

Original Paper

Web-Based Textual Analysis of Free-Text Patient Experience Comments From a Survey in Primary Care

Inocencio Daniel Maramba¹, BS, MD, MSc; Antoinette Davey¹, BA(Hons), MSc, MPhil; Marc N Elliott², BA, MA, PhD; Martin Roberts³, BA (Hons), Cert Ed, MSc; Martin Roland⁴, MA, BM, BCh, DM, FRCGP, FMedSci; Finlay Brown⁵, BMBS; Jenni Burt⁴, BA (Hons), MSc, PhD; Olga Boiko¹, MSc, MPhil, PhD; John Campbell¹, MBChB, MD, FRCGP

¹Primary Care, University of Exeter Medical School, University of Exeter, Exeter, United Kingdom

²RAND Corporation, Santa Monica, CA, United States

³Collaboration for the Advancement of Medical Education Research and Assessment (CAMERA), Peninsula Schools of Medicine and Dentistry, Plymouth University, Plymouth, United Kingdom

⁴Institute of Public Health, University of Cambridge, Cambridge, United Kingdom

⁵Peninsula College of Medicine and Dentistry, Universities of Plymouth and Exeter, Plymouth, United Kingdom

Corresponding Author:

Inocencio Daniel Maramba, BS, MD, MSc

Primary Care

University of Exeter Medical School

University of Exeter

JS01 Smeall Building

St Luke's Campus University of Exeter, Magdalen Road

Exeter, EX1 2LU

United Kingdom

Phone: 44 1392 722895

Fax: 44 1392 722894

Email: I.D.C.Maramba@exeter.ac.uk

Abstract

Background: Open-ended questions eliciting free-text comments have been widely adopted in surveys of patient experience. Analysis of free text comments can provide deeper or new insight, identify areas for action, and initiate further investigation. Also, they may be a promising way to progress from documentation of patient experience to achieving quality improvement. The usual methods of analyzing free-text comments are known to be time and resource intensive. To efficiently deal with a large amount of free-text, new methods of rapidly summarizing and characterizing the text are being explored.

Objective: The aim of this study was to investigate the feasibility of using freely available Web-based text processing tools (text clouds, distinctive word extraction, key words in context) for extracting useful information from large amounts of free-text commentary about patient experience, as an alternative to more resource intensive analytic methods.

Methods: We collected free-text responses to a broad, open-ended question on patients' experience of primary care in a cross-sectional postal survey of patients recently consulting doctors in 25 English general practices. We encoded the responses to text files which were then uploaded to three Web-based textual processing tools. The tools we used were two text cloud creators: TagCrowd for unigrams, and Many Eyes for bigrams; and Voyant Tools, a Web-based reading tool that can extract distinctive words and perform Keyword in Context (KWIC) analysis. The association of patients' experience scores with the occurrence of certain words was tested with logistic regression analysis. KWIC analysis was also performed to gain insight into the use of a significant word.

Results: In total, 3426 free-text responses were received from 7721 patients (comment rate: 44.4%). The five most frequent words in the patients' comments were "doctor", "appointment", "surgery", "practice", and "time". The three most frequent two-word combinations were "reception staff", "excellent service", and "two weeks". The regression analysis showed that the occurrence of the word "excellent" in the comments was significantly associated with a better patient experience (OR=1.96, 95%CI=1.63-2.34), while "rude" was significantly associated with a worse experience (OR=0.53, 95%CI=0.46-0.60). The KWIC

results revealed that 49 of the 78 (63%) occurrences of the word “rude” in the comments were related to receptionists and 17(22%) were related to doctors.

Conclusions: Web-based text processing tools can extract useful information from free-text comments and the output may serve as a springboard for further investigation. Text clouds, distinctive words extraction and KWIC analysis show promise in quick evaluation of unstructured patient feedback. The results are easily understandable, but may require further probing such as KWIC analysis to establish the context. Future research should explore whether more sophisticated methods of textual analysis (eg, sentiment analysis, natural language processing) could add additional levels of understanding.

(*JMIR Med Inform* 2015;3(2):e20) doi:[10.2196/medinform.3783](https://doi.org/10.2196/medinform.3783)

KEYWORDS

patient experience; patient feedback; free-text comments; quantitative content analysis; textual analysis

Introduction

Patient experience is an important component of quality of health care, and questionnaires capturing patient experience have been widely used to provide insight into the quality of primary health care provision [1-3]. Feedback from survey results has been proposed as a cost-effective method to support and facilitate quality improvement [4,5].

In addition to capturing responses via closed questionnaire items, open-ended questions eliciting free-text comments have also been widely adopted [6,7] as the exclusive use of quantitative data limits the potential of surveys to improve practice [8]. Analysis of free-text comments can provide deeper or new insight, identify areas for action, and initiate further investigation [9]. Also, they may be a promising way to progress from documentation of patient experience to achieving quality improvement [9,10]. Free-text comments have been evaluated using methods such as content analysis [11,12], thematic analysis [9,13,14], and the Holsti Method [15]. However, these approaches can be resource intensive [6,15,16]. To efficiently deal with a large amount of free-text, new methods of rapidly summarizing and characterizing the text are being explored [17].

Text clouds are visual representation of a body of text, where the more frequently occurring words appear larger in the "cloud"[18,19]. The first widespread use of text clouds was “tag clouds”, which originated as a representation of the "tags" or keywords that users would assign to a Web resource [20,21]. Tag clouds have been used in health related websites to counter biased information processing [22].

The same technology that creates tag clouds may also be used to create word clouds from texts and textual data in general [23]. Text clouds differ from tag clouds in that their purpose is predominantly comprehension of the text rather than navigation of webpages [23]. Text clouds can be used to rapidly summarize textual data, revealing textual messages in a pictorial form [24]. Text clouds may have utility in supporting searching and browsing of webpages, as well as impression formation and recognition/matching of textual data [25].

Web applications such as TagCrowd, Many Eyes, Wordle, and Tagxedo are commonly used to generate text clouds. The majority of these are free for nonprofit use.

Text clouds have been used in a wide range of health related areas, such as examining the differences between various versions of a General Medical Council document [24] as well as a UK Government White Paper [26], survey responses on ehealth [27], survey of pharmacists' perceptions [28], patients' use of online message forums [29], and to analyze the responses of multiple sclerosis sufferers to open-ended questions [30].

Other uses of computerized textual analysis in health include: automatic analysis of online discussions related to diabetes [31], content analysis of the free text comments in multi-source feedback about specialist registrars [32], automatic drug side effect discovery by analysis of online patient submitted reviews [33], keyword analysis of an online survey investigating nurses' perceptions of spirituality and spiritual care [34], and uncovering signs and symptoms of opiate exposure from comments posted on YouTube [35].

We aimed to investigate the feasibility of using Web-based textual analysis for extracting useful information from large amounts of free-text patient comments, and to identify key issues or topics that would be revealed by computerized text processing, using tools that are currently available at no cost on the Web.

