Quantitative Insights from Qualitative Data: An Example from the Healthcare Sector

Christine Pitt
Michael Mulvey
Jan Kietzmann

Abstract

Among the deluge of online data generated by users in the form of text on social media sites, health care reviews are among the most common, and potentially, the most insightful. Patients review and comment on the experiences with procedures as varied as hysterectomies, colonoscopies and chemotherapy. In their attempts to reduce the uncertainty associated with medical treatments, many patients nowadays also turn to social media, where they rely on the experiences articulated by other patients. In this study, IBM Watson is used to examine how knee replacement patients talk about their emotions and express sentiment through their comments online. Then a latent class cluster modelling procedure is used to segment these patients into distinct groups, according to their emotions (anger, disgust, fear, happiness, sadness and surprise), sentiment, and their overall satisfaction with knee replacement surgery. The findings show how qualitative online data can be transformed into quantitative insights regarding underlying market segments, which could then be targeted through different strategies by both marketers and healthcare practitioners.

Keywords: sentiment analysis; artificial intelligence; latent class cluster modelling; credence goods; health care

Introduction

In recent years, there has been an increased emphasis on the availability of online data and their potential value for the study of human social behavior (e.g., Chang, Kauffman & Kwon, 2014; Newman, 2016). Under the umbrella of “Big Data”, practitioners and academics use and discuss data generated from an increasing plurality of sources, including internet and mobile transactions, purposefully generated content through sensor networks or business transactions, as well as user-generated content (in the form of text, pictures and video) on social media (George, Haas & Pentland, 2014). Big Data’s value for researchers, particularly those in the social and behavioral sciences it is contended, comes “from the patterns that can be derived by making connections between pieces of data, about an individual, about individuals in relation to others, about groups of people, or simply about the structure of information itself.” (Boyd & Crawford, 2011, p. 2).

The “four Vs of Big Data” describe the complexity data analysts face (Dabirian, Kietzmann & Diba, 2017). Consider the sheer *volume* of data: by 2020, most companies in the U.S. will have 100 terabytes of data stored, and 2.3 trillion bytes of data will be created daily. Data come in an increasing *variety* of different formats and sources, including from organizational systems, but also from wearable devices, the Internet of Things, and social media. Only about 20% of our data are structured in a way that was once thought of as data, e.g., neatly assigning each customer name with a phone number and address that fit into a well-organized data table. The remaining 80% of all the world’s data are unstructured, including user-generated texts, photos, videos, social media updates, etc.; all highly qualitative formats that don’t follow well defined sets of rules. The *velocity* of data describes how quickly data are generated, collected and analyzed today. The New York Stock Exchange, for instance, creates 1 terabyte of trade information during each trading session, allowing traders to analyze data on-the-fly, as they are being generated. Messages on Facebook Messenger and WhatsApp combined are three times the global volume of SMS messages today. Ten years ago, in 2008, neither Facebook Messenger nor WhatsApp even existed. Lastly, the *veracity* of data addresses the accuracy and the degree of trustworthiness of big data - a major concern for research and practice (The Four V’s of Big Data, n.d.).

Considering the complexity of big data, the potential value can only be reaped when researchers take into account the intricate methodological processes that underlie the analysis of social data (Boyd & Crawford, 2011). The challenge for marketing is to make sense of the enormous volume, variety, velocity and veracity of unstructured, qualitative, user-generated data. The trouble is that, as Peter Norvig, Google’s director of research puts it: “We don’t have better algorithms. We just have more data.” (Cleland, 2011).

This paper takes up the challenge of making sense of online qualitative data. In order to contribute to the understanding of how qualitative data can be examined and made more meaningful (quantifiable) for marketers, a research context was chosen that was a) large in data volume, b) likely growing in terms of importance and velocity, c) unstructured in its format and d) highly subjective, populated by users and thus of uncertain value to the marketer. From the marketing literature, credence goods most certainly fit this bill. As this paper will show below, prosthetics are a highly appropriate research context. More specifically, the focus is on analyzing knee replacement patients’ online reviews about their highly emotional experiences in order to show why and how such reviews might be useful for practitioners and medical professionals.

