

How Patient Comments Affect Consumers' Use of Physician Performance Measures

David E. Kanouse, PhD,* Mark Schlesinger, PhD,† Dale Shaller, MPA,‡
Steven C. Martino, PhD,§ and Lise Rybowski, MBA||

Background: Patients' comments about doctors are increasingly available on the internet. The effects of these anecdotal accounts on consumers' engagement with reports on doctor quality, use of more statistically reliable performance measures, and ability to choose doctors wisely are unknown.

Objective: To examine the effects of providing patient comments along with standardized performance information in a web-based public report.

Design: Participants were randomly assigned to view 1 of 6 versions of a website presenting comparative performance information on fictitious primary care doctors. Versions varied by the combination of information types [Consumer Assessment of Healthcare Providers and Systems (CAHPS), Healthcare Effectiveness Data and Information Set (HEDIS), and patient comments] and number of doctors.

Participants: A random sample of working-age adults (N = 848) from an online panel representing the noninstitutionalized population of the United States.

Main Measures: Time spent and actions taken on the website, probing of standardized measures, and decision quality (chosen doctor rated highest on quantifiable metrics, chosen doctor not dominated by another choice). Secondary outcomes were perceived usefulness and trustworthiness of performance metrics and evaluations of the website.

Key Results: Inclusion of patient comments increased time spent on the website by 35%–42% and actions taken (clicks) by 106%–117% compared with versions presenting only CAHPS and HEDIS measures ($P < 0.01$). It also reduced participants' attention to standardized measures (eg, percentage of time probing HEDIS measures dropped by 67%, $P < 0.01$). When patient comments were present, fewer participants chose the doctor scoring highest on standardized metrics (44%–49% vs. 61%–62%, $P < 0.01$).

Conclusions: Including patient comments in physician performance reports enhances consumers' engagement but reduces their attention to standardized measures and substantially increases suboptimal choices. More research is needed to explore whether integrated reporting strategies could leverage the positive effects of patient comments on consumer engagement without undermining consumers' use of other important metrics for informing choice among doctors.

Key Words: performance measurement, patient engagement, patient satisfaction

(*Med Care* 2015;00: 000–000)

Websites and published reports presenting information on the performance of doctors and medical practices have become increasingly common.^{1,2} To help consumers make more informed decisions when selecting health care providers, these reports typically include standardized measures of patient experience, such as those derived from the Consumer Assessment of Healthcare Providers and Systems (CAHPS) surveys, and clinical process measures, such as those contained in the Healthcare Effectiveness Data and Information Set (HEDIS).³ These kinds of measures will soon be incorporated into the Physician Compare website mandated by the Patient Protection and Affordable Care Act of 2010 to support consumer choice of physicians.⁴

At the same time, websites conveying patients' anecdotal comments about their experiences with health care providers have proliferated.⁵ The increasing availability of patients' comments online may pose challenges to the use of more systematically gathered and standardized data on patients' encounters with medical care. Patients' comments often cover the same experiential domains as surveys,⁶ but in ways that may be easier to understand, more engaging, and more persuasive to consumers than statistically summarized information from a larger, more representative sample of patients.^{7–9}

To date, patients' comments have appeared largely in contexts that do not also present standardized performance metrics such as CAHPS or HEDIS. Indeed, internet searches are more likely to lead consumers to private websites that contain anecdotal information from patients than to government or community-sponsored websites that contain quantitative information.¹⁰ However, some have suggested that patient comments be added to sites that also convey standardized information, including a few proposals for the new Physician Compare website.^{11–14}

From the *RAND Corporation, Santa Monica, CA; †Yale University School of Public Health, New Haven, CT; ‡Shaller Consulting Group, Stillwater, MN; §RAND Corporation, Pittsburgh, PA; and ||The Severn Group, Ashburn, VA.

This study was funded by cooperative agreements U18HS016978 and U18HS016980 with the Agency for Healthcare Research and Quality. We thank Mike Cui of RAND for research assistance and Debra Dean of Westat for logistics support. Preliminary analyses of some of the data reported in this paper were presented at the 2012 Agency for Healthcare Research and Quality annual conference.

The authors declare no conflict of interest.

Reprints: David E. Kanouse, PhD, RAND Health, 1776 Main Street, P.O. Box 2107, Santa Monica, CA 90407-2138. E-mail: kanouse@rand.org.

