



Published in final edited form as:

*Health Aff (Millwood)*. 2016 April 1; 35(4): 697–705. doi:10.1377/hlthaff.2015.1030.

## What can Yelp teach us about measuring hospital quality?

**Benjamin L. Ranard, AB, Rachel M. Werner, MD, PhD, Tadas Antanavicius, H. Andrew Schwartz, PhD, Robert J. Smith, BS, Zachary F. Meisel, MD, MPH, MS, David A. Asch, MD, MBA, Lyle H. Ungar, PhD, and Raina M. Merchant, MD, MSHP**

The Perelman School of Medicine, University of Pennsylvania (BLR, RMW, RJS, ZFM, DAA, RMM), Penn Social Media and Health Innovation Lab, University of Pennsylvania (BLR, TA, HAS, RJS, DAA, LHU, RMM), The Leonard Davis Institute of Health Economics, University of Pennsylvania (BLR, RMW, ZFM, DAA, RMM), Division of General Internal Medicine, Department of Medicine, Perelman School of Medicine, University of Pennsylvania (RMW, DAA), Center for Health Equity Research and Promotion (CHERP), Philadelphia Veterans Affairs Medical Center (RMW, DAA); Department of Computer and Information Science, University of Pennsylvania (HAS, LHU); Department of Emergency Medicine, University of Pennsylvania (ZFM, RMM) – all located in Philadelphia, Pennsylvania, USA

### Abstract

Little is known about how real-time online rating platforms such as Yelp may complement the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) survey, the U.S. standard for evaluating patient experiences after hospitalization. We compared the content of Yelp narrative reviews of hospitals to the domains covered by HCAHPS. While the domains included in Yelp reviews covered the majority of HCAHPS domains, Yelp reviews covered an additional twelve domains not reflected in HCAHPS. The majority of Yelp topics most strongly correlated with positive or negative reviews are not measured or reported by HCAHPS. Yelp provides a large collection of patient and caregiver-centered experiences that can be analyzed with natural language processing methods to identify for policy makers what measures of hospital quality matter most to patients and caregivers while also providing actionable feedback for hospitals.

---

Since 2006,<sup>1, 2</sup> patient-reported experiences after hospitalization have been collected with the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) survey.<sup>1, 2</sup> HCAHPS survey results are publically reported on the Centers for Medicare & Medicaid Services (CMS) *Hospital Compare* website,<sup>1</sup> which rates all United States (U.S.) hospitals that receive Medicare payments on a variety of quality measures.<sup>3</sup> Currently, HCAHPS scores drive 30% of the financial incentives in the Medicare value-based purchasing program,<sup>4</sup> which will eventually penalize hospitals with poor performance by up to 2% of their Medicare payments.<sup>5</sup> Although HCAHPS is the current standard for patient experience of care data,<sup>1</sup> its origins derive from 1995.<sup>2</sup> In the intervening 20 years, the indications for and experience of hospitalization have changed greatly. Perhaps more importantly, 20 years ago patients were not spontaneously publishing their opinions about

health care facilities on social media sites where those opinions reach a public increasingly comfortable using them to inform their own decisions.

Evaluations like HCAHPS are the products of years of measurement research, are fielded and interpreted systematically,<sup>1, 2</sup> and collect a large number of patient responses per hospital.<sup>1</sup> However, they are expensive to deploy,<sup>6</sup> suffer from low response rates,<sup>7</sup> incur significant delay between hospitalization and public reporting of results<sup>8</sup> and, even if they can give an overall indication of patient satisfaction, rarely identify the source of perceived problems.<sup>9</sup>

In contrast, reviews on social media sites are organic, largely unstructured, essentially uncurated, and both seemingly haphazard and subject to gaming. Yet, the testimonials on social media sites are free, continuously updated,<sup>10, 11</sup> and often reveal exactly what the problem or positive occurrence was that impacted the patient's or family member's experience.<sup>9</sup>

Yelp (yelp.com) is a website where users submit star ratings and narrative reviews of local businesses (e.g., restaurants, retail stores, hotels), which are then posted for the public to view. It is the 28<sup>th</sup> most visited website in the U.S., the 129<sup>th</sup> most viewed site globally,<sup>12</sup> and has 142 million unique monthly visitors.<sup>13</sup> To date, Yelp is the most widely used, freely available commercial website in the U.S. for hospital ratings.<sup>10</sup> On the site, businesses are given a rating of one through five stars, which is displayed to the public.<sup>13</sup> The narrative component of reviews often reflects the features of a hospital experience most important to patients, providing information that structured surveys cannot.<sup>9</sup>

Previous work on online patient reviews of hospitals has focused on hospitals in the United Kingdom,<sup>14–16</sup> South Korea,<sup>17</sup> and Germany,<sup>18</sup> or has studied U.S. hospitals but has not examined the content of the online reviews.<sup>10, 19</sup> In England, the National Health Service (NHS) runs a website, NHS Choices, which allows structured patient reviews to be published;<sup>14, 15</sup> however, no such government run site currently exists in the U.S. Little is known about how alternative online sources of patient and caregiver experiences of hospitals in the U.S. may contribute to measuring hospital quality, about what information is contained in these online reviews, and how this information might be used to identify patient-centered outcomes for hospitalized patients.

We sought to compare the content of all Yelp narrative reviews of all U.S. hospitals with Yelp reviews to the domains covered by HCAHPS. Our secondary aim was to identify the Yelp topics most correlated with positive or negative Yelp review ratings of hospitals and to correlate Yelp ratings with HCAHPS overall rating.

## Study Data And Methods

### HCAHPS Data

We obtained HCAHPS survey data from the Hospital Compare dataset (available from data.medicare.gov, published April 2014).

