



2019

Beneath the surface of talking about physicians: A statistical model of language for patient experience comments

Taylor Turpen

NarrativeDx, tad@narrativedx.com

Lea Matthews MD

UCHealth, lea.c.matthews@gmail.com

Senem Guney PhD, CPXP

NarrativeDx, senem@narrativedx.com

Follow this and additional works at: <https://pxjournal.org/journal>

 Part of the [Computational Linguistics Commons](#), [Discourse and Text Linguistics Commons](#), [Health Communication Commons](#), [Health Information Technology Commons](#), [Health Services Administration Commons](#), and the [Health Services Research Commons](#)

Recommended Citation

Turpen, Taylor; Matthews, Lea MD; and Guney, Senem PhD, CPXP (2019) "Beneath the surface of talking about physicians: A statistical model of language for patient experience comments," *Patient Experience Journal*: Vol. 6 : Iss. 2 , Article 10.

DOI: 10.35680/2372-0247.1276

Available at: <https://pxjournal.org/journal/vol6/iss2/10>

This Research is brought to you for free and open access by Patient Experience Journal. It has been accepted for inclusion in Patient Experience Journal by an authorized editor of Patient Experience Journal.

Beneath the surface of talking about physicians: A statistical model of language for patient experience comments

Cover Page Footnote

The authors would like to thank Paul Rosen, MD, Associate Professor of Pediatrics, Jefferson Medical College for bringing us together as a research team to undertake this project. This article is associated with the Innovation & Technology lens of The Beryl Institute Experience Framework. (<http://bit.ly/ExperienceFramework>). You can access other resources related to this lens including additional PXJ articles here: http://bit.ly/PX_InnovTech

Beneath the surface of talking about physicians: A statistical model of language for patient experience comments

Taylor Turpen, *NarrativeDx*, tad@narrativedx.com

Lea Matthews, MD, *UCHealth*, lea.c.matthews@gmail.com

Senem Guney, PhD, *CPXP*, *NarrativeDx*, senem@narrativedx.com

Abstract

This study applies natural language processing (NLP) techniques to patient experience comments. Our goal was to examine the language describing care experiences with two groups of physicians: those with scores in the top 100 and those with scores in the bottom 100 among all physicians (n=498) who received scores from patient satisfaction surveys. Our analysis showed a statistically significant difference in the language used to describe care experiences with these two distinct groups of physicians. This analysis illustrates how to apply NLP techniques in categorizing and building a statistical model for language use in order to identify meaningful language and significant phrasing in a dataset of natural language. We provide a review of limited work at the intersection of language analysis and patient experience. We present our analysis and conclude with a discussion on what care providers and patient experience leaders can learn from language used in patient experience comments for the delivery of patient-centered care.

Keywords

Patient experience, language data, patient satisfaction surveys, natural language processing (NLP)

Introduction

Healthcare providers and patient experience leaders are very attentive to patients' responses to structured or closed-ended survey (check-box) questions that measure patients' perceptions of their care. Patient experience comments, however, provide a much richer source of data than surveys in helping us understand the patient's perspective on care experiences. The patient's perspective is not only necessary to understand and improve patient experience,¹ but it is also integral to providing patient-centered care.^{2,3}

Patient experience is a multidimensional concept. Wolf, Niederhauser, Marshburn, and LaVela provide a literature review of how patient experience has been defined.⁴ The Beryl Institute defines patient experience as "the sum of all interactions shaped by an organization's culture that influence patient perceptions across the continuum of care."⁵ Additional patient experience concepts in the literature include emotional and physical lived experience, personal interactions, importance of partnership/patient involvement, perception, information, and responsiveness.

Interactions between patients and their families and physicians are a significant component of patient experience. We undertook this study to reveal insights into the language used to describe care experiences with physicians. Specifically, we wanted to advance the

understanding of the language used in accounts of care experiences with two groups of physicians: those who performed comparatively well and those who performed comparatively poorly among all physicians who received scores from patient satisfaction surveys. We applied natural language processing techniques to test our hypothesis that patients use different language when describing their experiences with these two distinct groups of physicians. Our data set came from unstructured/open-ended responses in patient satisfaction surveys conducted at one of the nation's largest integrated pediatric health systems that covers five states in the US.