Copyright © 2015 Wolters Kluwer Health, Inc. All rights reserved.
ISSN: 0025-7079/15/000-000

As of 2012, fewer than 1 in 5 internet users reported having consulted online reviews of doctors or other providers.¹⁵ Across delivery modes and over time, consumers' use of performance metrics in selecting providers has been uncommon.^{16–18} This may partly reflect limitations in content, design, and implementation of reports that have been developed rather than indicating inherent limitations of report cards.^{19–21} Inclusion of patient comments could potentially help reduce the effect of some barriers, such as low numeracy, that have limited use. However, a potential drawback is that comments represent a much smaller sample of patients' experiences than are summarized in CAHPS scores, and are therefore less likely to be representative. Yet because people are willing to generalize from small amounts of data, they may fail to give sufficient weight to the quantitative information.²²

It is therefore important to gain a scientific understanding of whether patient comments have the potential to convey information about patient experience that consumers find important and engaging, and how this information influences the health care decisions of patients and those acting on their behalf. It is equally important to understand potential drawbacks of patient comments, such as whether they tend to displace thoughtful consideration of more objective and representative data.

In this article, we present the results of an experiment that assesses the effects of including patients' comments along with standardized performance information in a web-based public report. The experiment explores: (1) the impact of patients' comments on consumers' engagement with the information on the website and their choice of doctors, and (2) how the effects of comments vary with the cognitive burdens of choice (breadth of performance metrics presented and number of clinicians available).

METHODS

We designed an experiment in which a random sample of working-age adults with access to the internet were directed to a fictitious website containing comparative information on primary care doctors. Participants were randomly assigned to 1 of 6 versions of the website that varied by type of information presented and number of doctors from which they could choose. Their task was to review information on the site and select a preferred doctor; they then completed an online survey about their experience on the site.

Sample

Participants were recruited randomly from KnowledgePanel, a survey panel of about 50,000 members developed and maintained by Knowledge Networks (now GfK). This panel was constructed using a combination of random digit dialing and address-based sampling to represent the noninstitutionalized US population, including households with unlisted phone numbers, cellphone-only households, and nontelephone households.²³ We restricted eligibility to those who access the internet through computer as this group is more likely to search online for information about doctors. Of the 1757 panel members aged 25–64 who were invited to

participate, 48.3% provided informed consent. The research protocol was approved by the relevant institutional review boards.

The Select MD Website

We designed a website called SelectMD to display comparative information on the performance of fictitious doctors.²⁴ The website was designed to replicate basic content, presentation, functionality, and navigation features commonly found in contemporary web-based reports. The website sponsor was described as a nonprofit consumer group that is a trusted source for health information.²⁵

A "Performance Overview" page presented summary scores for "Service Quality" (CAHPS) and "Technical Quality" (HEDIS) for a set of either 12 or 24 clinicians. Summary scores were presented as 1–5 stars, where 3 stars was average for clinicians in that community. Physician performance within option sets varied over the full range. Participants were able to sort doctors by level of performance, "filter" based on sex or years of experience, and "drill down" to view the component measures underlying each summary score. A "scroll over" function allowed participants to learn more about how performance measures were defined. Selected experimental arms also contained patients' comments that were similar to those available on real-world websites. A tracking system (invisible to participants) recorded every click made by participants and the time spent on each page.

Experimental Design

We randomly assigned participants to 6 experimental arms that included different kinds and combinations of performance measures (Table 1). We assessed the implications of choice set size by presenting participants with information on either 12 or 24 doctors. In the 4 arms in which participants could choose among 12 doctors, the presence or absence of patients' comments was crossed with the complexity of standardized performance metrics (CAHPS alone vs. CAHPS combined with HEDIS measures). In the 2 remaining experimental arms, participants could choose among 24 doctors and were shown either CAHPS information combined with patients' comments or CAHPS and HEDIS information combined with patients' comments.

Developing Realistic Patients' Comments

We modeled the comments in SelectMD on actual patients' comments collected from websites reporting on physicians in Georgia, Missouri, New Jersey, and Oregon. On the basis of these real comments, we constructed a set of fictitious comments that contained between 1 and 3 statements, with each statement conveying 1 of 4 aspects of patient experience: doctor communication, access to needed tests or treatments, demonstration of care or concern, and courtesy/respect shown by office staff. Comments containing >1 statement could address multiple aspects of patient experience. Consistent with the comments harvested from websites, the comments in SelectMD were relatively brief, ranging from 10 to 75 words in length (mean = 37 words).

