Surgery 170 (2021) 764-768



Contents lists available at ScienceDirect

Surgery

journal homepage: www.elsevier.com/locate/surg

Presented at the Academic Surgical Congress 2020

Crowd-sourced hospital ratings are correlated with patient satisfaction but not surgical safety



Liam T. Synan, BS, MS^{a,b}, Mark A. Eid, MD^{a,c,d}, Casey R. Lamb, MD^a, Sandra L. Wong, MD, MS^{a,b,*}

^a Department of Surgery, Dartmouth-Hitchcock Medical Center and Geisel School of Medicine at Dartmouth, Lebanon, NH

^b Geisel School of Medicine at Dartmouth, Hanover, NH

^c VA Quality Scholars Program, White River Junction VAMC, White River Junction, VT

^d VA Outcomes Group, White River Junction VAMC, White River Junction, VT

ARTICLE INFO

Article history: Accepted 13 April 2021 Available online 20 May 2021

ABSTRACT

Background: Hospital reviews posted online by patients are unsolicited and less structured than Hospital Consumer Assessment of Healthcare Providers and Systems surveys. The differences between online review platforms and their degrees of correlation with validated satisfaction and safety measures are unknown.

Methods: We identified 515 large acute care teaching hospitals in the United States. We collected patient satisfaction results and postsurgical patient safety indicators from Hospital Compare. We also collected hospital star ratings (1-5) from Facebook, Google, and Yelp. Mean ratings were compared with paired *t* tests. Concordance between ratings websites, Hospital Consumer Assessment of Healthcare Providers and Systems scores, and surgical safety indicators were assessed with Pearson's correlation coefficient. *Results:* Mean Facebook ratings (3.81, interquartile range 3.5-4.3) were more favorable than Google (3.26, interquartile range 2.8-2.9). Facebook ratings were least strongly correlated with the Hospital Consumer Assessment of Healthcare Providers and Systems recommended hospital score ($\rho = 0.356$). Google was modestly correlated ($\rho = 0.479$), and Yelp was most strongly correlated ($\rho = 0.500$). The negative correlation between crowdsourced rating and composite safety indicator was too small to be meaningful on any platform.

Conclusion: There is variation between platforms in consumer ratings of hospitals. Ratings on Facebook are more favorable than Google or Yelp. These are independently correlated with Hospital Consumer Assessment of Healthcare Providers and Systems scores. These findings suggest that unstructured consumer reviews generally reflect similar directionality as Hospital Consumer Assessment of Healthcare Providers and Systems scores. Users should be aware of the significant difference between platforms. Consumer ratings platforms are not consistently correlated with postsurgical patient safety indicators, so online ratings may not reflect the safety of surgical care received.

© 2021 Elsevier Inc. All rights reserved.

Introduction

Crowd-sourced platforms (CSPs) such as Facebook, Google Reviews, and Yelp enable patients to offer open-ended and unsolicited reviews of hospitals. In contrast, responses to the structured Hospital Consumer Assessment of Healthcare Providers and Systems

https://doi.org/10.1016/j.surg.2021.04.011 0039-6060/© 2021 Elsevier Inc. All rights reserved. (HCAHPS) survey are solicited and adjusted for patient mix.^{1–3} However, patient awareness of structured comparative hospital information remains low.⁴ Patients may be unaware of structured hospital "report cards" because they are often poorly promoted.⁵ By contrast, CSPs are heavily promoted and are integrated into the daily lives of many users. As a result, web traffic on CSPs is significantly greater than on medicare.gov, the HCAHPS data host.^{6–9}

Therefore, it is relatively easy for patients to consider ratings from CSPs when they decide between clinicians and facilities.^{10–12} National studies have previously demonstrated correlation between unstructured hospital ratings on individual CSPs and HCAHPS scores.^{13–15} However, it is unclear whether those ratings are correlated with structured measures of surgical safety such as

SURGERY

Presented at the 16th Annual Academic Surgical Congress (ASC), Virtual Congress, February 2–4, 2021.

^{*} Reprint requests: Sandra L. Wong, MD, MS, Department of Surgery, Dartmouth-Hitchcock Medical Center, 1 Medical Center Drive, Lebanon, NH 03756.

