ORIGINAL RESEARCH

The Practice of Emergency Medicine

A cross-sectional analysis of Yelp and Google reviews of hospitals in the United States

Tony Zitek MD[1,2] [ORCID] | Joseph Bui BS[2] | Christopher Day BA[2] | Sara Ecoff BS, MA[3] | Brijesh Patel MD[1]

[1]Department of Emergency Medicine, Mount Sinai Medical Center, Miami Beach, Florida, USA

[2]Herbert Wertheim College of Medicine at Florida International University, Miami, Florida, USA

[3]Nova Southeastern University Dr. Kiran C. Patel College of Osteopathic Medicine, Fort Lauderdale, Florida, USA

Correspondence
Tony Zitek, MD, Department of Emergency Medicine, Mount Sinai Medical Center Miami Beach, FL, USA.
Email: zitek10@gmail.com

Abstract

Objective: Patient satisfaction is now an important metric in emergency medicine, but the means by which satisfaction is assessed is evolving. We sought to examine hospital ratings on Google and Yelp as compared to those on Medicare's Care Compare (CC) and to determine if certain hospital characteristics are associated with crowdsourced ratings.

Methods: We performed a cross-sectional analysis of hospital ratings on Google and Yelp as compared to those on CC using data collected between July 8 and August 2, 2021. For each hospital, we recorded the CC ratings, Yelp ratings, Google ratings, and each hospital's characteristics. Using multivariable linear regression, we assessed for associations between hospital characteristics and crowdsourced ratings. We calculated Spearman's correlation coefficients for CC ratings versus crowdsourced ratings.

Results: Among 3000 analyzed hospitals, the median hospital ratings on Yelp and Google were 2.5 stars (interquartile ratio [IQR], 2–3) and 3 stars (IQR, 2.7–3.5), respectively. The median number of Yelp and Google reviews per hospital was 13 and 150, respectively. The correlation coefficients for Yelp and Google ratings with CC's overall star ratings were 0.19 and 0.20, respectively. For Yelp and Google ratings with CC's patient survey ratings, correlation coefficients were 0.26 and 0.22, respectively. On multivariable analysis, critical access hospitals had 0.22 (95% confidence interval [CI], 0.14–0.30) more Google stars and hospitals in the West had 0.12 (95% CI, 0.05–0.18) more Google stars than references standard hospitals.

Conclusion: Patients use Google more frequently than Yelp to review hospitals. Median UnS hospital ratings on Yelp and Google are 2.5 and 3 stars, respectively. Crowdsourced reviews weakly correlate with CC ratings. Critical access hospitals and hospitals in the West have higher crowdsourced ratings.

Supervising Editor: Christian Tomaszewski, MD, MS.

- -

1 | INTRODUCTION

1.1 | Background

Some emergency physicians have expressed concern about the increased focus that hospital administrators have placed on patient satisfaction scores.[1] Although traditionally, patient satisfaction has been measured by formalized patient surveys, some hospitals now emphasize the importance of patient reviews on Yelp, Google, and other "crowdsourced" platforms (those that gather data from the public on the internet). As the gatekeepers to the hospital, emergency physicians play an important role in how hospitals are rated on these crowdsourced platforms.

1.2 | Importance

Although most people are familiar with Yelp and Google, the significance of medical reviews on these platforms is uncertain. Some prior studies that have assessed for correlations between samples of crowdsourced hospital reviews and ratings on Medicare's Care Compare (CC) website, which provides ratings for American hospitals based on both patient satisfaction surveys and clinical quality metrics.[2] These studies have found conflicting results but have generally shown a correlation between crowdsourced ratings and CC patient experience scores and a weaker correlation with clinical quality measures.[3–8] No prior studies have formally assessed how intrinsic hospital factors might affect ratings on crowdsourced review platforms.

1.3 | Goals of this investigation

We thus sought to update and expand upon prior works that have evaluated crowdsourced hospital ratings. In particular, we sought to determine the median Yelp and Google ratings throughout the United States and regionally. We also wanted to assess how intrinsic hospital factors such as location, size, and type of hospital are associated with crowdsourced ratings. Finally, we sought to evaluate how well-crowdsourced ratings correlate with the overall star and patient experience ratings on CC.

2 | METHODS

2.1 | Study design

We performed a cross-sectional analysis of Google and Yelp ratings as compared to the overall star ratings and patient survey ratings of hospitals in the United States as listed on Medicare's CC website. This study used all publicly available data and was thus exempt from review by our local institutional review board.