## Yelp Data

From the list of hospitals in the Hospital Compare dataset, we identified all hospitals with Yelp reviews posted as of July 15, 2014. We removed hospitals that did not have American Hospital Association (AHA) annual survey data. We also eliminated reviews “not recommended” by Yelp (a measure indicating a review is likely to be fake). Not recommended reviews are determined automatically by Yelp’s proprietary algorithm that considers a number of factors to try and remove fake reviews (e.g., one person posting many reviews from the same computer). According to Yelp, businesses that buy ads cannot influence which reviews are recommended.<sup>20</sup>

## Topic Generation

We used a type of natural language processing called Latent Dirichlet Allocation (LDA) to analyze the text of all Yelp narrative reviews of hospitals to produce 50 underlying topics. LDA is a widely used type of natural language processing that analyzes co-occurrences of words in text to produce a pre-determined number of topics.<sup>21</sup> Beyond specifying a number of topics to produce, LDA topic generation is fully computer-automated. Topics are groups of words (terms) that tend to co-occur. Those groups can then be labeled by human coders based on their content. For example, LDA generated a topic with co-occurring terms “pain, doctor, nurse, told, medication, meds, gave,” which was then labeled as “pain medications” by coders. We used the implementation of LDA provided by the MALLET package,<sup>22</sup> which has been used previously to analyze the content of other types of social media.<sup>23, 24</sup> LDA analysis has also been used previously to characterize online narrative reviews of physicians.<sup>25–27</sup>

## Number of Topics

In order to choose the number of topics to generate, LDA was used to generate 25, 50, 100 and 200 topics. Two reviewers (BLR, RJS) independently rated a blinded sample of each set of topics in order to determine which predetermined number of topics produced the most coherent topics. If too few topics are generated, the resulting topics are overly broad (e.g., from the 25 topic list: hospital, medical, care, center, health, people, hospitals). If there are too many topics generated, the topics become extremely specific, with many overlapping topics (e.g., from the 200 topic list: ultrasound, pregnant, baby, horrible, bleeding, weeks, miscarriage). After deciding on a predetermined number of topics to generate, we ascertained the significance of the topics by correlating the topics with high and low Yelp review ratings, and also by determining the prevalence of the topics in Yelp narrative reviews.

## Coding of Topics

LDA topics were labeled independently by co-authors (BLR and RJS) by viewing the top seven terms in each topic. Adjudication of discrepancies occurred via consensus with a third reviewer (RMM). When concordance was possible, LDA topics were assigned to HCAHPS domains (e.g., the topic “pain medications” was placed under the HCAHPS domain Pain Control). Otherwise, topics were assigned to new Yelp domains by co-authors (BLR and RMM). We assigned each LDA topic an example quote by selecting a quote from a list of

the ten Yelp reviews most associated with each LDA topic. Example quotes may have features of multiple topics, as topics are not mutually exclusive, and several topics may be represented in any single quote from a review.

### Hospital Characteristics

We used data from the AHA annual survey to describe hospitals by number of beds, region, teaching-status, and ownership. These hospital characteristics have previously been demonstrated to be associated with HCAHPS scores.<sup>28</sup>

### Statistical Analysis

We used summary statistics to describe hospital characteristics, number of Yelp reviews, and Yelp ratings for hospitals included in the study cohort. In order to determine which LDA Yelp topics were associated with high or low Yelp ratings, we correlated LDA Yelp topics with Yelp ratings by calculating the Pearson product-moment correlation coefficient (Pearson's  $r$ ) between each topic and Yelp review rating, and we calculated Bonferroni corrected two-tailed  $p$ -values. To determine the prevalence of Yelp topics, each Yelp review was assigned the 10 topics most correlated with that review according to the Pearson product-moment correlation coefficient. Prevalence was then calculated based on the percent of Yelp reviews that contained that topic (as measured by the topic appearing in the top 10 topics for a given Yelp review).

We correlated the mean Yelp review rating of each hospital with the HCAHPS overall hospital rating for each hospital and generated Pearson's  $r$ . The number of Yelp reviews per hospital likely affects the correlation; therefore, we calculated the correlation coefficient for different minimum cutoffs of Yelp reviews per hospital. The HCAHPS survey asks patients to give an overall hospital visit rating of 0 (worst) to 10 (best) and reports the results as percent of patients giving a rating of 9–10, percent giving 7–8 and percent giving less than or equal to 6. We generated a composite HCAHPS overall hospital score for each hospital by using a weighted average of the reported HCAHPS scores for overall hospital visit rating. This study was exempt by the University of Pennsylvania Institutional Review Board.

### Study Results

There were 4,681 hospitals in the Hospital Compare dataset and 1,451 of these hospitals had Yelp reviews (with a total of 18,058 Yelp reviews). Of the 4,681 hospitals in Hospital Compare, 4,360 hospitals had AHA data, and of these hospitals, 1,352 (31%) had Yelp reviews. In aggregate, our final cohort of 1,352 hospitals had 16,862 Yelp reviews with a median of four Yelp reviews per hospital (interquartile range (IQR) 2–13 Yelp reviews). The median date of Yelp review was 10/9/2012 (IQR 4/27/2011 to 10/19/2013). The reviews of hospitals displayed a bimodal distribution of star ratings: 31% of reviews gave one star and 33% gave five stars. This bimodal distribution contrasts with the favorably skewed distribution for all Yelp review ratings including non-healthcare reviews (Appendix Exhibit A1).<sup>29</sup> The median of average Yelp rating per hospital was 3.2 (IQR 2.47–4).

## Description of Hospital Cohort

The characteristics of the 1,352 hospitals with Hospital Compare data, AHA data, and Yelp reviews, as well as the number of Yelp reviews by hospital characteristic are reported in Exhibit 1.