E-mail address: Sandra.L.Wong@hitchcock.org (S.L. Wong); Twitter: @sandralwong

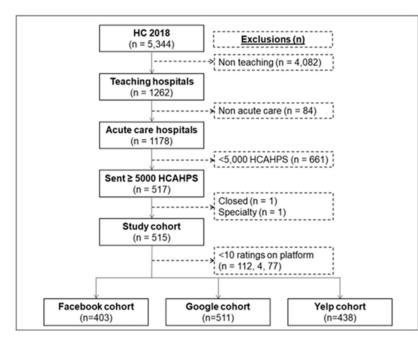


Fig. 1. Cohort selection. Five hundred and fifteen large, acute care teaching hospitals were included in this analysis. HC, hospital compare; HCAHPS, Hospital Consumer Assessment of Healthcare Providers and Systems.

the Agency for Healthcare Research and Quality Patient Safety Indicators (PSIs). Further, the relative strength of correlation between unstructured hospital ratings on different CSPs, structured satisfaction, and safety measures has not been evaluated in a head-tohead analysis.

We set out to determine whether there are significant differences between 3 CSPs (Facebook, Google, and Yelp) and whether ratings on these platforms are correlated with surgical safety. To accomplish this, we assessed mean ratings on CSPs and the strength of correlation between those ratings and structured HCAHPS scores. We also assessed the strength of correlation between hospitals ratings on each platform and a suite of structured surgical safety metrics for the same cohort.

Data and Methods

Data sources

We performed an ecological study using structured satisfaction and safety data from the Centers for Medicare and Medicaid Service Hospital Compare (HC) 2018 annual files and unstructured hospital ratings collected manually from Facebook, Google, and Yelp. We extracted HCAHPS patient satisfaction data and PSIs from the HC 2018 archive.¹⁶ We selected the HCAHPS "Would you recommend this hospital to your friends and family?" question because it most closely matches the stated purpose of ratings on CSPs. We also extracted the rates of surgical safety events at each hospital from the HC 2018 archive.¹⁶ A complete list of extracted measure IDs are shown in the data sources supplement (see Supplementary Materials).

We collected patient ratings for each hospital directly from their Facebook, Google, and Yelp pages. We collected the rating (1-5 scale, reported to 10th decimal place) and number of individuals contributing to that rating from each of the 3 platforms. In cases of ambiguity in the hospital name, we ensured that the address listed

on the CSP matched the address found in the HC file. After data were collected for each hospital in the cohort (L.S.), the cohorts were audited for accuracy (M.E., C.L.); no discrepancies were found.

Hospital cohort

HC 2018 includes 5,344 hospitals, and we selected a cohort of 515 according to the criteria presented in Fig 1 to focus this evaluation on large, acute care teaching hospitals. We selected hospitals that sent more than 5,000 HCAHPS surveys in 2018. For each platform, we excluded hospitals that were not represented on the specific platform, did not disclose reviews on the platform, and those with fewer than 10 reviews on that platform. Facebook, Google, and Yelp cohorts numbered 403, 511, and 438, respectively. Of 112 institutions excluded from the Facebook cohort, we were unable to find Facebook pages for 4 institutions; 97 had Facebook pages without reviews available; and 11 had fewer than 10 reviews. Of 4 7 institutions excluded from the Yelp cohort, we were unable to find Yelp pages for 2, and 75 had fewer than 10 reviews.

Outcome variables

The primary outcomes of the study were the difference in mean ratings between each of the 3 CSPs and the degree of correlation between hospital ratings on each CSP and the HCAHPS Recommended Hospital Linear Mean Score (LMS). The LMS is a calculated metric on a 0 to 100 scale; higher values indicate that patients are more likely to recommend the hospital to a family member or friend. The secondary outcome was the degree of correlation between unstructured hospital ratings on each CSP and surgical safety event rates based on PSIs.

Table I Review volume by platform						
Platform	Cohort size (N)	Mean numbe				

 Platform
 Cohort size (N)
 Mean number of reviews (IQR)

 Facebook
 403
 363 (141–476)

 Google
 511
 359 (166–438)

 Yelp
 438
 95 (27–117)

 HCAHPS
 515
 2,104 (1220–2556)

HCAHPS, Hospital Consumer Assessment of Healthcare Providers and Systems; *IQR*, interquartile range.