The paper proceeds as follows: After briefly introducing the context of prosthetics and knee replacement procedures as credence offerings, a review of the literature on emotions and sentiment is presented. Next, a study of the content in an online forum in which patients who have undergone knee replacement surgery talk about their experiences and how they feel is, described. Then, an explanation is given of how this content was analyzed using IBM Watson’s natural language processing capabilities,

so that scores could be obtained for the primary emotions expressed in each piece of text, as well as for the overall sentiment of the written word. Then follows a description of a latent class cluster modelling procedure that enables the patients to be segmented into distinct groups that differ according to emotions and sentiment, and also according to their overall satisfaction with knee replacement surgery. Finally, concluding remarks begin by acknowledging the limitations of the research, noting the implications for healthcare managers and marketers, and by identifying avenues for future research in this area.

Prosthetics and Knee Replacements

Prostheses, or artificial devices that replace body parts have been around for a long time. Long John Silver, the villain in Robert Louis Stevenson's 19th century adventure novel *Treasure Island* had lost his left leg at the hip, and under the left shoulder he carried a crutch to help him walk on his wooden "peg leg". Nowadays prostheses are far more common as they have come to enhance or replace parts of the human body, but most of them are hidden, or become part of the body itself. Examples include breast implants, pacemakers, and hip- and knee joint replacements.

In parallel to the increased affordability and availability of reliable prostheses, demand is also on the rise. Natarajan (2012), in his polemic of marketing in the 21st century, discusses how Aging, Culture, and Leisure, and their intersections, are key elements that set us apart from our predecessors. For one, people can expect to live longer, and their increased longevity means that especially retirees have changing ideas of how to spend their "sunset years". Consumer behavior, especially at an older age, starts to shift, frequently leading to an increase in demand for prosthetics and restorative care. Subsequently, the way in which aging consumers seek advice on these services and products is taking place online more often as the number of users over the age of 60 continues to increase (Nimrod, 2010).

From a consumer marketing perspective, undergoing a serious medical treatment is, for the average individual, a "credence" good (Shaffer & Sherrell, 1997). Nelson (1970) distinguished between "search" and "experience" offerings. Search offerings are those that the consumer can search for, and ascertain their quality during the search process; for example, the style of a dress. Experience goods, on the other hand are those offerings whose quality the consumer will only be able to judge after they have been purchased and consumed; for example, a meal at a restaurant. Darby and Karni (1973) add a third category of offering, that they term "credence" goods. Credence offerings are those for which the quality cannot be ascertained through search, and cannot even be realistically judged after their purchase and consumption. They offer as an example the removal of a patient's appendix: Most patients don't have the expertise to know whether their appendix even needed removing or not, and once it has been removed, most patients are not really in a position to judge whether the surgeon performed the surgery poorly or well. Knee replacement surgery presents an identical situation: The average individual is not well placed to know whether the solution to their problems is a knee replacement or some other form of treatment. After surgery, the patient is not really capable of judging whether the success of the procedure was due to the surgeon's skills and abilities, or whether their own and other environmental circumstances (their health, their pre- and post-surgery care, and so forth) were more instrumental to the success. Likewise, in the case of an unsuccessful knee replacement procedure, the patient is not really in a position to allocate blame to the surgeon, or to the fact that their own circumstances (poor general health, obesity etc.) or pre-and post- procedure care are at fault. Credence is indeed key: the consumers simply believe, or don't; or perhaps, they have to believe.

Since the first knee replacement surgery was carried out in 1968, total knee replacements have become not only one of the most common, but also most successful procedures in all of medicine (Trebše & Mihelič, 2012). Improvements in surgical materials and techniques have greatly increased the

procedure's effectiveness, and according to the Agency for Healthcare Research and Quality, more than 600,000 knee replacements are performed each year in the United States alone (see <http://orthoinfo.aaos.org/topic.cfm?topic=a00389>). Most of these procedures are carried out in order to alleviate the pain caused by osteoarthritis, and to afford the patient greater mobility. Yet knee surgery is not without its problems and complications. Apart from the severe pain suffered by many patients' post-surgery, there is also a marked risk of infection, and many patients are susceptible to blood clotting. There are also patients who are allergic to the morphine-based drugs used to treat the pain, and then need to ameliorate pain in other ways, and many of those who do take morphine for pain then endure extreme constipation as a result. In other words, knee replacements are an appropriate research context since a) they have been practiced for a long time, b) they are most certainly a credence offering, and c) recovery speed and success rates vary considerably, all leading to a trough of knee replacement relevant data online.

This provides an appropriate research context as knee replacement surgeries are a common procedure among people over the age of 50, mostly occasioned by the onset of osteoarthritis and the bone deterioration and subsequent discomfort and immobility that this causes. It can be expected that the demand for these procedures will continue to increase as the population ages and continues to live longer.

