Journal of Diabetes Science and Technology

DIABETES ON TWITTER: A SENTIMENT ANALYSIS

Gabarron E; Dorronzoro E; Rivera-Romero O; Wynn R

Authors:

Elia Gabarron, PhD.
Norwegian Centre for E-health Research, University Hospital of North Norway.

Enrique Dorronzoro, PhD.
Department of Electronic Technology. Universidad de Sevilla. Avenida de Reina Mercedes S/N, 41012 Sevilla, Spain.

Octavio Rivera-Romero, Prof.
Department of Electronic Technology. Universidad de Sevilla, Spain. Avenida de Reina Mercedes S/N, 41012 Sevilla, Spain.

Rolf Wynn, M.D., Ph.D., Professor.
Department of Clinical Medicine, Faculty of Health Sciences,
UiT - The Arctic University of Norway, N-9037 Tromsø, Norway.
And Division of Mental Health and Addictions, University Hospital of North Norway, N-9291 Tromsø, Norway.

Corresponding author:
Elia Gabarron
Norwegian Centre for E-health Research
University Hospital of North Norway
Sykehusveien, 23
Tromsø, 9019, Norway
E-Mail: elia.gabarron@ehealthresearch.no

Abbreviations:
API: Application Programming Interface
PHP: Hypertext Preprocessor
SD: Standard Deviation
T1D: Type 1 Diabetes
T2D: Type 2 Diabetes

Keywords: Diabetes; Sentiment Analysis; Social Media; Twitter; Type 1 Diabetes; Type 2 Diabetes

Figures and table count: 1 Table + 2 Figures

ABSTRACT

Background: Contents published on social media have an impact on individuals and on their decision-making. Knowing the sentiment towards diabetes is fundamental to understanding the impact that such information could have on people affected with this health condition and their family members. The objective of this study is to analyze the sentiment expressed in messages on diabetes posted on Twitter.

Method: Tweets including one of the following terms (“diabetes”, “t1d”, and/or “t2d”) were extracted for one week using the Twitter standard API. Only the text message and the number of followers of the users were extracted. The sentiment analysis was performed by using SentiStrength.

Results: A total of 67421 tweets were automatically extracted, of those 3.7% specifically referred to T1D; and 6.8% specifically mentioned T2D. One or more emojis were included in 7.0% of the posts. Tweets specifically mentioning T2D and that did not include emojis were significantly more negative than the tweets that included emojis (-2.22 vs. -1.48), $p<0.001$. Tweets on T1D and that included emojis were both significantly more positive and also less negative than tweets without emojis (1.71 vs. 1.49; and -1.31 vs. -1.50 respectively), $p<0.005$. The number of followers had a negative association with positive sentiment strength (r = -0.023, $p<0.001$) and a positive association with negative sentiment (r = 0.016, $p<0.001$).

Conclusion: The use of sentiment analysis techniques on social media could increase our knowledge of how social media impact people with diabetes and their families and could help to improve public health strategies.

INTRODUCTION

Sentiment analyses are natural language processing techniques that use computational algorithms to extract subjective information from written text and that can identify the strength of the positive and negative tone of the message [1-3]. These techniques have been used to analyze and predict people's behaviour regarding elections; to analyze the launch of new products or services [4, 5]; and have also been used in the health field [6-8].

The use of sentiment analysis in healthcare could help us to better understand how people talk about and feel with respect to specific health topics or health conditions.

However, analyzing mood is not a simple task, as emotions often are mixed or ambivalent. It is possible to describe mixed emotions as the simultaneous experience of different combinations of opposing emotions, and positive and negative emotions can occur more or less simultaneously [9]. This may represent a challenge when analyzing the emotions expressed in texts.

Recent studies have analyzed the sentiment expressed in text language including emojis [10-15]. 'Emoji' is a Japanese word defined as "*small digital image or icon used to express an idea or emotion*" [16]. Emojis help users to better express their views and emotions by using graphics (i.e.; facial expressions; people; animals and nature; food and drink; activities; travel and destinations; objects; symbols; and flags) [10].

Sentiment analysis techniques examining both text and emojis can provide an overview of the moods communicated through specific topics, and also have the advantage of covering larger populations, and providing immediate results [7]. The use of sentiment analysis is especially relevant in a social media context, as social media have become a natural environment where people can share and seek health information [17-19]. Social media contents have an impact on individuals and on their decision-making [20-23], not only due to the type of information that can be found [24] but also for the expressed tone (i.e. emotions) or the connotations of the message.

