The rate of reply and nature of responses to suicide-related posts on Twitter

Bridianne O'Dea[a,b,*], Melinda R. Achilles[a], Mark E. Larsen[a,b], Philip J. Batterham[c], Alison L. Calear[c], Helen Christensen[a,b]

[a] Black Dog Institute, Prince of Wales Hospital, Hospital Road, Randwick, NSW 2031, Australia
[b] Faculty of Medicine, University of New South Wales, Randwick, NSW 2031, Australia
[c] Centre for Mental Health Research, The Australian National University, Acton, ACT 2601, Australia

ABSTRACT

The social media platform Twitter has been used by individuals to communicate suicidal thoughts and intentions. Currently, the nature and rate of reply to this type of Twitter content is unknown. This brief report aimed to understand how Twitter users respond to suicide-related content as compared to non-suicide related content. Using a dataset of suicide and non-suicide related posts, replies, retweets and likes were analysed and compared. The content of the first replies to suicide-related posts were also reviewed. When compared to non-suicide related posts, those that were suicide-related received a significantly greater number of replies, with fewer retweets and likes. The rate of reply to the suicide-related posts was also significantly faster than that of the non-suicide related posts, with the average reply occurring within 1 h. Thematic analysis revealed that 62% of the first replies to suicidal posts were of a potentially helpful nature (e.g. discouraging suicide, caring, or clarifying), while 23% were dismissive or encouraging of the suicide. These findings indicate that Twitter users respond differently to suicidal content. Further research is needed to determine the effects of the replies on suicidal intentions or ideations, and whether this platform can be used to intervene, increase help-seeking, or provide anti-stigma campaigns.

1. Introduction

Twitter is a social media platform that allows users to broadcast short public messages known as a post or tweet, to a network of other users known as followers. These followers are then able to reply, retweet, and like each other's posts. A reply is a direct response to a post whereas a retweet involves another user re-posting a message within their network. A like is represented by a small heart icon and is used to show appreciation for a post. These varied sharing functions of Twitter are designed to allow information to spread quickly, be far-reaching, and to encourage interaction between users. Individuals have used Twitter to share their suicidal thoughts and intentions (O'Dea et al., 2017; O'Dea et al., 2015). Individuals experiencing suicidal ideation have an increased sense of hopelessness, a lack of belonging to others, and a perceived sense of burdensomeness (Joiner Jr et al., 2005). By conveying a sense of social connectedness and support, responses to suicidal posts on Twitter may be crucial for preventing deaths, while also reducing stigma and distress. On Facebook, users who received greater likes and comments to their content perceived their social

network to be stronger, with quicker reactions resulting in greater levels of support (Seo et al., 2016). In a study of suicide message boards, individuals experienced greater psychological improvement when responders gave constructive advice, active listening, sympathy, and presented alternatives to suicide (Niederkrotenthaler et al., 2016). The diffusion of suicidal posts within the Twitter network may also serve as an early identification or monitoring tool in which followers can be activated quickly in an emergency (Fu et al., 2013). As such, Twitter may offer an opportunity to reduce suicidal crises.

Currently, the nature and rate of reply to suicidal content on Twitter is unknown. It is likely that many suicide-related posts are not responded to as up to 70% of Twitter posts do not get any response and among those that do, replies take up to 1 h (Geere, 2010). Further, Naslund et al. (2016) outlined that facing hostile or derogatory comments is a possible risk of sharing mental health content on social media. A case study of an online suicide found that 20% of responses were cynical or indifferent (Fu et al., 2013). This is consistent with Li et al. (2015) who found that one third of responses to online suicides were stigmatising. Suicidal content has also been found to have a

greater spread than general Twitter content, potentially leading to contagion effects (Colombo et al., 2016). To better understand how users of Twitter respond to suicide-related posts, this brief report examines the time taken to reply to such posts in comparison to general Twitter posts, alongside differences in the number of replies, likes, and retweets. This paper also examines the nature of the first replies to suicide-related posts. It is anticipated that this research may inform best practice in future suicide prevention activities delivered within social media settings.