Patients facing knee replacement surgery, and those who have undergone the procedure undoubtedly experience a mix of emotions. They might be angry about having to undergo the operation, and understandably fearful. They might be sad about losing their mobility and their inability to enjoy aspects of life that they had previously taken for granted. They might feel disgust at some of the complications the procedure induces, but they might also feel joy over the successful outcome of their procedure. They will also hold some level of overall sentiment toward the procedure that will be positive or negative. In the following section, emotions and sentiment are discussed in more detail.

Emotions and Sentiment

Emotions

Cabanac (2002) defines an emotion as any mental experience that has high intensity and high hedonic content; in other words, it can be pleasurable or it can evoke unhappiness. There have been numerous attempts by social scientists to distinguish between, classify and define emotions. One of the best-known classifications of emotions is that of the psychologist Paul Ekman (1992; see also Handel, 2012), who distinguished six emotions as basic, namely (definitions taken from www.dictionary.com):

- *Anger*: a strong feeling of displeasure and belligerence aroused by a wrong.
- *Disgust*: a strong feeling of dislike for something that has a very unpleasant appearance, taste, smell, and so forth.
- *Fear*: a distressing emotion aroused by impending danger, evil, pain, etc.
- *Happiness* (an active or passive state of pleasure or pleasurable satisfaction) or *Joy* (the emotion evoked by well-being, success, or good fortune or by the prospect of possessing what one desires).
- *Sadness*: showing, expressing, or feeling sorrow or unhappiness.
- *Surprise*: occurs with a sudden feeling of wonder or astonishment, as through unexpectedness.

The research described here focuses on five of these emotions, all of which are pertinent to knee replacement surgery. A patient can feel *anger*, especially, for example, if they had to wait a long time for the surgery, or if they didn't receive the care they expected, or if the surgery was unsuccessful and left

them no better off than they were before. A patient can experience *disgust*, especially for example, if they experienced infection, or some other complication. A patient can experience *fear*, especially for example, if the surgery was unsuccessful, and there is the possibility that they will have to undergo it again, or if the healing is not going according to what they expected, and the future is uncertain. A patient can experience *joy*, especially for example, if they are able to walk again without pain, or because they are now able to do things they weren't able to do previously. And, a patient can experience *sadness*, especially for example, if the surgery didn't go according to plan, and they yearn for what might have been. The emotions of anger, disgust, fear, joy and sadness form the basis of analysis for the analysis tool used in this research (surprise is the exception), so that this is the classification of emotions that will be followed here.

Sentiment and Sentiment Analysis

The term “sentiment” is defined as “an attitude, thought, or judgment prompted by feeling” (“Sentiment,” n.d.). The advent of social media has meant that much of the data being generated today are in the form of text and is mostly qualitative in nature (comments on Facebook, tweets on Twitter, reviews on websites such as TripAdvisor, and blogs) (Weinberg, Davis & Berger, 2013). Sentiment analysis, also known as opinion mining, has developed as a tool to make quantitative sense out of what is essentially qualitative data. It aims to determine the “sentiment”, as defined above, of the speaker or author of a piece of text, and can range from negative to positive as scored on whatever scale the particular sentiment analysis software chooses to use. This is also referred to as the “polarity” of the document (e.g. Turney, 2002).

When considering the words used by consumers in their observations on a particular offering, key questions to be addressed include which emotions they emphasize, the polarity of the sentiment they express, and the extent to which the sentiment is influenced significantly by the individual's emotions. It is unlikely that all of the emotions impact the sentiment significantly, and so it is also worth knowing, for any consumer or group of consumers, which emotions expressed have the greatest impact on their sentiment. It is also important for marketing decision makers, in this case health care professionals, to be able to understand consumer sub-groups at a more fine-grained level, and to be able to develop strategies to target marketing efforts specifically to them. In order to investigate these questions from the viewpoint of knee replacement patients, a study was conducted that measured the extent of both the emotions they conveyed, and the overall sentiment they expressed. To do this, IBM's Watson Alchemy service was employed, which is described in the next section. Then, by means of a latent class cluster modelling, distinct sub-groups, or market segments of knee replacement patients were identified.