Methods

Participants and Procedure

The data was collected as part of the “Improve” study, a research program funded by the National Institute for Health Research (NIHR) [36], exploring various aspects of patient experience in primary care. One of the projects involved a post-consultation postal survey using a modified version of the English GP Patient Survey (GPPS) questionnaire. The GPPS is the largest survey program of patients registered with an English general practice. A random sample of patients from each English practice—~2.6 million patients each year in total—is invited to take part in the survey [37,38]. A particular change made to the GPPS questionnaire (at the request of participating practices) was the inclusion of a free-text comments question worded as follows: “Your [general] practice has asked that we collect any further comments you would like to make about the service they provide.”

Detailed survey methods have been previously reported in the paper by Roberts et al [39], which are briefly summarized here. Following a recent face-to-face consultation between November

2011 and June 2013 with one of 105 doctors from 25 practices in six areas of England (Cornwall, Devon, Bristol, Bedfordshire, Cambridgeshire, Peterborough, and North London), patients were sent a questionnaire regarding their experiences of care. One reminder was sent to nonrespondents. Free-text comments were anonymized during data entry, extracted from the database and exported to a text file. Approval for the study was obtained from the South West 2 Research Ethics Committee on January 28th 2011 (ref: 09/H0202/65). Return of a completed questionnaire was taken to indicate patient consent to participate in the study.

Textual Analysis Methods

Free-text comments were analyzed using three Web-based textual analysis tools: TagCrowd v.10/02/2011 [40], Many Eyes v.1.0 [41], and Voyant Tools v.1.0 [42], which were chosen for their ease of use and range of functionalities.

TagCrowd is a Web application for visualizing word frequencies in any text by creating what is popularly known as a word cloud, text cloud or tag cloud. We created text clouds based on an aggregated corpus of free-text patient comments.

We used the following parameters in TagCrowd: (1) frequently occurring English words and connectives (eg, “a”, “in”, “is”, “it”, and “you”) were ignored; (2) the tag cloud was created from the 50 most frequently occurring single words; (3) a stemming algorithm combined related words (eg, learn, learned, learning → learn). The 50 word limit was chosen as it has been used in previous work using text clouds to examine health information [24,26]. We also tried generating a 60 word text cloud but found the result to be difficult to read.

Many Eyes is a Web-based data visualization application created by IBM [41]. Fundamentally the software incorporates the capacity to create and view various forms of text visualization and representation. We chose to use Many Eyes because of its capability of creating text clouds from the most frequent two-word combinations. We hypothesized that two-word combinations might give a more nuanced insight into the meanings behind the most frequently used words as some of their associations would be preserved.

Voyant Tools is a Web-based reading and analysis environment for digital texts. It was created as part of a collaborative project to develop and theorize text analysis tools and text analysis rhetoric [42]. In addition to calculating word frequencies and creating text clouds, Voyant Tools performs other textual analysis functions, such as identifying distinctive words in the documents that make up a text corpus. To investigate the validity of the distinctive words component, we divided the comments into separate text files depending on whether the patients reported if they were either “satisfied” or “not satisfied” with their experience of care. The question was “In general, how satisfied are you with the care you get at this GP surgery or health center?”. Patients were given five options to rate their satisfaction with the practice: “very satisfied”, “fairly satisfied”, “neither satisfied nor dissatisfied”, “fairly dissatisfied”, and “very dissatisfied”. In this analysis, the “very satisfied” and “fairly satisfied” responses were recoded as “satisfied” and the last three options as “not satisfied”. We used the “distinctive

words” function to identify the words that occurred more frequently in comments originating from patients who were “satisfied” and words which occurred more frequently in comments from patients who were “not satisfied”.

Statistical Analysis

We used logistic regression to investigate the occurrence/nonoccurrence of words within individual patient comments. The words were selected from the results of the distinctive word analysis. We obtained and compared the frequency of use of the five most distinctive words from the comments classified as originating from either the “satisfied” or “not satisfied” patients (ten words in total).

Logistic regression was used to predict the presence or absence of each of these words in a comment from the standardized scores (z-scores) of the patients’ responses to the survey question on satisfaction. We used the following formula for the standardized scores (z-scores):

$$z=(x-\mu)/(\sigma)$$

Two additional models were run for each word, predicting its presence or absence from the z-scores of the patients’ ratings of their confidence and trust in the doctor and their ratings of the doctors’ communication skills. We derived these scores from the patients’ responses to two other structured questions that were asked in the questionnaire. These variables were chosen as we hypothesized that confidence and trust in the doctor, as well as the communication skills of the doctor could influence the words used by the patients in their comments. Statistical analyses were performed in STATA version SE13.1 for Windows. We then plotted the odds ratios for the selected words against their standardized frequencies. The standardized frequency is calculated in the same way as a z-score, where x is the frequency of a particular word, μ is the mean frequency of all words in the patient comments, and σ is their standard deviation.

Keyword in Context Analysis

Voyant Tools provides a Keyword in Context (KWIC) function. KWIC involves searching for a particular keyword in the text and analyzing its local meaning in relation to a fixed number of words immediately preceding and following it [43]. KWIC can help identify underlying connections that are being implied by the text [44]. KWIC analysis had been used in content analysis of blogs about female incontinence [45], as well as in content analysis of audiology service improvement documentation [46]. The KWIC function in Voyant tools can quickly display the KWIC for a selected keyword and the results can be exported to a format suitable for further analysis. For this analysis we selected 15 words that preceded and followed the word “rude”. The resulting text was then manually examined to determine the context of the use of “rude”, and the results were tabulated.

Results

Textual Analysis Methods

From 7721 respondents, we collected 3426 individual comments (comment rate: 44.4%). The comments came to a total of

150,699 words of which 6867 are unique words. The average length of response is 43.98 words. There are 273 instances of 90 unique, non-English terms (mostly misspellings). Figure 1 shows the text cloud resulting from all the free-text comments as generated by TagCrowd. The five most frequent words were: “doctor”, “appointment”, “surgery”, “practice”, and “time”. Included in the 50 most frequent words were those that have a positive connotation such as: “helpful” and “excellent”. Words with a negative connotation, such as “difficult” and “problem” were also present, but were less frequent.

The two-word text cloud generated by Many Eyes is shown in Figure 2, displaying the 200 most frequent two-word phrases

(bigrams). The five most frequent bigrams were: “reception staff”, “excellent service”, “two weeks”, “medical centre” and “good service”.

Figure 3 shows the results of the Voyant Tools distinctive words component when applied to comments categorized as originating from satisfied or dissatisfied patients. The words “surgery”, “excellent”, “service”, “good”, and “helpful” were the five most distinctive words from satisfied patients, while the words “doctor”, “feel”, “appointment”, “rude”, and “symptoms” were the five most distinctive words in the comments from dissatisfied patients.