This research illustrates a specific application of statistical natural language processing techniques to natural language (comment) data in patient experience as a specific domain. While probabilistic modeling of natural language requires a leap of faith to advance knowledge for a particular domain, a solid statistical analysis can be used to identify meaningful words and significant phrasing.⁶ The output of probabilistic modelling is usually a prediction that has obscure meaning, while a statistical analysis can be used empirically as a first step to understand natural language data better.

Our methodology followed two pre-processing steps and a two-step evaluation for determining whether or not people use specific wording when they describe their care experiences with physicians. To test our hypothesis, we

first evaluated whether the language we extracted through pre-processing was relevant to three categories of patient experience. Our second evaluation step answered whether the language used to describe care experiences was statistically correlated to a) the comments for physicians who scored comparatively well; b) the comments for physicians who scored comparatively poorly in patient satisfaction surveys among all physicians surveyed or c) the comments for both groups of physicians.

In the following sections, we will first go over the research to date that involves analysis of comments from surveys that collect patient feedback. Next, we will describe the data set we used and the methods of analysis we applied in this study. Following the description of our methodology, we will discuss our results, present our conclusions and suggest opportunities for future research.

Related Work in Evaluating Patient Experience Feedback and Natural Language Processing

Evaluating Patient Experience Feedback

Responses to open-ended questions (or comments) in patient satisfaction surveys provide a narrative outlet for patients to give detailed descriptions of their care experiences. While the sheer volume of information contained in these comments can be a deterrent to analysis, the use of technology can help streamline the analytical process.⁷ Besides comments in surveys, online review sites create a significant narrative outlet and resource, as 60% of people view user-generated information when searching the internet for healthcare information.⁸ For these reasons, efficient and consistent analysis of patient experience comments is foundational to the design and implementation of patient-centered programs for the improvement of patient experience and delivery of patient-centered care.⁹

Greaves et al. compared Likert-type scale ratings to patients' comments that were analyzed using sentiment analysis.¹⁰ For the categories of cleanliness, being treated with dignity, and overall recommendation of hospital, there was 81%, 84%, and 89% agreement respectively between scale ratings and sentiments extracted from comments.^{10,11} Wiseman et al. used framework analysis and two independent researchers to develop their specific positive and negative frameworks and themes to analyze 15,403 comments.¹² Their work was used to compare comments from two separate years in order to assess if there was an improvement in care.¹³ In the pediatric patient population, Espinel, Shah, Beach, and Boss analyzed 195 comments from pediatric otolaryngologists and surgical specialists using summative content analysis and a coding taxonomy developed by two investigators.¹⁴ Our research expands on this work with a larger sample size of patient experience comments from family members

of pediatric patients. Our analysis illustrates some of the methodological principles for building the analytical capability to detect linguistic nuances in patient experience comments.

The six core elements of providing high-quality care - safe, effective, patient-centered, timely, efficient, and equitable - have also been influential in defining patient experience.¹⁵ Our decision to analyze comments about care experiences under the categories of personality, interaction, and the composite category of quality, safety and access was grounded in these concepts and definitions from the literature.

Natural Language Processing

Natural language processing (NLP) is the process of converting raw, unstructured text or speech (in our case, comments from patient satisfaction surveys) into a structured representation of the natural language using some model of language. Frequently, the structured representation of meaning appears as an annotation or category assigned to the text by a human annotator. Comment categorization requires that annotators assign a category to each natural language sample based on their intuition of what the sample is about. To make categorization simpler, the comment is usually split up into sentences, phrases or word combinations known as ngrams. A single word combination is known as a unigram, two words together are known as a bigram and three words are known as a trigram. In this pre-processing of data, common words like "that," "these," "which," or "so" (stopwords) are usually not included in the ngram.

Natural language categorization and models can be improved by categorizing or modeling significant amounts of natural language data. Identifying patient experience language from a limited amount of language (beyond stopwords) is difficult. While the term frequency inverse document frequency (tf-idf) statistic is helpful on its own for the identification of significant terms in collections of documents,¹⁶ it has also been used to filter out insignificant features in classification tasks. For example, tf-idf, followed by the Pearson correlation statistic, was used to correlate the tweets with the real world experience of the Soccer World Cup 2014.¹⁷ Not only has tf-idf been useful for the identification of significant terms in different types of data, it can also be the first step for creating input into more advanced probabilistic models to achieve high predictive accuracy in document classification tasks.¹⁸ Our work applies tf-idf on the language that family members of pediatric patients use to describe care experiences. We then perform an in-depth statistical analysis on the tf-idf output. Our analysis illustrates how tf-idf could be used on similar datasets or as input to more advanced probabilistic models of language use.