Currently there are few publications studying the use of sentiment analysis related to diabetes [25-30]. Knowing the sentiment expressed by social media users towards diabetes is fundamental to understanding the impact that such information could have on people affected with this health condition and their family members. The objective of this study is to analyze the sentiment expressed in messages specifically focusing on diabetes posted on Twitter.

METHODS

Data collection

Aiming to use a sentiment analysis on tweets related with diabetes, we defined “diabetes”, “t1d”, and “t2d” as search terms. The term “diabetes” was chosen because it was the most commonly used term by the lead healthcare authorities and diabetes organizations in their Twitter profiles, such as the American Diabetes Association; CDC Diabetes; or the International Diabetes Association, among others. The terms “t1d” and “t2d” were chosen because they were very popular hashtags on Twitter referring to the two different types of diabetes. At the time of the study, and according to the Twitter API, the hashtag “t1d” had over 230 tweets daily; while “#t2d” had over 560 tweets. For one week (from May 23, 2018 to May 30, 2018), tweets containing at least one of these three keywords were collected using the standard API provided by Twitter. An API is a software that acts as intermediary between to applications allowing the communication between them. The standard API of Twitter provides a subset of the current tweets and their metadata. However it has several limitations compared with the premium one. Basically, the data obtained for every tweet is limited (i.e. replies to a tweet cannot be obtained with the standard API), the tweets obtained are a sample of the total tweets, the number of requests per minute is limited to 180, among other limitations. However, for this study the standard API functionality was considered adequate as the information retrieved and number of tweets retrieved was sufficient to conduct our study.

A software was developed to collect tweets using the scripting programming language PHP. This PHP software requests from Twitter, using Twitter API, the last tweets that contain one of the three keywords. These requests were done every 15 minutes for one week. For each tweet we retrieved: 1) The text of the tweet, including any emojis; and 2) the number of followers of the user that posted every extracted tweet. This information was stored in a MySQL database for sentiment analysis.

Tweets including any of the following words in the text were classified as referring specifically to T1D: “T1D”; “Type 1 Diabetes”; “Type 1”; or

“Type1Diabetes”. While tweets including any of the following words: “T2D”; “Type 2 Diabetes”; “Type 2”; or “Type2Diabetes” were identified and classified as tweets referring to T2D.

Anonymity and privacy

The data (i.e. tweets) used in this study had been made publicly available by being published openly on the Internet. Nevertheless, we wanted to take into consideration the privacy of the tweet emitters. Due to the impossibility of obtaining informed consent from all the tweet emitters -many of which did not use their actual name- we decided to extract only non-identifiable data. This means that no data identifying the emitters of the tweets was extracted or analyzed. The data that was analyzed was done so through the use of an automated process and thereafter aggregated, which further protected the privacy of the tweet emitters.

Analysis of tweets

The sentiment analysis of the extracted tweets was performed by using SentiStrength [2, 31, 32]. SentiStrength is a popular open-source software based on nonspecific messages designed to estimate the strength of positive and negative sentiment in short informal texts. Validation tests have shown that SentiStrength can detect positive emotion with 60.6% accuracy and negative emotion with 72.8% accuracy [2]. This tool has been widely used for Twitter analysis [2, 33-37]. The tool analyzes text and emoji independently. Firstly, the text contained in the tweet is analyzed using the lexicon sentiment and the sentiment values (positive and negative) of the sentence are calculated. The emoji sentiment is then added to the value derived from the sentiment lexicon, not exceeding -5 points as maximum (+5 in case of positive sentiment). Therefore, including a “negative” emoji in a tweet will impact on its overall negative sentiment value.

Statistics

Descriptive statistics were used to summarize absolute numbers, frequencies, means, and standard deviations (SD). Independent t-tests were used to compare the average positive and negative sentiment in tweets specifically referring to T1D; to T2D; and including or not including an emoji. Correlation analysis was used to examine the relationship between the sentiment and the number of followers. The data were analyzed with SPSS 25 for Mac (IBM Corp.).

RESULTS

A total of 67421 tweets were automatically extracted. Among these, a total of 2512 (3.7%) tweets specifically referred to T1D, and 4585 (6.8%) specifically mentioned T2D. At least one emoji was included in 4720 (7.0%) of the tweets.