IBM Watson's Alchemy Emotion and Sentiment analysis

IBM describes its Watson technology as a platform that uses natural language processing and machine learning to reveal insights from large amounts of unstructured data (<http://www.ibm.com/watson/what-is-watson.html>). The qualitative data that are available online covers a broad range of topics from entertainment to fashion to education and, as this example explores, health and medical advice. The available data exist in many different forms of user-generated content, including social media posts and online product reviews, and more professional content, such as curated content and corporate websites. While previously the creators of this content were predominantly tech savvy, younger users, nowadays, many of the content creators are over the age of 60. As the number of these Internet users continues to increase (Nimrod, 2010), more and more online content will indeed be generated by older Internet users.

There are a number of online forums where patients and medical professionals can share information and discuss various conditions. For example, Keeling, Laing and Newholm (2015) conducted their studies on four online communities focusing on breast cancer, prostate cancer, depression, and diabetes. The data in the current study was gathered in the form of 104 personal reviews of their knee replacement surgery by patients on the website Medicinenet.com (http://www.medicinenet.com/total_knee_replacement/patient-comments-1494.htm). In these, patients expressed in their own words how their surgery had proceeded, and how they felt about the experience. Each review was copied from the website, and pasted as text into the AlchemyLanguage tool on IBM's Watson Developer Cloud for analysis. The IBM Watson AlchemyLanguage service is a collection of text analysis functions that derive semantic information from any document, including text, HTML code, or the URL of a website. It exploits sophisticated natural language processing techniques to rapidly obtain a high-level of comprehension of the content. One of its main features is the ability to understand the emotions expressed in the language used in the text, and also to gain an overall indication of the sentiment of the document. The Alchemy tool was used to detect the emotions implied in the text of each patient's review, and also to uncover the overall sentiment of the review.

Results and Discussion

With regard to gender 59.6% of the respondents were female and 40.4% were male. With regard to age, 2.9% of the respondents were in the age group 35-44, 14.4% in the age group 45-54, 48.1% in the age group 55-64, 30.8% in the age group 65-74, and 3.8% in the age group 75 or older. The overall results of the study are summarized in Table 1 below. The average number of words used in a review by a knee replacement patient was 142.45, with the longest review analyzed being 397 words, and the shortest, 37 words. Watson scores the emotions on a scale ranging from 0 to 1, and sentiment is scored on a scale ranging from -1 to +1. One part of the analysis was indeed conducted manually: It was observed that a number of individuals had either stated that "knowing what they knew now, they would never have had the surgery (or something very similar)" (17.4%), or "I only wish I had done this earlier" (28.8%). The remaining individuals had not stated a preference either way (53.8%). These were all identified by means of a manual reading of the text.

One of the most important advantages of using Watson is the ability to go back to the data and identify who the respondents were that expressed the highest and lowest scores on the five emotions, as well as sentiment, and determine the precise words that they used.

Seven of these patients were as follows:

High anger (0.91), Patient 3: "Nothing works to take the swelling down. When I said something to him (the surgeon), he actually blamed me for overworking my knee! I no longer see him but I am unable to do very much because the swelling is so bad. I am getting so depressed and wish I never had this done."

High Disgust (0.72), Patient 101: "The worst thing was the nausea from the pain medicines they gave me the first two days."

High Fear (0.99), Patient 29: "I was down to where I could not walk 100 yards. I do really feel for patients that have had poor results and are still in pain. I did nothing special except maybe get lucky with my surgery and surgeon. I can tell you that this surgery hurts."

High Joy (0.75), Patient 95: "I am very happy to find this website and realize what I am going through is normal. I am 5 weeks out from bilateral knee replacement surgery and am apparently doing great!"

High Sadness (0.65), Patient 77: “I’ve been doing all my exercises but the last week my knee is swelling up and the pain is getting worse. I am starting to get a bit worried if this is normal. I’m back to see the consultant next week, I am starting to get worried.”

Not surprisingly, the overall sentiment expressed by most knee replacement patients is negative, with a mean sentiment of -0.39 for the sample. However, the range is considerable, with a low of -0.88, and a high positive sentiment of 0.24. *Patient 53 with the lowest sentiment of -0.88* says, “I’m so disappointed with my recovery that I cry every day. I can only straighten my knee to 4 degrees, but that is not what has me so stressed out. My greatest disappointment is that my range of motion (ROM) is only 85 degrees and when I am bent to that degree I am in the most excruciating pain. In most of the posts I have read on different sites, many patients have stated that they have their ROM of 115 in just a few weeks. My surgeon insists that I am not working hard enough, but believe me I am working to the best of my ability, but the pain is unbearable.” On the other hand, *Patient 80 with the highest, positive sentiment of .24* says, “I’m doing very well. My leg extension is at 0 degrees and flex is at 110! The surgery was done at 8:30 am and I walked out of the hospital at 1:00 pm. Yes, I had some pain but between the exercises I do daily, the walking I do without any devices, my physiotherapy assistance and little reliance on pain medicines, icing and using the CPM machine, this was a great experience. My surgeon was and is outstanding. One thing I would encourage everyone to do is really exercising the affected knee on a daily basis prior to the surgery. And finally, never miss a day of doing the exercising and walking!”

