Title

Geographic Clustering of Emergency Department Visits for Deliberate Self-Harm Injury in California

Author
Jensen, Jordan

Publication Date
2022-10-18

Data Availability
The data associated with this publication are available upon request.

Geographic Clustering of Emergency Department Visits for Deliberate Self-Harm Injury in California

Jordan Jensen

University of California, Merced

August 2022

Introduction

Suicide remains a major public health concern in the United States. In 2020, over 45,000 individuals took their own lives through suicide nationwide (Centers for Disease Control and Prevention, 2021b). Completed suicides are not the only cause for concern. It is estimated that approximately 500,000 individuals nationwide will present to an emergency department (ED) with injuries stemming from deliberate self-harm every year (Centers for Disease Control and Prevention, 2020). Deliberate self-harm has been difficult to measure as many injuries sustained through this behavior do not require visits to the ED (Bethell & Rhodes, 2009). Despite the inherent limitation that leads to underestimating the phenomena, ED visits still offer one of the best measures of the incidence and prevalence of deliberate self-harm (Chen & Aseltine, 2017).

An area of perennial interest in suicidology is the apparent clustering of suicide and self-harm. In research on the clustering of suicide and self-harm, two main types of clustering have been considered. The first type is known as a mass cluster. Mass clusters, also known as temporal clusters, occur when there is a greater than expected number of suicides or incidences of self-harm within a certain period (Centers for Disease Control and Prevention, 2021a). Point clusters, also known as spatial-temporal clusters, are like mass clusters in that they occur when there is a greater-than-expected number of suicides or incidences of self-harm within a certain period. However, point clusters differ from mass clusters in that they must also occur within a distinct geographical location (Centers for Disease Control and Prevention, 2021a). A third and less researched type of clustering is known as an echo cluster, which occurs at the same location as a past cluster but only after a specified amount of time (Hawton et al., 2020). The Centers for

Disease Control and Prevention (CDC) have outlined common criteria that clusters share. These criteria include a connection between cases (whether that be time, place, or a combination of the two), a greater than expected number of cases, a defined group of people, and occurring in a certain period (Centers for Disease Control and Prevention, 2021a). Current estimates indicate that approximately 1% to 5% of suicides in young adults occur in a cluster (Centers for Disease Control and Prevention, 2021a). Estimates of the percent of suicides that cluster in adults are not readily available.

The bedrock for research into the clustering of suicide and self-harm may be attributed to Joiner's research into the social contagion of suicide (Joiner, 1999). Joiner outlined through his Interpersonal Theory of Suicidal Behavior (IPTS) the antecedents of an individual's decision to die by suicide. These antecedents are, as the name suggests, socially derived, such as social isolation (thwarted belongingness) and perceived burdensomeness. A third antecedent, acquired capability, occurs when an individual loses their innate fear of death through repeated exposures to traumatic experiences (such as a peer's/relative's suicide) or physical pain. When these components are combined, an individual's risk of attempting suicide is heightened (Joiner, 2007; Van Orden et al., 2010). While not invoking clustering by name, the social aspects of this theory highlight how suicide and self-harming behavior may be socially contagious through exposure to suicide or other self-harming behavior.

Spatial analytical methods have been utilized by researchers to quantify and describe the presence of the clustering of suicide and suicidal behavior. In 1990 pioneering work conducted by Gould, Wallenstein, and Kleinman provided proof of the concept that the clustering of suicide could be analyzed through the use of spatial statistics and by using data from vital statistics registries (Gould et al., 1990). In 2007, Exeter and Boyle conducted an analysis to examine if the suicides of young adults clustered geographically, with an added goal of examining if these clusters persisted through time (Exeter & Boyle, 2007). Conducted in Glasgow, Scotland, the researchers used data from a death registry aggregated to an areal level known as a "consistent area through time" or CATT. Kulldorff's spatiotemporal scan statistic was then used to identify significant clusters of suicide deaths over three distinct time periods. Through this process, a significant cluster that persisted through time was identified in Glasgow and was highly correlated with a measure of social deprivation (Exeter & Boyle, 2007).