2. Method

2.1. Data collection

This study received ethics approval from University of New South Wales, Sydney, Australia. It is a secondary analysis utilising a dataset of suicide-related ($n = 127$) and non-suicide related Twitter posts ($n = 127$). Details of the collection and case/control matching of the posts is previously described elsewhere (O'Dea et al., 2017; O'Dea et al., 2015). In brief, the non-suicide related control dataset was matched, where possible, on time zone, day of the week, and time of the day of the post, to the nearest hour. Data was also matched on characteristics of the Twitter user including gender, number of followers, number of other users being followed (i.e. friends), and total number of previous posts, adjusted into quartiles. Checks for duplicate users were conducted, although none were found. The final dataset consisted of the posts, alongside the number of replies, retweets, and likes as well as time of first reply. Where possible, the first reply content was also collected. Four non-suicide related posts were removed due to outlying values of replies, retweets, and likes, as determined by standardised residual values greater than 3.30 (Tabachnick and Fidell, 2007). A total of 23 first replies were unable to be analysed due to the owners removing the content before data was extracted. The final dataset consisted of 127 suicide-related posts and 123 general posts, with content extracted for 87 and 31 first replies, respectively.

Given that data was not normally distributed, medians were reported. Non-parametric Mann Whitney U tests were used to explore the differences in replies, retweets, and likes, and time taken for first reply, between the suicide-related and non-suicide related posts. To explore how many posts received at least one reply, retweet, or like, these three variables were dichotomised (i.e. one or more vs. none). Although the sample was matched, additional testing was conducted to assess whether the size of the poster's network (e.g. number of friends and followers) was associated with receiving a reply, retweet, or like. Thematic analysis was used to determine the nature of the first replies. Using the guideline outlined by Braun and Clarke (2006), the replies were manually reviewed. Using an inductive approach, patterns and themes in the replies were identified and given codes. Two researchers (BOD and MA) refined the initial codes for cohesiveness, sorted to combine related concepts into encompassing main themes, and reached an agreement on the final themes. Using this framework, the replies were then reviewed, and final categorisations were made. The agreement rate between the two coders was 77%. Inconsistencies were identified, discussed, and resolved using consensus. Frequency counts and percentages were reported to highlight the representativeness of the themes. To protect Twitter users' privacy, specific examples of replies are not able to be given (Denecke et al., 2015).

3. Results

As outlined in Table 1, suicide-related posts received a significantly higher number of replies than non-suicide related posts, although significantly fewer retweets and likes. Suicide-related posts were also found to have a significantly quicker reply time when compared to non-suicide related posts. There were no significant differences in the number of friends or followers between the suicide and non-suicide

related posts ($P = .74$–.97).

Suicide-related posts were nearly 10 times more likely to receive at least one reply when compared to non-suicide related posts (81.1%, $n = 103/127$ vs. 30.1%, $n = 37/123$, $x^2 = 66.01$, OR: 9.97, 95% CI: 5.54–17.96, $P < .05$). No significant differences were found in the size of the network of those who received a reply (Mdn friends = 106.00; Mdn followers = 108.44) and those who did not (Mdn friends = 98.64; Mdn followers = 95.95; $P = .13$–.37). Suicide-related posts were significantly less likely to be retweeted at least once (20.5%, $n = 26/127$ vs. 44.7%, $n = 55/123$, $x^2 = 16.77$, OR: 0.32, 95% CI: 0.18–0.56, $P < .05$). The posts that were retweeted had a significantly higher number of followers (Mdn: 135.71 vs. 93.89, U = 2007.00, $P < .001$) and friends (Mdn: 119.68 vs. 98.05; U = 2680.50, $P < .05$) in their network when compared to those that were not. Suicide-related posts were also less likely to receive at least one like when compared to non-suicide related posts (32.3%, $n = 41/127$ vs. 56.9%, $n = 70/123$, $x^2 = 15.35$, OR: 0.36, 95% CI: 0.22–0.60, $P < .05$). The posts that were liked had a significantly higher number of followers (Mdn: 115.57 vs. 95.37; U = 3811.00 $P < .05$), but no significant difference in number of friends (Mdn: 103.90 vs. 101.74; U = 4651.50, $P = .80$).

