Early Performance Trends After the Public Posting of Ambulatory Patient Satisfaction Reviews

Paige G Wickner, MD, MPH[1,2,3], Christian Dankers, MD, MBA[1,3], Melanie Green, MHA[4], Hojjat Salmasian, MD, MPH, PhD[1,3], and Allen Kachalia, MD, JD[5,6]

Keywords
patient feedback, patient satisfaction, clinician–patient relationship, survey data

Introduction

Patients are increasingly turning to online physician reviews to guide their choice of physicians (1). In parallel, health-care organizations have started publicly posting physician reviews, as alternatives to independent review sites. In 2012, University of Utah Health began posting ambulatory patient reviews to foster transparency and trust with patients, provide clinician performance feedback, and demonstrate an institutional patient-centered focus (2,3). In 2016, for the same reasons, Brigham and Women's Hospital (BWH) started publicly posting ambulatory patient-experience reviews. As more health systems embark on this endeavor, there remains little published data about rating trends once reviews are publicly posted. In this report, we seek to share early data on physician ratings and reviews after a transition to public facing ambulatory patient experience comments. Our aims were to (a) determine whether ratings improved once they went public and (b) determine whether patients providing higher or lower ratings were more likely to leave comments.

Methods

Brigham and Women's Hospital's process takes star ratings and comments from 10 provider questions on our ambulatory survey and posts them online within the individual physician's hospital directory listing. We also post the comment posting and screening guidelines so that the public knows how the process works. All patient comments about their physician visits are posted unless they meet 1 of 4 limited criteria: offensive language, inflammatory or potentially libelous material, protected health information, or mentions other providers, trainees, or non-physician staff. Brigham and Women's Hospital utilizes the Press Ganey Ambulatory

eSurvey® to collect patient feedback after outpatient encounters. Press Ganey uses a star rating system from 1 to 5 for each question, where 1 is the lowest rating and 5 is the highest for all of their surveys. Scores from 10 questions about the physician are averaged into an overall star rating for the provider (4). We analyzed the data from surveys encompassing encounters from August 2012 until December 2017. This included data of 44 specialties (of which 22 went live with public comments during the study period), 1544 distinct providers, and 128,083 distinct encounters. We analyzed the data descriptively and by comparing the means in different groups using Wilcox rank sum test. For each specialty, the 1-year period before public comments was compared to the period after the specialty became public facing. Given that specialties went publicly facing at different times (i.e. a staged roll out), the "post" period ranged from 6 to 16 months based on the speciality. We also assessed the trends over time (from 2012 to 2017) using a general linear regression model. In the linear regression model, the star rating was the independent variable and the

[1] Harvard Medical School, Boston, MA, USA
[2] Division of Rheumatology, Allergy, and Immunology, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA
[3] Division of General Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA
[4] Press Ganey Associates, Inc. South Bend, IN, USA
[5] Division of General Medicine, Department of Medicine, Johns Hopkins School of Medicine, Baltimore, MD, USA
[6] Armstrong Institute for Patient Safety and Quality, Johns Hopkins Medicine, Baltimore, MD, USA

Corresponding Author:
Allen Kachalia, Johns Hopkins Medicine, 600 N. Wolfe Street, CMSC 131, Baltimore, MD 21287, USA.
Email: kachalia@jhu.edu

Table 1. Comparison of Ratings Before and After Public Posting of Patient Comments and Ratings, by Specialty.

Specialty	Public Ratings Period (months)	Mean Difference in Star Ratings	P Value	Number of Providers	Number of Surveys
Allergy and immunology	10	0.034	.044	19	1861
Anesthesia and pain management	6	0.034	.183	16	1593
Cardiac surgery	7	0.079	.562	6	220
Cardiovascular medicine	10	0.022	.047	86	7695
Dermatology	10	0.043	<.001	44	16505
Endocrine surgery	7	0.069	.411	2	118
Endocrinology, diabetes, and hypertension	10	0.045	.009	38	3236
Foot and ankle surgery	7	0.037	.042	4	994
Gastrointestinal and general surgery	7	0.011	.062	23	2721
Infectious disease	10	0.003	.081	28	862
Neurology	16	0.044	<.001	88	7683
Neurosurgery	16	0.029	.213	21	1928
Obstetrics/gynecology	10	0.019	.775	62	6701
Orthopedic surgery	16	0.025	.001	33	7502
Plastic surgery	7	0.036	.010	15	2312
Primary care	7	0.037	<.001	145	30815
Psychiatry	8	0.047	.035	42	1231
Renal disease	8	0.012	.155	33	1171
Rheumatology	10	0.004	.051	33	4889
Sleep medicine	10	0.181	.005	5	346
Thoracic surgery	7	0.046	.010	17	1380
Vascular and endovascular surgery	7	0.137	<.001	7	570

encounter date and the study arm (pre vs post) were the predictors, and specialty and time were used as dummy variables.