The Bottom Line

Emergency physicians are under increasing scrutiny, including performance with respect to patient satisfaction. Crowdsourced ratings for 3000 US hospitals were weakly correlated with Centers for Medicare and Medicaid Services data, with scores consistently lower.

2.2 | Selection of participants

We analyzed hospital-level data from the CC website as well as Google and Yelp ratings from their respective websites. We included all American hospitals that had both an "overall star rating" and a "patient survey rating" (defined below) listed on the CC website. We excluded psychiatric hospitals, children's hospitals, and hospitals that were closed at the time of data collection. All data were collected between July 8, 2021 and August 2, 2021.

2.3 | Measurements

For each hospital that met criteria for analysis, we collected the following data: overall star rating on CC, patient survey rating on CC, number of Yelp stars, number of Yelp reviews, number of Google stars, number of Google reviews, state in which the hospital resides, type of hospital control (for-profit, governmental, church-operated, or other not-for-profit), number of staffed beds, and whether the hospital is acute care or critical access. The variables collected are described in more detail as follows.

The Centers for Medicare and Medicaid Services (CMS) overall star ratings on CC are scored 1–5 stars with no partial stars. This single score represents an aggregate of seven areas of quality: mortality, safety of care, readmission, patient experience, effectiveness of care, timeliness of care, and efficient use of medical imaging.[9]

The CMS patient survey ratings on CC are scored 1–5 stars with no partial stars. These scores are based on the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) surveys. The HCAHPS surveys ask a random sample of discharged adult patients 29 questions about their recent hospital visit. CMS publishes participating hospitals' HCAHPS results on the CC website 4 times a year.[10–11]

Yelp, established in 2004, allows for reviews of businesses (hospitals included) from 1 to 5 stars. Individual reviews do not allow for partial stars, but the overall ratings for individual businesses are rounded to the nearest half star. Yelp applies a filter on reviews to proactively filter out false reviews.[12]

Reviews on Google started in 2007, at which point reviews were posted for businesses via Google Maps.[13] Google reviews allow users to post reviews of hospitals and businesses from 1 to 5 stars. Individual reviews do not allow for partial stars, but overall ratings for individual businesses are rounded to the nearest tenth of a star.[14]

Although data were collected from July 8, 2021 and August 2, 2021, the data abstracted from Yelp and Google represent an aggregate of months or years of data that began to be gathered when the hospital opened an account with Yelp or Google and ended the day the data were abstracted. Therefore, the period of time for which the data for each hospital was generated was highly variable, but the data points collected are exactly what a potential patient would see if they visited the Yelp or Google page for a hospital. Additionally, it should be noted that Yelp and Google reviews may not necessarily be completed by actual patients. They could be posted by family members or friends of patients, or they could be filled out by someone who simply wants to try to help or harm the hospital's reputation. As mentioned above, to some extent false reviews are filtered out, but not completely, and these issues should be considered when interpreting reviews of hospitals on crowdsourced platforms.

The data for this study were abstracted and input into a Google sheet by 3 medical students and 1 resident physician, all of whom were trained to perform data abstraction by the principal investigator. Overall star ratings and patient survey ratings were taken directly from the CC website (available at https://www.medicare.gov/care-compare/). Yelp stars and review numbers were found by searching the hospital name on Yelp's website or mobile app. Hospital names were searched in Google to find the Google reviews data. Lastly, each hospital was searched in the American Hospital Association directory (available at https://guide.prod.iam.aha.org/guide/) to get the number of staffed beds and type of control.

Twenty percent of the data were audited by the principal investigator. In 3 cases (out of 600 audited hospitals), the initial abstractor reported that there were no Yelp or Google reviews for the hospital when, in fact, reviews were available. The issue was that the hospitals had changed their names, and their reviews were listed under a different name. The data were corrected. Given the issue identified in the initial audit, one additional audit was performed by 1 of the medical students on the data from all hospitals for which the original abstractor reported there were either no Yelp or Google reviews. An additional 4 cases were found to have reviews under a changed hospital name, and the data were corrected. Otherwise, no errors in data abstraction were found.

2.4 | Outcomes

We had several goals for this study. First, we sought to determine the median number of reviews and stars for American hospitals on both Yelp and Google. We performed calculations on national, regional, and state levels to help determine what scores might be considered "good" relative to other hospitals.

Next, we sought to elucidate how certain intrinsic hospital characteristics might be associated with Yelp or Google ratings. In particular, we assessed whether region, number of staffed beds, type of hospital (critical access or acute care), or type of hospital control (for-profit, governmental, church-operated, or other not-for-profit) were associated with crowdsourced ratings.