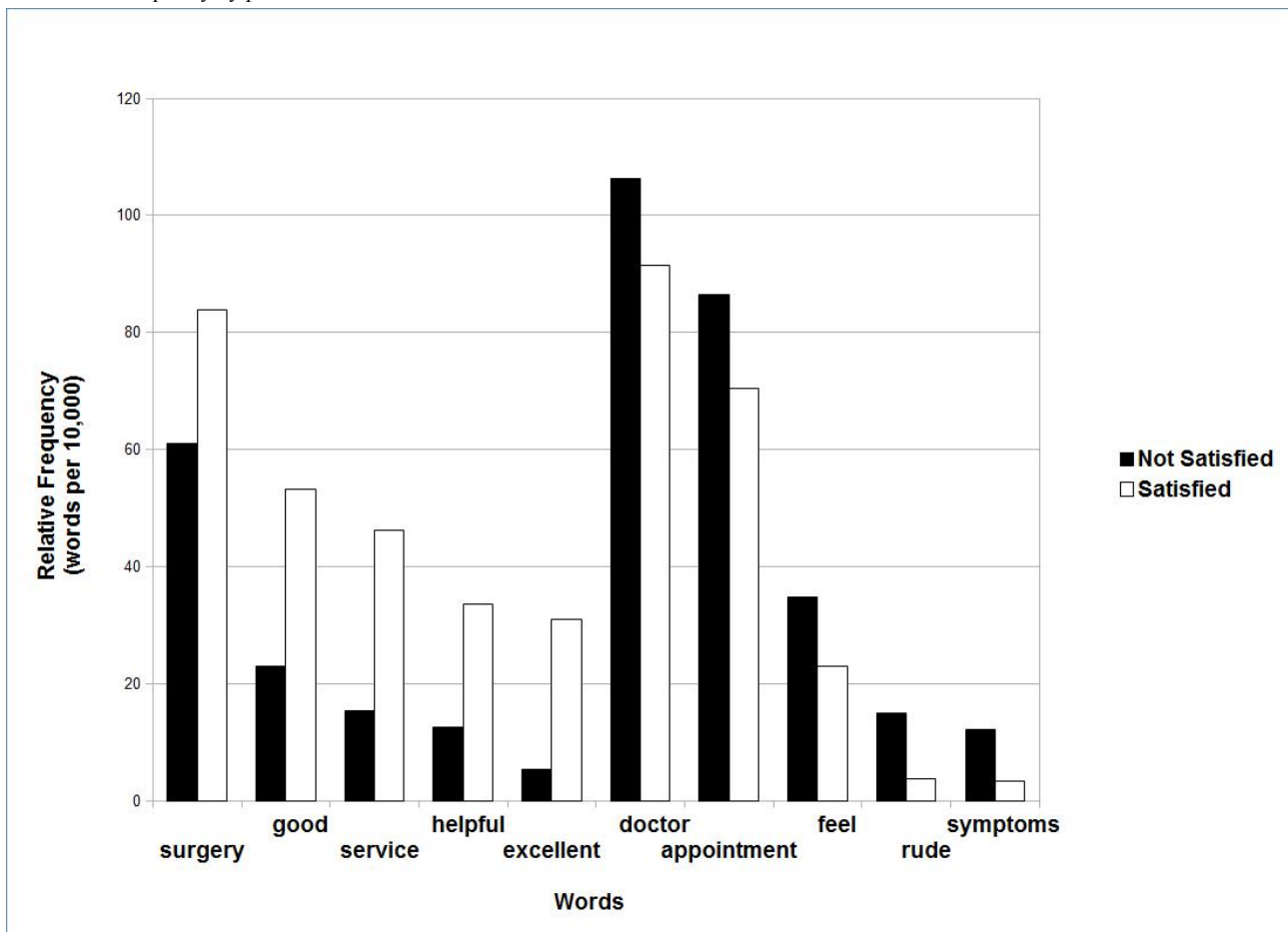
Figure 1. Single-word text cloud created in TagCrowd from all free text comments.



Figure 2. Two-word text cloud created in Many Eyes.



Figure 3. Word frequency by patient satisfaction.



Statistical Analysis

From the logistic regression models, odds ratios were calculated for the distinctive words. In this analysis, the odds ratio indicates the amount by which the odds of a particular word occurring at

least once in a comment are multiplied for every point increase in the z-score. Table 1 reports the results of the logistic regression for the 10 distinctive words and the scores for satisfaction, doctor-patient communication, and confidence and trust of the patient in the doctor.

Table 1. Odds ratio (95% CI, R² value) for the occurrence of distinctive words corresponding to a one standard deviation increase in measures of patient experience.

Word	Overall satisfaction N=3134 OR (CI, R ²)	Doctor's communication skills N=3062 OR (CI, R ²)	Confidence and trust in the doctor N=3066 OR (CI, R ²)
Service	1.39 (1.25-1.54, .02)	1.25 (1.13-1.39, .007)	1.29 (1.16-1.43, .009)
Good	1.11 (1.02-1.21, .002)	1.09 (0.99-1.2, .001)	1.06 (0.97-1.16, .0005)
Excellent	1.96 (1.63-2.34, .04)	2.09 (1.69-2.58, .09)	1.76 (1.45-2.15, .003)
Surgery	0.94 (0.88-1.01, .0008)	0.98 (0.90-1.05, .0001)	1.00 (0.93-1.08, .00)
Helpful	1.19 (1.07-1.32, .005)	1.24 (1.10-1.40, .006)	1.10 (0.99-1.22, .001)
Appointment	0.67 (0.63-0.72, .04)	0.77 (0.77-0.82, .01)	0.80 (0.74-0.86, .01)
Doctor	0.76 (0.71-0.81, .02)	0.81 (0.75-0.87, .01)	0.81 (0.76-0.88, .008)
Feel	0.79 (0.72-0.87, .01)	0.78 (0.71-0.86, .01)	0.77 (0.70-0.85, .01)
Rude	0.53 (0.46-0.60, .10)	0.63 (0.55-0.74, .04)	0.60 (0.51-0.70, .05)
Symptoms	0.60 (0.51-0.70, .06)	0.61 (0.52-0.72, .05)	0.64 (0.54-0.77, .03)