In 2013, Jones et al. conducted a similar study that attempted to identify suicide clusters in the neighboring country of Wales. Similarly, to Exeter and Boyle, the researchers used data on suicide deaths from a death registry. However, they differed from the previous research by examining the cases at the person level, using the individual's address of residence as the spatial unit of analysis. The researchers used Kulldorff's spatiotemporal scan statistic to identify spatiotemporal clusters of suicide deaths in Wales between 2000 and 2009. A single significant cluster was identified between late 2007 and early 2008 that contained ten deaths by suicide. Despite this study identifying such a specific cluster that correlated with media concern over a cluster, the researchers noted several limitations to their study that future research should consider. Some of these limitations included the limited ability to analyze clusters in a timely matter as data on suicide is often only produced on an annual basis and that community perceptions of an ongoing cluster may outweigh statistical evidence of one occurring (Jones et al., 2013).

A more recent study conducted by Kassem et al., in 2019, attempted to identify clusters of suicide in Idaho. The researchers used many of the same methods as previous researchers, such as retrieving suicide data from death registries, aggregating it to an areal unit, and employing a spatial scan statistic. A key difference was that the scan statistic used was a purely spatial version of Kulldorff's spatial scan statistic that does not include a measure for time. The researchers identified a single most likely cluster along with nine secondary clusters existing between 2010 and 2014. The researchers also compared locations within clusters to locations outside of clusters for demographic characteristics provided by the American Community Survey (ACS). The researchers noted that their method was a feasible option for public health departments interested in identifying spatial clusters of suicide (Kassem et al., 2019).

Researchers have used a broad range of different study windows in their analyses. Exeter and Boyle as well as Jones and colleagues analyzed whole countries for the presence of suicide clusters over time while Johnson, Too, and Torok, analyzed single states (Exeter & Boyle, 2007; Johnson et al., 2017; Jones et al., 2013; Too et al., 2017; Torok et al., 2017). Kassem and colleagues as well as Sy and colleagues analyzed clustering over several states with high levels of suicide deaths (Kassem et al., 2019; Sy et al., 2019). There is also a general use of aggregation to a lower resolution of data such as the census block level or the county level, although Jones and colleagues were able to conduct an analysis at the individual level based on the residence of the individual who died by suicide (Jones et al., 2013).

In the current literature, there is a lack of studies that examine *non-fatal* deliberate self-harm (DSH) as the explicit outcome of interest, often including it only as a secondary measure of suicide attempts. For example, research conducted by Too and colleagues used hospital data on DSH to study the clustering of suicide attempts in Western Australia, finding evidence of spatio-temporal clusters. Torok and colleagues completed a similar study in New South Wales, Australia. Using hospital and vital statistics data, the researchers conducted a time-independent spatial analysis examining the occurrence clusters of suicide attempts and completed suicides concurrently over time. DSH was used as a measure of suicide attempts in both of these cases. While these studies were pioneering in their own right for incorporating DSH, gaps still remain in regarding the use of spatial analysis to investigate this behavior. To our knowledge, there has been no study conducted within the United States that employs geospatial analysis to examine the clustering of DSH. Examining the clustering of DSH is of importance for several reasons. Self-harm can be highly traumatic to those who experience it and there is evidence to suggest that it may lead to a higher risk of eventual death by suicide (Bergen et al., 2012; Zahl & Hawton, 2004). There is also a need to investigate if the potential clustering of DSH exists along certain demographic characteristics such as age and sex to determine which groups are most at risk of being involved inside of a cluster. Through investigating DSH through spatial analyses, there may be an opportunity to inform prevention interventions with more specificity than current options which rely upon generalized prevention techniques.

Continuing with the use of the methodologies laid out by these previous studies, this study aims to identify and describe purely spatial clusters (i.e., time-independent) of high rates of ED visits for injuries sustained through deliberate self-harm within the state of California. A second aim is to identify subgroup-specific clusters regarding age and sex. To our knowledge, this study is the first to use geospatial techniques to investigate the clustering of self-harm within California whose large and diverse population provides a suitable platform to perform this research while also exhibiting a need for it.