Finally, to help assess the value of crowdsourced reviews, we determined correlation coefficients for Yelp and Google ratings with overall star and patient survey ratings on CC. We also calculated the correlation coefficient for Yelp ratings versus Google ratings.

2.5 | Data analysis

We calculated the median number of Yelp and Google reviews and stars with interquartile ranges for all included hospitals and each subgroup. We also calculated means and standard deviations for national data.

We used multivariable linear regression to test for an association between hospital characteristics and both Yelp and Google ratings. Because no there are no prior data on this subject, we chose covariates based on a priori hypothesis that the following may be associated with crowdsourced ratings: region, number of staffed beds, type of hospital, and type of hospital control. The variables were categorized as follows: region (Midwest, Northeast, South, or West), hospital type (acute care or critical access), hospital size (small, medium, or large), and type of hospital control (for-profit, governmental, church-operated, or other not-for-profit). States were divided into the four regions mentioned above as done by the Centers for Disease Control and Prevention.[15] We defined a small hospital as having <100 staffed beds, a medium hospital as having 100–499 staffed beds, and a large hospital as having 500 or more beds.

We used Spearman's correlation coefficient to assess for correlations among Yelp, Google, and the CMS ratings on CC. Given that hospitals with very low numbers of reviews are more likely to have inaccurate overall Yelp or Google scores, a sensitivity analysis was performed using only hospitals with at least five reviews. Spearman's correlation coefficients were recalculated with this sample.

Data were aggregated in Excel (version 16.60, Microsoft, Redmond, WA) and analyzed in R Studio (version 2022.02.2). Using two-sided hypothesis tests, p-value <0.05 was considered statistically significant.

3 | RESULTS

3.1 | Overall data

Between July 8 and August 2, 2021, there were 5165 hospitals listed on CC. As demonstrated in Figure 1, 3000 met our criteria for analysis. Among these hospitals, the median for both the overall star rating and the patient survey rating on CC was 3 stars (IQR, 2–4).

Regarding crowdsourced reviews, at least one Yelp review was present for 91.5% of hospitals. The median and mean numbers of Yelp reviews for each hospital were 13 (IQR, 3–46) and 45.6 (SD, 82.7), respectively. The median and mean numbers of Yelp stars for each hospital were 2.5 (IQR, 2–3) and 2.6 (SD, 0.87), respectively. The 10th percentile for Yelp ratings of hospitals was 1.5 stars whereas 3.5 stars was the 90th percentile.

In total, 99.9% of hospitals had at least one Google review. The median and mean numbers of Google reviews for each hospital were 151 (IQR, 65.8–301) and 271.1 (SD, 387.9), respectively. The median
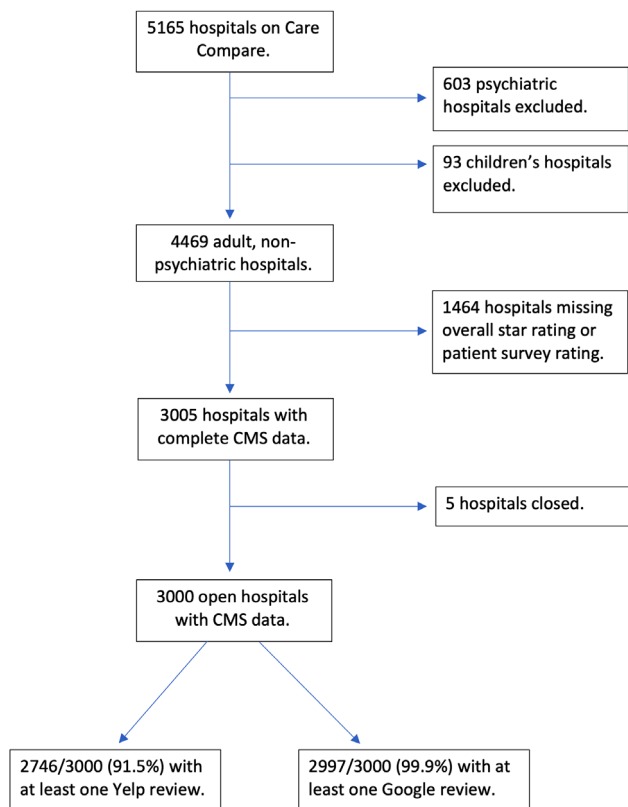
4 of 8 | **WILEY** | JACEP OPEN
JOURNAL OF THE AMERICAN COLLEGE OF EMERGENCY PHYSICIANS OPEN
A GLOBAL JOURNAL OF EMERGENCY MEDICINE