Table 1: Overall Results – Word Count, Emotions and Sentiment of Knee Replacement Patients

	Word Count	Anger (scored 0 to 1)	Disgust (scored 0 to 1)	Fear (scored 0 to 1)	Joy (scored 0 to 1)	Sadness (scored 0 to 1)	Sentiment (scored -1 to 1)
Mean	142.45	0.47	0.16	0.54	0.06	0.36	-0.39
Std. Dev.	69.24	0.21	0.13	0.19	0.09	0.14	0.24
Maximum	397.00	0.91	0.72	0.99	0.75	0.65	0.24
Minimum	37.00	0.09	0.02	0.11	0.01	0.07	-0.88

Latent Class Cluster Modelling

Consumers’ emotions regarding offerings tend to exert strong influences on their judgments. Accordingly, marketing decision makers are advised to attend to the sources of feelings that can be attributed to these offerings, especially for hedonic and experiential categories and settings (Pham, 2011). Understanding individual differences is paramount in marketing and consumer research. The qualitative findings presented illustrated the types of variance found in the sentiment and emotions expressed in patients’ reviews. Although theory and data support the idea that consumers’ emotional reactions often vary for the same offering, models provide a useful tool to evaluate the level of descriptive and predictive consistency of those differences.

In this section, a formal exploration of the extent to which the emotions expressed in patient reviews impact patient sentiment and satisfaction. Specifically, this analysis addresses the following research questions:

RQ₁: Are there different profiles of patient emotion expressed in the reviews?

RQ₂: How many profiles are there?

RQ₃: What are the profiles?

RQ₄: Are the profiles predictive of patient sentiment and patient satisfaction (in terms of whether they would or would not wish to undergo the procedure again, or are indifferent to it)?

	Patient Emotion Cluster								
	Cluster1	Cluster2	Cluster3	Intercepts			Clusters		
	<i>Baseline</i>	<i>Contented</i>	<i>Joyful</i>	Overall	Wald	p-value	Wald(=)	p-value	R ²
Model									
Intercept	0	-0.87	-3.43				32.8	0.00	
Indicators									
Anger	0	-0.26	-0.34	0.55	515.8	0.00	62.6	0.00	0.33
Disgust	0	-0.10	-0.14	0.19	125.8	0.00	25.0	0.00	0.13
Fear	0	0.02	-0.24	0.53	596.8	0.00	10.9	0.00	0.04
Joy	0	0.05	0.64	0.03	453.5	0.00	144.6	0.00	0.93
Sadness	0	-0.02	-0.20	0.37	486.1	0.00	19.2	0.00	0.05
Dependents									

Sentiment	0	0.29	0.52	-0.49	347.7	0.00	46.6	0.00	0.36
Satisfaction	0	2.02	3.49		21.6	0.00	16.0	0.00	0.25
wish I hadn't				0.00					
neither				0.94					
glad I did				-0.59					

Next, the Step3-Dependent submodule in LatentGOLD 5.1 (Vermunt & Magidson, 2015) was used to analyze the extent to which membership in the patient emotion clusters predicts the distal outcomes of patient sentiment (using the Watson score) and satisfaction (using the researcher-coded three-level ordinal outcome: wish I hadn't, neither, wish I had earlier). The Step-3 analysis corrects for classification error to prevent bias and applies the BCH adjustment, which requires no strong distributional assumptions about the distal outcomes (Bakk, Oberski, & Vermunt, 2016).

As presented in Table 3, patient sentiment was significantly related to the patient emotion clusters (Ward = 347.7, $p = 0.00$) and the level of sentiment varied across clusters (Ward(=) = 46.6, $p = 0.00$). Overall, 36 percent of the variance in sentiment was accounted for by the three patient emotion clusters. Specifically, sentiment was lowest among Cluster 1 members (beta = 0.00), was higher among Cluster 2 members (beta = 0.29), and highest among Cluster 3 members (beta = 0.52). This result affirms the role of basic emotions (and the three patient emotion clusters) in explaining variance in patients' attitudes toward knee replacement surgery.