## LDA Analysis of Yelp Reviews

After reviewing different potential numbers of topics to generate (as described in the methods), we assigned the LDA model to generate 50 topics from the 16,862 Yelp reviews. There was an average of 10.8 (SD 4.8) topics covered in a Yelp review. Of these 50 topics, seven (14%) were discarded because they were related to names and/or places (example discarded topic: st, hospital, sister, mary's, johns, staff, seton), and two (4%) were discarded because the terms could not be used to produce an obvious topic (example discarded topic: I'm, don't, time, people, give, you're, guy). The remaining 41 topics were classified as related to an HCAHPS domain (9 topics [22%], Exhibit 2) or a new domain not in HCAHPS (32 topics [78%], Exhibit 3). There were 12 new Yelp domains (generated from the 32 topics found in Yelp that lacked an analogous HCAHPS domain): *cost of hospital visit, insurance and billing, ancillary testing, facilities, amenities, scheduling, compassion of staff, family member care, quality of nursing, quality of staff, quality of technical aspects of care, and specific type of medical care.*

## LDA Topics Correlated to Yelp Ratings

Four of the top five (80%) Yelp topics most strongly associated with positive Yelp review ratings were not covered by HCAHPS domains. These topics were: *caring doctors, nurses, and staff* ( $r = 0.46$ ); *comforting* ( $r = 0.29$ ); *surgery/procedure and peri-op* ( $r = 0.23$ ); and *labor and delivery* ( $r = 0.20$ ). These topics relate to the interpersonal relationships of patients with physicians, nurses, and staff with regard to how “caring” or “comforting” they were, or the topics relate to specific service lines of the hospital. Two of the top five (40%) Yelp topics most strongly associated with negative Yelp review ratings, *insurance and billing* ( $r = -0.26$ ) and *cost of hospital visit* ( $r = -0.26$ ), were also not covered by HCAHPS domains (Exhibit 4).

## Prevalence of LDA Topics in Yelp Reviews

The prevalence of the LDA topics in the Yelp reviews is shown in Appendix Exhibit A2.<sup>29</sup> Nine of the top 15 (60%) most prevalent topics are not in HCAHPS. The top 5 most prevalent topics are *caring doctors, nurses and staff* (46% of Yelp reviews); *waiting for doctors and nurses* (43% of Yelp reviews); *friendliness of ED staff* (35% of Yelp reviews); *nice, friendly staff and nurses* (33% of Yelp reviews); and *patient treatment by physician and information provided* (32% of Yelp Reviews).

## Yelp Rating Correlated to HCAHPS overall rating

The mean Yelp rating per hospital was correlated with the HCAHPS overall rating for each hospital. Pearson's  $r$  was 0.50 using the 871 hospitals with a minimum of three reviews per hospital (Exhibit 5). Increasing values of  $r$  could be seen by increasing the minimum number

of Yelp reviews per hospital cutoff, e.g., >5 Yelp reviews per hospital ( $r = 0.56$ ), >15 Yelp reviews ( $r = 0.63$ ), and >25 reviews ( $r = 0.69$ ) (Appendix Exhibit A3).<sup>29</sup>

## Discussion

Seventy-two percent of American internet users report looking online for health information in 2012<sup>30</sup> and 42% report looking at social media for health-related consumer reviews.<sup>31</sup> Meanwhile, only 6% of Americans had heard of the Hospital Compare website as of 2008.<sup>32</sup> This discrepancy raises the possibility that there are important lessons from the information contained in online consumer reviews that can be used to improve current more formal rating systems and increase their use in consumer decision making.

This study has three main findings. First, hospitals with at least three Yelp reviews ( $n=871$ ) have mean Yelp ratings that correlate relatively strongly with an HCAHPS item asking about overall hospital rating ( $r = 0.50$ ). This builds on older data. Bardach et al. showed that high Yelp review ratings correlate with high HCAHPS overall ratings for 270 hospitals with greater than five Yelp reviews. Additionally, Greaves et al. has demonstrated a correlation in England between patient ratings collected on the NHS Choices website and traditional paper based survey responses.<sup>14</sup> Our study's correlation is notable considering we only required a minimum of three Yelp reviews per hospital. Second, Yelp reviews cover far more topics than HCAHPS does. While Yelp reviews include information about seven of the eleven HCAHPS domains, 12 additional Yelp domains not covered by HCAHPS were identified from Yelp topics. Third, many of the Yelp topics most associated with strongly positive or negative Yelp ratings are not covered by HCAHPS nor are a majority of the most prevalent Yelp topics.

The third finding suggests that current hospital ratings based on HCAHPS may be missing the major drivers of patients' overall experience of care. When Yelp users give an overall hospital rating, they presumably write a narrative review about aspects that went into their overall rating. When patients respond to the HCAHPS question asking about overall hospital rating, they may (or may not) be assigning an overall rating based on HCAHPS domains. The Yelp topics may provide a more nuanced view of aspects of hospital quality that patients value. For example, in consumer focus groups that were conducted as part of the development of HCAHPS to identify consumer preferences for hospital quality, "compassion and kindness" of staff was important.<sup>33</sup> Yet "compassion and kindness" is not an item in the HCAHPS survey but instead is broadly reported under "How often did doctors [or nurses] communicate well with patients?" Meanwhile, *compassion of staff* was an important Yelp domain that contained seven of the 41 Yelp topics. While it is possible a rating of communication may include an element of compassion, it is also possible to show a lack of compassion and empathy without traditional communication: "[My boyfriend]...was on the floor crying and screaming in pain. Nurses walked by as if we didn't exist!!! Finally...[t]hey stripped him naked and proceeded to do tests on him with very little empathy. The ER was freezing cold, they didn't even offer him a blanket." Patients who are looking for hospitals with empathetic, compassionate staff might rather visit Hospital Compare and view ratings for "How comforting and empathetic were the hospital staff?"

rather than, or in addition to, “How often did doctors [or nurses] communicate well with patients?”

This is the first study to use automated computational methods to create de novo topics from online hospital reviews and the first study to characterize the content of narrative online hospital reviews in the U.S. Two previous studies that characterized de novo topics from online reviews or tweets used qualitative coding to describe the content of 200 reviews from 20 NHS hospitals<sup>15</sup> or 1,000 tweets to English hospitals.<sup>16</sup> This study examined a large number of reviews from a large number of U.S. hospitals, used an automated method that allows for easier analysis of large datasets, and attempted to identify important domains of patients’ hospital experience by analyzing what patients and their caregivers write about when reviewing their care experiences.