To convey a specific affective tone, each statement mixed emotionally neutral words with adjectives or adverbs

TABLE 1. Experimental Design

Arm	N	Performance Measures	Size of Choice Set
1	129	CAHPS	12
2	125	CAHPS+HEDIS	12
3	152	CAHPS+patient comments	12
4	142	CAHPS+HEDIS+patient comments	12
5	155	CAHPS+patient comments	24
6	146	CAHPS+HEDIS+patient comments	24

CAHPS indicates Consumer Assessment of Healthcare Providers and Systems; HEDIS, Healthcare Effectiveness Data and Information Set.

documented to have a clear emotional valence.²⁶ We combined statements to create 160 comments with overall emotional valences that were strongly negative (20%), mildly negative (30%), mildly positive (30%), or strongly positive (20%). To assess the perceived informativeness, authenticity, and emotional valence of these comments, we conducted 3 rounds of pilot testing.²⁷ Comments judged by pilot subjects to be inauthentic or discordant from their intended emotional valence were discarded. Examples of patient comments are shown in the Appendix.

Assignment of Patient Comments to Select MD Doctors

Each doctor on the SelectMD website was assigned a “profile” of 4–6 fictitious patient comments. The modal valence of their comment profile was matched to their CAHPS score (number of stars), so that clinicians with higher CAHPS scores were assigned more positive comments. However, each clinician’s comment profile included at least 1 comment that ran counter to the modal emotional tone for that profile; thus, even a profile for a highly (poorly) rated clinician contained at least 1 comment that was weakly negative (positive). The specific comments assigned to each clinician were randomly drawn from the pool of appropriately valenced comments each time a participant logged onto the website.

Measures

Participant reactions were assessed in a postexposure survey. Other outcomes were constructed from tracking data or participants’ observed choices.

Engagement With the Website

We used tracking data to measure the participants’ amount of time and number of actions (clicks) on the website in general and probing of standardized measures specifically. Actions that could be taken included changing screens, highlighting clinicians, and applying filters (eg, by years of experience) to change the set of clinicians displayed.

Reactions to the Website

Participants were asked how easy or difficult it was to use the site and how satisfied they were with the choice of doctors available, both assessed using close-ended categorical response scales.

Perceptions of Information

For each of the 3 types of information (CAHPS, HEDIS, patient comments), participants were asked, using close-ended categorical response scales, how useful that information was in helping them select a doctor and how trustworthy they considered that information.

Decision Quality

We measured decision quality in 2 ways: first, we assessed whether the selected doctor was rated highest on CAHPS (for arms 1, 3, and 5) or rated highest for the combination of CAHPS plus HEDIS scores (for arms 2, 4, and 6), and if not rated the highest, then whether the selected doctor rated second best, or worse than second best, among the remaining options. Second, in arms 2, 4, and 6, we assessed whether the participant chose a doctor who was a dominated choice—that is, at least 1 other doctor scored as well or better than the chosen doctor on 1 standardized measure (CAHPS or HEDIS) and better than the chosen doctor on the other standardized measure. Considering only the standardized metrics, dominated choices represent poor decisions because the participant could have chosen better no matter what relative value the participant placed on CAHPS versus HEDIS measures.

Statistical Methods

Our primary analyses involved simple cross-arm comparisons to support causal inference making. Because measures of time spent on the site were positively skewed, we used log transformation to normalize distributions before conducting significance tests.²⁸ Because the valence of patient comments was correlated with the CAHPS measure, it was not included as a predictor in the model.

RESULTS

Sociodemographics of the Study Sample

As Table 2 shows, restricting study eligibility to those with internet access resulted in a slight skew in the characteristics of study participants as compared with the US working-age population: study participants were slightly older, less likely to be Hispanic, and more likely to report at least 1 doctor visit but less likely to report 10 or more visits.

Perceptions of Performance Measures

Participants’ perceptions of the 3 types of performance information presented on the SelectMD website did not vary significantly across experimental conditions, so we report combined results for everyone exposed to each type (Table 3). Over two thirds of participants reported that it was “very” or “somewhat” easy to use CAHPS, HEDIS, and patient comments to pick a doctor. CAHPS was seen as somewhat easier to use than HEDIS ($P < 0.01$) or patient comments ($P < 0.01$). Each type of performance information was modally described as only “somewhat” trustworthy, with no marked differences across type.

TABLE 2. Sample Characteristics Compared With US Working-Age Population

Characteristics	Sample [N (%)]	US Working-Age Population (%)*
Sex (male)	392 (46.2)	49.3
Age (y)		
25–33	160 (18.9)	23.1
34–42	182 (21.4)	22.4
43–51	226 (26.6)	24.5
52–64	281 (33.1)	30.0
Race/ethnicity		
White, non-Hispanic	627 (73.9)	66.8
Black, non-Hispanic	92 (10.8)	11.7
Other, non-Hispanic	33 (3.9)	5.8
Hispanic	74 (8.7)	14.6
2 or more races, non-Hispanic	23 (2.7)	1.1
Education		
Less than high school	76 (9.0)	11.0
High school	254 (29.9)	30.0
Some college	243 (28.6)	27.3
College and beyond	276 (32.5)	31.7
Census region		
Northeast	139 (16.4)	18.1
Midwest	202 (23.8)	21.6
South	333 (39.2)	36.8
West	175 (20.6)	23.4
Doctor visits in past 12 mo		
None	148 (17.5)	21.1
1	159 (18.8)	16.7
2–3	252 (29.7)	26.5
4–9	217 (25.6)	22.5
10 or more	72 (8.5)	13.2