Patient satisfaction was also significantly related to the patient emotion clusters (Ward = 21.6, $p = 0.00$) and the level of sentiment varied across clusters (Ward(=) = 16.0, $p = 0.00$). Overall, 25 percent of the variance in satisfaction was accounted for by the three patient emotion clusters. Satisfaction was lowest among Cluster 1 members (beta = 0.00), was higher among Cluster 2 members (beta = 2.02), and highest among Cluster 3 members (beta = 3.49). This result affirms the role of basic emotions and patients' emotion clusters in explaining variance in patients' satisfaction with the knee replacement surgery experience.

The model calculated cluster-specific means of the indicator emotions and dependent outcomes, and also estimated patient cluster membership probabilities based on their emotion mix. Table 4 provides a succinct representation of the results that can be used to compare the classification of the seven illustrative case patients. As would be expected, patients who exhibited the highest levels of negative emotions were all classified members of Segment 1: Patient 3 (highest anger), Patient 29 (highest fear), Patient 77 (highest sadness), and Patient 101 (highest disgust) as well as Patient 53 (lowest sentiment). Segment 2 includes Patient 80 (highest sentiment) and Segment 3 includes Patient 95 (highest joy). Interestingly, the patients characterized by high fear (Patient 29) and high sadness (Patient 77) shared membership between Segment 1 (the modal classification) and Segment 2.

Table 4: Patient Emotion Cluster Profiles

	Patient Emotion Cluster		
	Cluster1	Cluster2	Cluster3
	<i>Baseline</i>	<i>Contented</i>	<i>Joyful</i>
Cluster Size	68.9%	28.9%	2.2%
Indicators			
Anger	0.551	0.293	0.209
Disgust	0.193	0.093	0.056
Fear	0.535	0.557	0.292

Joy	0.030	0.080	0.665
Sadness	0.368	0.350	0.166
<i>Dependents</i>			
Sentiment	-0.487	-0.193	0.031
Satisfaction			
wish I hadn't	0.243	0.019	0.002
neither	0.622	0.372	0.123
glad I did	0.135	0.608	0.875
Mean	0.892	1.589	1.874

Membership Probabilities

<i>Patient</i>	<i>Scale (Highest Score)</i>			
3	Anger (0.91)	1.00	0.00	0.00
101	Disgust (0.72)	1.00	0.00	0.00
29	Fear (0.99)	0.61	0.39	0.00
95	Joy (0.75)	0.00	0.00	1.00
77	Sadness (0.65)	0.80	0.20	0.00
53	Sentiment – low (-0.88)	1.00	0.00	0.00
80	Sentiment – high (0.24)	0.23	0.77	0.00

One of the most powerful features of LC analysis is its ability to identify and accommodate “outliers” in the model, unlike conventional regression analysis, which stipulates that the data be free of outliers (Ding, 2008). In the present dataset, Cluster 3 members Patients 59 and 95 reported a unique level of joy and formed a distinct cluster, representing a small yet important subgroup among patients who have had knee replacement surgery. Interpretive consumer researchers seek to study extreme cases (participants with unique stories) as they are often the most interesting and informative in understanding consumer behavior (Belk, Fischer, & Kozinets, 2012). So, consistent with the qualitative research tradition, the authors find that LC analysis has the capacity to identify individuals that have less common, extreme, or deviant views.

Finally, a Step3-Covariate test to explore differences in demographics (age and gender) across patient emotion clusters was also conducted, however the effects were not statistically significant.

Discussion: Limitations, Managerial Implications and Future Research

This paper has sought to illustrate how the qualitative data contained in online reviews of a credence good, namely knee replacement surgery, can be analyzed in a way that converts them into quantitative data that can be used in statistical analysis. The analysis conducted here permitted the identification of different segments of knee replacement patients, that can inform the marketing decision making and strategies of medical care professionals. The study presented here has a number of limitations. First, the sample, while adequate for statistical analysis is still rather small, especially when one considers how many knee replacement procedures are conducted annually throughout the world. Second, there is the possibility that there is some kind of response bias in the data, and that there are extreme responses at both ends of the spectrum. Those who are delighted at the outcome of their surgery might want to share

their joy and optimism with a large online audience, and those who are in extreme pain, frustrated, simultaneously angry and disgusted and also sad, might want to have a virtual shoulder to cry on.