The analysis of the negative sentiment value showed that the negative strength of tweets specifically mentioning T2D was significantly more negative than the tweets not mentioning T2D (-2.20 vs. -1.65), $p<0.001$. Tweets referring to T2D without any emoji were significantly more negative than the ones including an emoji (-2.22 vs. -1.48), $p<0.001$. On the other side, analysis of the positive sentiment value showed that the positive strength of tweets on

T2D was significantly lower when the tweet included any emoji, than when no emoji was included (1.39 vs. 1.57), $p<0.001$. Tweets on T1D including emojis had the highest positive strength (1.71) and also the lowest negative strength (-1.31), as compared to tweets on T1D without emojis. The sentiment analysis of the whole sample and by categories is reported in Table 1.

Regarding the users' number of Twitter followers according to the sentiment, the analysis of the whole sample showed a significant negative association between a positive sentiment strength and the number of followers, $r = -0.023$, $p<0.001$ (See Figure 1); and a significant positive association between a negative sentiment and the number of followers, $r = 0.016$, $p<0.001$ (See Figure 2). Tweets explicitly referring to T2D had a significant negative correlation with the number of followers, $r = -0.042$, $p<0.005$; but there was no significant correlation with negative sentiment.

No significant correlation was found between positive or negative sentiment tweets and the number of followers in the subsample of messages including emojis; nor in the subsample of tweets specifically mentioning T1D.

DISCUSSION

Main findings

This could be one of the first studies analyzing the sentiment expressed and emojis on social media posts focusing specifically on diabetes. Tweets specifically mentioning T2D were the more negative, especially the ones that did not include any emoji. Tweets on T1D including emojis had a higher positive strength, and also a lower negative strength (when compared to tweets without emojis). Users posting tweets with a more neutral sentiment (less positive and less negative) had the highest number of followers.

Type 2 Diabetes on Twitter: the most negative sentiment

Tweets on T2D were significantly more negative than those on T1D. We do not have a good explanation of why so many of the tweets on the topic of diabetes were identified as communicating negative sentiment (i.e. emotion) or why this was more pronounced for T2D than for T1D. However, one explanation may be that the Twitter community could be posting more negative sentiment messages on T2D as the disease often is perceived as lifestyle-related. This means that some users affected with T2D could have negative feelings such as shame or guilt and therefore express themselves with negative sentiment. Other Twitter users could be blaming or shaming those affected by T2D by posting negative messages.

Prior studies have also shown that many health-related tweets contain a message that may be perceived of as negative [22, 23, 38]. In some instances, the negative content or sentiment may be explained by a misuse or misappropriation of the disease-term, for instance by its inclusion in a joke or some ironic expression (‘I will get diabetes if I eat this’). Many tweets on health-related issues are jokes [38]. Tweets that could be classified as jokes based on their tone and nature have been included in the analyzed sample. Most of them contained negative words and sometimes they also included negative emojis. Therefore, many of those tweets were associated with a negative sentiment. Although those tweets typically were inappropriate, many of them used sarcasm, which is not detected by SentiStrength. Additionally, the use of emojis complicated the sentiment analysis because emojis may express a different sentiment or emotion that the text itself. This may introduce emotional ambiguity, which is probably a desired effect in many tweets - for instance in cases of double entendre, irony or sarcasm.

Tweets including emojis are more positive and also less negative

While emojis can express both positive and negative moods or ideas, our results show that tweets including emojis are linked to both a more positive sentiment strength and to a less negative sentiment strength. In that sense, the use of emojis seems to increase the sentiment strength of the tweets. Our results are in accordance with the findings from a previous study [39] that analyzed the sentiment of tweets on two events, one positive and one negative.

Positive emotions are linked to a more cooperative behavior and better decision-making [40]. In our case, we found that tweets on T1D including emojis were especially positive. This could mean that users posting these kind of tweets could be individuals affected by T1D themselves, or love someone who has T1D, and their positivity could benefit the Twitter community that discusses diabetes. Emotions are contagious and likely to be modified, and this also applies to emotions expressed in social media [41-43]. Our study found that T2D in Twitter was mostly linked to a negative sentiment, but we also found that tweets containing emojis were more positive and less negative. Therefore, the Twitter content with a negative sentiment related to T2D could potentially be counteracted by increasing the number of posts with a positive tone, and by including emojis.