*Characteristics for the US working-age population (ages 25–64) are from the 2010 Current Population Survey,²⁹ except for number of doctor visits, which is from the National Health Interview Survey (2010).³⁰

Effects of Patient Comments

Table 4 shows the effects of including patients' comments on the website along with CAHPS (arms 3 and 5) or both CAHPS and HEDIS (arms 4 and 6). The relevant comparisons are across blocks (arms 3 and 5 vs. arm 1; arms 4 and 6 vs. arm 2). Comparisons involving arms 5 and 6 include the combined effect of patients' comments and a larger set of available doctors.

User Engagement

When patients' comments appeared on the website, participants spent a third more time on the site and performed

more than twice as many actions (differences statistically significant when logged to account for a long right-hand tail in the distribution). This increased engagement primarily involved exploring the content of the comments. It did not reflect either more drilling down to the components of CAHPS or HEDIS measures ("probing" reported in Table 4) or greater filtering of the physician choice set (results not reported). In fact, respondents who were given the opportunity to view patients' comments spent less time probing for detail on CAHPS and HEDIS measures (difference statistically significant only for arms 3 and 1), and they consistently spent a smaller proportion of time on the website probing for detail on these standardized measures (differences statistically significant for all comparisons except between arms 5 and 1 (Table 4).

Perceived Usefulness

The inclusion of patients' comments on the website did not significantly alter participants' perceptions of the overall usefulness of the site or their satisfaction with the choices they had among clinicians. Nor did it have any significant effects on participants' ratings of the usefulness or trustworthiness of either CAHPS or HEDIS measures.

Quality of Choice

In each of the 4 arms that included patients' comments, the percentage of participants who chose the best clinician available (ie, with the highest average score on standardized metrics) was significantly lower than in the corresponding arm without patients' comments. A similar finding emerged for selection of a lower performing clinician: the percentages of participants making a suboptimal choice more than doubled in arms 4 and 6 relative to arm 2. The addition of a larger choice set in arm 6 resulted in an even greater proportion of suboptimal choices compared with the smaller choice set in arm 4 ($P < 0.05$).

Table 5 shows that among participants who viewed both CAHPS and HEDIS scores, the presence of patient comments (arms 4 and 6) often led to choices that were worse than second best on standardized performance metrics; this was especially the case in arm 6, where there were 24 physicians to consider.

TABLE 3. Perceived Ease of Use and Trustworthiness of Standardized Performance Measures and Patient Comments

	Easy to Use to Pick Best Doctor (%)			Trustworthiness (%)		
	Very	Somewhat	Very + Somewhat	Very	Somewhat	Very + Somewhat
CAHPS	44.2	34.4	78.6	13.5	69.6	83.1
HEDIS	33.0	34.9	67.9**	15.1	70.2	85.3
Comments	32.4	38.8	71.3**	17.4	64.2	81.6

Entries are the percentages of all participants who judged each type of measure to be very or somewhat easy to use/trustworthy. Data were collected following exposure to the SelectMD website. Results were pooled across all experimental conditions where the measure was present (arms 1–6 for CAHPS, arms 2, 4, and 6 for HEDIS, arms 3, 4, 5, and 6 for comments).

CAHPS indicates Consumer Assessment of Healthcare Providers and Systems; HEDIS, Healthcare Effectiveness Data and Information Set.

**Statistically significant difference from the CAHPS measure ($P < 0.01$).

TABLE 4. Engagement, Perceived Usefulness of Comparative Performance Information, and Quality of Choices When Presented With and Without Patient Comments