Nevertheless, the findings of the research may provide insights to both marketers and healthcare practitioners. First, the methodology employed here enables qualitative data, to be manipulated in such a way that it sheds quantitative light onto issues that are important to decision makers. Second, the methodology is relatively simple and inexpensive. It merely requires copying and pasting text into a website that processes it at no cost, and entering the resulting scores into a spreadsheet for further manipulation. Third, the data obtained delivers real insights into the feelings and emotions of consumers who have endured a serious surgical procedure. It permits qualitative data to be transformed into quantitative insights into underlying market segments, which could then be targeted by means of different marketing strategies. It allows service providers to understand what angers, disgusts and saddens those who have endured a negative experience that impacts their overall sentiment toward a credence offering, while at the same time providing a lens into the emotions of those who have come away with a positive experience. Those who treat knee replacement patients, be they surgeons, nurses, physiotherapists or other care givers might use these insights either to better prepare patients for the knee replacement procedure, or to help them come to terms with their situation after the operation.

The research conducted in this study, and the findings, also suggest a number of avenues for future research. First, the texts used in this research were derived from relatively short reviews – the longest was only 397 words. This is adequate for sentiment analysis, but if longer reviews could be obtained, the IBM Watson text analysis suite can provide additional insight into the origin of the text. For example, it is also possible to gain considerable understanding of the individual personality and the values of the person whose words are of interest using the Watson Personality Insights tool (<http://www.ibm.com/watson/developercloud/personality-insights.html>). This tool ideally requires around 3,000 words to make a “strong” analysis. This type of data could be obtained by conducting in-depth semi-structured interviews with knee replacement patients, and the content transcribed into text files. Second, the text gathered in this exercise could also be analyzed using other content analysis tools, such as for example, DICTION, which uses Hart’s (1984a; 1984b) fundamental characteristics of text dimensions and calculated variables to shed light on the nature of a piece of text. These variables could be analyzed separately, or linked to the emotions and sentiment considered in this paper. Third, a host of other statistical tools could be used to explore the data further including structural equation modelling to establish causality, and correspondence analysis procedures to graphically assign respondents to more homogenous groups that could be targeted with marketing and create messages in more specific ways.

Knee replacement surgery is projected to be one of the most common surgical procedures to be performed in the near future, with the total number of procedures conducted per annum forecast to grow by 673% to 3.48 million procedures from 2005 to 2030 in the United States alone (Kurtz et al., 2007). But knee replacements are not the only prosthetic procedure that is on the rise. The growth of medical credence treatments can mainly be attributed to the demographics of an aging population, and to improvements and developments in surgical technique, that will give the option of a surgical solution to health problems to more people. Prosthetic surgeries are also not the only medical treatment or condition to have garnered patient comments on the website, MedicineNet.com. The site also contains reviews on many different medical conditions and treatments such as hysterectomies, colonoscopies, hip replacements, arthritis, ADHD, and hernias to name a few. A better understanding of the emotions and experiences of patients using qualitative text data such as reviews will, in the long run, undoubtedly result in the delivery of an improved standard of overall care.

In the past, marketing researchers have tended to find themselves in one of three camps, as their research objectives determine the data gathering and analysis methodologies they employ. Those who

strive for generalizability choose surveys, those who pursue precision and control opt for experiments, and those who seek context and richness utilize qualitative methods. As McGrath (1981) pointed out more than thirty years ago, social science research is never perfect because researchers are sitting on three-horned dilemmas – striving for one kind of objective and using one approach immediately cause the researcher to forego the other objectives that could have been achieved using their appropriate tools. The qualitative approach followed here has the potential to provide richness and context, which is especially important in light of the increasing volume, velocity, veracity and variability of user-generated content online. This paper demonstrates, that when used in conjunction with powerful natural language processing and machine learning technologies and statistical software, the marketing researcher now has the potential to offer more generalizability, and perhaps even, precision.