In addition to offering a view across broader domains, online consumer rating platforms offer several advantages: reviews are real-time, without the delay in reporting that HCAHPS experiences. Additionally, narrative reviews, can supplement data from HCAHPS by providing actionable insights for hospitals to improve patient satisfaction. Finally, Yelp is easy to use and consumers are already comfortable with the platform. The over 140 million users who already look up restaurants and businesses on Yelp may be more likely to turn to services such as Yelp to look up hospitals ratings rather than to Hospital Compare.

Despite these advantages, concerns remain about using information from platforms such as Yelp. Unlike HCAHPS, Yelp does not solicit reviews from a random selection of hospitalized patients. Yelp reviews are likely more common from younger and more educated patients and caregivers. Second, Hospitals may have difficulty deciding how to incorporate unsolicited online reviews and how to respond to them. In response to sites like Yelp, several hospitals have begun allowing patients to review doctors directly on the hospital’s website. Hospitals argue that this provides transparency and information to patients and allows hospitals to verify reviews are from actual patients. Critics argue that hospitals are not in a neutral position to determine which reviews to post.<sup>34</sup> Finally, not all hospitals are represented by Yelp reviews, and half of hospitals have four or fewer reviews, which limits usefulness. This concern may be ameliorated in the future if the use of Yelp and similar platforms to rate hospitals continues to increase.

Those who publicly report quality could learn from consumer review websites and consumer review websites could do more to incorporate systematic assessments of providers. A major barrier to the public reporting movement has been engaging consumers in responding to publicly reported quality information.<sup>35</sup> If features from popular review websites such as Yelp could be replicated in more systematic assessments of providers, there is potential to increase patient engagement with them. In fact, the NHS in England has already taken steps in this direction and currently allows consumers to post narrative reviews and star ratings on its public website.<sup>15</sup> Furthermore, Hospital Compare just began displaying star ratings for HCAHPS survey results to make their website more engaging for consumers.<sup>36</sup> Lessons could also be learned from Amazon (amazon.com) on how to curate large numbers of consumer reviews by highlighting the most helpful positive and negative reviews. Collection of the narrative reviews themselves could also be made more systematic.<sup>9</sup> For example, a

random selection of patients could be emailed or texted after discharge to ask for a narrative review of the hospital visit. Natural language processing approaches could then be used in real-time to analyze and incorporate these reviews into online hospital ratings. Consumer websites could also do more to incorporate information from systematic assessments. A step in this direction, Yelp recently partnered with ProPublica to display certain information from Hospital Compare on the Yelp pages of medical facilities.<sup>37</sup>

## Limitations

There are several limitations to this study. Yelp has inherent selection bias in the reporting of hospital experiences; however, our primary goal was to characterize the content of Yelp reviews. Only 1,352 hospitals had Yelp reviews. It is possible that were more reviews posted on hospitals not currently reviewed (primarily small, non-teaching, and Midwestern or Southern hospitals), different domains of patient satisfaction might have been identified. With a median of four reviews per hospital, Yelp reviews for many hospitals are currently too sparse to support robust consumer assessment of an individual hospital's quality. However, analysis of aggregated hospital reviews can still reveal what consumers generally write about (and value) when reviewing hospitals.

## Conclusion

Online consumer-review platforms such as Yelp can supplement information provided by more traditional patient experience surveys and contribute to our understanding and assessment of hospital quality. The content of Yelp narrative reviews mirrors many aspects of HCAHPS, but also reflects new areas of importance to patients and caregivers that may have important implications for policy makers seeking to measure patient experience of hospital quality and hospitals attempting to improve patient satisfaction.