Results

In the 1 year before public display of reviews at BWH, our average physician rating was 4.77 of 5. The first 4 specialties posted their physician ratings publicly in July of 2016; in 2017, 18 additional specialties made their reviews public facing in a staged fashion. Comparing the average hospital-wide ratings before and after the inception of the program, we saw a small, but statistically significant upward trend (from 4.77 to 4.80, P value < .0001). A smaller, but also statistically significant, upward trend was also observed for those specialties that were not publicly displayed (on average, scores increased by a factor of 1.00033 each year, P value < .0001), and a similar baseline trend was observed for those specialties that chose to display the comments publicly. Overall, the public posting of reviews was associated with a statistically significant rise in the improvement trends (odds ratio = 1.03, P value < .0001) when compared to specialties that did not post publicly. Every specialty that went public had a higher average overall rating afterward (Table 1) with 13 of 22 having statistically significant changes.

We also evaluated whether dissatisfied patients were more likely to leave reviews, an action that could negatively skew their physicians' online profiles. There were 40,093 five-star ratings, with 5128 (12.79%) associated comments and 132 1-star ratings with 10 associated

Table 2. Physician Ratings and Comments Provided by Patients.[a]

Star Rating	Total Ratings	Ratings With Comment (%)	Percentage of Overall Comments
5	40 093	5128 (12.79%)	94.3%
4	4841	206 (4.26%)	3.78%
3	938	62 (6.61%)	1.14%
2	365	31 (8.49%)	0.57%
1	132	10 (7.58%)	0.18%

[a]As a note, a 1.99 was treated as a 1 star and a 4.99 was a 4 star; 1, lowest rating and 5, highest rating.

comments (7.58%; Table 2). Patients who assigned 5 stars to their encounter were significantly more likely to leave comments compared to those who assigned a 1-4 star rating (P value < .0001; Table 2). In addition, the overall number of patients who gave a high rating (4 or 5) far outweighed the number of patients who gave a low star (1 or 2) rating, both before and after the public display of the data.

Discussion

To our knowledge, this is the first analysis reporting trends in specialty-specific star rating comparisons with a transition to public facing ratings and comments. We also found that positive ratings far exceeded negative ratings, and patients who give a 5 star rating provided comments significantly more often than those giving a lower star rating.

While our early experience has given us important and high-level information on the nature of the data and pace of

change, other questions—which are harder to answer—remain. Will providers use these comments for improvement, and if so, how? It has been suggested that anonymous patient feedback cannot improve physician performance due to its lack of specific context for the provider and may in fact be falsely reassuring or inaccurately alarming (5). Our oversight group has indeed debated how best, at a practice and individual level, to interpret "subjective" responses. However, as a result of our process, we have identified individual physicians with patient experience comment and score patterns that have led to personalized training efforts. In the future, we hope to develop improved tools for providers with lower scores and automatic methods of rating and comment analysis to alert us when a provider is receiving concerning patterns of stars and comments.

Although we believe that these data are important to transparently share with patients, how patients should best utilize this information remains unknown. The interpretation of potentially conflicting closed question ratings and open-ended narratives and their contribution to clinician score variation is critical to understand (6). Equally important is how we then help educate patients on optimal use.

In addition to the issues identified above, there are other limitations to our analysis. We have 1 year of unadjusted data, and the longer term effects remain unknown. One of the challenges in analyzing rating data is that for physicians who are part time clinicians or practice in low-volume specialties, there may not be sufficient data for pattern observation. We did not separately account for physicians who were either in only the pre- or the post-group. We also did not have the ability to account for other simultaneous departmental or divisional patient experience improvements in our analysis that may have affected our scores. It is likely that the small numerical difference we see before versus after public-facing comment transition is statistically significant due to the large sample size. It remains unclear whether seeing a physician who has a rating of 4.75 of 5, ensure a noticeably different patient experience than a 4.8. However, we saw the improvement across specialties with different sample sizes, suggesting that the change is not exclusively attributable to randomness. Physician awareness and education of questions being asked on the survey and engagement in the comment process may likely have contributed to the positive score improvement seen in our data.