Experimental Measures	Experimental Arm					
	Without Comments (12 Doctors)		With Comments (12 Doctors)		With Comments (24 Doctors)	
	Arm 1 (CAHPS)	Arm 2 (CAHPS + HEDIS)	Arm 3 (CAHPS)	Arm 4 (CAHPS + HEDIS)	Arm 5 (CAHPS)	Arm 6 (CAHPS + HEDIS)
Time spent on website [s (ln s)]	311.4 (5.45)	345.8 (5.52)	421.5 (5.72)**	489.6 (5.77)*	514.7 (5.87)**	517.7 (5.72)***
No. actions on website	3.6	5.2	7.8*	10.7*	9.0*	9.1*
Time spent probing standardized measures [s (ln s)]	42.5 (4.14)	60.6 (4.36)	22.6 (3.56)**	45.5 (4.21)	30.6 (3.83)	50.8 (4.10)
Actions probing standardized measures	0.7	1.4	0.6	1.1	0.7	1.2
Percentage of time on website probing CAHPS	11.8	9.0	5.4**	5.8	9.4	5.5
Percentage of time on website probing HEDIS	—	10.1	—	3.3**	—	5.8**
Usefulness of website [†]	78.3	74.4	Perceived usefulness			
Satisfaction with choice of clinicians [‡]	67.4	69.6	80.9	81.0	83.2	77.4
Usefulness of CAHPS [‡]	89.9	85.6	72.4	67.6	77.4	71.9
Usefulness of HEDIS [‡]	—	89.6	84.2	86.6	87.1	84.9
Trustworthiness of CAHPS [‡]	82.2	83.2	—	88.0	—	84.2
Trustworthiness of HEDIS [‡]	—	84.8	80.3	82.4	83.0	84.9
Selected best clinician [§]	61.2	61.6	—	83.8	—	86.3
Selected lower performing (dominated) clinician	—	17.6	49.3*	43.7*	34.6*	37.0*
			—	37.3*	—	50.0*

[†]Entries are the percentages of participants who report that they definitely or probably would use.[‡]Entries are the percentages of participants who respond very or somewhat satisfied/trustworthy/useful.[§]Best = highest average score on standardized performance metrics.

CAHPS indicates Consumer Assessment of Healthcare Providers and Systems; HEDIS, Healthcare Effectiveness Data and Information Set.

*Statistically significant difference from the parallel arm (1 or 2) without patient comments ($P < 0.05$).**Statistically significant difference from the parallel arm (1 or 2) without patient comments ($P < 0.01$).***Statistically significant difference from the parallel arm (1 or 2) without patient comments ($P < 0.001$).

TABLE 5. Clinician Choices, Compared with the Highest Scoring Star Rating in the Choice Set

Arm in Experiment	Percent Selecting		
	Top Star Rated	Second Best	Worse Than Second Best
Arms with CAHPS Only			
Arm 1			
No comments, 12 MDs	61	24	15
Arm 3			
Comments, 12 MDs	49	34	17
Arm 5			
Comments, 24 MDs	35	50	15
Arms with CAHPS+HEDIS			
Arm 2			
No comments, 12 MDs	62	30	8
Arm 4			
Comments, 12 MDs	43	35	22
Arm 6			
Comments, 24 MDs	37	12	51

CAHPS indicates Consumer Assessment of Healthcare Providers and Systems; HEDIS, Healthcare Effectiveness Data and Information Set.

DISCUSSION

Our findings reveal a paradox in providing consumers with comments from patients. On the one hand, these comments galvanize consumers' attention and increase their engagement (time and extent of interaction) with reports on clinician performance. On the other hand, their inclusion dramatically reduces consumers' attention to standardized performance metrics and substantially increases the likelihood of selecting doctors who perform less well on those measures. In the most complex choice sets, lower performing clinicians represent half of all selected clinicians, which is nearly 3 times the level found for the simple choice sets without patient comments. The finding that the presence of patient comments leads to especially large reductions in choice quality in more complex choice sets could reflect cognitive overload, or it could be because patient comments are less likely to be congruent with HEDIS scores than with CAHPS scores, thereby forcing consumers to make greater tradeoffs. Our data do not illuminate this question, which is deserving of further research.

These findings should be interpreted in light of certain methodological considerations. We studied choices among primary care doctors; the emotional richness of patients' comments may make them more salient for these choices, for which trust and caring are vital,³¹ than they would be for choices among hospitals or health plans.^{32,33} Although participants in this study chose among doctors in the realistic setting of their own homes, their choices were hypothetical. This may have limited participants' engagement, although as this was equally true across all experimental arms, it should not distort the comparisons presented here. Although there is some evidence that this type of stated choice experiment can yield results in other domains that are similar to real-world decisions,^{34,35} few validation studies have been performed, and generalizability of results may be affected by such things as similarity of the experimental environment to real-world choice environments and whether participants are included who would not be making such choices in the real world.³⁶ This study was conducted online, and results may differ for performance

information consumers receive in print. Finally, the study population was limited to working-age Americans with internet access, so findings may not generalize to older people or those who have little experience using the internet. Nevertheless, our sample likely represents a broad segment of the consumers most likely to encounter patient comments on the web.^{37,38}