ZITEK ET AL.


```
┌──────────────────────┐
│ 5165 hospitals on Care│
│      Compare.         │
└──────────────────────┘
            │         ┌──────────────────────┐
            ├────────▶│ 603 psychiatric      │
            │         │ hospitals excluded.  │
            │         └──────────────────────┘
            │         ┌──────────────────────┐
            ├────────▶│ 93 children's        │
            │         │ hospitals excluded.  │
            ▼         └──────────────────────┘
┌──────────────────────┐
│ 4469 adult, non-      │
│ psychiatric hospitals.│
└──────────────────────┘
            │         ┌──────────────────────┐
            ├────────▶│ 1464 hospitals missing│
            │         │ overall star rating or│
            │         │ patient survey rating.│
            ▼         └──────────────────────┘
┌──────────────────────┐
│ 3005 hospitals with   │
│ complete CMS data.    │
└──────────────────────┘
            │         ┌──────────────────────┐
            ├────────▶│ 5 hospitals closed.  │
            ▼         └──────────────────────┘
┌──────────────────────┐
│ 3000 open hospitals   │
│ with CMS data.        │
└──────────────────────┘
       ┌──────┴──────┐
       ▼             ▼
┌──────────────┐ ┌──────────────┐
│2746/3000     │ │2997/3000     │
│(91.5%) with  │ │(99.9%) with  │
│at least one  │ │at least one  │
│Yelp review.  │ │Google review.│
└──────────────┘ └──────────────┘
```

FIGURE 1 Included and excluded hospitals.

and mean numbers of Google stars were 3.0 (IQR, 2.7–3.5) and 3.1 (SD, 0.62), respectively. The 10th percentile for Google ratings of hospitals was 2.4 stars, and 4.0 stars was the 90th percentile.

3.2 | Data stratified by region, state, and hospital type

We analyzed data by hospital characteristics. Results are demonstrated in Table 1. Notably, hospitals in the Western region of the United States, critical access hospitals, large hospitals (≥500 staffed beds), and for-profit hospitals had higher crowdsourced ratings compared to others.

Table 2 demonstrates a comparison of the scores on the different rating platforms, stratified by state (using the 10 states with the most included hospitals). Notably, there was a wide range in the median number of Yelp reviews with California having a median of 174.5 per hospital as compared to Ohio, which had a median of only 5 per hospital. Google reviews were more numerous than Yelp reviews in every state.

3.3 | Regression analysis of hospital characteristics

We performed a multivariable linear regression analysis using region, acute care versus critical access type, hospital size, and type of control as variables. As shown in Table 3, factors found to have indepen-

dent associations with higher Yelp scores were location in the West, small size, and being a critical access hospital. Factors found to have independent associations with lower Yelp scores were location in the Northeast and governmental control. There were no statistically significant associations for the other assessed factors.

Factors found to have independent associations with higher Google scores were location in the West, being a critical access hospital, large size, and being for-profit. There were no statistically significant associations for the other assessed factors.

3.4 | Correlation coefficients

Finally, we assessed for correlations between the CC ratings and crowdsourced ratings. The Spearman correlation coefficient between overall star ratings and Yelp ratings was 0.19 (95% CI, 0.15–0.23). Between overall star ratings and Google ratings, it was 0.20 (95% CI, 0.17–0.23).

The Spearman correlation coefficient between patient survey ratings and Yelp ratings was 0.26 (95% CI, 0.23–0.29). For patient survey ratings and Google ratings, it was 0.22 (95% CI, 0.19–0.25).

The Spearman correlation coefficient between Yelp ratings and Google ratings of hospitals was 0.27 (95% CI, 0.24–0.30).

3.5 | Sensitivity analysis

The correlation coefficients calculated above were repeated using only hospitals with at least five reviews. Of the 3000 initially analyzed hospitals, 99.1% had at least five Google reviews and 67.6% had at least five Yelp reviews. The Spearman correlation coefficients for the full data and the adjusted data using only hospitals with at least five reviews are demonstrated in Table 4.

4 | LIMITATIONS

When interpreting this study, there are some limitations to consider. First, this study was intended to compare the hospital ratings on CC to those on Yelp and Google for all adult hospitals in the United States, but 1464 hospitals listed on CC either did not have an overall star rating or a patient survey rating, resulting in their exclusion from analysis. Some of these hospitals were veterans affairs hospitals, for which exclusion from analysis is appropriate because these hospitals are not open to the public. However, other hospitals were simply missing data, and it is unclear how their inclusion would have affected our results. Nonetheless, this study remains the most comprehensive assessment of Yelp and Google reviews of hospitals to date.