## References

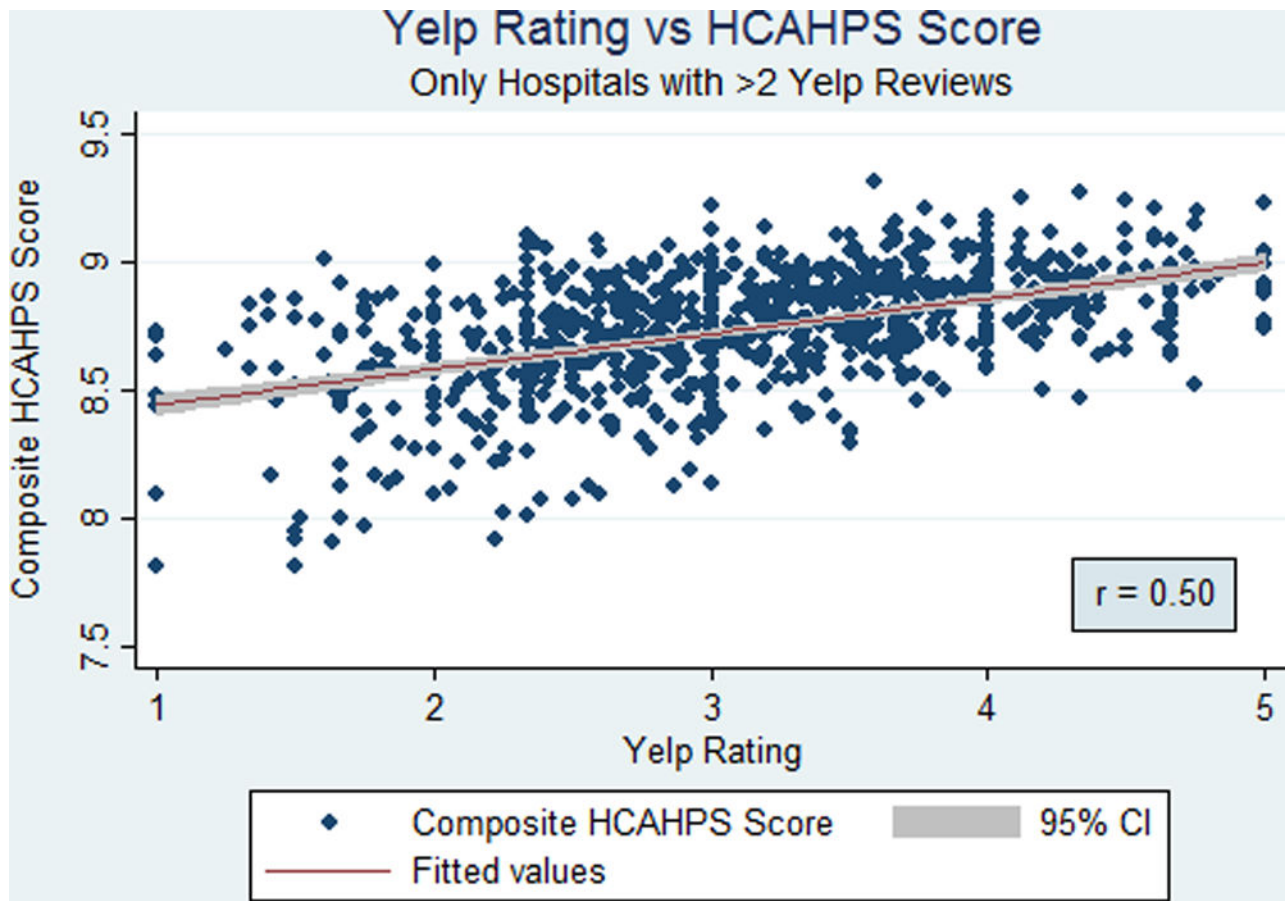
1. Giordano LA, Elliott MN, Goldstein E, Lehrman WG, Spencer PA. Development, implementation, and public reporting of the HCAHPS survey. *Med Care Res Rev*. 2010; 67(1):27–37. [PubMed: 19638641]
2. Goldstein E, Farquhar M, Crofton C, Darby C, Garfinkel S. Measuring hospital care from the patients' perspective: an overview of the CAHPS Hospital Survey development process. *Health Serv Res*. 2005; 40(6 Pt 2):1977–95. [PubMed: 16316434]
3. Centers for Medicare & Medicaid Services. What is Hospital Compare?. Baltimore, MD: Centers for Medicare & Medicaid Services; 2014. [cited 2014 Sep 30]; Available from: <http://www.medicare.gov/hospitalcompare/About/What-Is-HOS.html>
4. 78 FR 50700 (Aug. 19, 2013).
5. 78 FR 51023 (Aug. 19, 2013).
6. Jordan, H.; White, A.; Joseph, C.; Carr, D. Costs and benefits of HCAHPS: final report. Cambridge, MA: Abt Associates Inc; 2005. [Internet][cited 2015 Jan 1]. Available from: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/downloads/HCAHPSCostsBenefits200512.pdf>
7. Centers for Medicare & Medicaid Services. Summary of HCAHPS Survey Results. Baltimore, MD: 2014. [cited 2014 Sep 30]; Available from: [http://www.hcahponline.org/files/Report\\_July\\_2014\\_States.pdf](http://www.hcahponline.org/files/Report_July_2014_States.pdf)



8. Centers for Medicare & Medicaid Services. HCAHPS Public Reporting. Baltimore, MD: Centers for Medicare & Medicaid Services; [cited 2015 Jan]; Available from: [http://www.hcahpsonline.org/files/HCAHPS%20Public%20Reporting\\_April2015\\_Dec2016.pdf](http://www.hcahpsonline.org/files/HCAHPS%20Public%20Reporting_April2015_Dec2016.pdf)
9. Schlesinger M, Grob R, Shaller D, Martino SC, Parker AM, Finucane ML, et al. Taking Patients' Narratives about Clinicians from Anecdote to Science. *N Engl J Med*. 2015; 373(7):675–9. [PubMed: 26267629]
10. Bardach NS, Asteria-Penalosa R, Boscardin WJ, Dudley RA. The relationship between commercial website ratings and traditional hospital performance measures in the USA. *BMJ quality & safety*. 2013; 22(3):194–202.
11. Griffis HM, Kilaru AS, Werner RM, Asch DA, Hershey JC, Hill S, et al. Use of social media across US hospitals: descriptive analysis of adoption and utilization. *J Med Internet Res*. 2014; 16(11):e264. [PubMed: 25431831]
12. Alexa Internet Incorporated. yelp.com. San Francisco, California: Alexa Internet Incorporated; 2014. [cited 2014 Sep 29]; Available from: <http://www.alexa.com/siteinfo/yelp.com>
13. Yelp Incorporated. Overview. San Francisco, CA: Yelp Incorporated; 2014. [cited 2015 Aug 16]; Available from: <http://www.yelp-press.com/>
14. Greaves F, Pape UJ, King D, Darzi A, Majeed A, Wachter RM, et al. Associations between Internet-based patient ratings and conventional surveys of patient experience in the English NHS: an observational study. *BMJ quality & safety*. 2012; 21(7):600–5.
15. Lagu T, Goff SL, Hannon NS, Shatz A, Lindenauer PK. A mixed-methods analysis of patient reviews of hospital care in England: implications for public reporting of health care quality data in the United States. *Jt Comm J Qual Patient Saf*. 2013; 39(1):7–15. [PubMed: 23367647]
16. Greaves F, Laverty AA, Cano DR, Moilanen K, Pulman S, Darzi A, et al. Tweets about hospital quality: a mixed methods study. *BMJ quality & safety*. 2014; 23(10):838–46.
17. Jung Y, Hur C, Jung D, Kim M. Identifying key hospital service quality factors in online health communities. *J Med Internet Res*. 2015; 17(4):e90. [PubMed: 25855612]
18. Dreves F, Hinz V. Who chooses, who uses, who rates: the impact of agency on electronic word-of-mouth about hospitals stays. *Health Care Manage Rev*. 2014; 39(3):223–33. [PubMed: 23727786]
19. Glover M, Khalilzadeh O, Choy G, Prabhakar AM, Pandharipande PV, Gazelle GS. Hospital Evaluations by Social Media: A Comparative Analysis of Facebook Ratings among Performance Outliers. *J Gen Intern Med*. 2015
20. Yelp Incorporated. Recommended Reviews. 2014. [cited 2014 Sep 9]; Available from: [http://www.yelp-support.com/Recommended\\_Reviews/](http://www.yelp-support.com/Recommended_Reviews/)
21. Blei DM, Ng AY, Jordan MI. Latent Dirichlet Allocation. *Journal of Machine Learning research*. 2003; 3:993–1022.
22. McCallum, AK. MALLET: A Machine Learning for Language Toolkit [Internet]. 2002. Available from: <http://mallet.cs.umass.edu>
23. Schwartz HA, Ungar LH. Data-Driven Content Analysis of Social Media: A Systematic Overview of Automated Methods. *The ANNALS of the American Academy of Political and Social Science*. 2015; 659(1):78–94.
24. Schwartz HA, Eichstaedt JC, Kern ML, Dziurzynski L, Ramones SM, Agrawal M, et al. Personality, gender, and age in the language of social media: the open-vocabulary approach. *PLoS ONE*. 2013; 8(9):e73791. [PubMed: 24086296]
25. Brody S, Elhadad N. Detecting salient aspects in online reviews of health providers. *AMIA Annu Symp Proc*. 2010; 2010:202–6. [PubMed: 21346969]
26. Paul, MJ.; Wallace, BC.; Dredze, M., editors. What affects patient (dis) satisfaction? analyzing online doctor ratings with a joint topic-sentiment model. 2013.
27. Wallace BC, Paul MJ, Sarkar U, Trikalinos TA, Dredze M. A large-scale quantitative analysis of latent factors and sentiment in online doctor reviews. *J Am Med Inform Assoc*. 2014; 21(6):1098–103. [PubMed: 24918109]
28. Lehrman WG, Elliott MN, Goldstein E, Beckett MK, Klein DJ, Giordano LA. Characteristics of hospitals demonstrating superior performance in patient experience and clinical process measures of care. *Med Care Res Rev*. 2010; 67(1):38–55. [PubMed: 19638640]
- 29.