Methods

Data

Data on visits to California-licensed emergency departments were obtained from the California Department of Health Care Access and Information (HCAI). These data included visits by individuals five years or older with a California residential zip code and ranged from January 1, 2009, to December 31, 2013. Data were aggregated to the zip code level. For individuals with more than one visit to the emergency department, only their initial visit was retained for analysis to ensure the assumption of independent observations. This visit is hereafter referred to as the index visit. These data were then linked to zip code level data from the 2010 census which provided data on overall population counts, demographic-specific population counts by age and sex, and the spatial geometry for each zip code (United States Census Bureau, 2021).

The dataset initially contained 171,817 deliberate self-harm visits. After removing repeated visits by the same patients, this decreased to 118,194 index visits. A total of 1,769 California zip codes were present in the census data. Of these, 35 were excluded from our analytic dataset: 15 because they were not observed in the HCAI dataset, 17 because they were geographic islands (defined as zip codes that had no neighbor that shared a common border such as the zip code that contains the city of Lone Pine which is surrounded by wilderness), and 10 because the census dataset recorded a population of 0. Geographic islands were excluded from the analysis as the analyses used in this paper incorporate and require information from spatial neighbors. Zip codes that have 0 population recorded are often federal buildings (e.g., post offices), military installations (e.g., naval shipyards), or universities. Following the removal of these zip codes, there were 1,742 total zip codes in the dataset.

Measures

The primary measure of interest was the rate of ED visits for self-harm injuries observed in each zip code, calculated as the number of ED visits divided by the census-reported population size. All visits with an International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) external cause-of-injury code (E-code) E950.0-958.x, in any diagnostic position, were defined as deliberate self-harm visit. These E-codes are assigned to individuals who present to the emergency department with evidence of a suicide attempt or self-inflicted injury through various means (cutting/piercing injuries, poisoning, strangulation, etc.). Other patient characteristics assessed at the time of the visit were gender (male, female) and age group (5 to 14, 15 to 29, 30 to 44, 45 to 59, 60 to 74, and $\geq$ 75). Age was recorded as continuous in the ED data and aggregated into categories aligned with the 2010 census measure of age.

Statistical Analyses

Initial summary statistics were computed to capture the phenomenon of ED visits for deliberate self-harm in California between 2009 and 2013. This was completed by tabulating the overall counts of the incidence of ED self-harm visits by demographic factors such as sex, ethnicity, age, and method of DSH. The overall rate of index ED visits due to deliberate self-harm was calculated by dividing the total number of these visits over the census population for California in 2010. Zip code-specific rates were generated through the same process using the zip code-level data for visits and population size.

Regarding our primary analysis, we attempted to identify geographic areas of high incidence rates for ED self-harm visits, using the purely spatial form of Kulldorff’s scan statistic (Kulldorff, 1997, 2005). Kulldorff’s discrete spatial scan statistics are calculated by moving circular windows of pre-defined maximum radii over the centroids for every zip code in the analysis. Each window defined by the scan is considered to be a potential cluster that is based around the centroid of each zip code. A likelihood ratio test compares the rates of deliberate self-harm inside and outside of each window. Thus, a finding of no significant clusters would demonstrate that no geographic regions were identified as having large incidence rates. As even non-fatal deliberate self-harm is relatively rare, each zip code level count of deliberate self-harm events is assumed to be independently Poisson distributed. The spatial window estimated as having the highest likelihood ratio is reported as the most likely cluster.

The primary analysis defined distance using a network file. This network file defined neighbors as those zip codes that shared at least one geographic border. These first-order neighbors assigned the distance between each first-order neighbor to be one. The decision to incorporate this network file was done to provide consistency with any future spatial regression analyses that will be conducted, which commonly incorporate a network file rather than other distance metrics. As a sensitivity analysis, the centroid-to-centroid distance was also considered, a more commonly used practice. The significance of spatial clusters was assessed using Monte Carlo replications to construct a p-value. The maximum spatial cluster size was set at the default level of 50% of the underlying population at risk and a circular spatial window was selected.