The closer to neutral sentiment, the more followers

In a previous publication, Berguerisse-Díaz et al [44] found that 9 of the top 10 authorities ranked were institutional accounts (3 stockmarket-listed commercial ventures; 3 national or international diabetes associations; 3 non-profit organizations funded by people who have experienced T1D) and the remaining one belonged to an individual blogger with T1D involved with a number of diabetes advocacy organizations. All of the top 10 authorities posted messages frequently related to health information (public health messages; links to articles, blogs and studies about risks, treatments and cure; population health fears; publicity about outreach and awareness events and activities; advice about diabetes management and diagnosis; lifestyle, diet and cookery tips, news and links; life stories and experiences; dangers of sugar, sugar replacements and/or soda). Messages related to those themes were largely irregular and variable over time [44]. In a previous study that analyzed tweets on T1D it was found that non-governmental organizations, communication media, and people with T1D were the users with the highest number of followers and that the tweets posted by the patients were the most retweeted [45]. The analysis of tweet contents posted by a diabetes patients' organization found that almost half of the messages were on diabetes awareness (recipes, celebrations, celebrities, jokes, etc.) [46]. Despite the variability, many of these themes may be treated using neutral tweets, contrasting with our findings.

Our findings have implications for public health interventions as we showed that neutral tweets on diabetes were associated with a higher number of followers. Public health promoters and other stakeholders on Twitter that aim to increase their number of followers and the impact of their promotion should consider posting messages that have a neutral sentiment, i.e. that do not express strong emotions.

Limitations and future directions

Our study has several limitations. Although we extracted a large amount of tweets focusing on diabetes, this is a random one-week sample and might therefore not be representative. We might have missed relevant messages by limiting our search strategy to the words “diabetes”, “t1d”, and “t2d”. Further, not all tweets were indexed or made available via the standard API search interface. Tweets were collected in a small time window since they were published, due to the standard API limitations, which prevented us from performing analysis of likes or retweets. Additional research could consider including other words related to this health condition, expand the search by including other languages, include further social media channels, and analyze the users’ profiles. We used SentiStrength to analyze the sentiment of tweets on diabetes. This tool has not yet been validated for use in the health domain. However, the tool was created using posts on "MySpace", and therefore it has a good performance with informal texts, such the ones that can be found on Twitter.

CONCLUSIONS

The use of sentiment analysis techniques helped to identify that tweets on diabetes type 2 more often had a negative sentiment; while posts on type 1 diabetes more often were associated with a positive mood, especially if they included emojis. Tweets on diabetes closer to neutral sentiment were associated with a higher number of followers. These observations might be relevant for developing better public health strategies and for promoting a positive and constructive attitude among people that read and discuss about the illness on social media.

FUNDING SOURCES

This work is supported by the Northern Norway Regional Health Authority (Helse Nord RHF), grant number HNF1370-17. EDZ is supported by the V Plan Propio de Investigación of the Universidad de Sevilla.

DISCLOSURES

The authors declare that they have no conflict of interest.

REFERENCES

1. Beaunoyer, E., M. Arsenault, A.M. Lomanowska, et al., *Understanding online health information: Evaluation, tools, and strategies.* Patient Educ Couns, 2017. **100**(2): p. 183-189. doi:10.1016/j.pec.2016.08.028
2. Gohil, S., S. Vuik, and A. Darzi, *Sentiment Analysis of Health Care Tweets: Review of the Methods Used.* JMIR Public Health Surveill, 2018. **4**(2): p. e43. PMID: PMC5938573. doi:10.2196/publichealth.5789
3. Zorman, M. and M. Verlic, *Explanatory approach for evaluation of machine learning-induced knowledge.* J Int Med Res, 2009. **37**(5): p. 1543-51. doi:10.1177/147323000903700532
4. Borras-Morell, J.E., *Data mining for pulsing the emotion on the web.* Methods Mol Biol, 2015. **1246**: p. 123-30. PMID: doi:10.1007/978-1-4939-1985-7_8