Statistical analysis

Statistical testing was performed using R version 4.0.2 in R Studio.¹⁷ Mean ratings difference between each CSP pair was assessed with a paired *t* test. The correlation between hospital ratings on each CSP and HCAHPS LMS was assessed with the Pearson product moment correlation coefficient. The correlation between hospital ratings on each CSP and surgical safety event rates was also assessed using the Pearson product moment correlation coefficient. Linear regression was also performed for both correlation analyses, and all results were plotted using the ggplot2 package.¹⁸ No human subjects were involved in this research and the study is considered IRB-exempt

Results

Cohort characteristics

Table I summarizes the cohort size and number of individual reviews on each CSP. There were significantly more HCAHPS survey responses than reviews posted on any of the CSPs. Facebook and Google had a similar number of reviews, and each had significantly more reviews than Yelp. The Facebook and Yelp cohorts were significantly smaller than the Google cohort, largely because 97 (87% of exclusions) Facebook pages did not disclose ratings and 75 (97% of exclusions) Yelp pages had fewer than 10 reviews.

Table II summarizes the geographic footprint and ownership status for our hospital cohort. In comparison with the 5,344 hospitals included in the national HC database, our cohort included relatively more hospitals from the Northeast (18% vs 13%) and West (23% vs 20%) and relatively fewer from the Midwest (23% vs 29%) census regions.¹⁶ Our cohort also had relatively more nonprofit hospitals (70% vs 56%) and relatively fewer government (14% vs 23%) and privately owned (15% vs 20%) hospitals.¹⁶ The regional and ownership differences between cohorts for each CSP were minimal.

Table I	I
Cohort	characterist

Primary results

Fig 2 summarizes the mean rating and strength of correlation with the HCAHPS Recommended Hospital LMS for the 3 CSPs. Mean Facebook ratings (3.81, IQR 3.5–4.3) were more favorable than those found on Google (3.26, IQR 2.8–3.6) or Yelp (2.59, IQR 2.3–2.9), with statistically significant pairwise comparisons (P < .001). Facebook ratings were least strongly correlated with the HCAHPS recommended hospital LMS ($\rho = 0.356$). Google correlation strength with the same HCAHPS metric was greater ($\rho = 0.479$), and Yelp was most strongly correlated ($\rho = 0.500$).

Secondary results

Table III summarizes the strength of correlation between measures of patient satisfaction on each platform and each of the Agency for Healthcare Research and Quality PSIs. Statistically significant correlations (bolded) were all negative, though none were strong. For the composite indicator PSI 90, Facebook had the weakest correlation ($\rho = -.0647$), and it was not significant. The other platforms all had modest, but significant, negative correlations with the composite indicator. Google took on an intermediate value, with a slight negative correlation ($\rho = -.1237$), and Yelp had the strongest correlation ($\rho = -.1639$) of the CSPs. The HCAHPS recommended hospital LMS was most strongly correlated ($\rho = -.2431$).

Discussion

We found that Facebook mean ratings were the most positive, and ratings on that platform were least correlated with the HCAHPS recommended hospital LMS. Google fell in the middle for both ratings and the strength of correlation between those ratings and the HCAHPS recommended hospital LMS. The CSP with the lowest ratings, Yelp, was the most highly correlated with HCAHPS recommended hospital LMS.

Conversely, correlations between ratings on all platforms and a suite of surgical safety indicators were generally insignificant, but those that were significant were all modest and negative. We paid particular attention to the strength of correlation between patient ratings and the composite indicator of safety, as the composite score should give the most optimal global representation of safety. The HCAHPS recommended hospital rating had the strongest negative correlation. Of the CSPs, Facebook had the weakest negative correlation with the rate of surgical safety events and Yelp the strongest.