Separate datasets were prepared for further stratified analyses by gender and age distribution. Stratified analyses were conducted through the same method as the primary analysis. Zip code-level population data from the 2010 Census was used to define estimated gender- and age-specific population denominators for each zip code. This information was then merged with zip code-level counts of ED self-harm visits, stratified by the same categories, to generate gender and age-specific self-harm visit rates in each zip code. Cleaning, analysis, and mapping of the data were completed using STATA 17.0, R 4.1.3, and SaTScan 10.0.

Results

Between 2009 and 2013 in California, the overall visit rate for index ED self-harm injuries was 369 per 100,000 people. There were 137,479 index visits to the ED for self-harm recorded between January 1, 2009, and December 31, 2013. Most of these visits (n = 79,928, 58.1%) were made by female patients. There were four visits assigned to individuals whose gender was reported as unknown. Regarding ethnicity: 4.4% of visits were from Asian patients, 9.1% of visits were from non-Hispanic Black patients, 24.1% of visits were from Hispanic patients, 57% of visits were from non-Hispanic White patients, and 5.5% of visits were from those of unknown race/ethnicity. Regarding age: 3.6% of visits were from 5-14-year-old patients, 41% of visits were from 15-29-year-old patients, 26.1% of visits were from 30-44-year-old patients, 21.6% of visits were from 45-59-year-olds, 5.6% of visits were from 60-74-year-old patients, and 2.1% of visits were from 75+-year-old patients. The most common methods used for self-harm were poisoning via a solid or liquid substance (65.7%), or, cutting or piercing (21.1%), unspecified means (7.8%), hanging or suffocation (2.06%), firearms (0.7%), and falling or jumping (0.7%) making up the remainder. Refer to Table 1 for more information on patient characteristics.

Table 1 Summary of Statewide Self-Harm Visits to the ED, 2009-2013 (n = 137,479 index visits)

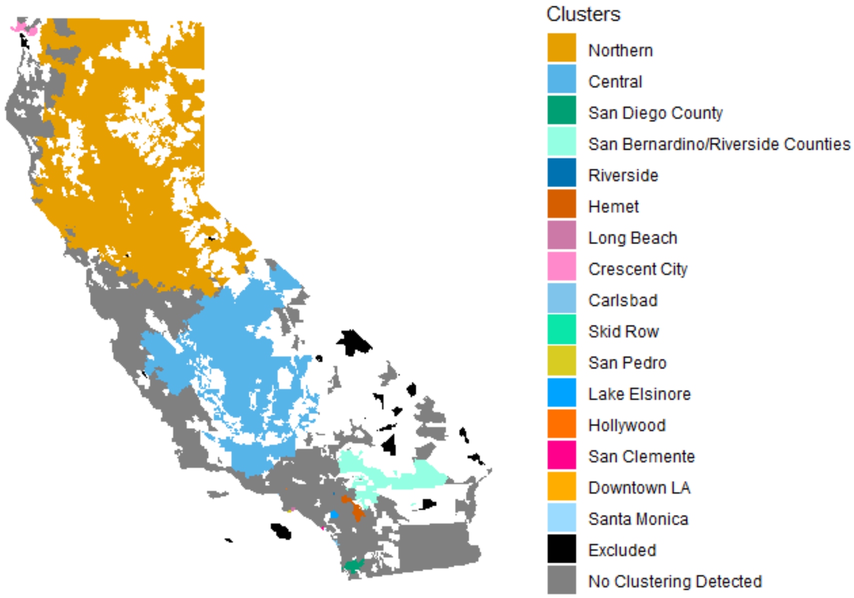
Patient Characteristic	Count (%)
Sex	

Female	79,928 (58.1%)
Male	57,547 (41.9%)
Unknown	4 (< 1%)
Race/Ethnicity	
Non-Hispanic Asian/Pacific Islander	6,017 (4.4%)
Non-Hispanic Black	12,504 (9.1%)
Hispanic	33,094 (24.1%)
Non-Hispanic White	78,354 (57.0%)
Non-Hispanic Other	7,510 (5.5%)
Age in years	
5-14	4,997 (3.6%)
15-29	56,390 (41.02%)
30-44	35,897 (26.11%)
45-59	29,642 (21.56%)
60-74	7,700 (5.6%)
75+	2,853 (2.1%)
Method	
Poisoning via solid or liquid substances	90,314 (65.7%)
Poisoning via gas	44 (0.03%)
Poisoning via other gas or vapors	556 (0.40%)
Hanging or suffocation	2,829 (2.06%)
Submersion or drowning	114 (0.08%)
Firearm	1,010 (0.73%)
Cutting or piercing	28,965 (21.1%)
Falling or jumping	989 (0.72%)
Late effects of injury	5 (< 0.01%)
Other unspecified means	10,676 (7.77%)
Unknown	1,977 (1.44%)

