved any differences to ensure consistency. The authors independently categorized all responses into 1 of 1 themes, allowing for a primary and secondary theme when deemed necessary. Any disagreements in coding were reconciled by a third author (V.C.). This ensured reliability through comparison of data coding by multiple members of the research team.[21]

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 assessment. The Kolmogorov-Smirnov and Shapiro-Wilk tests were used to evaluate normality of continuous variables. We assessed for associations with demographics via t tests, Pearson correlations, or Spearman correlations. Comprehension of unadjusted data was compared to risk-adjusted data using Wilcoxon signed-rank test. We tested for an association between comprehension and perceived reliability of quality metric data using Pearson correlation coefficient.

Analyses were conducted in STATA MP software version 14.0 (StataCorp, College Station, TX); $P < .05$ was considered statistically significant. No adjustments were made for multiple comparisons. The study was designated exempt by the Institutional Review Board of University of Michigan (no. HUM0012090).

RESULTS

Of 132 eligible participants, 67 SHEA SRN members responded (response rate, 51%), and 54 respondents (41% of total sample) attempted all questions on the assessment. Doctors comprised 81% (n = 48) of the cohort, while the remainder included infection control nurses and epidemiologists (Table 1).

Interpretation of CLABSI Data

Overall, the mean percent correct on the assessment was 73% (95% CI, 69%–77%). No significant association was observed between performance on the survey and respondent demographics such as education, years in practice, and hospital size (Table 1). However, some questions were answered correctly

TABLE 1.   Respondent Characteristics and Associations with Comprehension[a]

| Category | Total | Association With Comprehension[b] | |
		Statistical Data	P Value
Country, No. (%)			
USA	43 (75)	Mean, 75	.36
Other	14 (25)	Mean, 71	
Profession, No. (%)			
Doctor:	48 (81)	Mean, 75	.31
Other	11 (19)	Mean, 70	
Years with SRN, mean (SD)[b]	4 (3)	Rho = −0.07	.62
Years practicing, mean (SD)	17 (11)	R = −0.04	.79
No. of beds in home institution, median (IQR)	400 (212–800)	Rho = 0.19	.15
Perceived reliability, mean (SD)	61 (21)	R = −0.28	.03

NOTE. SRN, Society for Healthcare Epidemiology of America (SHEA) Research Network; SE, standard deviation.
[a]N = 67; not all respondents answered each demographic question.
[b]The 2-sample t test was used to measure associations regarding both country and profession. Continuous variables underwent testing to assess for normal distribution via Kolmogorov-Smirnov and Shapiro-Wilk tests. Based on these results, Spearman's Rho or Pearson's correlation assessed for an association with comprehension.

TABLE 2.   Complete List of Questions with Respondent Performance

Question	Cohort % Correct
Unadjusted Data Questions	
If hospital G's number of actual infections doubled, what would its CLABSI rate be?	92
Which hospital has the lowest CLABSI rate?	90
Which hospital uses the most central lines?	90
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?	73
Risk-Adjusted Data Questions	
Which hospital is most effective at preventing CLABSI?	85
Which hospital's patients are the most predisposed to developing CLABSI?	83
If hospital B had its number of projected infections halved, what is its SIR?	70
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?	69
Suppose hospitals A and H have the exact same CLABSI prevention practices. Which hospital will have the higher number of CLABSI?	48
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?	38

more often than others—with a range of 38% to 92% for correct answers (Table 2).

When evaluating responses to questions that presented unadjusted data, mean percent correct was 86% (95% CI, 81%–90%). There were 4 questions in this category, with a performance range of 73%–92%. For example, when presented with data regarding CLABSI incidence and asked, "Which hospital has the lowest CLABSI rate?" 90% of respondents answered correctly. Conversely, when asked, "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?" only 73% of respondents answered correctly.

