

# “Told”: the Word Most Correlated to Negative Online Hospital Reviews

Anish K. Agarwal, MD MPH<sup>1,2,3</sup>, Arthur P. Pelullo, MS<sup>3</sup>, and Raina M. Merchant, MD MSHP<sup>1,2,3</sup>

<sup>1</sup>Department of Emergency Medicine, University of Pennsylvania, Philadelphia, PA, USA; <sup>2</sup>National Clinician Scholars Program, University of Pennsylvania, Philadelphia, PA, USA; <sup>3</sup>, Penn Medicine, Center for Digital Health, Philadelphia, PA, USA.

J Gen Intern Med 34(7):1079–80  
DOI: 10.1007/s11606-019-04870-6  
© Society of General Internal Medicine 2019

## INTRODUCTION

Online review platforms have the ability to amplify an individual's voice. The emergence of these platforms provides an open and public forum for patients to express their opinions which may help redefine how health systems approach the patient experience.

With over 140 million reviews, Yelp! ([www.yelp.com](http://www.yelp.com)), the online review platform, continues to grow in scale and popularity. Patients have begun to share their opinions and ratings with over 8.4 million healthcare related reviews.<sup>1, 2</sup> Though online reviews of individual providers may be low yield,<sup>3</sup> these reviews may expose blind spots not traditionally considered by hospitals. It is the observability and scale of these online reviews that provides an opportunity for both innovation and promise.<sup>4</sup>

Our objective was to analyze hospital reviews posted on Yelp through the application of machine learning and natural language processing. We hypothesized there would be common language correlated with negative and positive reviews.

## METHODS

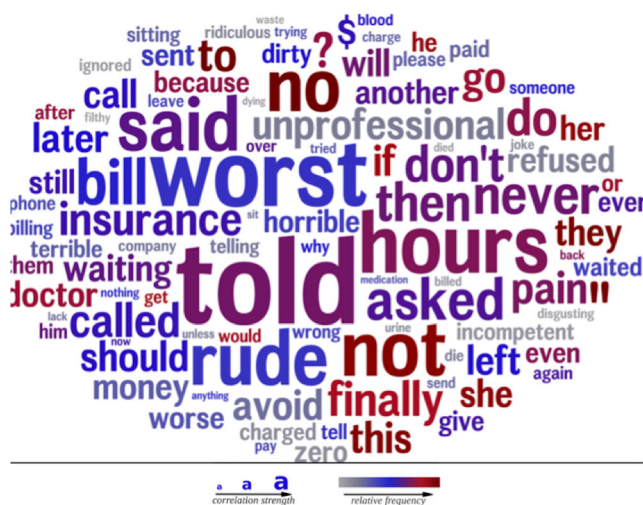
This was a retrospective analysis of US hospital reviews on Yelp over 12 years (May 2005–March 2017). Hospital reviews were identified using the facilities reporting data to the American Hospital Association. Each review included narrative text, date, and star rating. Reviews were of facilities and not individual providers. This study was considered exempt by the University of Pennsylvania Institutional Review Board. A machine learning tool, Differential Language Analysis (DLA) was used to identify word frequency and correlation with 1-star (negative) or 5-star (positive) reviews. DLA provides a method of open-vocabulary analysis to determine correlation, using frequency and Pearson's correlation, among individual words and star rating.

## RESULTS

Over 12 years, there were 51,376 reviews from 1566 hospitals. 41.9% were 1-star and 31.6% were 5-stars (mean 2.84 stars). “Told” was the single most correlated word among 1-star reviews, present in 9578 reviews (18.6%) with an average of 1.79 stars. Figure 1 displays a word cloud of all words correlated with 1-star reviews with color representing frequency and size representing correlation. An example of narrative text, “I constantly **told** them that none of that was true and the nurse there wouldn't believe me.” The words “great” and “friendly” were the words most correlated with 5-star reviews (Fig. 2). “Friendly” was present in 5594 reviews (10.9%) with an average of 4.02 stars. An example of narrative text, “the doctors, nurses and x-ray technician who helped me out were all so cool and friendly. It really restored my faith in humanity after I got hit on my bike.”