5. Hirschberg, J. and C.D. Manning, *Advances in natural language processing.* Science, 2015. **349**(6245): p. 261-6. doi:10.1126/science.aaa8685
6. D'Alfonso, S., O. Santesteban-Echarri, S. Rice, et al., *Artificial Intelligence-Assisted Online Social Therapy for Youth Mental Health.* Front Psychol, 2017. **8**: p. 796. PMID: PMC5454064. doi:10.3389/fpsyg.2017.00796
7. Greaves, F., D. Ramirez-Cano, C. Millett, et al., *Use of sentiment analysis for capturing patient experience from free-text comments posted online.* J Med Internet Res, 2013. **15**(11): p. e239. PMID: PMC3841376. doi:10.2196/jmir.2721
8. Mazzocut, M., I. Truccolo, M. Antonini, et al., *Web Conversations About Complementary and Alternative Medicines and Cancer: Content and Sentiment Analysis.* J Med Internet Res, 2016. **18**(6): p. e120. PMID: PMC4929351. doi:10.2196/jmir.5521
9. Larsen, J.T. and A.P. McGraw, *Further evidence for mixed emotions.* J Pers Social Psych, 2011. **100**: p. 1095-1110. doi:10.1037/a0021846
10. Kralj Novak, P., J. Smailovic, B. Sluban, et al., *Sentiment of Emojis.* PLoS One, 2015. **10**(12): p. e0144296. PMID: PMC4671607. doi:10.1371/journal.pone.0144296
11. Rodrigues, D., M. Prada, R. Gaspar, et al., *Lisbon Emoji and Emoticon Database (LEED): Norms for emoji and emoticons in seven evaluative dimensions.* Behav Res Methods, 2018. **50**(1): p. 392-405. doi:10.3758/s13428-017-0878-6
12. Sheth, A., S. Perera, S. Wijeratne, et al., *Knowledge will Propel Machine Understanding of Content: Extrapolating from Current Examples.* Proc IEEE WIC ACM Int Conf Web Intell Intell Agent Technol, 2017. **2017**: p. 1-9. PMID: PMC6021355. doi:10.1145/3106426.3109448
13. Skiba, D.J., *Face with Tears of Joy Is Word of the Year: Are Emoji a Sign of Things to Come in Health Care?* Nurs Educ Perspect, 2016. **37**(1): p. 56-7.
14. Troiano, G. and N. Nante, *Emoji: What does the scientific literature say about them?-A new way to communicate in the 21th century.* Journal of Human Behaviour in the Social Environment, 2018. **28**(4): p. 528-33. PMID: doi:10.1080/10911359.2018.1437103
15. Wijeratne, S., L. Balasuriya, A. Sheth, et al., *EmojiNet: Building a Machine Readable Sense Inventory for Emoji.* Proc Int Workshop Soc Inform, 2016. **10046**: p. 527-541. PMID: PMC5519135. doi:10.1007/978-3-319-47880-7_33
16. Oxford Living Dictionaries. *Emoji*. 2018; Available from: https://en.oxforddictionaries.com/definition/emoji.
17. Gabarron, E., A.Y. Lau, and R. Wynn, *Is There a Weekly Pattern for Health Searches on Wikipedia and Is the Pattern Unique to Health Topics?* J Med Internet Res, 2015. **17**(12): p. e286. PMID: PMC4704932. doi:10.2196/jmir.5038
18. Li, Y., X. Wang, X. Lin, et al., *Seeking and sharing health information on social media: A net valence model and cross-cultural comparison.* Technological Forecasting and Social Change, 2018. **126**: p. 28-40. PMID: doi:10.1016/j.techfore.2016.07.021