We are encouraged by the positive ratings and comments that we believe helps highlight the quality and skill of our physicians. Though the scores have only improved since the inception of the program, there remain areas for improvement. How to pair this practice with meaningful quality measurements and how best to educate patients on interpretation merit further investigation. Receiving negative feedback, even in the context of an overall positive star rating and a majority of positive comments, can be deflating. The contribution of provider satisfaction ratings on provider perceptions, wellness, and practice with complex interactions also need further serious consideration (7–9). We remain optimistic that institutions can meet these challenges and create a transparent system that benefits patients and physicians alike.

Authors' Note

At the time of this analysis and initial submission, Allen Kachalia was at Brigham and Women's Hospital in Boston, MA. On December 1, 2018, he started a new position at Johns Hopkins Medicine in Baltimore, MD. Melanie Green has changed positions to work at Press Ganey Associates, Inc. At the time the data was extracted for analysis, she was in the Department of Analytics, Planning and Process Improvement, Brigham and Women's Hospital, Boston, MA.

Declaration of Conflicting Interests

Funding

References

1. Burkle CM, Keegan MT. Popularity of internet physician rating sites and their apparent influence on patients' choices of physicians. BMC Health Serv Res. 2015;15:416.
2. Lee V. Transparency and trust—online patient reviews of physicians. N Engl J Med. 2017;376:197-9.
3. Lee VS, Miller T, Daniels C, Paine M, Gresh B, Betz AL. Creating the exceptional patient experience in one academia health system. Acad Med. 2016;91:338-44.
4. Press Ganey: Star ratings calculation methodology. 2015. Retrieved from: http://www.pressganey.com/docs/default-source/default-document-library/star_calculation_methodology.pdf?sfvrsn=0. Accessed July 2018.
5. McCartney M. Will doctor rating sites improve the quality of care? No. BMJ. 2009;338:b1033.
6. Martino SC, Shaller D, Schlesinger M, Parker AM, Rybowski L, Grob R, et al. CAHPS and comments: how closed-ended survey questions and narrative accounts interact in the assessment of patient experience. J Patient Exp. 2017;4:37-45.
7. Zgierska A, Rabago D, Miller MM. Impact of patient satisfaction ratings on physicians and clinical care. Patient Prefer Adherence. 2014;8:437-46.
8. Baines R, Regan de Bere S, Stevens S, Read J, Marshall M, Lalani M, et al. The impact of patient feedback on the medical performance of qualified doctors: a systematic review. BMC Med Educ. 2018;18:173.
9. Holiday AM, Kachalia A, Meyer GS, Sequist TD. Physician and patient views on public physician rating websites: a cross sectional study. J Gen Intern Med. 2017;31:626-31.

Author Biographies

Paige G Wickner is a medical director for the Department of Quality and Safety and the Executive Director for the Center for Patients and Families at Brigham and Women's Hospital. She is an

allergist and immunologist at Brigham and Women's Hospital, and has research and clinical interests in patient experience, quality and safety and drug allergies.

Christian Dankers is a practicing hospitalist at Brigham and Women's Hospital and worked on patient experience initiatives as the Associate Chief Quality Officer while at Brigham and Women's Hospital. He now serves as the Associate Chief Quality Officer at Partners Healthcare.

Melanie Green is an advisor with Press Ganey Associates. Melanie is responsible for providing guidance, training, leadership and expertise in the arena of the patient experience. In her Advisor role at Press Ganey, she serves various hospitals and systems primarily in the New England Region.

Hojjat Salmasian is a physician specialized in clinical informatics, and a researcher in health care quality and safety and patient experience. He currently serves as the Medical Director for Data Science and Analytics at the Brigham and Women's Hospital.

Allen Kachalia is Senior Vice President of Patient Safety and Quality and Director of the Armstrong Insitute for Patient Safety and Quality at Johns Hopkins Medicine. He is an academic hospitalist in the Division of General Medicine at the Johns Hopkins School of Medicine in Baltimore, MD.