A total of 16 distinct, non-overlapping clusters were identified from the primary analysis. The largest cluster contained 498 zip codes in northern California covering largely rural areas as well as the urban areas of Sacramento and Santa Rosa. The second largest cluster contained 164 zip codes and encapsulated most of the San Joaquin Valley. Five smaller clusters contained zip codes in Los Angeles County, San Bernardino and Riverside Counties, Orange County, and San

Diego County, respectively. Nine clusters consisted of single zip codes. Of the 1742 zip codes included in the analysis, 41.7% were identified as being within a cluster. Refer to Figure 1 for a graphical display of the primary analysis.

Figure 1 (n = 1742)



Similar to the primary analysis, 16 clusters were identified when the data was restricted to visits by male patients. The largest cluster contained 372 zip codes and contained largely rural areas in northern California and more urban areas surrounding Sacramento. The second largest cluster contained 222 zip codes and encompassed a substantial portion of the San Joaquin Valley, as well as sections of the Pacific Coastal Ranges. 5 smaller clusters contained zip codes in San Francisco County, Los Angeles County, San Bernardino, and Riverside Counties, and San Diego County. Nine clusters consisted of single zip codes. When restricted to visits by female patients, 10 clusters were identified. The largest cluster contained 497 zip codes and spanned the northern rural regions to the northern tip of the San Joaquin Valley. The second largest cluster contained 455 zip codes and spanned from the center of the San Joaquin Valley to Los Angeles

County. Four smaller clusters contained zip codes in Los Angeles County, San Bernardino County, Orange County, and San Diego County. Two clusters consisted of single zip codes.

Four clusters were identified when the data were restricted to ages under 5 to 14. The largest cluster contained 503 zip codes and spanned from the northern rural regions in the north to the northern tip of the San Joaquin Valley. The second largest cluster contained 128 zip codes and stretched from the center of the San Joaquin Valley to its southernmost extent. The remaining clusters contained zip codes in San Bernardino County and San Diego County.

Eight clusters were identified when the data were restricted to ages 15 to 29. The largest cluster contained 600 zip codes spanned from the rural northern regions to the northern tip of the San Joaquin Valley. The second largest cluster contained 257 zip codes and spanned from the center of the San Joaquin Valley to Los Angeles County. Three smaller clusters contained zip codes in Del Norte County, San Bernardino County/Riverside County, and San Diego County. Three clusters contained single zip codes.

Eight clusters were identified when the data were restricted to ages 30 to 44. The largest cluster contained 494 zip codes and spanned from the northern rural region to the northernmost tip of the San Joaquin Valley. The second largest cluster contained 462 zip codes and spanned from the center of the San Joaquin Valley to Los Angeles County. Five smaller clusters contained zip codes in Del Norte County, Los Angeles County, San Bernardino/Riverside County, and San Diego County. One cluster contained a single zip code.

12 clusters were identified when the data were restricted to ages 45 to 59. The largest county contained 494 zip codes and spanned from the northern rural region to Sacramento. The

second largest cluster contained 71 zip codes and contained zip codes from the Western desert regions. Six smaller clusters contained zip codes in Fresno County, Monterey/San Luis Obispo County, Los Angeles County, San Bernardino County, and San Diego County. Four Clusters contained single zip codes.

Eight clusters were identified when the data were restricted to ages 60 to 74. The largest cluster contained 502 zip codes and spanned from the northern rural region to the northernmost tip of the San Joaquin Valley. Three smaller clusters contained zip codes in Fresno County, San Bernardino County, and San Diego County. Four Clusters contained single zip codes There was only a single cluster identified for ages 75 and up which was located in San Diego County.