Methodology

Data Collection

Our data for this study came from 25,161 anonymous patient satisfaction surveys that were collected throughout 2014 via mail and email, representing 498 providers in 42 specialties from 252 office departments/locations affiliated with the pediatric health system. Given our objective to learn about the language used to describe care experiences with physicians, we sampled the data based on the criteria that 1) the provider was a physician and 2) the comment section was not blank and was eligible for analysis. 18,401 out of 25,161 surveys met these criteria. We then categorized the surveys by their “top box scores.” Top box scores represent the percentage of respondents who give the highest score possible on the survey scale. This metric has been influential to understand patients’ perceptions of care experiences, including patients’ and family members’ experiences with physicians. For this reason, in our comparative analysis, we used top box scores to organize comments about physicians into two groups: comments for “top performing physicians” and comments for “bottom performing physicians.” For the purposes of this analysis, we defined top performing physicians as those whose top box scores placed them in the ranking of the top 100 among all the physicians surveyed (n=498). Similarly, bottom performing physicians received top box scores that placed them in the ranking of the bottom 100 among all the physicians surveyed.

Data Analysis

We conducted this study to test whether or not family members of pediatric patients use specific language in their comments about their care experiences with physicians. The assumption underlying our research design was that care experiences shaped the language used in patient experience comments. Our experimental setup assumed that if the language from the comments about top performing physicians differed significantly from the language about bottom performing physicians, then we could conclude that patient and family experience had some effect on the language used.

Our methodology involved two pre-processing steps to prepare the data for annotation. As the first pre-processing step, we extracted a list of word combinations (ngrams - excluding stopwords) from the comments about top and bottom performing physicians. Next, we filtered down the list of ngrams using the tf-idf statistic with respect to the comments about top and bottom performing physicians, putting them into a bag of ngrams. Tf-idf takes the term frequency (the number of times that a term appeared in a comment) of term i in document j and multiplies it by the log of the inverse ratio of the number of times that the term appeared in a document, yielding the formula:

$$w_{i,j} = tf_{i,j} \times idf_i$$

The bag of ngrams consisted of the top 50 unigram (one-word), bigram (two-word) and trigram (three-word) combinations based on the tf-idf statistic. This resulted in a total of 300 ngrams to be categorized by annotators.

Next, two annotators independently reviewed all the ngrams and assigned them to one of the three patient experience categories of personality, interaction, and the composite category of quality, safety and access. The descriptions for these categories and guidelines for the annotators were developed by the authors with extensive domain expertise and understanding of the relevant literature. The annotators were provided with annotation guidelines including the following definitions: Personality is a character trait that describes the care provider’s attitude in the care encounter. Interaction is commonly a past-tense description of a care encounter that happened between the provider and the patient and/or family member. Quality, safety and access is a composite category that covers care encounters specifically about quality, safety, or access issues in care experiences. If the ngram could not be assigned to any of these patient experience categories, the annotators were told to assign it to other. The two annotators adjudicated the annotations for those ngrams that got conflicting categorizations.

We then calculated the number of times that the 300 ngrams occurred in all the comments and the percentages shown in Table 1. Next, we evaluated whether or not the subset of ngrams that belonged to the patient experience categories of personality, interaction, and quality, safety and access occurred more often than those ngrams that belonged to other.

Because a large percentage of the patient experience ngrams were relevant to the patient experience, we proceeded with a statistical significance test of all ngrams from the comments about the top and bottom performing physicians. For the 300 ngrams, we performed a Pearson chi-square statistical test for ngram relevance based on the number of times that a ngram appeared in the comments for top performing physicians versus the number times that it appeared in the comments for bottom performing physicians. Those ngrams with significance (p) values less than .001 and corresponding X2 values were considered to be significant given 1 degree of freedom. We determined that a significance level of .001 was appropriate given the large number of unique ngrams and ngram counts.