When consumers are faced with more complex information than they have the capacity or willingness to process, they may respond by reducing the amount of information they consider,^{24,39–41} which can adversely affect decision quality.^{42–44} In this case, they may simply have paid more attention to the comments. The inclusion of patients' comments did not reduce participants' assessments of the usefulness or trustworthiness of CAHPS or HEDIS measures. Nonetheless, they relied less on those measures in choosing among clinicians, perhaps because their attention had been redirected or they had difficulty integrating information from patient narratives with standardized metrics.⁴⁴

Our findings cannot explain why integration proves difficult. But it is demonstrably not simply because consumers were overloaded by too much information. Were that the case, one would expect a similar decline in decision quality when HEDIS is added to CAHPS, as it also represents additional data—and is viewed by consumers as equally easy to process. However, as can be seen by comparing the difference between arms 1 and 3 with the difference between arms 1 and 2 (Table 4), the quality of selection declines in the first case but not the second. Why comments disrupt people's use of standardized measures—and how those effects might be ameliorated—are essential questions for subsequent research.

Patient comments often capture and convey patients' experiences in ways that other consumers find informative and useful, complementing standardized metrics. Comments may also help consumers envision what the differences in numeric scores on patient experience or performance measures might mean experientially, thereby enhancing the affective salience of this information. However, when the presence of comments curtails attention to standardized performance measures so that those measures are poorly

understood or incompletely considered, consumers' choices and their understanding of physician quality are compromised. The contemporary proliferation of patient comments thus poses a real threat to the infrastructure of standardized performance reports constructed by public and private sponsors over the past 15 years.^{1,16} Concerns about the impact of unfavorable patient comments on physicians' reputations have led to calls for suppressing (through legal means) the dissemination of patient comments or otherwise "inoculating" consumers against their purportedly pernicious influence.⁴⁵ Our findings could be seen as providing yet 1 more reason for sponsors of public reports on health care quality to eschew patient comments.

We favor a more cautiously constructive approach for several reasons. The suppression of patient narratives is neither feasible nor morally acceptable. Efforts to limit the diffusion of information over the internet are likely to fail. Moreover, sites presenting patients' comments are proliferating precisely because they serve a need that current public reports with standardized measures of clinical encounters do not adequately address. Websites populated with patients' comments convey what many consumers value most about other patients' experiences: an understanding of what rendered those experiences positive or negative, a feel for the emotional content of encounters, and insights into what the consumer can expect from a particular clinician. Understanding these aspects of patient experience is vital for many consumers; understanding this importance to consumers is equally essential for report sponsors and researchers.

We believe that patient comments can play a vital and distinctive role in helping consumers understand and assess health care if they can be made more representative of patient experience than are the haphazardly volunteered comments currently available online. It is therefore incumbent on those who seek to empower medical consumers to explore new ways of eliciting narratives in a systematic way from a representative set of patients. It is equally important to develop and test new ways of reporting patient narratives and standardized performance metrics in an integrated manner so that they complement rather than substitute for one another. This could be done in several ways. Patient comments could be used to help define the meaning of standardized metrics, for example, by illustrating how specific CAHPS component measures connect with patients' concrete experiences. In addition, comments could help consumers decide how much to value the difference between 4 stars and 5 stars and thus how willingly they would trade-off this difference for other aspects of quality or other attributes of clinicians' practices (eg, cost, accessibility).

Including narratives that help consumers understand and augment the experiences captured more abstractly in standardized performance measures could increase the currently limited use of quality information in decision making.¹⁷ Ultimately, quality reports will be valued by consumers only insofar as they describe meaningful clinical experiences and clarify medical choices. The substantial impact of patient comments—for both better and worse—highlights a crucial gap in our knowledge regarding public reporting, one that should be rectified to make the US health care system more responsive to patients.⁴⁶

APPENDIX

TABLE A1. Example Comments

Strong positive

Very caring about his patients and interested in getting to know them. Great office staff as well!

Takes his time with every patient (which explains the wait) but is worth it. Leaves no stone left unturned. Very easygoing.

I had a real bad allergy attack last spring. They worked me right into the schedule. The doctor asked my lots of questions about my allergies, referred me to an allergist and gave me sample meds that have worked wonders.

Mild positive

Thought we'd be in the waiting room forever. But got in to see doctor pretty fast. Receptionist kept apologizing. Asking us if we needed anything. Made me not mind as much. Would probably go back.

The doctor came when my dad was brought into the emergency room. She stayed a while and consoled us. Later she called to check on him.

I started having these spells last year I think they was hot flashes. Dr G took some tests told me all the various ways to treat them. He was cool that I didn't want to take hormones right away. Even suggested different herbs I could try.