Additionally, this study used CMS data on CC to assess the utility of reviews on Yelp and Google, but the data from CMS may not be good measures of quality or patient experience. Indeed, HCAHPS surveys have been criticized for having low response rates, which may result in negatively skewed data.[16] Moreover, some authors have noted that the measures CMS uses to determine quality of care may not be scientifically sound.[17] At this point, there is no gold standard measure to

TABLE 1 Median CMS hospital ratings and crowdsourced hospital ratings stratified by hospital characteristics[a].

Hospital characteristic	Overall star rating	Patient survey rating	Yelp stars	Google stars
Region				
Midwest (n = 797)	4 (3–4)	3 (3–4)	2.5 (2–3)	3.0 (2.7–3.4)
Northeast (n = 493)	3 (2–4)	3 (2–4)	2.5 (2–3)	3.0 (2.6–3.3)
South (n = 1094)	3 (2–4)	3 (3–4)	2.5 (2–3)	3.0 (2.6–3.6)
West (n = 616)	3 (2–4)	3 (2–4)	2.5 (2.5–3)	3.2 (2.8–3.7)
Hospital type				
Acute care (n = 2693)	3 (2–4)	3 (3–4)	2.5 (2–3)	3.0 (2.7–3.5)
Critical access (n = 307)	3 (3–4)	4 (4–4)	3.0 (2–4)	3.1 (2.8–3.6)
Staffed beds				
<100 (n = 987)	3 (3–4)	4 (3–4)	2.5 (2–3.5)	3.0 (2.7–3.5)
100–499 (n = 1728)	3 (2–4)	3 (2–3)	2.5 (2–3)	3.0 (2.7–3.5)
≥500 (n = 285)	3 (2–4)	3 (3–3)	2.5 (2–3)	3.2 (2.8–3.6)
Type of control				
For-profit (n = 520)	3 (2–4)	3 (2–3)	2.5 (2–3)	3.4 (2.8–4.0)
Government (n = 410)	3 (2–4)	3 (3–4)	2.5 (2–3)	3.0 (2.6–3.4)
Church (n = 377)	4 (3–4)	3 (3–4)	2.5 (2–3)	3.0 (2.7–3.5)
Other nonprofit (n = 1693)	3 (3–4)	3 (3–4)	2.5 (2–3)	3.0 (2.7–3.4v)

Abbreviation: CMS, Centers for Medicare and Medicaid Services.
[a]Data are reported as median (interquartile range).

TABLE 2 A comparison of Care Compare overall star ratings, Care Compare patient survey ratings, Yelp reviews, and Google reviews for the 10 states with the most hospitals in the United States[a]

State	Included hospitals	Overall star rating	Patient survey rating	Yelp reviews	Yelp stars	Google reviews	Google stars
California	272	3 (2–4)	3 (2–3)	174.5 (85.8–312)	2.5 (2.5–3)	188.5 (109.8–348)	3.3 (2.8–3.7)
Texas	213	3 (3–4)	3 (3–4)	19 (5–50)	2.5 (2–3)	248 (94–439)	3.2 (2.8–3.8)
Florida	158	3 (2–4)	3 (2–3)	38 (18–61.8)	2.5 (2–3)	483.5 (268.2–951)	3.4 (2.9–3.9)
New York	137	2 (1–3)	2 (2–3)	29 (6–67)	2.5 (2–2.5)	168 (89–333)	2.8 (2.5–3.3)
Pennsylvania	134	4 (3–4)	3 (3–4)	7 (3–26.8)	2.5 (2–3)	138 (58.3–223.8)	2.9 (2.5–3.3)
Ohio	127	4 (3–5)	3 (3–4)	5 (3–15)	2.5 (2–3)	104 (64–213)	3.0 (2.6–3.3)
Illinois	127	3 (3–4)	3 (3–4)	16 (3–67)	2.5 (2–3)	162 (54–249.5)	3.0 (2.7–3.2)
Michigan	97	4 (3–4)	3 (3–4)	6 (2–25)	2.5 (2–3)	139 (62–271)	2.9 (2.6–3.2)
Indiana	89	3 (3–4)	3 (3–4)	6 (2–13)	2.5 (2–3.5)	104 (55–178)	3.1 (2.8–3.3)
North Carolina	88	3 (3–4)	3 (3–4)	6.5 (3–22.3)	2.5 (2–3)	162 (72–269)	2.7 (2.4–3.2)

[a]Values are reported as median (interquartile range).

confidently assess quality of patient care. The CC overall star rating is easy to understand, accessible, and at least partially based on medical evidence. For these reasons, we used it as our measure of quality.