Results

Of the 100 annotated unigrams that occurred 26,664 times in the comments, 24.86% belonged to one of the three patient experience categories. Of the 100 bigrams that occurred 2,269 times in the comments, 86.21% belonged to the patient experience categories. Of the 100 trigrams

Table 1. Ngram Occurrence Percentages by Category

Totals	Unigram	Bigram	Trigram
Other	75.14%	13.79%	7.49%
Interaction	12.52%	24.20%	37.31%
Quality, Safety and Access	10.83%	41.30%	20.05%
Personality	1.50%	20.71%	35.15%

that occurred 788 times in the comments, 92.51% belonged to the patient experience categories. The ngram occurrence percentages by category are shown in Table 1. When using the ngram counts from the comments about top performing physicians, 52 ngrams were considered to be significant ($p < .001$). When using the counts from the comments about bottom performing physicians, 49 ngrams were considered to be significant ($p < .001$). There was an overlap of 36% between the ngrams that were significantly correlated to the comments for top performing physicians and the ngrams that were significantly correlated to the bottom performing physicians. Overall, 63 ngrams out of the 300 were uniquely correlated to comments for either bottom or top performing physicians ($p < .001$). Those ngrams are shown in Table 2.

The annotators were able to assign a significant number of ngrams to the patient experience categories. The total ngram occurrences for these categories are shown in Table 2. Only the *other* category was significantly correlated to comments for both the top and bottom performing physicians ($p < .001$). Roughly a quarter of the unigram occurrences were relevant to patient experience. The vast majority of the annotated bigram and trigram occurrences, 86.21% and 92.51% respectively, were relevant to patient experience. The tf-idf statistic could bias the Pearson correlation by

selecting for ngrams that are already statistically significant. We determined that filtering through tf-idf on unique ngrams from the comments about top and bottom performing physicians correctly selected for word combinations relevant to our patient experience categories, because the bag of ngrams was annotated without any consideration for physician performance in patient satisfaction surveys. Our results also showed that 63 of the ngrams were exclusively correlated to language from the comments about either top or bottom performing physicians. Because many of the ngrams were uniquely correlated to the comments about top and bottom performing physicians at the .001 level, we conclude that language use in patient experience comments is specific to care experiences with physicians.

Discussion

The reimbursement for services based on the quality of care has accelerated initiatives within healthcare organizations to provide patient-centered care. As part of these initiatives, patient experience has become a critical component of measuring care quality.³ Standardized surveys with structured/closed-ended (check-box) questions helped the healthcare industry to establish a common platform for measuring performance in patient experience.

Table 2. Ngrams that are Significantly Correlated with Comments for Top and Bottom Performing Physicians

Bottom	Top
years, health, makes, takes, treatment, enough, experience staff, good thanks, appointment day, wait exam room, called schedule appointment, nurse brought us, 30 minutes waiting, staff always nice, always professional caring, wait time exam, nice great job, would recommend anyone, could ask better, one best doctors, make feel like, thorough explained everything, everyone always friendly, office staff friendly, wonderful care provider, love love love, spends lot time, thorough took time, make child feel, god bless always, visit student comes	left, still, seemed, wasn, issues, everyone nice, staff amazing, back exam room, made eye contact, called us back, get called back, never feel like, always call back, arrived early taken, every time go, _____office, brought us back, ladies front desk, look forward seeing, time exam room, remember name great, offered wheelchair daughter, head neurology department, everyone office always, staff made sure, department today disappointed, anything else needed, staff friendly good, took time us, would recommend practice, staff friendly helpful, ¹

Patients also provide their feedback about their care experiences through comments on surveys and reviews on social media and doctor review sites. It is imperative for providers, and specifically for patient experience leaders, to process and learn from large volumes of this particular type of natural language (verbatim) feedback for multiple reasons. First, patients who give perfect top box scores to their providers in surveys might give negative feedback when given the opportunity to describe their care experiences in their own words.^{19,20} Gallan et al show the significant contradictions between quantitative scores and qualitative responses in patient experience data. They argue that patient experience professionals should factor the analysis of both types of data into their improvement efforts for issues specific to different phases of care experiences. Also, regardless of potential discrepancies between quantitative and qualitative patient experience data, only through the analysis of verbatim patient feedback can providers gain insights into the root causes behind why they receive the scores they do in surveys.⁹