The results of the cluster analyses indicated that the use of the network file led to the identification of more specific clusters. In each of the analyses that used the network file, either one less cluster was identified, or fewer locations were contained within identified clusters than in the same analyses that did not use the network file.

Discussion

Using data on non-fatal deliberate self-harm ED visits and census population counts in California, we attempted to identify spatial clustering using a spatial scan statistic in the overall population, and by age and sex. Through these analyses, three persistent and time-independent clusters were identified among almost all stratifications of the data. While the size of these clusters changed through each stratification, they generally spanned the same geographic area. The first persistent cluster was located in northern California, typically ranging from the border with Oregon at its northern extent to Sacramento in the south. The Northern cluster was geographically close to the second and contained a sizable portion of the San Joaquin Valley. The San Joaquin Valley cluster generally began in the city of Modesto in the north and ended in the city of Santa Clarita to the south. The final cluster that was present in all stratifications of the analyses was centered in San Diego County. While not as geographically expansive as the other persistent clusters, it is striking that the San Diego cluster remained present no matter the stratification. The sensitivity analyses revealed only a minimal difference in cluster detection between analyses that used a network file and analyses that did not.

This study was similar to others that have used spatial scan statistics to identify clustering of deaths by suicide and suicide attempts in that it was able to identify clustering. However, it did differ in that the size of the clusters identified in the analyses were especially large in comparison. This may be due in part to how the spatial scan statistic operates through using a circular window which may fail to identify smaller clusters that are irregularly shaped, leading to the identification of single large clusters instead of several smaller clusters (Zhang et al., 2010).

This study is limited in its findings by several factors. Aggregating individual-level data to zip code-level data lowers the overall resolution of the data as well as the strength of the findings. Traditionally, clustering is thought to occur in either mass clusters, which can span a large geographic area, or point clusters, which take place in the same geographic area (Centers for Disease Control and Prevention, 2021a). By aggregating to the zip code level, the analysis is unable to attribute a cluster to a specific location within a zip code. Only the zip code centroid is reflected in the output. Another limitation is that visits to the ED for self-harm may be misattributed to a zip code through entry error at the point of service. This would result in a proportion of cases being analyzed as part of a zip code to which they have no relation.

This study is also limited by the fact that EDs cannot capture a complete picture of suicidal and self-harm behavior throughout the entirety of California. Many instances of self-harm are not severe enough to require a visit to the emergency department (Bethell & Rhodes, 2009). This may imply that the data used in the analyses consists of only the most severe cases of self-harm. Another limitation related to the source data is that of repeated visits. Only the index visit was kept in the dataset to ensure independence in the observations. This eliminates any data from individuals who visited the ED for another self-harm event. This is not an uncommon occurrence and methods have been proposed to include these visits in spatial analyses, although it was outside of the scope of this study (Torok et al., 2017).

Conclusion

By using scan statistics and ED visit data on non-fatal deliberate self-harm in California, this study identified three time-indented clusters located in northern California, central California, and San Diego County. To the researcher's knowledge, this is the first study to employ scan statistics to identify clustering of self-harm that incorporates a comprehensive view of California. Future research should prioritize sourcing data at a higher resolution and incorporating time into the analyses. The use of aggregate level data, while usually the only feasible source, may limit the findings of this kind of research in its ability to discern exactly where clustering is taking place. There is also a need for an investigation of the use of scan statistics that do not rely on a circular window which may result in the identification clustering at a more precise level. This higher level of precision may allow for prevention efforts to be conducted more effectively by identifying the areas with the most need.

References

Bergen, H., Hawton, K., Waters, K., Ness, J., Cooper, J., Steeg, S., & Kapur, N. (2012). How do methods of non-fatal self-harm relate to eventual suicide? *Journal of Affective Disorders*, *136*(3), 526–533. https://doi.org/10.1016/j.jad.2011.10.036

Bethell, J., & Rhodes, A. E. (2009). Identifying deliberate self-harm in emergency department data. *Health Reports / Statistics Canada, Canadian Centre for Health Information = Rapports Sur La Santé / Statistique Canada, Centre Canadien d'information Sur La Santé*, *20*(2), 35–42.