19. Nath, C., J. Huh, A.K. Adupa, et al., *Website Sharing in Online Health Communities: A Descriptive Analysis.* J Med Internet Res, 2016. **18**(1): p. e11. PMID: PMC4730108. doi:10.2196/jmir.5237
20. Chmiel, A., J. Sienkiewicz, M. Thelwall, et al., *Collective emotions online and their influence on community life.* PLoS One, 2011. **6**(7): p. e22207. PMID: PMC3144870. doi:10.1371/journal.pone.0022207
21. Laranjo, L., A. Arguel, A.L. Neves, et al., *The influence of social networking sites on health behavior change: a systematic review and meta-analysis.* J Am Med Inform Assoc, 2015. **22**(1): p. 243-56. PMID: PMC4433372. doi:10.1136/amiajnl-2014-002841
22. Oyeyemi, S.O., E. Gabarron, and R. Wynn, *Ebola, Twitter, and misinformation: a dangerous combination?* Bmj, 2014. **349**: p. g6178. doi:10.1136/bmj.g6178
23. Wynn, R., S.O. Oyeyemi, J.A. Johnsen, et al., *Tweets are not always supportive of patients with mental disorders.* Int J Integrated Care, 2017. **17**(3): p. A149.
24. Lau, A.Y., E. Gabarron, L. Fernandez-Luque, et al., *Social media in health--what are the safety concerns for health consumers?* Health Inf Manag, 2012. **41**(2): p. 30-5.
25. Gabarron, E., M. Bradway, L. Fernandez-Luque, et al., *Social media for health promotion in diabetes: study protocol for a participatory public health intervention design.* BMC Health Serv Res, 2018. **18**(1): p. 414. PMID: PMC5989446. doi:10.1186/s12913-018-3178-7
26. Huh, J., M. Yetisgen-Yildiz, and W. Pratt, *Text classification for assisting moderators in online health communities.* J Biomed Inform, 2013. **46**(6): p. 998-1005. PMID: PMC3874858. doi:10.1016/j.jbi.2013.08.011
27. Liu, X., M. Sun, and J. Li, *Research on gender differences in online health communities.* Int J Med Inform, 2018. **111**: p. 172-181. doi:10.1016/j.ijmedinf.2017.12.019
28. Lu, Y., Y. Wu, J. Liu, et al., *Understanding Health Care Social Media Use From Different Stakeholder Perspectives: A Content Analysis of an Online Health Community.* J Med Internet Res, 2017. **19**(4): p. e109. PMID: PMC5400888. doi:10.2196/jmir.7087
29. Pai, R.R. and S. Alathur, *Assessing mobile health applications with twitter analytics.* Int J Med Inform, 2018. **113**: p. 72-84. doi:10.1016/j.ijmedinf.2018.02.016
30. Salas-Zarate, M.D., J. Medina-Moreira, K. Lagos-Ortiz, et al., *Sentiment Analysis on Tweets about Diabetes: An Aspect-Level Approach.* Comput Math Methods Med, 2017. **2017**: p. 5140631. PMID: PMC5337803 doi:10.1155/2017/5140631
31. Thelwall, M., K. Buckley, G. Paltoglou, et al., *Sentiment Strength Detection in Short Informal Text.* J Am Soc Inf Sci Tech, 2010. **61**(12): p. 2544-2558. doi:10.1002/asi.21416
32. Thelwall, M., K. Buckley, and G. Paltoglou, *Sentiment in twitter events.* J Am Soc Inf Sci Tech, 2010. **62**(2): p. 406-418. doi:10.1002/asi.21462
33. Bhattacharya, S., P. Srinivasan, and P. Polgreen, *Engagement with health agencies on twitter.* PLoS One, 2014. **9**(11): p. e112235. PMID: PMC4224440. doi:10.1371/journal.pone.0112235