Finally, some factors that might influence hospital ratings on Yelp and Google were not assessed in this study, such as socioeconomic status, area deprivation index, racial segregation, and population density. Therefore, further analysis as to the factors that affect crowdsourced hospital ratings is warranted.

5 | DISCUSSION

Given the increasing emphasis that hospital administrators are placing on crowdsourced reviews, emergency physicians and administrators should understand the significance of these reviews. This study is the most comprehensive assessment of Yelp and Google reviews of hospitals to date. Our analysis identified several important results.

TABLE 3 Results of a multivariable linear regression analysis assessing the association among hospital characteristics and crowdsourced hospital ratings

Hospital characteristic	Expected change in rating for characteristic compared to reference	
	Yelp stars	Google stars
Region		
Midwest (n = 797)	Reference	Reference
Northeast (n = 493)	−0.12 (95% CI, −0.23 to −0.02)	−0.05 (−0.12 to 0.02)
South (n = 1094)	−0.003 (95% CI, −0.09 to 0.08)	−0.02 (95% CI, −0.08 to 0.02)
West (n = 616)	0.17 (95% CI, 0.07 to 0.26)	0.12 (95% CI, 0.05 to 0.18)
Hospital type		
Acute care (n = 2693)	Reference	Reference
Critical access (n = 307)	0.37 (95% CI, 0.23 to 0.50)	0.22 (95% CI, 0.14 to 0.30)
Staffed beds		
<100 (n = 987)	0.10 (95% CI, 0.02 to 0.18)	−0.05 (95% CI, −0.10 to 0.002)
100–499 (n = 1728)	Reference	Reference
≥500 (n = 285)	0.004 (95% CI, −0.10 to 0.11)	0.19 (95% CI, 0.12 to 0.27)
Type of control		
For-profit (n = 520)	−0.06 (95% CI, −0.15 to 0.03)	0.35 (95% CI, 0.29 to 0.41)
Government (n = 410)	−0.11 (95% CI, −0.21 to −0.01)	−0.05 (95% CI, −0.11 to 0.02)
Church (n = 377)	0.04 (95% CI, −0.07 to 0.14)	0.06 (95% CI, −0.003 to 0.13)
Other nonprofit (1693)	Reference	Reference

Abbreviation: CI, confidence interval.

TABLE 4 The Spearman correlation coefficients between the hospital ratings on the CMS Care Compare website versus those from Yelp and Google[a]

	Yelp stars (all hospitals)	Google stars (all hospitals)	Yelp stars (≥5 reviews)	Google stars (≥5 reviews)
Overall star rating	0.19 (0.15–0.23)	0.20 (0.17–0.23)	0.25 (0.21–0.29)	0.20 (0.17–0.23)
Patient survey rating	0.26 (0.23–0.29)	0.22 (0.19–0.25)	0.36 (0.32–0.40)	0.22 (0.19–0.25)

Abbreviations: CI, confidence interval; CMS, Centers for Medicare and Medicaid Services.
[a]Data are presented as the Spearman correlation coefficient followed by (95% CI).

First, Yelp is used much less frequently for hospital reviews than Google; this is true in all regions of the United States. However, in California, Yelp reviews are nearly as common as Google reviews with the median number of Yelp reviews per hospital being 174.5 compared to 188.5 Google reviews. The popularity of Yelp for hospital reviews in California may be related to Yelp's origin and headquarters being in California, but Google's headquarters are also in California. In other states, like Ohio (where the median number of Yelp reviews per hospital is 5), Yelp data are not likely to be useful in comparing hospitals given the very small number of reviews. This finding that Google reviews outnumber Yelp reviews for hospitals confirms the reports from prior studies,[5,8] but the discrepancy has become larger, suggesting increased recent use of Google for hospital reviews.

Second, the median number of Yelp and Google stars for hospitals throughout the country is 2.5 and 3.0, respectively. Yelp has not publicly reported a median rating for all businesses, but they did report that 55% of businesses had 4 or 5 stars, overall, as of September 30, 2021.[18]

One study reported the average Google rating for all businesses was 4.1 stars.[19] Thus, it appears that hospital ratings are substantially lower than other businesses on both Yelp and Google. Although there is some regional variation in crowdsourced ratings, none of the analyzed areas had Yelp or Google ratings that approached the expected values for other business types. Nationally, and in many areas of the country, a hospital with 3.5 Yelp stars or 4.0 Google stars is in the top 10% of hospitals.