Cohort characteristics					
Platform	Facebook	Google	Yelp	HCAHPS	All hospitals
Region, n (%)					
Territories	0 (0%)	0 (%)	0 (%)	0 (%)	73 (1%)
Midwest	82 (20%)	114 (22%)	88 (20%)	117 (23%)	1,532 (29%)
Northeast	72 (18%)	91 (18%)	85 (19%)	91 (18%)	675 (13%)
South	147 (36%)	188 (37%)	149 (34%)	188 (37%)	2,008 (38%)
West	102 (25%)	118 (23%)	116 (26%)	119 (23%)	1,056 (20%)
Ownership, n (%)					
Government	51 (13%)	71 (14%)	61 (14%)	72 (14%)	1,225 (23%)
Physician	1 (<1%)	2 (<1%)	2 (<1%)	2 (<1%)	72 (1%)
Proprietary	72 (18%)	79 (15%)	70 (16%)	79 (15%)	1,064 (20%)
Tribal	0 (0%)	0 (0%)	0 (0%)	0 (0%)	9 (<1%)
Nonprofit	279 (69%)	359 (70%)	305 (70%)	362 (70%)	2974 (56%)
Total	403	511	438	515	5,344

HCAHPS, Hospital Consumer Assessment of Healthcare Providers and Systems.

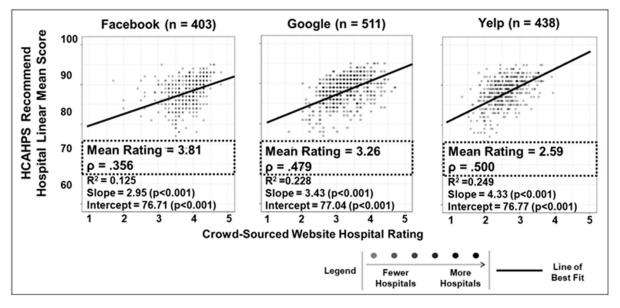


Fig. 2. Mean ratings and strength of correlation between crowd-sourced platforms and HCAHPS scores. HCAHPS, Hospital Consumer Assessment of Healthcare Providers and Systems.

Table III Correlation between hospital ratings and structured safety indicators

Measure Name	Measure ID	Facebook P (P value)	Google P (P value)	Yelp P (P value)	HCAHPS P (P value)
Pressure sores	PSI_3	.01 (.83)	-0.13 (<.01)	-0.10 (.04)	-0.11 (.01)
Deaths among surgical inpatients with serious treatable complications	PSI_4	-0.04 (.43)	-0.17 (<.01)	-0.017 (<.01)	-0.10 (.04)
Collapsed lung	PSI_6	0.07 (.14)	-0.03 (.47)	0.07 (.16)	0.03(.49)
Fall with hip fracture	PSI_8	-0.08 (.13)	-0.11 (.01)	-0.04 (.44)	-0.07 (.10)
Perioperative hemorrhage	PSI_9	0.03 (.55)	-0.09 (.05)	0.01 (.85)	0.07 (0.11)
Postoperative acute kidney injury	PSI_10	-0.03 (.60)	- 0.02 (<. 0 1)	0.05 (0.35)	0.01 (0.80)
Postoperative respiratory failure	PSI_11	-0.23 (<.01)	-0.10 (.02)	-0.17 (<.01)	-0.27 (<.01)
Postoperative PE/DVT	PSI_12	0.04 (0.37)	- 0.17 (< .01)	-0.08 (0.08)	-0.19 (<.01)
Postoperative sepsis	PSI_13	-0.05 (.29)	-0.07 (.11)	-0.13 (<.01)	-0.15 (<.01)
Postoperative wound dehiscence	PSI_14	0.04 (.39)	-0.05 (.28)	-0.03 (.47)	-0.01 (.80)
Accidental lacerations	PSI_15	0.05 (.31)	-0.09 (.03)	0.06 (.22)	0.03 (.46)
Serious complications	PSI_90	-0.06 (.20)	-0.12 (<.01)	-0.16 (<.01)	-0.24 (<.01)

Bold indicates P < .05. ρ designates Pearson correlation.

DVT, deep vein thrombosis; HCAHPS, Hospital Consumer Assessment of Healthcare Providers and Systems; PE, pulmonary embolus.

A prior analysis of the relationship between Facebook ratings and HCAHPS scores in 136 New York hospitals found a significant positive association between Facebook ratings and 10 of the 12 HCAHPS star ratings categories.¹³ A national analysis of the relationship between Yelp ratings and HCAHPS scores found a positive correlation between overall star rating and the Yelp rating in 871 US hospitals.¹⁴ We extended these analyses by considering ratings on multiple platforms jointly.