Mild negative

I can't say anything bad, but I can't say anything good either. She was punctual. She listened somewhat. I almost felt like I got out everything I needed to say. Follow-up and after-care was very slow and difficult.

Dr S's office people could be a little more accommodating, you know, try a little harder to give people appointments on the same day if they're sick. Or at least the next. Dr S is kind of snobby and cold but he's supposed to be competent.

I went to Dr T for my allergies. This new nurse had tons of perfume on. I sneezed the whole time I waited. I complained to the doctor but he just ignored me. Just asked why I still had my cat.

Strong negative

Very patronizing; ordered a bunch of lab tests, but simply mailed me the results and refused to discuss them with me. Left me alone without a clue and feeling pretty mad about it.

Office Staff always losing charts. Then they make excuses for it. Doctor hurried. He spends too little time with patients. Mostly runs in and out. It's hard to get appointments when you need them.

Wanted to see him that day for a bad allergy attack. Only sees patients a few days a week. Wasn't too concerned when I finally got an appointment. Kind of laughed me off. People in the office aren't much better. Not overly friendly.

REFERENCES

1. Consumer Health Ratings. com: your guide to online healthcare ratings. Available at: <http://www.consumerhealthratings.com>. Accessed March 17, 2015.
2. Schlesinger M, Grob R, Shaller D, et al. Taking patients' narratives about clinicians from anecdotes to science. *New Engl J Med*. 2015;373:675–679.
3. Agency for Healthcare Research and Quality. Public Reports on Provider Performance for Consumers. Available at: <http://www.ahrq.gov/professionals/quality-patient-safety/quality-resources/tools/public-reporting/index.html>. Accessed March 17, 2015.
4. Sinaiko AD, Eastman D, Rosenthal MB. How report cards on physicians, physician groups, and hospitals can have greater impact on consumer choices. *Health Aff*. 2012;31:602–611.
5. Lagu T, Hannon NS, Rothberg NB, et al. Patients' evaluations of health care providers in the era of social networking: an analysis of physician-rating websites. *J Gen Intern Med*. 2010;25:942–946.
6. López A, Detz A, Ratanawongsa N, et al. What patients say about their doctors online: a qualitative content analysis. *J Gen Intern Med*. 2012; 27:685–692.
7. Huppertz JW, Carlson JP. Consumers' use of HCAHPS ratings and word-of-mouth in hospital choice. *Health Serv Rev*. 2010;45: 1602–1613.