While standardized surveys have been a great step forward in the reliable measurement of patient experience for patient-centered care, these surveys present significant limitations for providers to engage in effective and timely service recovery. As healthcare providers continue to deal with the limitations of CAHPS (Consumer Assessment of Healthcare Providers and Systems) surveys, they look for complementary mechanisms to collect and learn from patient feedback. Social media provide opportunities for patients and families to share their feedback in their own words, which offsets the limitations of long, structured surveys to understand patients' perceptions of care.²¹ Research has also shown how Yelp reviews can provide not only sufficient but even better insights than CAHPS surveys into patient experience.²² Patients and family members increasingly use social media sites like Yelp and physician review sites as outlets for sharing verbatim feedback.^{21,22,23,24} This has led to research that lays the scientific foundation for how to incorporate open-ended questions for collecting verbatim patient feedback within the bounds of standardized surveys.²⁵ It has also increased the need for the analytical capabilities necessary to manage and learn from natural language or narrative data.^{25,26,27} As the emphasis shifts from check-box questions to comments for collecting patient feedback, patient experience leaders find themselves with the responsibility to identify meaningful trends and insights from large volumes of patient experience comments for data-driven decision making.

This study showed that patients use specific language to describe care experiences. Our analysis focused on how patients phrase their accounts of care experiences with top performing physicians differently as compared to their accounts about bottom performing physicians. Based on our findings, what types of insights can be gained into the

care experiences of patients? The following part of our discussion will highlight these insights.

Building Blocks for an Evidence Basis of Language Data

Our study described a statistically sound approach for the analysis of natural language (comment) data in patient experience. Our analysis revealed a significant correlation between patients' experiences with physicians and how patients talk about physicians. Our study showed the fundamental methodology of validating categories for the NLP-based analysis of language data and resulted in the validation of the three high-level categories of personality, interaction, and the composite category of quality, safety and access. Patient experience leaders can use this knowledge to evaluate what analytical capabilities they need in order to have an evidence basis of patients' language about care experiences. In order to process verbatim patient feedback from multiple, disparate platforms, this work needs to be extended to many more categories of patient experience.

In future applications of the analytical technique described in this paper, a domain expert can evaluate ngrams for categorization, meaning, and phrasing. A domain expert/patient experience leader can also use our categorization, meaning and phrasing as a statistically informed starting point for evaluating verbatim feedback from patients. By using our automated analysis, providers can efficiently and accurately determine the statistically significant mentions of issues about care services. For example, the trigram "office staff friendly" in Table 2 was found to be significant in the comments for bottom performing physicians and it was annotated under the category of personality. This indicates that when a physician is underperforming from the patient's perspective, patients tend to talk more about the personality of the physician's staff rather than the personality of the physician himself or herself. Patient experience leaders can use insights like this to give evidence-based coaching on patient-provider communication.

Taking Improvement Actions based on What Patients Say

Our analysis illustrates how to focus on the most significant phrasings in comments that belong to distinct patient experience categories. In our dataset, mentions of quality, safety and access ngrams like "30 minutes waiting" and "wait exam room" indicate that wait times are more important to patients of the bottom performing physicians. Patients' language use implies that their tolerance for wait times goes up if they perceive their physicians to provide above average care experiences. To further confirm this observation, it is important to note that none of the unique ngrams from the comments about the top performing physicians mention waiting. Such

understanding of how patients describe their experiences would guide patient experience professionals to 1) know how to look for evidence (by knowing significant phrasings) in patients' language about specific aspects of patient experience, in this instance the experience of waiting, and 2) to take improvement actions that are grounded in this evidence.

It should also be noted that several ngrams from the comments about the bottom performing physicians were positive in sentiment, such as "love love" or "wonderful care provider." While a sentiment analysis of the data was beyond the scope of this paper, we observed that patients use language with positive sentiment associations in their comments even for bottom performing physicians. It is also helpful to note that the phrasing "wonderful care provider" is the most significant and distinct usage of many possible ngram utterances with similar meaning, such as "wonderful provider," "great care provider," "provider wonderful," etc.