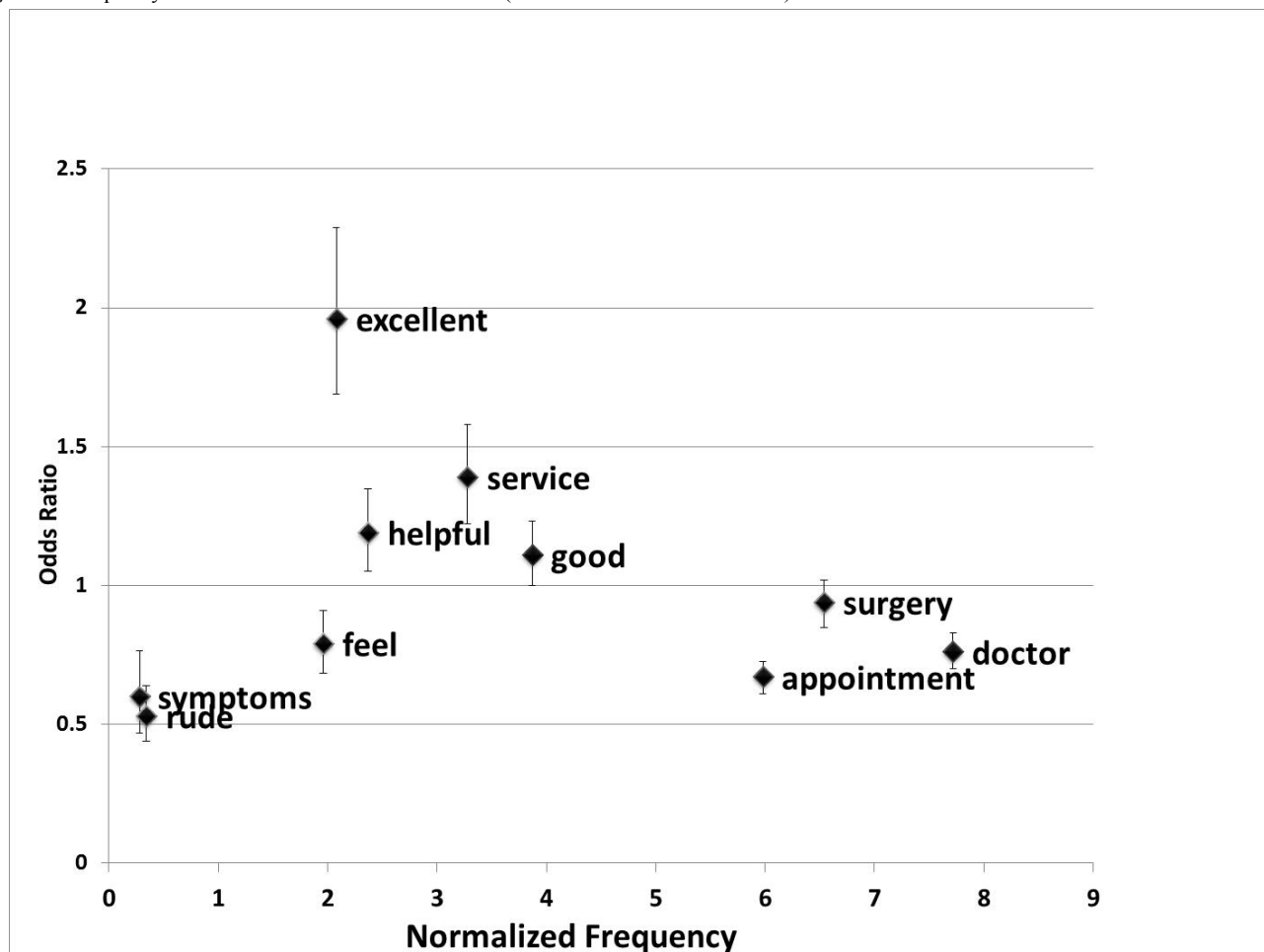
As shown in the table, the regression for the word “excellent” results in an OR of 1.96, for the bivariate model for patient satisfaction; that is, an increase of one standard deviation in the patient satisfaction is associated with almost twice the odds of the word “excellent” occurring in the comments. There is also a significant association of the occurrence of “excellent” in the comments with the z-scores for doctor communication skills and confidence in the doctor which have odds ratios of 2.09 and 1.76 respectively.

In contrast, the word “rude” has an OR of 0.53 in the bivariate model for patient satisfaction, indicating that an increase in one standard deviation in the satisfaction score almost halves the odds that the word “rude” will appear in the comments. The OR is also significantly lower for the occurrence of “rude” when scores for doctor communication skills or confidence in the

doctor are higher. To summarize, the words “service” and “excellent” had a significant positive association for overall satisfaction, doctor's communication skills scores, and confidence and trust in the doctor scores. The word “helpful” had a significant positive association for overall satisfaction and doctor's communication skill scores, but was not significant for confidence and trust. The words “rude” and “symptoms” had a significant negative association with all three scores.

Figure 4 shows a plot of the odds ratios for the occurrence of a word due to a one standard deviation increase in patient satisfaction score as calculated by the bivariate model. The odds ratios for ten distinctive words (five most distinctive words each from satisfied and dissatisfied patients) are plotted against their standardized frequencies (x-axis).

Figure 4. Frequency of selected words and the odds ratios (with 95% confidence intervals) associated with the z-scores for satisfaction.



Keywords in Context (KWIC) Analysis

We chose to look at the context of the usage of the word “rude” using Keyword in Context (KWIC) analysis. We examined 15 words to the left and right of the keyword in question because we felt that was the minimum amount where we could satisfactorily establish the context of use of the word in question. Earlier attempts using 10 words on each side gave results wherein the context was still ambiguous in some of the comments. We manually reviewed the output from the KWIC tool, and established the context of the various instances of the

adjective “rude”. We then constructed a table listing the sources of the rude behavior and their frequency of mention (Table 2).

Overall, “reception staff” was mentioned in 63% of the occurrences of the word “rude”, while doctors accounted for 22% of occurrences. Among the patients who were dissatisfied, the proportion of doctors being associated with occurrences of the word “rude” increased to 30% compared to 15% in satisfied patients. Reception staff had a larger proportion of association with occurrences of the word “rude” among satisfied patients, 72%, than in patients who were dissatisfied: 54%.

Table 2. Keyword in Context (KWIC) analysis for the word “rude”. Frequency of occurrence (% within patient type) by patient satisfaction and subject.

	Subject of adjective “rude”						
	Frequency of occurrence (% within patient type)						
Patients	Doctor	Nurse	Practice manager	Reception staff	Staff	Patient	Total
Not satisfied	11 (30)	1 (3)	2 (5)	20 (54)	3 (8)	0 (0)	37
Satisfied	6 (15)	3 (7)	1 (2)	29 (72)	1 (2)	1 (2)	41
All	17 (22)	4 (5)	3 (4)	49 (63)	4 (5)	1 (1)	78

Discussion

Principal Results

We found the three textual analysis tools easy to use and the results were generated very quickly, considering the volume of

text that was processed (approximately 150,699 words). The tools used the standard ASCII text file format, which most data analysis software can easily export to. The text clouds did give a concise summary of what the majority of comments were about, but it was difficult to establish the exact context of the use of the most frequent words. The distinctive word analysis

gave more insight by showing the different usage of words by differing sources of comments (satisfied vs dissatisfied patients). Some of the words in the output of the distinctive word analysis showed significant associations with satisfaction scores, notably, “excellent” and “rude”. The word “excellent” was associated with high patient satisfaction scores. The high frequency of the two-word phrase “excellent service” showed that patients used it in relation to their rating of the quality of service they received.

The word “rude” occurred much more frequently in comments from dissatisfied patients, suggesting that rude behavior encountered by patients may trigger dissatisfaction. The KWIC analysis showed that the word “rude”, was most commonly associated with the reception staff. Receptionists have been recognized as crucial members of the primary health care team [47], and recent work has suggested that the historical perception of the receptionist as a “dragon behind the desk” has been getting in the way of understanding the role of receptionists and thus improving patient care [48]. Also worth noting is that the proportion of rude actions being attributed to doctors was higher amongst the patients who were dissatisfied with their practice. Winsted has identified rudeness from doctors and other forms of negative behavior as being a “dissatisfier” in medical encounters [49].

The increased relative frequency of the word “feel” in the comments from dissatisfied patients might indicate an emotional reaction being a component of patient dissatisfaction. However, the word may also be used in other contexts, for instance “I feel that the doctor should”, as opposed to “I feel disappointed.” The recurrence of the word “symptoms” in the comments from dissatisfied patients could indicate a relationship between dissatisfaction and the perception of poor health, as has been reported previously by Xiao et al [50]. It may also provide a comment on the perceived thoroughness of the clinical encounter. One point of interest is that the words positively associated with patient satisfaction focus on the system (eg, “excellent service”), while those associated with dissatisfaction highlight some of the interpersonal aspects of care (eg, “rude”, “feel”).