Sources

- Bakk, Z., Oberski, D. L., & Vermunt, J. K. (2016). Relating latent class membership to continuous distal outcomes: improving the LTB approach and a modified three-step implementation. *Structural Equation Modeling: A Multidisciplinary Journal*, 23(2), 278-289.
- Belk, R., Fischer, E., & Kozinets, R. V. (2012). *Qualitative Consumer and Marketing Research*. Sage.
- Boyd, D., & Crawford, K. (2011). Six provocations for big data. In *A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society* (Vol. 21). Oxford: Oxford Internet Institute.
- Cabanac, M. (2002). What is emotion? *Behavioural Processes*, 60(2), 69-83.
- Chang, R. M., Kauffman, R. J., & Kwon, Y. (2014). Understanding the paradigm shift to computational social science in the presence of big data. *Decision Support Systems*, 63, 67-80.
- Cleland, S. (2011, October). Google's "infringenovation" secrets. *Forbes*. Retrieved from <http://www.forbes.com/sites/scottcleland/2011/10/03/googles-infringenovation-secrets>
- Dabirian, A., Kietzmann, J., & Diba, H. (2017). A great place to work!? Understanding crowdsourced employer branding. *Business Horizons*, 60(2), 197-205.
- Darby, M. R., & Karni, E. (1973). Free competition and the optimal amount of fraud. *The Journal of Law and Economics*, 16(1), 67-88.
- DeSarbo, W. S., Kamakura, W. A., & Wedel, M. (2006). Latent structure regression. In R. Grover & M. Vriens (Eds.), *The Handbook of Marketing Research: Uses, Misuses, and Future Advances* (pp. 394-417). Thousand Oaks CA: Sage.
- Ding, C. S. (2008). Enhancing accuracy in research using regression mixture analysis. In J. W. Osborne (Ed.), *Best Practices in Quantitative Methods* (pp. 410-421). Thousand Oaks, CA: Sage Publications.
- Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, 6(3-4), 169-200.
- George, G., Haas, M. R., & Pentland, A. (2014). Big data and management. *Academy of Management Journal*, 57(2), 321-326.
- Handel, S. (2012). Classification of Emotions. Retrieved from <http://www.theemotionmachine.com/classification-of-emotions>
- Hart, R. P. (1984a). Systematic analysis of political discourse: The development of diction. *Political Communication Yearbook*, 1, 97-134.
- Hart, R. P. (1984b). *Verbal style and the presidency: A computer-based analysis*. Academic Pr.
- Keeling, D. I., Laing, A., & Newholm, T. (2015). Health communities as permissible space: supporting negotiation to balance asymmetries. *Psychology & Marketing*, 32(3), 303-318.
- Kurtz, S., Ong, K., Lau, E., Mowat, F., & Halpern, M. (2007). Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. *The Journal of Bone and Joint Surgery*, 89(4), 780-785.
- McGrath, J. E. (1981). Dilemmatics: The study of research choices and dilemmas. *American Behavioral Scientist*, 25(2), 179-210.
- Natarajan, R. (2012). The countenance of marketing: a view from the 21st century ivory tower. *Quantitative Marketing and Marketing Management*, 485-499.
- Nelson, P. (1970). Information and consumer behavior. *Journal of Political Economy*, 78(2), 311-329.
- Newman, B. I. (2016). Reinforcing lessons for business from the marketing revolution in US presidential politics: A strategic triad. *Psychology & Marketing*, 33(10), 781-795.
- Nimrod, G. (2009). Seniors' online communities: A quantitative content analysis. *The Gerontologist*, 50(3), 382-392.
- Pham, M. T. (2011). The influence of feelings and emotions on consumers' judgments. In J. W. Alba (Ed.), *Consumer Insights: Findings from Behavioral Research* (pp. 69-70). Cambridge, MA: Marketing Science Institute.
- Sentiment. (n.d.). Retrieved from <https://www.merriam-webster.com/dictionary/sentiment>
- Shaffer, T. R., & Sherrell, D. L. (1997). Consumer satisfaction with health-care services: The influence of involvement. *Psychology & Marketing*, 14(3), 261-285.
- The Four V's of Big Data. (n.d.). Retrieved from <http://www.ibmbigdatahub.com/infographic/four-vs-big-data>
- Trebše, R., & Mihelič, A. (2012). Joint Replacement: historical overview. In *Infected Total Joint Arthroplasty* (pp. 7-11). Springer London.
- Turney, P. D. (2002). Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics* (pp. 417-424). Association for Computational Linguistics.
- Vermunt, J. K., & Magidson, J. (2015). Upgrade manual for Latent GOLD 5.1. Belmont, MA: Statistical Innovations.
- Weinberg, B. D., Davis, L., & Berger, P. D. (2013). Perspectives on big data. *Journal of Marketing Analytics*, 1(4), 187-201.