For questions related to interpretation of risk-adjusted data, the average percent correct was significantly lower than for unadjusted data (65%; 95% CI, 60%–70%; P < .001) (Figure 1). The range of percent correct responses for risk-adjusted data varied from 38% to as high as 85%. For example, when asked, "If hospital B had its number of projected infection halved, what is its SIR?" 70% of respondents answered correctly. However, when asked, "Suppose hospitals A and H have the exact same CLABSI prevention practices. Which hospital will have the higher number of CLABSI?" only 48% of respondents answered correctly. Responses and explanations for each question can be found in the Supplement online.
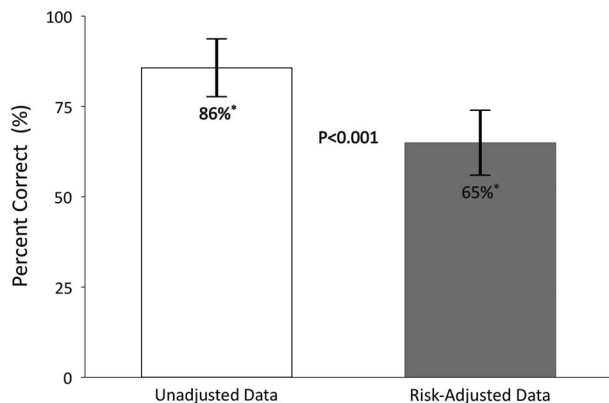
FIGURE 1. Performance of experts on risk-adjusted versus unadjusted CLABSI data. Questions are divided into 2 categories: Unadjusted data (those pertaining to non-risk-adjusted data, eg, unadjusted metrics) and risk-adjusted data (those pertaining to risk-adjusted data, eg, adjusted metrics). CLABSI, 73% (95% CI, 69%–77%); unadjusted, 86% (95% CI, 81%–90%); risk adjusted, 65% (95% CI, 60%–70%).
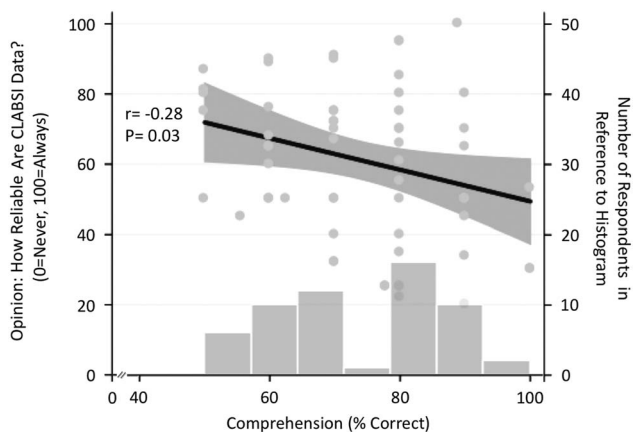

FIGURE 2. Comprehension of CLABSI data is associated with perceptions on how reliable CLABSI data are. A histogram of comprehension is also shown. Respondents were asked: "In your opinion, how reliable are quality metric data about CLABSI?" 95% confidence intervals are shown in grey.

Perceptions Regarding CLABSI Data

When asked, "In your opinion, how reliable are quality metric data about CLABSI?" the average score was 61 (0, never; 100, always; 95% CI, 56–67). While no respondent demographics were statistically associated with this response, perceived reliability showed a significant, inverse relationship with comprehension (r = −0.28 P = .03). That is, the better a respondent appeared to comprehend the data, the less likely they thought the data were reliable (Figure 2).

A total of 131 responses to the question, "In your opinion, what are the 3 biggest problems for reliability of quality metric

data at your hospital?" were provided. Overall, 51 experts provided at least 1 response, 45 experts provided 2 responses, and 35 experts provided 3 responses. A full listing of the responses can be found in the Supplement online. Each response was categorized into 8 themes related to concerns about quality metric data (Table 3). Experts listed definition defects as the most common theme. For example, many respondents expressed "difficulty with applying surveillance definitions" as a main problem. In order of frequency, the next 2 recurring themes were information bias, such as errors in counting central line days, and concerns regarding risk-adjustment, including "inadequate risk-adjustment" or "risk-stratification."