8. Finucane ML, Alhakami A, Slovic P, et al. The affect heuristic in judgments of risks and benefits. *J Behav Decis Making*. 2000;13:1–17.
9. Satterfield T, Slovic P, Gregory R. Narrative valuation in a policy judgment context. *Ecol Econ*. 2000;34:315–331.
10. Marshall M, McLoughlin V. How do patients use information on providers? *BMJ*. 2010;341:1255–1257.
11. Sick B, Abraham JM. Seek and ye shall find: consumer search for objective health care cost and quality information. *Am J Med Quality*. 2011;26:433–440.
12. Hibbard JH, Peters E. Supporting informed consumer health care decisions: data presentation approaches that facilitate the use of information in choice. *Annu Rev Public Health*. 2003;24:413–433.
13. Lagu T, Lindenauer PK. Putting the public back in public reporting of health care quality. *JAMA*. 2010;304:1711–1712.
14. Findlay S, Lansky D. Numbers to crunch. Better public data needed to help patients compare and choose physicians. *Mod Healthc*. 2011;41:26.
15. Fox S, Duggan M. Information triage. Pew Research Center Internet and American Life Project. 2013. Available at: <http://www.pewinternet.org/2013/01/15/information-triage/>. Accessed August 6, 2015.
16. Christianson JB, Volmar KM, Alexander J, et al. A report card on provider report cards: current status of the health care transparency movement. *J Gen Intern Med*. 2010;25:1235–1241.
17. Faber M, Bosch M, Wollersheim H, et al. Public reporting in health care: how do consumers use quality-of-care information? a systematic review. *Med Care*. 2009;47:1–8.
18. Kolstad JT, Chernew ME. Quality and consumer decision making in the market for health insurance and health care services. *Med Care Res Rev*. 2009;66(suppl):28S–52S.
19. Hibbard JH. What can we say about the impact of public reporting? Inconsistent execution yields variable results. *Ann Intern Med*. 2008;148:160–161.
20. Hussey PS, Luft HS, McNamara P. Public reporting of provider performance at a crossroads in the United States: summary of current barriers and recommendations on how to move forward. *Med Care Res Rev*. 2014;71(5 suppl):5S–16S.
21. Robinowitz DL, Dudley RA. Public reporting of provider performance: can its impact be made greater? *Annu Rev Public Health*. 2006;27:517–536.
22. Tversky A, Kahneman D. Belief in the law of small numbers. *Psychol Bull*. 1971;76:105–110.
23. Chang L, Krosnick JA. National surveys via RDD telephone interviewing versus the internet: comparing sample representativeness and response quality. *Public Opin Quart*. 2009;73:641–678.
24. Schlesinger M, Kanouse DE, Rybowski L, et al. Consumer response to patient experience measures in complex information environments. *Med Care*. 2012;50(Suppl):S56–S64.
25. Blendon R, Brodie M, Benson J, et al. *American Public Opinion and Health Care*. Washington, DC: CQ Press; 2010.
26. Bradley MM, Lang PJ. Affective norms for English words (ANEW): instruction manual and affective ratings. Technical report C-1. Center for Research in Psychophysiology, University of Florida. Available at: <http://www.uvm.edu/~pdodds/teaching/courses/2009-08UVM-300/docs/others/everything/bradley1999a.pdf>. Accessed March 17, 2015.
27. Martino SC. Consumers' appraisal of anecdotal accounts of patient experience. Presented at AHRQ 2012 Annual Conference, September 9–11, 2012, Bethesda, MD. Session 69.
28. Von Hippel PT. Normalization. In: Lewis-Beck M, Bryman A, Liao TF, eds. *Encyclopedia of Social Science Research Methods*. Thousand Oaks, CA: Sage; 2004:746–748.
29. Data extracted from the US Census Bureau, Current Population Survey, 2010 Annual Social and Economic Supplement. Available at <http://www.census.gov/cps/data/>. Accessed April 11, 2014.
30. Schiller JS, Lucas JW, Ward BW, et al. Summary health statistics for U.S. adults: National Health Interview Survey, 2010. *Vital Health Stat*. 2012;(252):1–207.
31. Mechanic D, Schlesinger M. The impact of managed care on patients' trust in medical care and their physicians. *JAMA*. 1996;275:1693–1697.
32. Kleimann Communication Group and Consumers Union. Choice architecture: design decisions that affect consumers' health plan choices. Rockville, MD, 2012. Available at: http://consumersunion.org/pdf/Choice_Architecture_Report.pdf. Accessed March 17, 2015.
33. Kaiser Family Foundation. 2008 Update on Consumers' Views of Patient Safety and Quality Information. Available at: <http://kaiserfamilyfoundation.files.wordpress.com/2013/01/7819.pdf>. Accessed March 17, 2015.
34. Burke RR, Harlam BA, Kahn BE, et al. Comparing dynamic consumer choice in real and computer-simulated environments. *J Consumer Res*. 1992;19:71–82.
35. Carson R, Louviere JJ, Anderson D, et al. Experimental analysis of choice. *Market Lett*. 1994;5:351–367.
36. Dhar R, Simonson I. The effect of forced choice on choice. *J Market Res*. 2003;40:146–160.
37. Gao GG, McCullough JS, Agarwal R, et al. A changing landscape of physician quality reporting: analysis of patients' online ratings of their physicians over a 5-year period. *J Med Internet Res*. 2012;14:e38.
38. Mata R, Nunes L. When less is enough: cognitive aging, information search, and decision quality in consumer choice. *Psychol Aging*. 2010;25:289–298.
39. Spranca MD, Elliott MN, Shaw R, et al. Disenrollment information and Medicare plan choice: Is more information better? *Health Care Financ Rev*. 2007;28:47–59.
40. Payne JW, Bettman JR, Schkade DA. Measuring constructed preferences: towards a building code. *J Risk Uncertainty*. 1999;19:243–270.
41. Finucane ML, Mertz CK, Slovic P, et al. Task complexity and older adults' decision-making competence. *Psychol Aging*. 2005;20:71–84.
42. Lee BK, Lee WN. The effect of information overload on consumer choice quality in an on-line environment. *Psychol Market*. 2004;21:159–183.
43. Hwang MI, Lin JW. Information dimension, information overload and decision quality. *J Inform Sci*. 1999;25:213–218.
44. Schlesinger M, Kanouse DE, Martino SC, et al. Complexity, public reporting, and choice of doctors: a look inside the blackest box of consumer behavior. *Med Care Res Rev*. 2014;71(suppl):38S–64S.
45. McCartney M. Will doctor rating sites improve the quality of care? No. *BMJ*. 2009;338:b1033.
46. Grob R, Schlesinger M. Epilogue: principles for engaging patients in US health care and policy. In: Hoffman B, Tones N, Grob R, Schlesinger M, eds. *Patients as Policy Actors*. New Brunswick, NJ: Rutgers University Press; 2011:278–291.