34. Charlton, N., C. Singleton, and D.V. Greetham, *In the mood: the dynamics of collective sentiments on Twitter.* R Soc Open Sci, 2016. **3**(6): p. 160162. PMID: PMC4929909. doi:10.1098/rsos.160162
35. Iosub, D., D. Laniado, C. Castillo, et al., *Emotions under discussion: gender, status and communication in online collaboration.* PLoS One, 2014. **9**(8): p. e104880. PMID: PMC4139304. doi:10.1371/journal.pone.0104880
36. Rodrigues, R.G., R.M. das Dores, C.G. Camilo-Junior, et al., *SentiHealth-Cancer: A sentiment analysis tool to help detecting mood of patients in online social networks.* Int J Med Inform, 2016. **85**(1): p. 80-95. doi:10.1016/j.ijmedinf.2015.09.007
37. Suvakov, M., M. Mitrovic, V. Gligorijevic, et al., *How the online social networks are used: dialogues-based structure of MySpace.* J R Soc Interface, 2013. **10**(79): p. 20120819. PMID: PMC3565699. doi:10.1098/rsif.2012.0819
38. Gabarron, E., J.A. Serrano, R. Wynn, et al., *Tweet content related to sexually transmitted diseases: no joking matter.* J Med Internet Res, 2014. **16**(10): p. e228. PMID: PMC4210955. doi:10.2196/jmir.3259
39. Shiha, M.O. and S. Ayvaz, *The effects of emoji in sentiment analysis.* Int J Comp Elect Eng, 2017. **9**(1): p. 360-369. doi:10.17706/ijcee.2017.9.1.360-369
40. Rand, D.G., G. Kraft-Todd, and J. Gruber, *The collective benefits of feeling good and letting go: positive emotion and (dis)inhibition interact to predict cooperative behavior.* PLoS One, 2015. **10**(1): p. e0117426. PMID: PMC4308081. doi:10.1371/journal.pone.0117426
41. Ferrara, E. and Z. Yang, *Measuring Emotional Contagion in Social Media.* PLoS One, 2015. **10**(11): p. e0142390. PMID: PMC4636231. doi:10.1371/journal.pone.0142390
42. Kramer, A.D., J.E. Guillory, and J.T. Hancock, *Experimental evidence of massive-scale emotional contagion through social networks.* Proc Natl Acad Sci U S A, 2014. **111**(24): p. 8788-90. PMID: PMC4066473. doi:10.1073/pnas.1320040111
43. van Kleef, G.A., M.W. Heerdink, and A.C. Homan, *Emotional influence in groups: the dynamic nexus of affect, cognition, and behavior.* Curr Opin Psychol, 2017. **17**: p. 156-161. doi:10.1016/j.copsyc.2017.07.017
44. Beguerisse-Diaz, M., A.K. McLennan, G. Garduno-Hernandez, et al., *The 'who' and 'what' of #diabetes on Twitter.* Digit Health, 2017. **3**: p. 2055207616688841. PMID: PMC6001201. doi:10.1177/2055207616688841
45. Gabarron, E., A. Makhlysheva, and L. Marco, *Type 1 Diabetes in Twitter: Who All Listen To?* Stud Health Technol Inform, 2015. **216**: p. 972.
46. Gabarron, E., M. Bradway, and E. Arsand, *What are diabetes patients discussing on social media?* Int J Integrated Care, 2016. **5**: p. S14.

TABLES AND TABLE LEGENDS

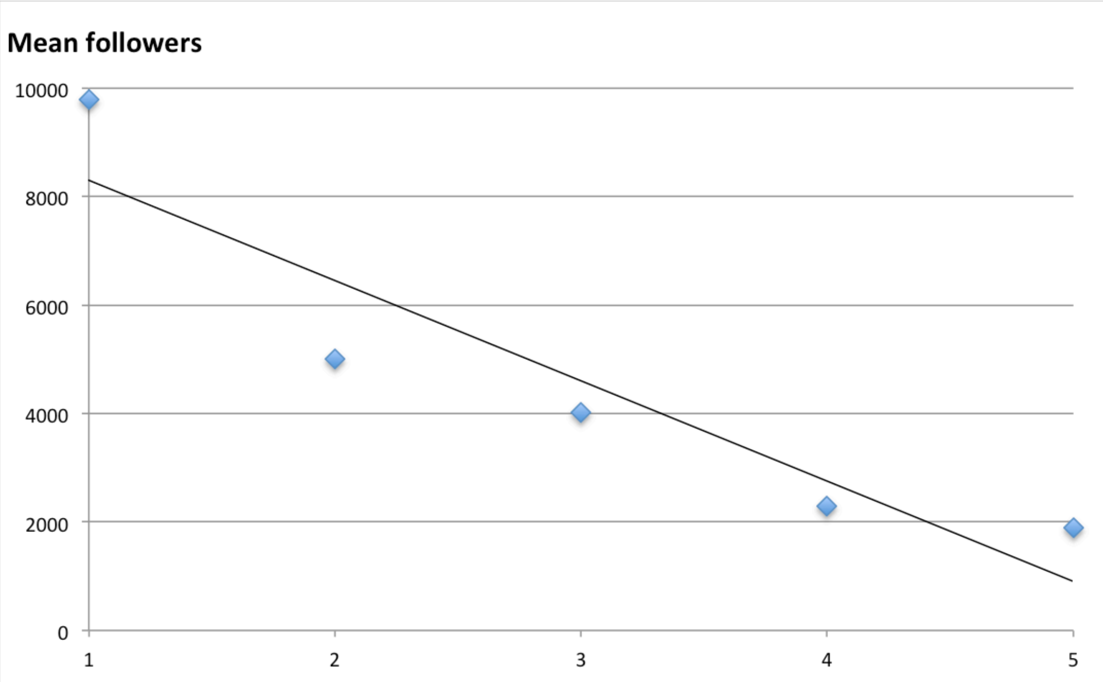
Table 1. Average sentiment of the tweets