Thematic analysis revealed that the nature of the first replies ($n = 87$) to the suicide-related posts could be classified as: "discouraging" (i.e. direct requests for the person to stop their suicide, $n = 31$, 35.6%); "caring" (i.e. emotionally supportive words with a reassuring tone and reasons for living may have been provided, $n = 18$, 20.7%); "dismissive" (i.e. dismissive words suggesting the post was as a joke, or non-threatening, $n = 14$, 16.1%); "encouraging" (i.e. direct requests for the person to suicide or approval of suicide, $n = 6$, 6.9%); "clarifying" (i.e. requests for more information, asking if the person is okay, or what had happened, $n = 5$, 5.7%); or "unclear" (i.e. replies with unclear intent or meaning, $n = 13$, 14.9%). Overall, two thirds of the replies to the suicide-related posts (62%, $n = 54/87$) were of a potentially helpful nature, with nearly one in four (23%, $n = 20/87$) dismissive or encouraging of the suicide.

4. Discussion

This study compares the rates of reply to suicide-related Twitter posts to non-suicide related posts. We found that suicide-related posts were significantly more likely to receive a reply. The average time taken to reply to a suicide-related post was just over 1 h, although the median reply time was less than 5 min. This was significantly faster than the reply to time to non-suicide related posts. These differences may be due to the fatal nature of suicide, such that the network is motivated to respond more frequently and with haste to posts of this kind. Importantly however, the results indicate that there is significant variance in reply time to suicide-related content. We do not know if the reply times found in this study are adequate for preventing self-harm or death. Further, one in five suicide-related posts did not receive any reply. This could be due to a myriad of factors, although the size of one's Twitter network was not found to be significant. This suggests that increasing the number of followers or friends may not lead to a greater likelihood of reply, although this finding needs to be replicated in a larger study. The effects of nonresponse are also not yet known but may inflate feelings of hopelessness and isolation. Alternatively, it may have a null-effect. Given that past research has found the reply time to impact Facebook users' perceived social support and loneliness (Seo et al., 2016), further research is needed to determine the nature of this effect across platforms and within the context of suicide.

Most notably, this study confirms that many of the individuals who posted suicide-related content received a first reply that was potentially helpful by discouraging their suicide attempt, providing support, or attempting to understand what was happening for them. Based on Niederkrotenthaler et al.'s (2016) prior observations, these types of responses may reduce distress. However, nearly one in four replies were dismissive and pro-suicide, and none of the replies were found to offer

Table 1
Differences in the number of replies, retweets, and likes.

Response type	Range	Suicide-related posts (n = 127)			Range	Non-suicide related posts (n = 123)			U (z)
		Mdn	Mean rank	Mean (SD)		Mdn	Mean rank	Mean (SD)	
Replies	0–14	1.00	147.66	1.39 (1.79)	0–15,704	0	102.62	9.49 (37.66)	4996 (−5.27)*
Retweets	0-69	0	107.95	1.44 (7.34)	0–138,233	0	143.62	535.20 (1938.34)	10,139 (4.69)*
Likes	0–100	0	107.17	2.13 (9.70)	0–166,763	1.00	144.43	500.08 (1542.72)	10,139 (4.48)*
First reply	1.00–1473.78[a]	3.80[a]	51.97[b]	1.09[b] (3.73)	1.12–470,683.07[a]	59.01[a]	78.52[b] hr	35.26[b] (96.90)	1938 (3.74)*
Friends	37–9627	391.00	102.62	838.82 (1385.16)	5–25,356	395.00	102.32	962.54 (2849.06)	4987 (−0.36)
Followers	39–45,554	469.50	103.61	2600.78 (7149.90)	46–37,733	503.50	100.85	1691.90 (4489.23)	4866 (−0.328)

[*] $p < .001$.
[a] Minutes.
[b] Hours.

informational support. This is consistent with past research, demonstrating that the sharing of suicidal content on social media may carry a negative risk for the poster (Li et al., 2015; Naslund et al., 2016). Although retweets were higher for non-suicide related Twitter posts, one in five suicide-related posts were re-shared. Our results suggest that the size of one's Twitter network may influence the retweeting rate and number of likes received. The meaning of these actions in relation to suicide-related content remains unclear. The current study could not determine the effects of retweeting and liking. It may be speculated that individuals who have their posts retweeted may be more likely to receive intervention or support (Fu et al., 2013). Conversely, retweeting may increase the risk of negative responses and may contribute to suicide contagion (Colombo et al., 2016). Future research is needed to measure the degree to which types of responses, or lack thereof, influences individuals' mood and help-seeking for suicidal ideation. Other factors that are important to understanding the nature and impact of responses but were not assessed in this study include the context of the original suicide-related post, the genuine intent of the post, and the posting history of the Twitter user. Further studies are needed to evaluate the effects of sharing suicide related content on social media on both the individual and their network.

