Geospatial suicide clusters and emergency responses: An analysis of text messages to a crisis service

Mark E. Larsen, *Member, IEEE,* Michelle Torok, Kit Huckvale, Bilal Reda, Sofian Berrouiguet, Helen Christensen

Abstract— **Suicide is a leading cause of death globally, and certain locations experience clusters of increased frequencies of suicidal behaviours. Prevention efforts are warranted in these locations to prevent contagion effects, and there is increasing interest in technology-supported suicide prevention interventions. Crisis support services are also implementing online and mobile health support. This study investigated the relationship between geospatial suicide clusters in the US and service use, and emergency responses initiated by, a text message-based crisis support service. 103,570 conversations involving 64,391 unique users over a two-year period were de-identified, analysed, and mapped to the state and county level. Moderate correlations were observed between service user rate and suicide mortality (ρ=0.53), and active rescues and suicide mortality (ρ=0.46). Suicide clusters were associated with a non-significant increase in service use (p=0.06) and active rescues (p=0.48). While it was not possible to observe significant cluster effects within this dataset, future analysis involving data collected through mobile health platforms may lead to better identification of risk at an individual level.**

I. Introduction

Suicide is a leading cause of death globally [1] – there are approximately 800,000 deaths by suicide each year, and it is notably the second most common cause of death for young people aged 15-29 [2]. Crisis services are designed to provide access to immediate, short-term support and resources to reduce stress, improve coping strategies, and help individuals to manage future crises. Traditionally crisis support is delivered over the phone (for example, 'Lifeline' in Australia, the 'Samaritans' in the UK, and the 'National Suicide Prevention Lifeline' in the US), however new ehealth and mhealth modes of engagement via online chat, social media, and short message service (SMS) text messaging are being introduced [3]. One such service is Crisis Text Line – a text-message based crisis service available across the US [4].

Geospatial suicide clusters are specific, usually public, and easily accessible sites which provide either the means or opportunity for suicide, and are therefore associated with an increased frequency of suicide deaths [5]. Prevention efforts at these sites are particularly important to avoid *contagion* effects [6], where knowledge or media reporting about one suicidal act may increase the likelihood that others will attempt suicide at the site. While prevention activities in these sites have typically focused on physical means restriction, such as installing fencing, there is increasing interest in technology-supported suicide prevention strategies including dedicated crisis hotline communication infrastructure [5,7], and CCTV video monitoring and detection [8,9].

Services such as Crisis Text Line may be more likely to be used by individuals in high stress or crisis situations, such as those who have visited, or are planning to visit, specific sites with intent to attempt suicide. This paper examines the relationships between geospatial suicide clusters and patterns of engagement with the Crisis Text Line SMS service in the US. We originally aimed to conduct two analyses: firstly, whether *mentions* of suicide clusters were associated with an increased rate of emergency response; and secondly, whether suicide clusters were *geographically* associated with increased crisis contacts to the service. Changes to the data sharing platform during the initiation of the project removed access to the linguistic content of the messages, therefore it was not possible to conduct the first planned analysis. The results of the geographic analysis are reported here.

II. Methods

This project was approved by the University of New South Wales Human Research Ethics Committee (HC16809) and the Crisis Text Line ethics review board. A waiver of signed consent was granted due to the secondary analysis of de-identified, routinely collected data.

A. Datasets

1) Crisis conversation data

Details of SMS contacts to Crisis Text Line were provided via the service's Data Enclave [10]. The Data Enclave is a platform designed to allow researchers access to curated, de-identified sections of the service's database.

De-identified details of conversations with the service over the period 25 August 2013 to 2 September 2017 were provided. These details included:

- A unique conversation-level and user-level identifier.
- Conversation start and end timestamps.
- Whether an active rescue was initiated via the emergency services (yes/no).
- The registered location of the service user (based on the NPA-NXX dialling code) and mapped to a state and county where possible via a Federal Information Processing Standard (FIPS) code.

Access to identifiable details such as the service user's name, date of birth, phone number, or the contents of the text messages were not provided. Demographic data were

M. E. Larsen, M. Torok, K. Huckvale, B. Reda and H. Christensen are with the Black Dog Institute, University of New South Wales, Randwick, NSW 2031, Australia (corresponding author: mark.larsen@blackdog.org.au).

S. Berrouiguet is with the University Hospital of Brest, Boulevard Tanguy Prigent, 29609 Brest, France.

available for a subset of users who agreed to these being collected through a post-conversation survey.

2) Population and mortality data

The populations of each county and state were obtained from the United States Census Bureau [11], to allow population rates to be calculated. The number of suicide-related deaths for each state were obtained from the Centers for Disease Control and Prevention (CDC) WONDER database [12]. Mortality data were suppressed for a large number of counties, due to low count numbers risking possible individual reidentification. County-level death data were therefore not analysed further.