DISCUSSION

This survey of SRN members found significant variability in how infection control experts performed when asked to evaluate and interpret CLABSI metrics. As in the case of non-experts,[17] expert respondents in this study also performed significantly lower when presented with risk-adjusted data versus unadjusted data. While no measured demographic factors were associated with performance, we noted that perceived reliability of CLABSI metrics was inversely correlated with performance. Experts also expressed concerns regarding the reliability of quality metric data, principal among them difficulties in understanding and applying definitions to cases, as well as concerns regarding adequacy or transparency of risk adjustment. These findings have important implications for public policy efforts that seek to leverage quality metric data to preventing HAIs such as CLABSI.

While substantial progress in CLABSI has been made,[7] thousands of cases still occur in the United States each year.[22] Despite being one of the most costly HAIs,[5] between 65% and 70% of cases of CLABSI are potentially preventable.[23] In an effort to move the needle, CLABSI outcomes continue to be publicly reported, linked to performance measurement, and tied to reimbursement.[24] However, inconsistent interpretation of data within an expert population should cause concern about this paradigm. The inherent complexity of how the data are shared, presented, and interpreted might weaken the usefulness of CLABSI reporting in its current state. Supportively, the top 3 categories of expert concerns, accounting for >82% of all responses, highlighted complexities and/or deficiencies of the data fed back to hospitals. Future studies are needed to define which aspects of quality metric data are most challenging and how these may be mitigated. With ongoing changes to definitions and recalibration of what may or may not "count" as an infection,[11,25] this issue should be prioritized among policy makers.

Given that performance measures like CLABSI are increasing in influence,[12] experts should not ignore the challenge associated with summarizing, sharing, and acting on data. Data collection and processing is perhaps only half the battle; data delivery and messaging must also be considered. In this context, important lessons from the patient decision-making literature are relevant.[26,27] The method of delivering data to

TABLE 3.   Concerns for Experts Regarding Quality Metric Data: Themes and Example Responses[a]

Theme	No.	%	Sample Response
1. Definition defects	45	34	"Difficulty with applying surveillance definitions"
2. Information bias	37	28	"Training of data collector"
3. Concerns regarding risk-adjustment	26	20	"Lack of risk-adjustment for comorbid conditions"
4. Gaming	7	5	"Predicted infections based on gamed national data"
5. Insufficient resources	5	4	"Limited staffing"
6. Inability to draw conclusions	4	3	"Interpretation of the metrics from non-epidemiology trained leadership…"
7. Ineffective interventions	2	2	"Reasonable QI measures may or may not affect rates."
8. Other	5	4	

[a]Total no. of responses, 131.

patients has been shown to impact both objective decision-making and subjective perceptions.[28–30] For instance, Halvorsen et al.[29] showed a 24% difference in patient consent rates to receive a cardiovascular medication based only on altering how the data are displayed. Hawley et al[30] demonstrated that data presentation impacts patient knowledge of clinical information, and that increased knowledge was associated with making a superior treatment choice. It is plausible, then, that similar effects may also apply to stakeholders within the performance measurement world.[12] Future studies to test different modalities of data delivery and messaging related to performance measurement are needed.

Our study has several limitations. First, this is a survey-based project limited to a select group of expert providers; thus, generalizability to other quality measures, conditions, or providers may be limited. As with all surveys, nonrespondent bias may have influenced our findings. However, our response rate is comparable to other surveys within this sample of medical experts.[31,32] Second, we are unable to assess whether differences in interpretation are related to the survey instrument and how data were shown or to factors such as provider knowledge or experience. As the instrument underwent rigorous development, pilot testing, dissemination to other physicians and input from experts, we believe our results reflect clinician uncertainty when interpreting these data.[17] Finally, we could not evaluate whether interpretation or subjective perceptions of CLABSI data affect downstream outcomes, such as CLABSI rates. Only future studies that measure these outcomes can assess this perspective.