To further illustrate the importance of knowing how to seek and identify significant phrasings in patients' language, we can point to the use of "wasn" as a unique interaction ngram in the comments about top performing physicians. This particular ngram indicated what the top performing physicians were not, such as "My doctor wasn't rushed." It is an important finding that this ngram occurred more frequently in the comments about top performing physicians compared to the comments about bottom performing physicians. This shows us that patients tend to describe their interactions with top performing physicians in terms of "what the experience was not like." This finding opens up the possibility of analyzing positive comments as a potentially significant source of insights into the types of care experiences patients would not find satisfactory. Patient experience professionals can leverage such details understanding of what patients say to design and implement improvement actions that better meet the patients' needs and expectations.

Deciding When Technology Might be Necessary

Patients tell providers in their own words what they need and expect in their care services when they are given the opportunity. While patient experience surveys have been the most common source for patient feedback, today's patients do not limit themselves to standardized surveys to offer feedback, especially their natural language (verbatim) feedback. Processing large volumes of patient feedback from disparate sources requires significant analytical capabilities. Our study illustrates the scientific groundwork for NLP-based analyses to extract valuable insights from verbatim feedback about care experiences, which is the most accurate reflection of the patient's voice for the design and delivery of patient-centered care.

Faced with increasing volumes of verbatim feedback, patient experience professionals need to either develop analytical capabilities themselves or implement NLP-based technologies, so they can ground their improvement actions in the evidence basis of the patient's voice. Patient experience leaders need to be informed of the fundamental techniques behind the automated analysis of natural language to decide whether they will develop these capabilities themselves or adopt third-party solutions for their analytical needs. In either case, patient experience professionals need to evaluate the solutions they implement. We hope that our analysis and categorization provide some foundational knowledge to help with such evaluation.

Future Research at the Intersection of Language Analysis and Patient Experience

The methods applied in this study can be used to test similar correlations in the sentiment of patients' language. This study was limited to the analysis of patients' language describing their experiences with one type of provider - physicians. Future research can analyze patients' language describing many different aspects of care delivery, including patients' perceptions of care experiences with other types of providers, such as nurses. A complete understanding of patients' language use is necessary for the development of care quality assessment tools based on language data. We see the results of our study as an important first step towards a comprehensive research agenda in the analysis of patient feedback, including descriptive and predictive modeling of patients' language in their accounts of care experiences.

Conclusion

Steady progress in the delivery of patient-centered care requires that patient experience leaders understand what patients say in their own words about care services. Patient experience leaders can understand what patients say if they use scientifically rigorous methods to analyze patients' language. It is important for researchers and practitioners alike to apply rigorous methods in their efforts to assess and improve patient experience. Our study revealed the possibilities for new insights to be found beneath the surface of patient experience comments about physicians. This study uncovers a segment of the statistical work necessary to gain meaningful insights about a very specific component of patient experience - family members' perceptions of the services that physicians provide in a pediatric care setting. In this study, we aimed to show the value of a statistical approach to analyze patients' language about all aspects of care experiences across all possible care settings. Future research can build on our findings to develop comprehensive programs for providers to act on the evidence basis of language data in their efforts to deliver patient-centered care with excellent patient experience.