3) Geospatial clusters

An online search was conducted by the research team to identify published academic literature, grey literature, and media reports which named specific geospatial suicide clusters or 'hotspots' across the US. Sites were considered if estimated or actual suicide death data were reported, and their geographic locations were mapped to the county-level. Ten such locations were identified for analysis. As best practice, to reduce the risk of contagion effects, the locations have not been named in this manuscript, however are available upon request from the corresponding author.

4) Analysis

As the Crisis Text Line service was soft-launched in targeted geographic regions in 2013, we excluded the first 12 months' service data to allow for uptake to propagate across the country. To allow for full year comparisons with population and mortality data, we further restricted the analysis period to the full 2015 and 2016 calendar years.

The number of unique service users were identified at the state and county-levels, and the user contact rate calculated per 100,000 population, per annum. The active rescue rate was calculated as number of conversations where an active rescue was initiated, per 100 conversations. The crude suicide mortality rate was calculated as the number of suicide deaths, per 100,000 population, per annum. Due to the non-normalcy of the data, pairwise rank correlations between the user contact rate, active rescue rate, and suicide mortality rate were performed, and the Spearman correlational (ρ) calculated.

As the identified geospatial clusters are sparsely distributed across the country, and knowledge about their location is unlikely to be restricted by arbitrarily-defined county boundaries, we estimated a locality user contact rate. The locality for each county included those counties immediately adjacent, as defined in the United States Census Bureau County Adjacency File [13]. For each locality, the user contacts, active rescues and populations were summed across the adjacent counties. The rates of user contacts and active rescues were then compared for localities surrounding clusters vs other localities. Due to the non-normalcy of the data, Wilcoxon ranked sum tests were performed to compare service use rates between localities centred on a geospatial suicide cluster vs other localities.

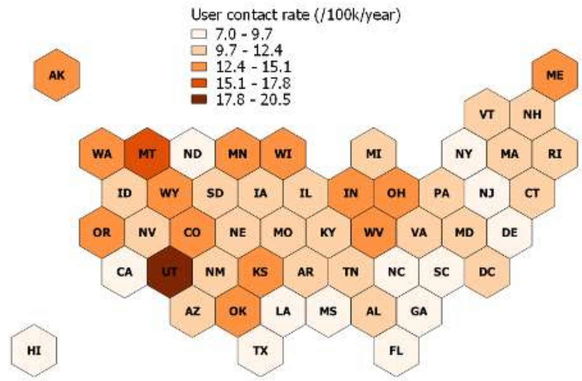

Figure 1. Cartogram (one hexagonal cell per state) illustrating the variation in user contact rate at a state level.

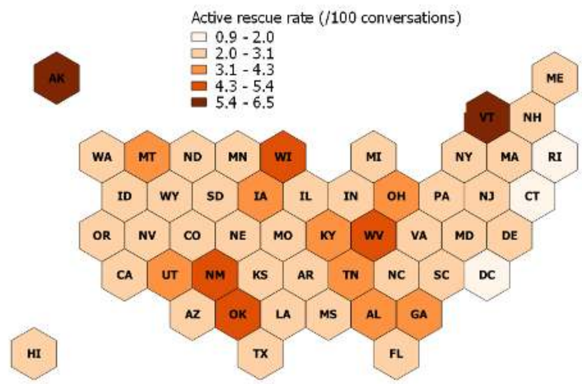

Figure 2. Cartogram illustrating the rate of emergency service active rescues at a state level.

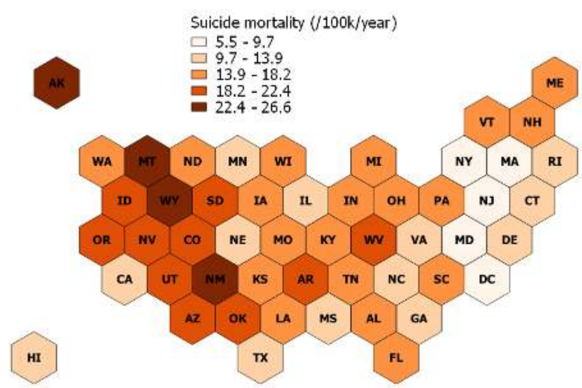

Figure 3. Cartogram illustrating the suicide mortality rate at a state level.

III. RESULTS

De-identified data were obtained through the Crisis Text Line Data Enclave for 211,258 conversations from 127,443 unique service users for the period 25 August 2013 to 2 September 2017. Users typically engaged in a single conversation (median: 1; interquartile range, IQR: 1-1; range: 1-139) lasting just over an hour (median: 63.5 minutes; IQR:

42.3-91.9 minutes; range: 61 seconds – 382.2 days). Age data were provided by 10% of users, with 48% aged 14-17 years old. Gender was reported for 12% of users, 73% of whom identified as female, 16% identified as male, and 11% provided another gender identity.