Sentiment analysis	Positive Mean (SD)	Negative Mean (SD)
Tweets specifically mentioning T1D*		
Yes (n=2512)	1.51 (0.7)	-1.48 (0.8)
No (n=64909)	1.47 (0.7)	-1.70 (0.9)
Tweets specifically mentioning T1D**		
Including emoji (n=219)	1.71 (0.8)	-1.31 (0.7)
Without emoji (n=2293)	1.49 (0.7)	-1.50 (0.8)
Tweets specifically mentioning T2D*		
Yes (n=4585)	1.56 (0.8)	-2.20 (1.2)
No (n=62836)	1.46 (0.7)	-1.70 (0.9)
Tweets specifically mentioning T2D*		
Including emoji (n=126)	1.39 (0.7)	-1.48 (0.7)
Without emoji (n=4459)	1.57 (0.8)	-2.22 (1.2)
Tweets including emoji*		
Yes (n=4720)	1.62 (0.8)	-1.36 (0.8)
No (n=62701)	1.46 (0.7)	-1.72 (0.9)
Whole sample (n=67421)	1.47 (0.7)	-1.69 (0.9)

** t-test, p<0.001; **p<0.005*

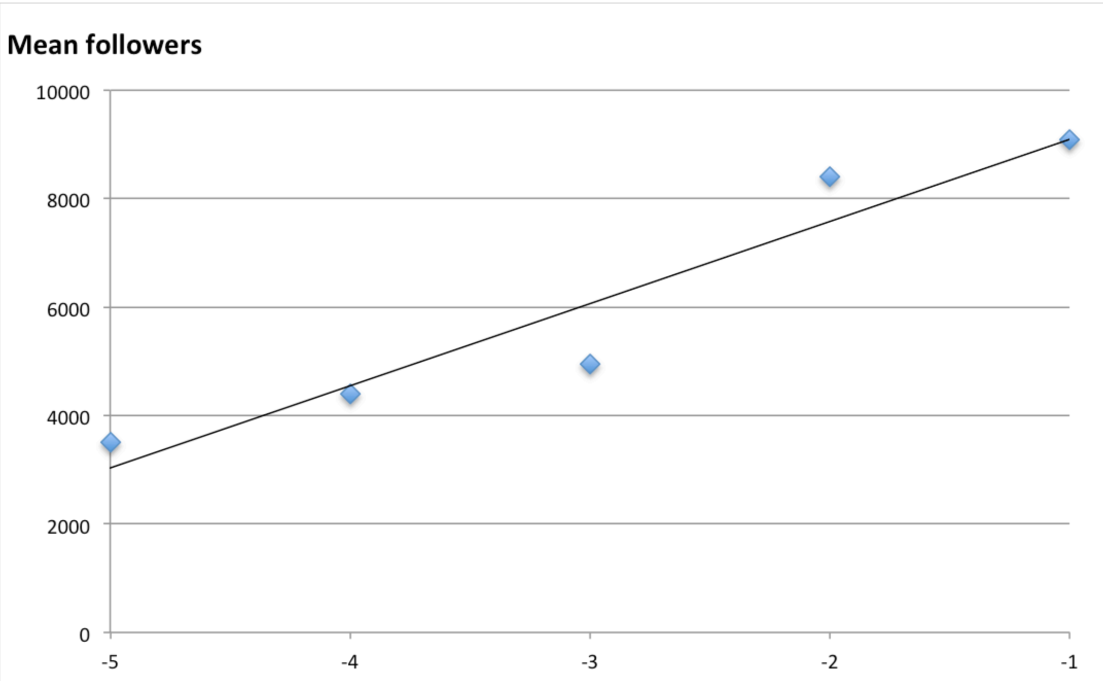
FIGURES AND FIGURES LEGENDS

Figure 1. Positive sentiment and number of followers



Pearson correlation r = -0.023, *p<0.001*

Figure 2. Negative sentiment and number of followers



Pearson correlation r = -0.042, p<0.005