References

1. Shaller Consulting Group. A tale of three practices: how medical groups are improving the patient experience 2011. https://www.ahrq.gov/sites/default/files/wysiwyg/c/ahps/quality-improvement/reports-and-case-studies/CaseStudyPatientExperience_July2011FINAL_revised20110728.pdf Accessed January 26, 2018.
2. Becker's Hospital Review. The importance of patient experience for hospitals: why it pays to excel 2011. <https://www.beckershospitalreview.com/hospital-management-administration/the-importance-of-patient-experience-for-hospitals-why-it-pays-to-excel.html> Accessed January 25, 2018.
3. Brandon C, Cordina J, Gretz W, Neher K. Measuring the patient experience: lessons from other industries. McKinsey on Healthcare 2015. <https://healthcare.mckinsey.com/measuring-patient-experience-lessons-other-industries> Accessed January 25, 2018.
4. Detsky AS, Krumholz HM. Reducing the trauma of hospitalization. *JAMA*. 2014;311(21): 2169-2170.
5. Goldman DP, Vaiana M, Romley JA. The emerging importance of patient amenities in hospital care. *New England Journal of Medicine*. 2010;363(23): 2185-2187.
6. Fox S. The social life of health information. Pew Research Center. <http://www.pewinternet.org/2011/05/12/the-social-life-of-health-information-2011/> 2014. Accessed January 25, 2018.
7. Doing-Harris K, Mowery DL, Daniels C, Chapman WW, Conway M. Understanding patient satisfaction with received healthcare services: A natural language processing approach. *AMIA Annu Symp Proc*. 2016: 524-533.
8. López A, Detz A, Ratanawongsa N, Sarkar U. What patients say about their doctors online: A qualitative content analysis. *Journal of General Internal Medicine*. 2012;27(6):685-692.
9. Cognetta-Rieke C, Guney S. Analytical insights from patient narratives: the next step for better patient experience. *Journal of Patient Experience*. 2014;1(1):22-24.
10. Greaves F, Ramirez-Cano D, Millett C, et al. Use of sentiment analysis for capturing patient experience from free-text comments posted online. *Journal of Medical Internet Research*. 2013;15(11):e239.
11. Greaves F, Ramirez-Cano D, Millett C, Darzi A, Donaldson L. Harnessing the cloud of patient experience: using social media to detect poor quality healthcare. *BMJ Quality & Safety*. 2013;22(3):251-255.
12. Wiseman T, Lucas G, Sangha A, et al. Insights into the experiences of patients with cancer in London: framework analysis of free-text data from the national cancer patient experience survey 2012/2013 from the two London integrated cancer systems. *BMJ Open*. 2015;5:e007792.
13. Detsky AS, Krumholz HM. Reducing the trauma of hospitalization. *JAMA*. 2014;311(21): 2169-2170.
14. Espinel AG, Shah RK, Beach MC, Boss EF. What parents say about their child's surgeon. *JAMA Otolaryngology-Head & Neck Surgery*. 2014;140(5):397-402.
15. Crossing the quality chasm: a new health system for the 21st century. Institute of Medicine (US) Committee on Quality of Health Care in America 2001. <https://www.ncbi.nlm.nih.gov/pubmed/25057539> Accessed January 25, 2018.
16. Salton G, Wong A, Yang CS. A vector space model for automatic indexing. *Communications of the ACM*. 1975;18(11):613-620.
17. Barnaghi P, Ghaffari P, Breslin JG. Opinion mining and sentiment polarity on twitter and correlation between events and sentiment. 2016 IEEE Second International Conference on Big Data Computing Service and Applications (BigDataService).
18. Manevitz LM, Yousef M. One-class SVMs for document classification. *Journal of Machine Learning Research* 2. 2001: 139-154.
19. Gallan AS, Marina G, Roxana G. Perfect ratings with negative comments: learning from contradictory patient survey responses. *Patient Experience Journal*. 2017;4(3):15-28.26. Schlessinger M, Grob R, Shaller D, et al. Taking patients' narratives about clinicians from anecdote to science. *New England Journal of Medicine*. 2015;373:675-679.
20. Gallan AS, Guney S. The interplay between qualitative and quantitative data in patient experience metrics and measurement 2017. <https://info.narratedx.com/ndx-webinar-gallan-qualitative-quantitative-data-px-1> Accessed August 16, 2018.
21. Rafferty MO, Grey K. Beyond patient experience surveys: leveraging social media to glean patient feedback. *Nurse Leader*. 2014;12(3):31-35.
22. 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 Affairs*. 2016;35(4):697-705.
23. Smailhodzic E, Hooijsma W, Boonstra A, Langley DJ. Social media use in healthcare: A systematic review of effects on patients and on their relationship with healthcare professionals. *BMC Health Services Research*. 2016;16(1):442.
24. Hawkins CM, DeLao AJ, Hung C, Eng P. Social media and the patient experience. *Journal of the American College of Radiology*. 2016;13(12PB):1615-1621.
25. Grob R, Schlessinger M, Parker AM, et al. Breaking narrative ground: innovative methods for rigorously

- eliciting and assessing patient narratives. *Health Services Research*. 2016;51(3): 1248-1272.
26. Schlesinger M, Grob R, Shaller D, et al. Taking patients' narratives about clinicians from anecdote to science. *New England Journal of Medicine*. 2015;373:675-679.
 27. Schlesinger M, Grob R, Shaller D, et al. A rigorous approach to large-scale elicitation and analysis of patient narratives. *Medical Care Research and Review* 2018. <https://doi.org/10.1177/1077558718803859>