To access the Appendix, click on the Appendix link in the box to the right of the article online.

30. Fox, S.; Duggan, M. Health Online 2013. Washington, D.C.: Pew Research Center's Internet & American Life Project; 2013. [cited 2014 Jul 30]; Available from: <http://pewinternet.org/Reports/2013/Health-online.aspx>
31. PricewaterhouseCoopers Health Research Institute. Social media "likes" healthcare: From marketing to social business. Delaware: PricewaterhouseCoopers LLP; 2012. [cited 2014 Jul 30]; Available from: <http://www.pwc.com/us/en/health-industries/publications/health-care-social-media.jhtml>
32. Kasier Family Foundation. 2008 Update on Consumers' Views of Patient Safety and Quality Information: Summary & Chartpack KFF. 2008. [updated 2014 Dec 8]; Available from: <http://kaiserfamilyfoundation.files.wordpress.com/2013/01/7819.pdf>
33. Sofaer S, Crofton C, Goldstein E, Hoy E, Crabb J. What do consumers want to know about the quality of care in hospitals? Health Serv Res. 2005; 40(6 Pt 2):2018–36. [PubMed: 16316436]
34. Ramey, C. Long Island Hospital Posts Doctor Ratings: North Shore-LIJ is the first in New York area, and among the few in the U.S., to do so. The Wallstreet Journal. 2015 Aug 26. Available from: <http://www.wsj.com/>
35. Ketelaar NA, Faber MJ, Flottorp S, Rygh LH, Deane KH, Eccles MP. Public release of performance data in changing the behaviour of healthcare consumers, professionals or organisations. Cochrane Database Syst Rev. 2011; (11):CD004538. [PubMed: 22071813]
36. Centers for Medicare & Medicaid Services. Technical Notes for HCAHPS Star Ratings. 2015. [cited 2015 Oct 19]; Available from: [http://www.hcahpsonline.org/files/HCAHPS\\_Stars\\_Tech\\_Notes\\_9\\_17\\_14.pdf](http://www.hcahpsonline.org/files/HCAHPS_Stars_Tech_Notes_9_17_14.pdf)
37. Yelp Incorporated. Yelp's Consumer Protection Initiative: ProPublica Partnership Brings Medical Info to Yelp. 2015. [cited 2015 Oct 19]; Available from: <http://officialblog.yelp.com/2015/08/yelps-consumer-protection-initiative-propublica-partnership-brings-medical-info-to-yelp.html>



**EXHIBIT 5. Mean Yelp review rating correlated with HCAHPS overall hospital rating by hospital**

SOURCE Authors' analysis of Yelp reviews, American Hospital Association, and Hospital Compare data. NOTES Mean Yelp review rating is correlated with HCAHPS overall rating for hospitals with more than two Yelp reviews ( $n=871$ ),  $p < 0.001$ . Sensitivity analysis with different cutoffs of Yelp reviews per hospital is shown in Appendix Exhibit A3.<sup>29</sup>