Limitations

While the textual analysis applications are easy to use and give results quickly, one limitation is that an internet connection is required for all the software tools to work. However, a high-speed connection is not necessary, and the software runs on any modern operating system with an updated Web browser.

When we attempted to identify the messages contained in the text cloud, we found it difficult to ascertain the significance of the high frequency of the words “doctor”, “practice”, “surgery”, “appointment”, and “time”. This is due to the text cloud showing the words dissociated from their original context, making it difficult to discern the meaning behind the high frequencies of these words. This loss of context due to the dissociation of the words from one another is a major limitation in the interpretation of the results of the text cloud. When words are separated from one another, and only their frequencies rather than their relationships are scrutinized, there is a danger of overlooking subtle and important nuances and meanings formed by the

synergy of the words [24]. The software tools are also limited in that they are unable to group together words that are synonymous, (eg, “doctor”, “dr”, and “gp”), unless the software is specifically instructed to group these synonyms. In addition to the individual words, meaning is also conveyed by the patterns that words form.

Another consequence of dissociation is that our method does not automatically deal with negation of terms. However, for the words “good”, “excellent”, “helpful” “rude”, and “symptoms”, we examined the results of the keyword in context extraction to see if they contained instances of negation. We were satisfied that all mentions of those words did not contain instances of negation. A more sophisticated approach using natural language processing and machine learning is required to automatically deal with negation. Sentiment analysis, which is a more sophisticated textual analysis technique, is one method that takes the patterns of words, and not just their frequencies, into account. Research reports are emerging in which sentiment analysis has been used to examine free text comments from patients [51-57].

A further limitation of this study is related to the nature of the question being presented to elicit the comments. The very broad nature of the request for comments means that the patients’ responses were, almost inevitably, quite varied. The wide spectrum of issues raised in the comments make them quite difficult to neatly categorize and characterize. The quality of information gleaned from patient responses could be improved by focusing the wording of the request for comments to address central issues of interest [9]. This focus of interest could also be coupled to particular quantitative questions, to give insight into why the patient answered the question in that particular way.

Further Research

Web-based textual analysis shows promise as a means of rapidly summarizing the messages contained in free-text comments from primary care patients. Text clouds are a feasible means of presenting the most frequent words used in free-text comments from patients. However, text clouds are limited by an inability to provide a contextually meaningful summary of the original corpus of comments. This is commonly encountered when relying primarily on a simple, mechanistic algorithm, in this case, word frequency. Words convey meaning by working together, and there is a synergy created through the combination of various words [24]. A more accurate way of capturing the messages contained in the free-text comments by a computer mediated approach is through KWIC analysis. The use of more sophisticated technologies, such as machine learning, natural language processing, neural networks, and sentiment analysis may address some of these shortcomings. Future research needs to be done around generating sentence level summarization using the techniques from the NLP community [58], such as latent Dirichlet allocation [59,60]. For a wider uptake, a user-friendly (preferably open source) application needs to be developed to fill this gap. This would enable practices to make better use of the large amounts of free-text feedback that they have collected. In addition, careful attention needs to be paid to formulating focused and precise requests for comments which

might be expected to yield feedback that could provide a substantial basis for computer mediated textual analysis. Finally, mixed methods approaches as well as sociocybernetics methods have also been proposed as a way of completing the picture of patient experience [61,62].

Conclusions

Our study has shown that by using Web-based text processing tools to extract information from patient comments, we can discover words that the patients have used in their comments that have significant associations with quantitative measurements of patient experience. The logistic regression revealed strong positive and negative associations between the satisfaction scores and the occurrence of certain words. KWIC analysis was

then used to examine the context of the uses of words, which yielded useful information; for example, the sources of rude behavior that is associated with patient dissatisfaction. This approach could help practices in formulate policies to increase patient satisfaction. Sequential use of these methods may prove useful in documenting how patients' experience of care changes over time, similar to the method used by Gill et al in revealing the longitudinal changes in the document "Good Medical Practice" produced by the General Medical Council [24]. An approach that examines the key words in the context is useful in deriving insights from the free-text comments. Further research is necessary in refining these methods, so that the results would be comparable to traditional techniques of content analysis.

Acknowledgments

This paper presents independent research funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research Programme (Grant Reference Number RP-PG-0608-10050). The views expressed are those of the authors and not necessarily those of the National Health Service England (NHS England), the NIHR or the Department of Health.

We would like to thank the patients, practice managers, GPs, and other staff of the general practices who kindly agreed to participate in this study and without whom the study would not have been possible. Thanks also to Gary Abel, Natasha Elmore, Emily Taylor, Jenny Newbould, Emma Whitton, Amy Gratton, Charlotte Paddison, Mary Carter, and Dawn Swancutt for invaluable help with study set-up, practice recruitment, data collection, and data entry. We confirm that all personal identifiers have been removed or disguised so that study participants are not identifiable and cannot be identified through the provided details.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshots of Web based textual analysis tools.

[[PDF File \(Adobe PDF File\), 1MB - medinform_v3i2e20_app1.pdf](#)]

Multimedia Appendix 2

Survey questionnaire.

[[PDF File \(Adobe PDF File\), 290KB - medinform_v3i2e20_app2.pdf](#)]

References

1. Roland M, Elliott M, Lyratzopoulos G, Barbiere J, Parker R, Smith P, et al. Reliability of patient responses in pay for performance schemes: analysis of national General Practitioner Patient Survey data in England. *BMJ* 2009 Sep 29;339(sep29 3):b3851-b3851 [FREE Full text] [doi: [10.1136/bmj.b3851](https://doi.org/10.1136/bmj.b3851)]
2. Jenkinson C, Coulter A, Bruster S. The Picker Patient Experience Questionnaire: development and validation using data from in-patient surveys in five countries. *Int J Qual Heal Care Internet* 2002;14(5):353-358 [FREE Full text]
3. Giordano LA, Elliott MN, Goldstein E, Lehrman WG, Spencer PA. Development, implementation, and public reporting of the HCAHPS survey. *Med Care Res Rev* 2010 Feb;67(1):27-37. [doi: [10.1177/1077558709341065](https://doi.org/10.1177/1077558709341065)] [Medline: [19638641](https://pubmed.ncbi.nlm.nih.gov/19638641/)]
4. Cleary PD. The increasing importance of patient surveys. *Qual Health Care* 1999 Dec;8(4):212 [FREE Full text] [Medline: [10847881](https://pubmed.ncbi.nlm.nih.gov/10847881/)]
5. Cleary P. Expanding the use of patient reports about patient-centered care. *Isr J Health Policy Res Internet Israel Journal of Health Policy Research* 2013 Jan [FREE Full text]
6. Wright C, Richards SH, Hill JJ, Roberts MJ, Norman GR, Greco M, Taylor Matthew R S, et al. Multisource feedback in evaluating the performance of doctors: the example of the UK General Medical Council patient and colleague questionnaires. *Acad Med* 2012 Dec;87(12):1668-1678. [doi: [10.1097/ACM.0b013e3182724cc0](https://doi.org/10.1097/ACM.0b013e3182724cc0)] [Medline: [23095930](https://pubmed.ncbi.nlm.nih.gov/23095930/)]
7. Graham C, Woods P. Patient experience surveys. In: Ziebland S, Coulter A, Calabrese JD, Locock L, editors. *Understanding and Using Health Experiences: Improving patient care*. Oxford: Oxford University Press; 2013.