During the two-year analysis period, 103,570 conversations with 64,391 users were mapped to US counties through the NPA-NXX and FIPS codes. The national user contact rate was 10.4 users/100k/year, which ranged from 7.0 users/100k/year in Florida to 20.5 users/100k/year in Utah. The variation in user contact rates is plotted as a cartogram in Fig. 1, with each hexagonal cell representing a state. Nationally, 2.9% of conversations resulted in an active rescue, ranging from 0.9% in Rhode Island to 6.5% in Alaska (see Fig. 2). The average suicide mortality rate was 13.8 deaths/100k/year, ranging from 5.5 deaths/100k/year in the District of Columbia to 26.6 deaths/100k/year in Alaska (see Fig. 3). Pairwise correlations indicated moderate associations between the three calculated rates, as shown in Table I.

TABLE I. CORRELATION COEFFICIENTS BETWEEN THE CALCULATED USER CONTACT, ACTIVE RESCUE, AND SUICIDE MORTALITY RATES

Correlation coefficients (ρ)	Active rescue rate	Suicide mortality rate
User contact rate	0.41	0.53
Active rescue rate	–	0.46

After pooling adjacent county-level data, the median user contact rate was 9.9 users/100k/year, and the median active rescue rate was 4.0%. Fig. 4 shows scatterplots illustrating the contact rate and rescue rate for each county, characterised by the county population. The contact rate was higher for geospatial suicide clusters than other localities (median: 12.2 vs 9.9 users/100k/year), although this difference did not reach significance (p=0.06). Similarly, active rescues were more common in clusters (median: 4.3 vs 4.0 per 100 conversations), but this difference was not significant (p=0.48).

IV. Discussion

These results have demonstrated that there are moderate correlations, at a national level, between engagement with a text message suicide crisis service and suicide mortality: higher mortality rates are associated with higher user contact rates (ρ=0.53), although causality cannot be inferred these results. User contact rates were higher in locales associated with suicide clusters, however this difference did not reach significance. No significant difference was observed in the proportion of conversations requiring active responses from the emergency services. While some of these results are suggestive, there is no clear evidence that engagement with this mobile health crisis provider is associated with geospatial suicide clusters.

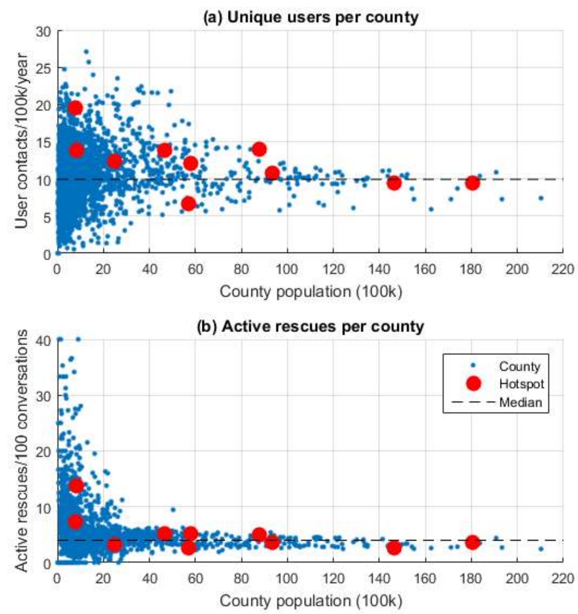

Figure 4. Scatterplots illustrating (a) the user contact rate, and (b) the active rescue rate for each county, as a function of the county population. Data plotted following pooling across adjacent counties. Each small blue dot represents an individual county, and each large red circle represents a county containing a suicide cluster. The median value is shown as a dashed black line.

There are a number of methodological challenges which may have contributed to the lack of positive findings. Firstly, the location of a service user was mapped to a county using their mobile phone area code – this is a poor representation of a user's actual location and their position relative to a specific cluster. Alternative modes of engagement may allow more accurate location data to be examined – for example, using location or GPS services on an app [14]. The identification of cluster locations was also a challenge, and identifying these through online sources may not have accurately identified true clusters of suicidal behaviour. These data are particularly challenging to obtain from medical or police records due to poor coding of self-harm and suicide incidents [15], and inaccurate identification of specific incident locations. Better data collection and analysis systems may therefore help improve identification of high-risk locations. Furthermore, this study was limited to 10 sites, therefore severely limiting the statistical power to detect differences.

It should also be noted that, while these sites represent a geospatial clustering of suicidal behaviour at a population level, only a small proportion of individuals in crisis will make or enact plans involving one of these sites. Therefore, it was not possible to reliably detect the signal of interest in this study. While it was not possible to access the content of the text messages for this analysis, natural language processing has been applied to social media posts and shows promise for detection of risk at an individual level [16,17]. Furthermore, other data modalities have been successfully examined via mobile health platforms in the context of individual mental health, including GPS location [14], accelerometery [18], and markers of social connectivity [19]. Together with improved coding of suicidal behaviour within electronic health records

– these emerging datasets offer the potential to improve identification of those who may be in crisis.