**Exhibit 1**

## Characteristics of hospitals with Yelp reviews

	<b>0 Yelp Reviews</b> (N=3008)	<b>1–4 Yelp Reviews</b> (N=689)	<b>5+ Yelp Reviews</b> (N=663)	<b>Total</b> (N=4360)
<b>Bed Size – no. (%)</b>				
0–49	1173 (39)	121 (18)	26 (4)	1320 (30)
50–199	1218 (40)	275 (40)	217 (33)	1710 (39)
200–399	434 (14)	206 (30)	252 (38)	892 (20)
400+	183 (6)	87 (13)	168 (25)	438 (10)
<b>Region – no. (%)</b>				
Northeast	332 (11)	116 (17)	121 (18)	569 (13)
South	1379 (46)	330 (48)	228 (34)	1937 (44)
Midwest	862 (29)	119 (17)	27 (4)	1008 (23)
West	435 (14)	124 (18)	287 (43)	846 (19)
<b>Teaching – no. (%)</b>				
Yes	109 (4)	46 (7)	124 (19)	279 (6)
No	2899 (96)	643 (93)	539 (81)	4081 (94)
<b>Ownership – no. (%)</b>				
Government	857 (28)	112 (16)	76 (11)	1045 (24)
Non-Profit	1697 (56)	430 (62)	477 (72)	2604 (60)
Physician <sup>†</sup>	103 (3)	25 (4)	8 (1)	136 (3)
Proprietary	351 (12)	122 (18)	102 (15)	575 (13)

SOURCE Authors' analysis of Yelp reviews and American Hospital Association data.

NOTES Percentages may not sum to 100 due to rounding.

<sup>†</sup>Physician includes both non-profit and proprietary hospitals.

Percentages may not sum to 100 due to rounding

**Exhibit 2**

Yelp LDA topics associated with HCAHPS domains, presented by HCAHPS domain

HCAHPS Domain (HCAHPS Questions)	Yelp Topic (topic terms)	Example Quote
<p><b>Communication with Nurses</b></p> <ul style="list-style-type: none"> <li>How often did nurses treat you with courtesy and respect?</li> <li>How often did nurses listen carefully to you?</li> <li>How often did nurses explain things in a way you could understand?</li> </ul>	<p><b>Rude doctor/nurse communication</b> * (nurse, asked, told, didn't, doctor, questions, rude)</p>	<p>"I asked a question...[the nurse] snapped at me with a raised voice and very condescendingly asked me if there was anything else I did not understand. No one deserves to be treated like that for a question that is asked to their health care PROFESSIONAL"</p>
<p><b>Communication with Doctors</b></p> <ul style="list-style-type: none"> <li>How often did doctors treat you with courtesy and respect?</li> <li>How often did doctors listen carefully to you?</li> <li>How often did doctors explain things in a way you could understand?</li> </ul>	<p><b>Patient treatment by physician and information provided</b> (patient, medical, hospital, information, treatment, physician, case)</p>	<p>"For my first visit with [the doctor,] I was given a 3 part brochure listing his education, qualifications, his expectations of you as a patient and what you should expect from him as your surgeon. I have never had any Dr., dentist or other health professional give me any information approaching the thoroughness of that brochure."</p>
<p><b>Responsiveness of Hospital Staff</b></p> <ul style="list-style-type: none"> <li>How often did you get help as soon as you wanted after you pressed the call button?</li> <li>How often did you get help in getting to the bathroom or in using a bedpan as soon as you wanted?</li> </ul>	<p><b>Nursing responsiveness</b> (nurse, room, bed, hours, minutes, nurses, hour)</p>	<p>"[My sister] rang for the nurse b/c she had to go to the bathroom-nurse didn't come for 40 min!!! Even after we went to get her she still took another 20 min!!!"</p>
	<p><b>Waiting for doctors and nurses</b> (minutes, room, waiting, doctor, wait, hours, nurse)</p>	<p>"Checked in @ 7:30 AM no procedure until 12:30 PM. Designated echo doctor changed 3 times. Procedures started @ 1:30 PM took 2 hrs, 4.5 hrs for recovery. The wait was tough..."</p>
<p><b>Pain Control</b></p> <ul style="list-style-type: none"> <li>How often was your pain well controlled?</li> <li>How often did the hospital staff do everything they could to help you with your pain?</li> </ul>	<p><b>Pain medications</b> (pain, doctor, nurse, told, medication, meds, gave)</p>	<p>"...[D]o not come here if you have a severe migraine because you will be pumped up full of narcotics, which is not the appropriate treatment. "</p>
<p><b>Cleanliness of Hospital Environment</b></p> <ul style="list-style-type: none"> <li>How often were your room and bathroom kept clean?</li> </ul>	<p><b>Clean, private, nice hospital rooms</b> (room, hospital, rooms, nice, staff, private, clean)</p>	<p>"...the common areas and patient room were clean and decorated smartly (laminated wood floors and tempur pedic mattress for sofa bed!)"</p>
<p><b>Overall Hospital Rating</b></p> <ul style="list-style-type: none"> <li>What number would you use to rate this hospital during your stay?</li> </ul>	<p><b>Horrible hospital</b> (hospital, place, people, don't, worst, treated, horrible)</p>	<p>"This place is terrifying. I'd rather bleed to death."</p>
	<p><b>Great care</b> (wife, hospital, staff, care, good, great, time)</p>	<p>"My wife gave birth here and the attention and service was excellent..."</p>
<p><b>Likelihood to Recommend</b></p> <ul style="list-style-type: none"> <li>Would you recommend this hospital to your friends and family?</li> </ul>	<p><b>Reviewing hospital experience</b> (hospital, review, experience, ive, im, years, time)</p>	<p>"If you have a choice, go anywhere else. The people, the food, the quality of care is abysmal."</p>