Overall, these findings suggest that many Twitter users can be relied upon to provide helpful responses when replying to a suicide-related post. However, in the instances where suicide-related posters do not receive support from their network, Twitter may be a setting for the delivery of automated help-seeking and anti-stigma campaigns. Developments in machine learning and automation may assist in administering such campaigns in real-time and at scale. Health organisations are currently using Twitter to deliver both information and supportive messages, although effectiveness remains unknown (Rui et al., 2013). Further, the ethical implications of this type of monitoring and content targeting are yet to be fully outlined and may include breaches of privacy. Future research is needed before direct intervention can take place.

Conflict of interest

The authors declare that there is no conflict of interest.

Funding

Acknowledgments

References

Braun, V., Clarke, V., 2006. Using thematic analysis in psychology. Qual. Res. Psychol. 3, 77–101. https://doi.org/10.1191/1478088706qp063oa.

Colombo, G.B., Burnap, P., Hodorog, A., Scourfield, J., 2016. Analysing the connectivity and communication of suicidal users on twitter. Comput. Commun. 73, 291–300. https://doi.org/10.1016/j.comcom.2015.07.018.

Denecke, K., Bamidis, P., Bond, C., Gabarron, E., Househ, M., Lau, A., Mayer, M.A., Merolli, M., Hansen, M., 2015. Ethical issues of social media usage in healthcare. Yearb. Med. Inform. 10, 137. https://doi.org/10.15265/IY-2015-001.

Fu, K.W., Cheng, Q., Wong, P.W., Yip, P.S., 2013. Responses to a self-presented suicide attempt in social media. Crisis. https://doi.org/10.1027/0227-5910/a000221.

Geere, D., 2010. It's not just you: 71 percent of tweets are ignored [online]. Available: www.wired.com/2010/10/its-not-just-you-71-percent-of-tweets-are-ignored, Accessed date: 21 August 2017.

Joiner Jr., T.E., Brown, J.S., Wingate, L.R., 2005. The psychology and neurobiology of suicidal behavior. Annu. Rev. Psychol. 56, 287–314. https://doi.org/10.1146/annurev.psych.56.091103.070320.

Li, A., Huang, X., Hao, B., O'dea, B., Christensen, H., Zhu, T., 2015. Attitudes towards suicide attempts broadcast on social media: an exploratory study of Chinese microblogs. PeerJ 3, e1209. https://doi.org/10.7717/peerj.1209.

Naslund, J., Aschbrenner, K., Marsch, L., Bartels, S., 2016. The future of mental health care: peer-to-peer support and social media. Epidemiol. Psychiatr. Sci. 25, 113–122. https://doi.org/10.1017/S2045796015001067.

Niederkrotenthaler, T., Gould, M., Sonneck, G., Stack, S., Till, B., 2016. Predictors of psychological improvement on non-professional suicide message boards: content analysis. Psychol. Med. 46, 3429–3442. https://doi.org/10.1017/S003329171600221X.

O'Dea, B., Wan, S., Batterham, P.J., Calear, A.L., Paris, C., Christensen, H., 2015. Detecting suicidality on Twitter. Internet Interv. 2, 183–188. https://doi.org/10.1016/j.invent.2015.03.005.

O'Dea, B., Larsen, M.E., Batterham, P.J., Calear, A.L., Christensen, H., 2017. A linguistic analysis of suicide-related Twitter posts. Crisis. https://doi.org/10.1027/0227-5910/a000443.

Rui, J.R., Chen, Y., Damiano, A., 2013. Health organizations providing and seeking social support: a Twitter-based content analysis. Cyberpsychol. Behav. Soc. Netw. 16, 669–673. https://doi.org/10.1089/cyber.2012.0350.

Seo, M., Kim, J., Yang, H., 2016. Frequent interaction and fast feedback predict perceived social support: using crawled and self-reported data of Facebook users. J. Comput.-Mediat. Commun. 21, 282–297. https://doi.org/10.1111/jcc4.12160.

Tabachnick, B.G., Fidell, L.S., 2007. Using Multivariate Statistics. MC, Pearson Education Inc, Boston.