Next, there are some intrinsic hospital characteristics associated with higher and lower crowdsourced reviews. For one, hospitals in the Western United States tend to have higher crowdsourced ratings than hospitals in other areas. Because CC ratings are not higher in the West, it is unclear why this is.

Additionally, critical access hospitals were found to have higher Yelp and Google ratings than acute care hospitals. Critical access hospitals also have higher patient survey ratings on CC. Thus, in general, patient satisfaction measures seem to be higher in critical access hospitals. We

are unaware of any previous publication that has reported this finding. One might hypothesize that it is easier for smaller hospitals to provide a better patient experience, but this association was inconsistent in our study (with large hospitals having higher Google ratings). Additionally, prior data have found lower patient satisfaction in rural hospitals than urban ones.[20] As such, the reason for higher patient satisfaction measures at critical access hospitals is uncertain.

One other notable hospital characteristic associated with crowdsourced ratings is the type of hospital control. In particular, we found that government-control was significantly associated with lower Yelp ratings and tended to be associated with lower Google ratings (although not significantly so). On the other hand, for-profit hospitals had higher Google ratings. Although the relationship between the type of hospital control and crowdsourced ratings has not been previously elucidated, one prior study found that government hospitals were more likely to have superior patient experience scores using HCAHPS data.[21] As such, it is possible that Yelp and Google reviews identify issues with government-controlled hospitals that are not captured in HCAHPS survey data. Moreover, higher Google ratings of for-profit hospitals that are not associated with higher CMS patient survey data could suggest that for-profit hospitals are actively trying to improve their Google reviews perhaps because online reputation is associated with hospital revenue.[22]

The last important results from this study relate to the correlation between the crowdsourced ratings and the CMS ratings on CC. As mentioned in the Results section, there are statistically significant associations between both Yelp and Google ratings and the overall star rating, but the correlation is weak. The correlation is slightly stronger between Yelp and Google ratings and patient survey ratings, but it is still modest. The weak correlations we found between crowdsourced rating platforms and patient survey ratings as well as the weaker correlation between the crowdsourced platforms and care quality are consistent with prior published reports.[5,8,23–25] Interestingly and not previously evaluated, Yelp and Google ratings only weakly correlate with each other. This may suggest that Yelp and Google each provide different information about patient satisfaction, or perhaps this means that both are poor indicators of patient satisfaction.

In summary, Google is used much more frequently than Yelp for reviewing hospitals. There are regional variations in hospital ratings on Yelp and Google, but nationally, median ratings are 2.5 and 3.0 stars, respectively, which are lower than other businesses. Yelp and Google reviews positively correlate with CMS overall star and patient survey ratings on CC, but the correlations are weak. Emergency physicians and hospital administrators should recognize that intrinsic hospital characteristics influence crowdsourced hospital ratings. In particular, critical access hospitals and hospitals in the West tend to have higher ratings on Yelp and Google. Given our findings, the utility of crowdsourced reviews for assessing hospitals is uncertain.

AUTHOR CONTRIBUTIONS

Tony Zitek and Joseph Bui conceived the study and designed the study. Tony Zitek and Brijesh Patel supervised the conduct of the data collection. Joseph Bui, Christopher Day, Sara Ecoff, and Brijesh Patel collected the data, including quality control. Tony Zitek provided statistical advice and analyzed the data. Tony Zitek drafted the manuscript and all authors contributed substantially to its revision. Tony Zitek takes responsibility for the article as a whole.

CONFLICTS OF INTEREST STATEMENT

The authors declare no conflicts of interest.