SOURCE Authors' analysis of Yelp reviews

NOTES

\* Also under communication with doctors

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

## Exhibit 3

Yelp LDA topics fitting under new domains not found in HCAHPS \*

Yelp Domains	Yelp Topic (topic terms)	Example Quote
Cost of hospital visit	<b>Cost of hospital visit</b> (insurance, bill, pay, cost, charge, visit, hospital)	"Ask what you will be charged before you allow this place to touch you! ... [T]heir charges are excessive. In hospital 18 hours and got a 36k bill for knee surgery that cost 1/3 the amount in Los Angeles."
Insurance and Billing	<b>Insurance and billing</b> (insurance, billing, bill, hospital, department, company, paid)	"Billing department is awful. They billed the wrong insurance company twice and then finally had the correct one but billed it incorrectly... Instead of contacting me they sent the bill to a collection agency..."
Ancillary testing	<b>Ancillary testing and results</b> (blood, test, doctor, results, tests, lab, work)	"Service way too slow...for relatively standard testing...90 mins so far and sample hasn't even been collected because doc needs lab to tell him what tests need to be ordered."
Facilities	<b>Facility</b> (place, big, hand, time, TV, watch, cool)	"This place is awesome...seek out the quite courtyards for some solitude."
	<b>Security and front desk</b> (front, security, desk, room, friend, back, waiting)	"There was no sign that said to check-in, so I didn't bother with going to the information desk...The security guard who told us to check-in was... rude..."
Amenities	<b>Parking</b> (parking, hospital, building, free, lot, nice, valet)	"2 hours free parking with validation... I like it how they are building a garden between the parking structure and the hospital."
	<b>Hospital food</b> (food, cafeteria, hospital, good, eat, coffee, order)	"This review is for the cafeteria, prices are super cheap and the burger and nachos were terrific. Chili cheese fries were soggy..."
Scheduling	<b>Scheduling appointments</b> (appointment, call, doctor, called, phone, time, back)	"The scheduling office will NOT schedule you for an appointment. They ask you to leave a message on the voice mail of the department. The department REDIRECTS you back to scheduling. Four hours of my life spent trying to make appointments."
Compassion of Staff	<b>Comforting</b> (time, made, feel, make, felt, didn't, comfortable)	"Best... staff ever! Treated us Like family! Thank you to all the gals that took away the anxiety fear and pain when I welcomed my princess into the world!"
	<b>Caring doctors, nurses, and staff</b> (care, staff, nurses, hospital, doctors, great, caring)	"From the cleaning crew to aides to nurses to room service to midnight vampires to doctors, everyone was fabulous. Always cheerful, kind, warm, gentle hands and hearts and good ear."
	<b>Wonderful medical staff and care</b> (dr, nurse, great, staff, care, wonderful, amazing)	"Wonderful staff who took great care of mom while comforting and reassuring me mom will be ok"
	<b>Friendliness of ED staff</b> (er, staff, friendly, emergency, visit, room, experience)	"ER took me over 6 hours on quiet Sunday afternoon. The people here do not value your time and are not friendly. Avoid!"
Family Member Care	<b>Family member care</b> (mom, hospital, mother, dad, father, family, care)	"Came here a few times but only to visit my grandma. From staying there, the nursing staff was very friendly and helpful to my grandma and knowing this I knew my grandma would recover a lot quicker."
Quality of Nursing	<b>Nursing tasks</b> (nurse, room, iv, blood, floor, put, finally)	"Trash and blood was all over the floor. The nurse came back in the room and picked up the trash on the floor (no gloves) then came to draw blood WITHOUT gloves on."
Quality of staff	<b>Nursing and staff patient-care</b> (nurses, staff, family, patient, patients, care, member)	"...[T]here are NO...certified social workers aboard on weekends, but scores of...un-certified social workers on weekends...These young and very inexperienced[d]...un-certified social workers interact with patients by interrogating patients' personal details as they try to gain patients' friendship when they'll disappear never returning to your unit for follow-up visits..."
Quality of technical aspects of care	<b>Doctor care/treatment</b> (doctor, doctors, dr, care, clinic, treatment, specialist)	"My surgeon only performs thyroidectomies (that's a rare, awesome thing in the ENT world) and I actually get stopped by other ThyCa survivors, curious about my neat, discreet scar."
Specific type of medical care	<b>Pediatric</b> (son, daughter, hospital, child, kids, time, children's)	"My goddaughter was treated at [hospital] for a congenital heart defect... [The hospital] has gone above and beyond for my family in every respect possible, and its doctors and nurses have... the utmost care and professionalism."

Yelp Domains	Yelp Topic (topic terms)	Example Quote
	<b>Labor and delivery</b> (baby, birth, nurses, labor, delivery, experience, nurse)	“Excellent for having natural childbirth, they respected all my wishes. I had a midwife ... and my own doula.”
	<b>Surgical management and physical therapy</b> (dr, physical, back, surgery, mri, therapy, year)	“This is one of the best rehab hospitals in the world. First class service by therapists, doctors, nurses, aids dietary, housekeeping. They taught me to walk again after total knee replacement.”
	<b>Wait time in ED</b> (room, emergency, er, waiting, hours, wait, time)	“They suck!!!!!! They will make u wait for 3 hrs in the emergency waiting room even though the waiting room is empty!!!!!!”
	<b>ED imaging after injury</b> (er, broken, back, accident, x-ray, foot, car)	“Their ER is awesome! I came in with a possible broken toe. They took me in right away, x-rayed my foot and diagnosed it (yup, broken toe)... 1.5 hours from arrival to dismissal. Best ER service Ive ever experienced.”

SOURCE Authors’ analysis of Yelp reviews.

NOTES

\* Topics may fit under multiple domains, and only select topics are shown for each domain.

**Exhibit 4**

Top 5 LDA yelp topics correlated to high and low yelp ratings \*

Topic Name	Correlation <sup>†</sup> , r	Covered by HCAHPS
Yelp topics most correlated with positive Yelp ratings		
Caring doctors, nurses, and staff	0.46	No
Comforting	0.29	No
Clean, private, nice hospital rooms	0.25	Yes
Surgery/procedure and peri-op	0.23	No
Labor and delivery	0.20	No
Yelp topics most correlated with negative Yelp ratings		
Horrible hospital	-0.33	Yes
Rude doctor/nurse communication	-0.29	Yes
Pain Control	-0.28	Yes
Insurance and billing	-0.26	No
Cost of hospital visit	-0.26	No

SOURCE Authors' analysis of Yelp reviews.

## NOTES

\* If a topic was similar to an already listed topic it was removed and replaced by the next most associated topic.

<sup>†</sup>All correlation coefficients are statistically significant at  $P < 0.0001$ .