8. Asprey A, Campbell JL, Newbould J, Cohn S, Carter M, Davey A, et al. Challenges to the credibility of patient feedback in primary healthcare settings: a qualitative study. *Br J Gen Pract* 2013 Mar;63(608):e200-e208 [[FREE Full text](#)] [doi: [10.3399/bjgp13X664252](https://doi.org/10.3399/bjgp13X664252)] [Medline: [23561787](https://pubmed.ncbi.nlm.nih.gov/23561787/)]
9. Riiskjær E, Ammentorp J, Kofoed P. The value of open-ended questions in surveys on patient experience: number of comments and perceived usefulness from a hospital perspective. *Int J Qual Health Care* 2012 Oct;24(5):509-516 [[FREE Full text](#)] [doi: [10.1093/intqhc/mzs039](https://doi.org/10.1093/intqhc/mzs039)] [Medline: [22833616](https://pubmed.ncbi.nlm.nih.gov/22833616/)]
10. Liu GC, Harris MA, Keyton SA, Frankel RM. Use of unstructured parent narratives to evaluate medical student competencies in communication and professionalism. *Ambul Pediatr* 2007;7(3):207-213. [doi: [10.1016/j.ambp.2007.03.001](https://doi.org/10.1016/j.ambp.2007.03.001)] [Medline: [17512880](https://pubmed.ncbi.nlm.nih.gov/17512880/)]
11. Shuyler K, Knight K. What are patients seeking when they turn to the Internet? Qualitative content analysis of questions asked by visitors to an orthopaedics Web site. *J Med Internet Res* 2003 Oct 10;5(4):e24 [[FREE Full text](#)] [doi: [10.2196/jmir.5.4.e24](https://doi.org/10.2196/jmir.5.4.e24)] [Medline: [14713652](https://pubmed.ncbi.nlm.nih.gov/14713652/)]
12. Tang PC, Black W, Young CY. Proposed criteria for reimbursing eVisits: content analysis of secure patient messages in a personal health record system. *AMIA Annu Symp Proc* 2006:764-768 [[FREE Full text](#)] [Medline: [17238444](https://pubmed.ncbi.nlm.nih.gov/17238444/)]
13. Burner ER, Menchine MD, Kubicek K, Robles M, Arora S. Perceptions of successful cues to action and opportunities to augment behavioral triggers in diabetes self-management: qualitative analysis of a mobile intervention for low-income Latinos with diabetes. *J Med Internet Res* 2014;16(1):e25 [[FREE Full text](#)] [doi: [10.2196/jmir.2881](https://doi.org/10.2196/jmir.2881)] [Medline: [24476784](https://pubmed.ncbi.nlm.nih.gov/24476784/)]
14. Ashley L, Jones H, Thomas J, Newsham A, Downing A, Morris E, et al. Integrating patient reported outcomes with clinical cancer registry data: a feasibility study of the electronic Patient-Reported Outcomes From Cancer Survivors (ePOCS) system. *J Med Internet Res* 2013;15(10):e230 [[FREE Full text](#)] [doi: [10.2196/jmir.2764](https://doi.org/10.2196/jmir.2764)] [Medline: [24161667](https://pubmed.ncbi.nlm.nih.gov/24161667/)]
15. Richards SH, Campbell JL, Walshaw E, Dickens A, Greco M. A multi-method analysis of free-text comments from the UK General Medical Council Colleague Questionnaires. *Med Educ* 2009 Aug;43(8):757-766. [doi: [10.1111/j.1365-2923.2009.03416.x](https://doi.org/10.1111/j.1365-2923.2009.03416.x)] [Medline: [19659489](https://pubmed.ncbi.nlm.nih.gov/19659489/)]
16. Frohna A, Stern D. The nature of qualitative comments in evaluating professionalism. *Med Educ* 2005 Aug;39(8):763-768. [doi: [10.1111/j.1365-2929.2005.02234.x](https://doi.org/10.1111/j.1365-2929.2005.02234.x)] [Medline: [16048618](https://pubmed.ncbi.nlm.nih.gov/16048618/)]
17. Verhoef LM, Van de Belt TH, Engelen L, Schoonhoven L, Kool RB. Social media and rating sites as tools to understanding quality of care: a scoping review. *J Med Internet Res* 2014;16(2):e56 [[FREE Full text](#)] [doi: [10.2196/jmir.3024](https://doi.org/10.2196/jmir.3024)] [Medline: [24566844](https://pubmed.ncbi.nlm.nih.gov/24566844/)]
18. Hearst M, Rosner D. Tag Clouds: Data Analysis Tool or Social Signaller? : Data Analysis Tool or Social Signaller? Proc 41st Annu Hawaii Int Conf Syst Sci (HICSS 2008) Internet Ieee; 2008 Presented at: Hawaii International Conference on System Sciences, Proceedings of the 41st Annual; 7-10 Jan. 2008; Waikoloa, HI p. 160 URL: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4438863>
19. Sinclair J, Cardew-Hall M. The folksonomy tag cloud: when is it useful? *Journal of Information Science* 2007 May 31;34(1):15-29 [[FREE Full text](#)] [doi: [10.1177/0165551506078083](https://doi.org/10.1177/0165551506078083)]
20. Thomas M, Caudle D, Schmitz C. To tag or not to tag? *Library Hi Tech* 2009 Sep 04;27(3):411-434 [[FREE Full text](#)] [doi: [10.1108/07378830910988540](https://doi.org/10.1108/07378830910988540)]
21. Schweier R, Romppel M, Richter C, Hoberg E, Hahmann H, Scherwinski I, et al. A web-based peer-modeling intervention aimed at lifestyle changes in patients with coronary heart disease and chronic back pain: sequential controlled trial. *J Med Internet Res* 2014;16(7):e177 [[FREE Full text](#)] [doi: [10.2196/jmir.3434](https://doi.org/10.2196/jmir.3434)] [Medline: [25057119](https://pubmed.ncbi.nlm.nih.gov/25057119/)]
22. Schweiger S, Oeberst A, Cress U. Confirmation bias in web-based search: a randomized online study on the effects of expert information and social tags on information search and evaluation. *J Med Internet Res* 2014;16(3):e94 [[FREE Full text](#)] [doi: [10.2196/jmir.3044](https://doi.org/10.2196/jmir.3044)] [Medline: [24670677](https://pubmed.ncbi.nlm.nih.gov/24670677/)]
23. Lamantia J. Internet 2007 cited. 2014 Apr 16. Text clouds: A new form of tag cloud URL: http://www.joelamantia.com/blog/archives/tag_clouds/text_clouds_a_new_form_of_tag_cloud.html [accessed 2014-08-13] [[WebCite Cache ID 6RnT6KYBL](#)]
24. Gill D, Griffin A. Good Medical Practice: what are we trying to say? Textual analysis using tag clouds. *Med Educ* 2010 Mar;44(3):316-322. [doi: [10.1111/j.1365-2923.2009.03588.x](https://doi.org/10.1111/j.1365-2923.2009.03588.x)] [Medline: [20444063](https://pubmed.ncbi.nlm.nih.gov/20444063/)]
25. Rivadeneira AW, Gruen DM, Muller MJ, Millen DR. Getting our head in the clouds: toward evaluation studies of tagclouds. USA: ACM; 2007 Presented at: CHI '07 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; 28 April - 3 May, 2007; San Jose, California p. 995-998.
26. Park S, Griffin A, Gill D. Working with words: exploring textual analysis in medical education research. *Med Educ* 2012 Apr;46(4):372-380. [doi: [10.1111/j.1365-2923.2011.04184.x](https://doi.org/10.1111/j.1365-2923.2011.04184.x)] [Medline: [22429173](https://pubmed.ncbi.nlm.nih.gov/22429173/)]
27. Moen A, Hackl W, Hofdijk J, Van Gemert-Pijnen L, Ammenwerth E, Nykänen P, et al. eHealth in Europe - Status and Challenges. *Yearb Med Inform* 2013;8(1):59-63. [Medline: [23974549](https://pubmed.ncbi.nlm.nih.gov/23974549/)]
28. Al Hamarneh Yazid N, Rosenthal M, McElnay JC, Tsuyuki RT. Pharmacists' perceptions of their practice: a comparison between Alberta and Northern Ireland. *Int J Pharm Pract* 2012 Feb;20(1):57-64. [doi: [10.1111/j.2042-7174.2011.00163.x](https://doi.org/10.1111/j.2042-7174.2011.00163.x)] [Medline: [22236181](https://pubmed.ncbi.nlm.nih.gov/22236181/)]