ORCID

Tony Zitek MD https://orcid.org/0000-0002-4357-6611

REFERENCES

1. First Person: Why I'm Leaving Emergency Medicine. *Emergency Medicine News.* 2022;44(9C).
2. Rau J, Only 251 Hospitals Score Five Stars in Medicare's New Ratings. *Kaiser Health News.* 2015. Accessed January 19, 2022. https://khn.org/news/only-251-hospitals-score-five-stars-in-medicares-new-ratings/
3. Bardach NS, Asteria-Penaloza R, Boscardin WJ, et al. The relationship between commercial website ratings and traditional hospital performance measures in the USA. *BMJ Qual Saf.* 2013;22(3): 194-202.
4. Chakraborty S, Church EM. Social media hospital ratings and HCAHPS survey scores. *J Health Organ Manag.* 2020;34(2):162-172.
5. Synan LT, Eid MA, Lamb CR, et al. Crowd-sourced hospital ratings are correlated with patient satisfaction but not surgical safety. *Surgery.* 2021;170(3):764-768.
6. Campbell L, 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(2): 119-129.
7. Huppertz JW, Otto P. Predicting HCAHPS scores from hospitals' social media pages: a sentiment analysis. *Health Care Manage Rev.* 2018;43(4):359-367.
8. Perez V, Freedman S. Do crowdsourced hospital ratings coincide with hospital compare measures of clinical and nonclinical quality? *Health Serv Res.* 2018;53(6):4491-4506.
9. Overall star rating for hospitals. Accessed January 19, 2022. https://www.medicare.gov/care-compare/resources/hospital/overall-star-rating
10. HCAHPS: Patients' Perspectives of Care Survey. Accessed January 19, 2022. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalHCAHPS
11. Patient-mix coefficients for October 2020 (1Q19 through 4Q19 discharges) Publicly Reported HCAHPS Results. Accessed January 19, 2022. https://hcahpsonline.org/globalassets/hcahps/mode-patient-mix-adjustment/october_2020_pma_web_document.pdf Posted July 16, 2020
12. Kamerer D. Understanding the Yelp review filter: an exploratory study. First Monday. 2014;19(9).
13. Goldman J. Add your reviews to businesses on Google Maps. The official blog for Google Maps. Accessed January 19, 2022. https://maps.googleblog.com/2007/06/add-your-reviews-to-businesses-on.html Posted June 19, 2007
14. Sprague DJ. The history of online reviews and how they have evolved. Shopper Approved. Accessed January 19, 2022. https://5stars.shopperapproved.com/the-history-and-evolution-of-online-reviews/Posted Posted December 20, 2019
15. Census regional trends for common human coronaviruses. Centers for disease control and prevention. Accessed August 10, 2021. https://www.cdc.gov/surveillance/nrevss/coronavirus/region.html. Last reviewed July 27, 2021.

16. Godden E, Paseka A, Gnida J, et al. The impact of response rate on Hospital Consumer Assessment of Healthcare Providers and System (HCAHPS) dimension scores. *Patient Exp J.* 2019;6(1):105-114.

17. MacLean CH, Kerr EA, Qaseem A. Time out: charting a path for improving performance measurement. *N Engl J Med.* 2018;378(19):1757-1761.

18. An introduction to Yelp metrics as of September 30, 2021. Yelp Newsroom: Fast Facts. Accessed January 19, 2022. https://www.yelp-press.com/company/fast-facts/default.aspx

19. Sterling G, The average GMB profile has 73 reviews, 45 photos, 5 Posts and a star rating of 4.08. Accessed January 19, 2022. https://searchengineland.com/the-average-gmb-profile-has-73-reviews-45-photos-5-posts-and-a-star-rating-of-4-08-339849. Posted August 26, 2020

20. Kang Y, Tzeng H-M, Zhang T. Rural disparities in hospital patient satisfaction: multilevel analysis of the Massachusetts AHA, SID, and HCAHPS data. *J Patient Exp.* 2020;7(4):607-614.

21. Lehrman WG, Elliot MN, Goldstein E, et al. Characteristics of hospitals demonstrating superior performance in patient experience and clinical process measures of care. *Med Care Res Rev.* 2010;67(1):38-55.

22. Heath S. Online reputation, provider reviews tied to hospital revenue. PatientEngagementHIT. Accessed January 19, 2022. https://patientengagementhit.com/news/online-reputation-provider-reviews-tied-to-hospital-revenue. Posted March 21, 2019.

23. Placona AM, Rathert C. Are online patient reviews associated with health care outcomes? A systematic review of the literature? *Med Care Res Rev.* 2021:10775587211014534.

24. Hefele JG, Li Y, Campbell L, et al. Nursing home facebook reviews: who has them, and how do they relate to other measures of quality and experience? *BMJ Qual Saf.* 2018;27(2):130-139.

25. Emmert M, Meszmer N, Schlesinger M. A cross-sectional study assessing the association between online ratings and clinical quality of care measures for US hospitals: results from an observational study. *BMC Health Serv Res.* 2018;18(1):82.

AUTHOR BIOGRAPHY

Tony Zitek, MD, is the Director of Research and Clinical Assistant Professor in the Department of Emergency Medicine at Mount Sinai Medical Center in Miami Beach, Florida.