The most comprehensive analysis of crowd-sourced ratings to date examined a national cohort of 2,995 US hospitals. In that study, investigators separated those hospitals into 3 categories within a hospital referral region (HRR) based on crowd-sourced ratings: best, worst, and mid-level. They then calculated the percentage of hospitals reported as the best or worst by the crowdsourced ratings that were also the best or worst according to HCAHPS data and found that they matched more often than not. Finally, they found weak or no correlation between crowdsourced ratings and clinical quality measures.¹⁵ The comparison of HHRs in this study was an elegant method for replicating the geographically constrained choices of individual patients. However, limiting the classification

to "best" or "worst" within a region limits the specificity with which the authors were able to compare platforms. We chose to compare CSP ratings to the HCAHPS LMS, which offers higher ratings resolution than simple star ratings in order to facilitate a more specific comparison between these platforms from a consumer perspective.

Of the CSPs we studied, only Facebook allows significant engagement by the rated institution. On the other platforms, the organization may respond to the review or even request its removal, but organizations are not able to directly publish content on Google and Yelp. We restricted our analysis to large academic centers, which tend to have a more active social media presence than their smaller or non-teaching counterparts, and this may be responsible for our finding that Facebook ratings are highest.¹⁹ Facebook is also the only CSP we studied that allows institutions to disable reviews entirely, so the mean Facebook rating in our study may have been driven higher by otherwise low-rated organizations disabling their reviews. We examined the relationship between patient ratings and surgical safety, but we did not take any steps to determine whether reviews were left by patients who underwent surgery or received another type of care. It is possible that if we were able to flag patients that received surgical care, perhaps based on keyword analysis of free-text reviews, this subset of patient reviews may have had a greater correlation with surgical safety rates.

Despite these limitations, we believe that our findings can inform patients, physicians, and hospital systems, particularly around patient satisfaction ratings. We found significant differences in ratings between platforms, so patients and referring providers should use the same source when evaluating multiple facilities. Physicians and hospital systems should be aware that, in aggregate, CSP ratings capture directionally similar public opinions as the better validated HCAHPS scores, though there were no meaningful correlations between CSP ratings and surgical safety scores.

CSP ratings are an important component of the decision-making process for surgical patients as they decide where to seek care. Between 36% and 51% of surgical patients used the internet to research hospital and surgeon quality.^{20,21} A systematic review of factors affecting surgeon choice found that hospital reputation was the most important hospital-related factor for patients choosing a surgeon. Further, patients may value their peers' opinions of hospitals and surgeons more highly than the opinions of their referring physician.²² These findings suggest that patients strongly consider CSP ratings when deciding where to seek surgical care.

This review also found that surgeon reputation was the most important factor affecting the patient's choice of surgeon.²² We did not evaluate the concordance between CSP ratings and validated measures of individual clinicians such as the Clinician and Group Consumer Assessment of Healthcare Providers and Systems because clinician-level data was far less robust and complete than institution-level data. However, there is evidence suggesting individual clinician reviews are similarly important to patients. The results of a forced choice experiment suggest that patients value CSP ratings more highly than validated measures such as Clinician and Group Consumer Assessment of Healthcare Providers and Systems scores.²³ These findings suggest that individual surgeons should consider monitoring their online presence since patients are likely to look them up online and generally consider the information they find to be credible. Further, online ratings are updated more guickly than the annual HCAHPS survey and allow patients to report their experiences in free text. Hospital administrators may take advantage of these attributes to get an earlier and more specific indication of how patients view their hospitals. As the volume of online ratings continues to grow, we believe that they will become an ever more frequently used source of information for patients to evaluate their care options and for healthcare institutions to evaluate their own performance in the satisfaction realm.

Funding/Support

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest/Disclosure

The authors have no financial disclosures and no conflicts of interest.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [https://doi.org/10.1016/j.surg.2021. 04.011].