References

[1] World Health Organization, "Disease and injury regional mortality estimates, 2000–2012," 2012. Available from http://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.html

[2] World Health Organization, "Suicide data," 2017. Available from http://www.who.int/mental_health/prevention/suicide/suicideprevent/en/

[3] M. E. Larsen, N. Cummins, T. W. Boonstra, B. O'Dea, J. Tighe, J. Nicholas, F. Shand, J. Epps, and H. Christensen, "The use of technology in suicide prevention," *Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 7316-7319. 2015.

[4] Crisis Text Line, "Crisis Text Line", 2018. Available from https://www.crisistextline.org/

[5] J. Pirkis, L. S. Too, M. J. Spittal, K. Krysinska, J. Robinson, and Y. T. Cheung, "Interventions to reduce suicides at suicide hotspots: A systematic review and meta-analysis," *Lancet Psychiatry*, vol. 2, no. 11, pp. 994-1001. 2015.

[6] A. Beutrais, "Suicide by jumping," *Crisis*, vol. 28, no. S1, pp. 58-63. 2007.

[7] G. R. Cox, C. Owens, J. Robinson, A. Nicholas, A. Lockley, M. Williamson, Y. T. D. Cheung, and J. Pirkis, "Interventions to reduce suicides at suicide hotspots: A systematic review," *BMC Public Health*, vol. 13, pp. 214. 2013.

[8] J. Lee, C. Lee, and N. Park, "Application of sensor network system to prevention suicide from the bridge," *Multimedia Tools and Applications*, vol. 75, no. 22, pp. 14557-14568. 2016.

[9] A. Lockley, Y. T. D. Cheung, G. Cox, J. Robinson, M. Williamson, M. Harris, A. Machlin, C. Moffat, and J. Pirkis, "Preventing suicide at suicide hotspots: A case study from Australia," *Suicide and Life-Threatening Behavior*, vol. 44, no. 4, pp. 392-407. 2014.

[10] A. R. Pisani, N. Kanuri, B. Filbin, C. Gallo, M. Gould, L. S. Lehmann, R. Levine, J. E. Marcotte, B. Pascal, D. Rousseau, S. Turner, S. Yen, and M. L. Ranney, "Protecting user privacy and rights in academic data-sharing partnerships: Principles from a pilot program at Crisis Text Line," *Journal of Medical Internet Research*, vol. 21, no, 1, pp. e11507. 2019.

[11] United States Census Bureau, "Annual Resident Population Estimates, Estimated Components of Resident Population Change, and Rates of the Components of Resident Population Change for States and Counties: April 1, 2010 to July 1, 2016". Available from https://www2.census.gov/programs-surveys/popest/datasets/2010-2016/counties/totals/

[12] Centers for Disease Control and Prevention, "CDC WONDER," 2017. Available from https://wonder.cdc.gov

[13] United States Census Bureau, "County Adjacency File". Available from https://www.census.gov/geo/reference/county-adjacency.html

[14] S. Saeb, M. Zhang, M. Kwasny, C. J. Karr, K. Kording, and D. C. Mohr, "The relationship between clinical, momentary, and sensor-based assessment of depression," *2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, pp. 229-232. 2015.

[15] J. R. Randall, L. L. Roos, L. M. Lix, L. Y. Katz, and J. M. Bolton, "Emergency department and inpatient coding for self-harm and suicide attempts: Validation using clinician assessment data," *International Journal of Methods in Psychiatric Research*, vol. 26, no. 3, pp. e1559. 2017.

[16] T. Nguyen, B. O'Dea, M. Larsen, D. Phung, S. Venkatesh, and H. Christensen, "Using linguistic and topic analysis to classify sub-groups of online depression communities," *Multimedia Tools and Applications*, vol. 76, no. 8, pp. 10653-10676. 2017.

[17] B. O'Dea, S. Wan, P. J. Batterham, A. L. Calear, C. Paris, and H. Christensen, "Detecting suicidality on Twitter," *Internet Interventions*, vol. 2, no. 2, pp. 183-188. 2015.

[18] O. Demasi, A. Aguilera, and B. Recht, "Detecting change in depressive symptoms from daily wellbeing questions, personality, and activity," *2016 IEEE Wireless Health (WH)*, pp. 1-8. 2016

[19] T. W. Boonstra, A. Werner-Seidler, B. O'Dea, M. E. Larsen, and H. Christensen, "Smartphone app to investigate the relationship between social connectivity and mental health," *Proceedings of the 39th Annual IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 287-290, 2017.