29. O'Grady L, Wathen CN, Charnaw-Burger J, Betel L, Shachak A, Luke R, et al. The use of tags and tag clouds to discern credible content in online health message forums. *Int J Med Inform* 2012 Jan;81(1):36-44. [doi: [10.1016/j.ijmedinf.2011.10.001](https://doi.org/10.1016/j.ijmedinf.2011.10.001)] [Medline: [22030035](https://pubmed.ncbi.nlm.nih.gov/22030035/)]
30. Osborne LA, Noble JG, Lockhart-Jones HM, Middleton R, Thompson S, Maramba IDC. Sources of discovery, reasons for registration, and expectations of an internet-based register for Multiple Sclerosis: Visualisations and explorations of word uses and contexts. *International Journal of Healthcare Information Systems and Informatics (IJHISI)* 2012;7(3):27-43.
31. Hamon T, Gagnayre R. Improving knowledge of patient skills thanks to automatic analysis of online discussions. *Patient Educ Couns* 2013 Aug;92(2):197-204. [doi: [10.1016/j.pec.2013.05.012](https://doi.org/10.1016/j.pec.2013.05.012)] [Medline: [23769423](https://pubmed.ncbi.nlm.nih.gov/23769423/)]
32. Archer J, McGraw M, Davies H. Assuring validity of multisource feedback in a national programme. *Arch Dis Child* 2010 May;95(5):330-335. [doi: [10.1136/adc.2008.146209](https://doi.org/10.1136/adc.2008.146209)] [Medline: [20457700](https://pubmed.ncbi.nlm.nih.gov/20457700/)]
33. Liu J, Li A, Seneff S. Automatic Drug Side Effect Discovery from Online Patient-Submitted Reviews: Focus on Statin Drugs. 2011 Presented at: IMMM 2011, The First International Conference on Advances in Information Mining and Management; October 23, 2011 to October 29, 2011; Barcelona, Spain p. 91-96 URL: http://www.thinkmind.org/index.php?view=article&articleid=immm_2011_5_10_20050
34. McSherry W, Jamieson S. The qualitative findings from an online survey investigating nurses' perceptions of spirituality and spiritual care. *J Clin Nurs* 2013 Nov;22(21-22):3170-3182. [doi: [10.1111/jocn.12411](https://doi.org/10.1111/jocn.12411)] [Medline: [24118520](https://pubmed.ncbi.nlm.nih.gov/24118520/)]
35. Chary M, Park EH, McKenzie A, Sun J, Manini AF, Genes N. Signs & Symptoms of Dextromethorphan Exposure from YouTube. *PLoS One Internet* 2014 Jan [[FREE Full text](#)] [doi: [10.1371/journal.pone.0082452](https://doi.org/10.1371/journal.pone.0082452)]
36. Improve Study Team. 2014. Improve Study Website Internet URL: <https://improve.exeter.ac.uk/> [accessed 2014-08-15] [[WebCite Cache ID 6RqYPViDo](#)]
37. Campbell J, Smith P, Nissen S, Bower P, Elliott M, Roland M. The GP Patient Survey for use in primary care in the National Health Service in the UK--development and psychometric characteristics. *BMC Fam Pract Internet* 2009 Jan [[FREE Full text](#)] [doi: [10.1186/1471-2296-10-57](https://doi.org/10.1186/1471-2296-10-57)]
38. gp-patient. uk. co URL: <https://gp-patient.co.uk/about> [accessed 2014-08-13] [[WebCite Cache ID 6RnSCcflL](#)]
39. Roberts MJ, Campbell JL, Abel GA, Davey AF, Elmore NL, Maramba I, et al. Understanding high and low patient experience scores in primary care: analysis of patients' survey data for general practices and individual doctors. *BMJ* 2014 Nov 11;349(nov11 3):g6034-g6034 [[FREE Full text](#)] [doi: [10.1136/bmj.g6034](https://doi.org/10.1136/bmj.g6034)]
40. Steinbock D. 2014. What is TagCrowd? Internet URL: <http://www.tagcrowd.com/blog/about/> [accessed 2014-08-13] [[WebCite Cache ID 6RnSQLTlu](#)]
41. IBM. 2014. Many Eyes Internet URL: <http://www-958.ibm.com/software/analytics/manyeyes/page/Tour.html> [accessed 2014-08-13] [[WebCite Cache ID 6RnSWkipS](#)]
42. Sinclair S. ca – The Rhetoric of Text Analysis Internet. Voyeur Tools: See Through Your Texts | Hermeneuti URL: <http://hermeneuti.ca/voyeur> [accessed 2014-08-13] [[WebCite Cache ID 6RnSnhJw9](#)]
43. Baskarada S, Koronios A. Data, Information, Knowledge, Wisdom (DIKW): A Semiotic Theoretical and Empirical Exploration of the Hierarchy and its Quality Dimension. *Australasian Journal of Information Systems* 2013 Jun 6;18(1) [[FREE Full text](#)]
44. Leech NL, Onwuegbuzie AJ. An array of qualitative data analysis tools: A call for data analysis triangulation. *School Psychology Quarterly* 2007;22(4):557-584. [doi: [10.1037/1045-3830.22.4.557](https://doi.org/10.1037/1045-3830.22.4.557)]
45. Saiki LS, Cloyes KG. Blog Text About Female Incontinence: Presentation of Self, Disclosure, and Social Risk Assessment. *Nurs Res Internet*(2) 2014;63 [[FREE Full text](#)] [doi: [10.1097/NNR.0000000000000016](https://doi.org/10.1097/NNR.0000000000000016)]
46. Barker F, de LS, Baguley D, Gagne JP. An evaluation of audiology service improvement documentation in England using the chronic care model and content analysis. *Int J Audiol* 2014 Jun;53(6):377-382. [doi: [10.3109/14992027.2013.860242](https://doi.org/10.3109/14992027.2013.860242)] [Medline: [24313709](https://pubmed.ncbi.nlm.nih.gov/24313709/)]
47. Arber S, Sawyer L. The role of the receptionist in general practice: A 'dragon behind the desk'? *Social Science & Medicine* 1985 Jan;20(9):911-921 [[FREE Full text](#)] [doi: [10.1016/0277-9536\(85\)90347-8](https://doi.org/10.1016/0277-9536(85)90347-8)]
48. Hammond J, Gravenhorst K, Funnell E, Beatty S, Hibbert D, Lamb J. Slaying the dragon myth: an ethnographic study of receptionists in UK general practice. *British Journal of General Practice* 2013 [[FREE Full text](#)] [doi: [10.3399/bjgp13X664225](https://doi.org/10.3399/bjgp13X664225)]
49. Frazer Winsted K. Patient satisfaction with medical encounters – a cross-cultural perspective. *International Journal of Service Industry Management* 2000 Dec;11(5):399-421. [doi: [10.1108/09564230010360137](https://doi.org/10.1108/09564230010360137)]
50. Xiao H, Barber JP. The effect of perceived health status on patient satisfaction. *Value Health* 2008 Jul;11(4):719-725. [doi: [10.1111/j.1524-4733.2007.00294.x](https://doi.org/10.1111/j.1524-4733.2007.00294.x)] [Medline: [18179667](https://pubmed.ncbi.nlm.nih.gov/18179667/)]
51. Alemi F, Torii M, Clementz L, Aron DC. Feasibility of real-time satisfaction surveys through automated analysis of patients' unstructured comments and sentiments. *Qual Manag Health Care* 2012;21(1):9-19. [doi: [10.1097/QMH.0b013e3182417fc4](https://doi.org/10.1097/QMH.0b013e3182417fc4)] [Medline: [22207014](https://pubmed.ncbi.nlm.nih.gov/22207014/)]
52. Greaves F, Ramirez-Cano D, Millett C, Darzi A, Donaldson L. Use of sentiment analysis for capturing patient experience from free-text comments posted online. *J Med Internet Res* 2013;15(11):e239 [[FREE Full text](#)] [doi: [10.2196/jmir.2721](https://doi.org/10.2196/jmir.2721)] [Medline: [24184993](https://pubmed.ncbi.nlm.nih.gov/24184993/)]