References

- Centers for Medicare and Medicaid Services. HCAHPS fact sheet. https://www. hcahpsonline.org/globalassets/hcahps/facts/hcahps_fact_sheet_october_2019. pdf. Accessed August 12, 2020.
- Centers for Medicare and Medicaid Services. HCAHPS star ratings technical notes. https://hcahpsonline.org/globalassets/hcahps/star-ratings/tech-notes/ hcahps_stars_tech_notes_july_2020.pdf. Accessed July 6, 2020.
- Centers for Medicare and Medicaid Services. Patient-mix coefficients for July 2019 (4017 through 3018 Discharges) publicly reported HCAHPS results. https://www.hcahpsonline.org/globalassets/hcahps/mode-patient-mixadjustment/july_2019_mode-patient-mix-adj_pma.pdf. Accessed July 6, 2020.
- Shi Y, Scanlon DP, Bhandari N, Christianson JB. Is anyone paying attention to physician report cards? the impact of increased availability on consumers' awareness and use of physician quality information health services research. *Heal Serv Res.* 2017;52:1570–1589.
- Bhandari N, Scanlon DP, Shi Y, Smith RA. Why do so few consumers use health care quality report cards? a framework for understanding the limited consumer impact of comparative quality information. *Med Care Res Rev.* 2019;76: 515–537.
- Similar Web. Facebook.com analytics: market share stats and traffic ranking. https://www.similarweb.com/website/facebook.com/. Accessed May 1, 2020.
- Similar Web. Google.com analytics: market share stats and traffic ranking. https://www.similarweb.com/website/google.com. Accessed January 5, 2021.
- Similar Web. Yelp.com analytics: market share stats and traffic ranking. https:// www.similarweb.com/website/yelp.com/#referrals. Accessed May 1, 2020.
 Similar Web. Medicare.gov analytics: market share stats and traffic ranking.
- https://www.similarweb.com/website/medicare.gov. Accessed January 5, 2021. 10. Hong YA, Liang C, Radcliff TA, Wigfall LT, Street RL. What do patients say about
- doctors online? A systematic review of studies on patient online reviews. J Med Internet Res. 2019;21.
- 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.
- Hedges L, Couey C. How patients use online reviews. https://www. softwareadvice.com/resources/how-patients-use-online-reviews/#back. Accessed January 5, 2021.
- Campbell L, Li Y, Li Y. Are Facebook user ratings associated with hospital cost, quality and patient satisfaction? A cross-sectional analysis of hospitals in New York state. *BMJ Qual Saf.* 2018;27:119–129.
- Ranard BL, Werner RM, Antanavicius T, et al. Yelp reviews of hospital care can supplement and inform traditional surveys of the patient experience of care. *Health Aff*. 2016;35:697–705.
- Perez V, Freedman S. Do crowdsourced hospital ratings coincide with hospital compare measures of clinical and nonclinical quality? *Health Serv Res.* 2018;53: 4491–4506.
- Centers for Medicare and Medicaid Services. [Dataset] Hospital Compare. Published 2018. https://data.cms.gov/provider-data/#2018-annual-files. Accessed January 5, 2021.
- 17. R Core Team 2020. R version 4.0.3. https://www.r-project.org/. Accessed January 5, 2021.
- R Core Team 2020. R version 4.0.3 ggplot2. https://ggplot2.tidyverse.org/. Accessed January 5, 2021.
- Griffis HM, Kilaru AS, Werner RM, et al. Use of social media across US hospitals: descriptive analysis of adoption and utilization. J Med Internet Res. 2014;16: e264.
- Kurup V, Considine A, Hersey D, et al. Role of the Internet as an information resource for surgical patients: a survey of 877 patients. *Br J Anaesth*. 2013;110: 54–58.
- Fraval A, Chong YM, Holcdorf D, Plunkett V, Tran P. Internet use by orthopaedic outpatients: current trends and practices. *Australas Med J.* 2012;5:633–638.
- Yahanda AT, Lafaro KJ, Spolverato G, Pawlik TM. A systematic review of the factors that patients use to choose their surgeon. World J Surg. 2016;40(1): 45–55.
- Yaraghi N, Wang W, Gao GG, Agarwal R. How online quality ratings influence patients' choice of medical providers: controlled experimental survey study. *J Med Internet Res.* 2018;20.