## DISCUSSION

This study has two key findings related to online hospital reviews and their content. First, the single word most



**Fig. 1** Words most correlated with negative (1-star) hospital reviews on Yelp. Using Differential Language Analysis (DLA), hospital reviews posted on Yelp were analyzed. The color gradient displays weak or strong correlations to the negative reviews. Gray words have weak correlation with a negative review and darker words, as displayed in gradient key, indicate higher correlation. Text size indicates frequency of word use within reviews—larger text size represents a word used more frequently



Fig. 2 Words most correlated with positive (5-star) hospital reviews on Yelp. Using Differential Language Analysis (DLA), hospital reviews posted on Yelp were analyzed. The color gradient displays weak or strong correlations to the negative reviews. Gray words have weak correlation with a positive review and darker words, as displayed in gradient key, indicate higher correlation. Text size indicates frequency of word use within reviews—larger text size represents a word used more frequently

correlated to negative reviews was “told.” Second, the most correlated words with positive reviews were “great” and “friendly.” Machine learning and natural language processing offer an approach to analyzing the large amount of narrative text within reviews and begin to shed light on issues around communication and interactions between patients and providers.

The growth and expansion of online narrative reviews provides a rich and large data source. There are clearly marked differences and limitations between the standardized national surveys and unfiltered online reviews. Online reviews are not validated and individuals may have external motivations to intentionally drive ratings of a hospital or organization. Furthermore, these reviews are applied at the hospital level and thus cannot be generalized to the individual provider.

Distilling the vast amount of narrative responses from millions of reviews would prove to be onerous and impractical. Though individuals are reviewing and commenting on their unique encounters, the collective themes within a region, system, or specialty may highlight shared experiences.

As healthcare transforms, there has been a focus on becoming more patient-centered. Patient-centered care requires providers to effectively communicate with their patients: to ask, to listen, and to explain.<sup>5</sup> Providers must be cognizant of the words and language they use.<sup>6</sup> Teaching the ability to communicate through caring language is crucial for healthcare leaders and educators.

Patients have already thought differently about how to make their voices heard by taking to online platforms to review their experiences. In an effort to become patient-centered, providers must begin to think differently about how to listen and communicate. Capturing digital patient generated data through methods such as machine learning can provide an approach to improving the patient experience.

**Acknowledgments:** We thank Lyle Ungar, PhD, for assistance with machine learning analyses.

**Corresponding Author:** Anish K. Agarwal, MD MPH; National Clinician Scholars Program, University of Pennsylvania, 423 Guardian Drive, 1310 Blockley Hall, Philadelphia, PA 19104, USA (e-mail: anish.agarwal@uphs.upenn.edu).

**Compliance with Ethical Standards:**

**Conflicts of Interest:** The authors declare that they do not have a conflict of interest.

**Publisher's Note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## REFERENCES

1. Factsheet [Internet]. Yelp. [cited 2017 Nov 21]; Available from: <https://www.yelp.com/factsheet>. Last accessed 9 Nov 2018.
2. Merchant RM, Volpp KG, Asch DA. Learning by Listening-Improving Health Care in the Era of Yelp. *JAMA* 2016;316(23):2483-4.
3. Lagu T, Metayer K, Moran M, et al. Website Characteristics and Physician Reviews on Commercial Physician-Rating Websites. *JAMA* 2017;317(7):766-8.
4. 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 Proj Hope* 2016;35(4):697-705.
5. Selby JV, Lipstein SH. PCORI at 3 years—progress, lessons, and plans. *N Engl J Med* 2014;370(7):592-5.
6. The Learning Healthcare Project. The Learning Health Care Project [Internet]. [cited 2017 Nov 30]; Available from: <http://www.learninghealthcareproject.org/section/background/learning-healthcare-system>. Last access 9 Nov 2018.