53. Cambria E, Hussain A, Durrani T, Havasi C, Eckl C, Munro J. Sentic Computing for patient centered applications. : IEEE; 2010 Presented at: IEEE 10th Int Conf SIGNAL Process Proc Internet; 2010; Beijing p. 1279-1282 URL: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5657072> [doi: [10.1109/ICOSP.2010.5657072](https://doi.org/10.1109/ICOSP.2010.5657072)]
54. Sokolova M, Bobicev V. Sentiments and Opinions in Health-related Web Messages. In: Proceedings of Recent Advances in Natural Language Processing. 2011 Presented at: RANLP 2011; 12-14 September 2011; Hissar, Bulgaria p. 132-139.
55. Sokolova M, Bobicev V. What Sentiments Can Be Found in Medical Forums? In: Proceedings of Recent Advances in Natural Language Processing 2013. 2013 Presented at: RANLP; 7-13 September 2013; Hissar, Bulgaria p. 639.
56. Bobicev V, Sokolova M, Jafer Y, Schramm D. Learning Sentiments from Tweets with Personal Health Information. : Springer Berlin Heidelberg; 2012 Presented at: 25th Canadian conference on Advances in Artificial Intelligence; May 28-30 2012; Toronto, Canada p. 37-48. [doi: [10.1007/978-3-642-30353-1_4](https://doi.org/10.1007/978-3-642-30353-1_4)]
57. Bobicev V, Sokolova M, Oakes M. Recognition of Sentiment Sequences in Online Discussions. In: Proceedings of the Second Workshop on Natural Language Processing for Social Media (SocialNLP). 2014 Presented at: SocialNLP; August 24 2014; Dublin, Ireland p. 44-49 URL: http://www.google.ca/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CCMQFjAA&url=http%3A%2F%2Fwww.aclweb.org%2Fanthology%2FW14-5907&ei=jfGvKjALaxggTw&Ggbw&usq=AFQjCNGx-VGrlkO6RT6d6H_rP_VJOA&sig2=ICX_gu8OwGseFFNBmMoFw
58. Das D, Martins AFT. A survey on automatic text summarization. Lit Surv Lang Stat II course C 2007;4:192-195 [FREE Full text]
59. Blei DM, Ng AY, Jordan MI. Latent dirichllocation. Journal of Machine Learning Research 3 2003:993-102 [FREE Full text]
60. 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):1098-1103. [doi: [10.1136/amiajnl-2014-002711](https://doi.org/10.1136/amiajnl-2014-002711)] [Medline: [24918109](https://pubmed.ncbi.nlm.nih.gov/24918109/)]
61. Lyles CR, Sarkar U. Additional considerations for 'Harnessing the cloud of patient experience'. BMJ Qual Saf 2013 Aug;22(8):698 [FREE Full text] [doi: [10.1136/bmjqs-2013-001893](https://doi.org/10.1136/bmjqs-2013-001893)] [Medline: [23476069](https://pubmed.ncbi.nlm.nih.gov/23476069/)]
62. Dijkum C, Lam N, Verheul W. The challenge of modeling complexity in the social sciences illustrated with an example: the communication between GP and Patients. Methodol Rev Appl Res Internet 2013 [FREE Full text]

Abbreviations

ASCII: American Standard Code for Information Interchange

CI: confidence interval

GPPS: General Practice Patient Survey

KWIC: keyword in context

NHS England: National Health Service England

NIHR: National Institute for Health Research

NLP: natural language processing

OR: odds ratio

Edited by G Eysenbach; submitted 22.08.14; peer-reviewed by M Sokolova, WY Wang; comments to author 21.10.14; revised version received 02.12.14; accepted 20.12.14; published 06.05.15

Please cite as:

Maramba ID, Davey A, Elliott MN, Roberts M, Roland M, Brown F, Burt J, Boiko O, Campbell J

Web-Based Textual Analysis of Free-Text Patient Experience Comments From a Survey in Primary Care

JMIR Med Inform 2015;3(2):e20

URL: <http://medinform.jmir.org/2015/2/e20/>

doi: [10.2196/medinform.3783](https://doi.org/10.2196/medinform.3783)

PMID: [25947632](https://pubmed.ncbi.nlm.nih.gov/25947632/)

©Inocencio Daniel Maramba, Antoinette Davey, Marc N Elliott, Martin Roberts, Martin Roland, Finlay Brown, Jenni Burt, Olga Boiko, John Campbell. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 06.05.2015. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.