However, our study has several strengths. First, we drew a sample from a population of infection prevention experts. Although we expected experts to consistently interpret CLABSI data across a range of questions, we did not find this to be the case. These findings are not only relevant not only to CLABSI but also to other quality improvement initiatives where deficiencies in data interpretation and application may blunt impact. Second, we employed both quantitative and qualitative methods to assess performance and impressions related to CLABSI reporting. This approach allows for a deeper perspective in defining barriers to understanding CLABSI data.[33] Finally, our results are novel and indicate that cognitive decision making might be highly relevant in the quest to make performance measures more effective.

Future studies that test different methods to deliver these data and outcomes appear necessary.

Performance measurement is a fundamental component of quality improvement. Without effective performance feedback, stakeholders are flying blind as they combat HAIs such as CLABSI.[34,35] However, these measures have grown increasingly complex because of the efforts to standardize performance via risk adjustment.[36] It is now unclear whether these painstakingly collected measures actually provide actionable feedback for decision makers.[12,37] It is vital that policy experts understand how to account for data complexity in the delivery of quality measures as feedback agents. Without attention to such change, performance measures may remain unreliable or ineffective in motivating clinicians to change practice.

Address correspondence to Sushant Govindan, MD, Taubman Center FL 3 Rm 3920, 1500 E Medical Center Dr, SPC 5360, Ann Arbor, MI 48109 (sushantg@med.umich.edu).

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/ice.2017.243

REFERENCES

1. Scott RD. *The Direct Medical Costs of Healthcare-Associated Infections in US Hospitals and the Benefits of Prevention*. Atlanta, GA: Centers for Disease Control and Prevention; 2009.
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. Chopra V KS, Olmsted RN, Safdar N, Saint S. Chapter 10: Prevention of central line-associated bloodstream infections: brief update review. In UoCSF-SE-BP, eds *Making Health Care Safer II: an Updated Critical Analysis of Patient Safety Practices*. Rockville, MD: Agency for Healthcare Research and Quality, U.S. Dept. of Health and Human Services; 2013:88–110.

4. Stevens V, Geiger K, Concannon C, Nelson RE, Brown J, Dumyati G. Inpatient costs, mortality and 30-day re-admission in patients with central-line-associated bloodstream infections. *Clin Microbiol Infect* 2014;20:O318–O324.

5. Zimlichman E, Henderson D, Tamir O, et al. Health care-associated infections: a meta-analysis of costs and financial impact on the US health care system. *JAMA Intern Med* 2013;173:2039–2046.

6. *Action Plan to Prevent Healthcare-associated Infections*. Washington, DC: US Department of Health and Human Services; 2009.

7. *National and State Healthcare-Associated Infections Progress Report*. Atlanta, GA: Centers for Disease Control and Prevention; 2015.

8. Render ML, Hasselbeck R, Freyberg RW, et al. Reduction of central line infections in Veterans Administration intensive care units: an observational cohort using a central infrastructure to support learning and improvement. *BMJ Qual Saf* 2011;20:725–732.

9. Cherifi S, Mascart G, Dediste A, et al. Variations in catheter-related bloodstream infections rates based on local practices. *Antimicrob Resist Infect Control* 2013;2:10.

10. Worth LJ, Brett J, Bull AL, McBryde ES, Russo PL, Richards MJ. Impact of revising the National Nosocomial Infection Surveillance System definition for catheter-related bloodstream infection in ICU: reproducibility of the National Healthcare Safety Network case definition in an Australian cohort of infection control professionals. *Am J Infect Control* 2009;37:643–648.

11. Corley A, Cantara M, Gardner J, Trexler P, Rock C, Maragakis LL. Central line-associated bloodstream infection rate elevation: attributable to National Healthcare Safety Network surveillance definition changes, ongoing opportunities for infection prevention, or both? *Am J Infect Control* 2017;45:1030–1032.

12. Austin JM, McGlynn EA, Pronovost PJ. Fostering transparency in outcomes, quality, safety, and costs. *JAMA* 2016;316:1661–1662.

13. Epstein AM. Will pay for performance improve quality of care? The answer is in the details. *N Engl J Med* 2012;367:1852–1853.

14. Ivers N, Jamtvedt G, Flottorp S, et al. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database Systemat Rev* 2012;6:CD000259.

15. 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;29:1534–1541.

16. Ryan A, Blustein J. Making the best of hospital pay for performance. *N Engl J Med* 2012;366:1557–1559.

17. Govindan S, Chopra V, Iwashyna TJ. Do clinicians understand quality metric data? An evaluation in a Twitter-derived sample. *J Hosp Med* 2017;12:18–22.

18. 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 Pub Health Informat* 2013;5:218.

19. Collins D. Pretesting survey instruments: an overview of cognitive methods. *Qual Life Res* 2003;12:229–238.

20. Boyatzis RE. *Transforming Qualitative Information: Thematic Analysis and Code Development*. London: SAGE; 1998.

21. Silverman D. *Interpreting Qualitative Data: Methods for Analyzing Talk, Text and Interaction*. 3rd ed. London: SAGE; 2006.

22. Centers for Disease Control and Prevention. Vital signs: central line-associated blood stream infections—United States, 2001, 2008, and 2009. *Morb Mortal Wkly Rep* 2011;60:243–248.

23. Umscheid CA, Mitchell MD, Doshi JA, Agarwal R, Williams K, Brennan PJ. Estimating the proportion of healthcare-associated infections that are reasonably preventable and the related mortality and costs. *Infect Control Hosp Epidemiol* 2011;32:101–114.

24. Pronovost PJ, Goeschel CA, Wachter RM. The wisdom and justice of not paying for "preventable complications. *JAMA* 2008;299:2197–2199.

25. National Health Safety Network. *NHSN Newsletter: Patient Safety*. vol. 9, 2014. Atlanta, GA: Centers for Disease Control and Prevention.

26. Fagerlin A, Zikmund-Fisher BJ, Ubel PA. Helping patients decide: ten steps to better risk communication. *J Natl Cancer Inst* 2011;103:1436–1443.

27. Hesse BW, Beckjord E, Rutten LJ, Fagerlin A, Cameron LD. Cancer communication and informatics research across the cancer continuum. *Am Psychol* 2015;70:198–210.

28. Fagerlin A, Wang C, Ubel PA. Reducing the influence of anecdotal reasoning on people's health care decisions: is a picture worth a thousand statistics? *Med Decis Making* 2005;25:398–405.

29. Halvorsen PA, Selmer R, Kristiansen IS. Different ways to describe the benefits of risk-reducing treatments: a randomized trial. *Ann Intern Med* 2007;146:848–856.

30. 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;73:448–455.

31. Drees M, Pineles L, Harris AD, Morgan DJ. Variation in definitions and isolation procedures for multidrug-resistant gram-negative bacteria: a survey of the Society for Healthcare Epidemiology of America Research Network. *Infect Control Hosp Epidemiol* 2014;35:362–366.

32. Masnick M, Morgan DJ, Wright MO, et al. Survey of infection prevention informatics use and practitioner satisfaction in US hospitals. *Infect Control Hosp Epidemiol* 2014;35:891–893.

33. Creswell JW KA, Plano Clark VL, Smith KC. Best practices for mixed methods research in the health sciences. *National Institutes of Health* 2011:2094–2103.

34. Purbey S MK, Bha C. Performance measurement system for healthcare processes. *Int J Productiv Perform Manage* 2007;56:241–251.

35. Alvero AM, Bucklin BR, Austin J. An objective review of the effectiveness and essential characteristics of performance feedback in organizational settings (1985–1998). *J Organ Behav Manage* 2001;21:3–29.

36. Loeb JM. The current state of performance measurement in health care. *Int J Qual Health Care* 2004;16(Suppl 1):i5–i9.

37. Chassin MR, Loeb JM, Schmaltz SP, Wachter RM. Accountability measures—using measurement to promote quality improvement. *N Engl J Med* 2010;363:683–688.