Centers for Disease Control and Prevention. (2020). *Injury prevention & control: Nonfatal injury data*. Nonfatal Injury Reports. https://www.cdc.gov/injury/wisqars/nonfatal.html

Centers for Disease Control and Prevention. (2021a). *Suicide, suicide attempt, or self-harm clusters*. https://www.cdc.gov/suicide/resources/suicide-clusters.html#:~:text=In the United States%2C approximately, examination of age-specific effects.

Centers for Disease Control and Prevention. (2021b, May 22). *Fast facts*. Centers for Disease Control and Prevention. https://www.cdc.gov/suicide/facts/index.html

Chen, K., & Aseltine, R. H. (2017). Using hospitalization and mortality data to identify areas at risk for adolescent suicide. *Journal of Adolescent Health*, *61*(2), 192–197. https://doi.org/10.1016/j.jadohealth.2017.02.020

Exeter, D. J., & Boyle, P. J. (2007). Does young adult suicide cluster geographically in Scotland? *Journal of Epidemiology and Community Health*, *61*(8), 731–736. https://doi.org/10.1136/jech.2006.052365

Gould, M. S., Wallenstein, S., & Kleinman, M. (1990). Time-space clustering of teenage suicide. *American Journal of Epidemiology*, *131*(1), 71–78. https://doi.org/10.1093/oxfordjournals.aje.a115487

Hawton, K., Hill, N. T. M., Gould, M., John, A., Lascelles, K., & Robinson, J. (2020). Clustering of suicides in children and adolescents. *The Lancet Child and Adolescent Health*, *4*(1), 58–67. https://doi.org/10.1016/S2352-4642(19)30335-9

Johnson, A. M., Woodside, J. M., Johnson, A., & Pollack, J. M. (2017). Spatial patterns and neighborhood characteristics of overall suicide clusters in Florida from 2001 to 2010. *American Journal of Preventive Medicine*, *52*(1), e1–e7. https://doi.org/10.1016/j.amepre.2016.07.032

Joiner, T. (1999). The clustering and contagion of suicide. *Current Directions in Psychological Science*, *8*(3), 89–92.

Joiner, T. (2007). *Why people die by suicide*. Harvard University Press. https://doi.org/10.2307/j.ctvjghv2f

Jones, P., Gunnell, D., Platt, S., Scourfield, J., Lloyd, K., Huxley, P., John, A., Kamran, B., Wells, C., & Dennis, M. (2013). Identifying probable suicide clusters in Wales using national mortality data. *PLoS ONE*, *8*(8). https://doi.org/10.1371/journal.pone.0071713

Kassem, A. M., Carter, K. K., Johnson, C. J., & Hahn, C. G. (2019). Spatial clustering of suicide and associated community characteristics, Idaho, 2010-2014. *Preventing Chronic Disease*, *16*(3). https://doi.org/10.5888/pcd16.180429

Kulldorff, M. (1997). A spatial scan statistic. *Communications in Statistics - Theory and Methods*, *26*(6), 1481–1496. https://doi.org/10.1080/03610929708831995

Kulldorff, M. (2005). *SaTScan - Software for the spatial, temporal, and space-time scan statistics*. https://www.satscan.org

Sy, K. T. L., Shaman, J., Kandula, S., Pei, S., Gould, M., & Keyes, K. M. (2019). Spatiotemporal clustering of suicides in the US from 1999 to 2016: a spatial epidemiological approach. *Social Psychiatry and Psychiatric Epidemiology*, *54*(12), 1471–1482. https://doi.org/10.1007/s00127-019-01736-4

Too, L. S., Pirkis, J., Milner, A., & Spittal, M. J. (2017). Clusters of suicides and suicide attempts: Detection, proximity and correlates. *Epidemiology and Psychiatric Sciences*, *26*(5), 491–500. https://doi.org/10.1017/S2045796016000391

Torok, M., Konings, P., Batterham, P. J., & Christensen, H. (2017). Spatial clustering of fatal, and non-fatal, suicide in New South Wales, Australia: Implications for evidence-based prevention. *BMC Psychiatry*, *17*(1). https://doi.org/10.1186/s12888-017-1504-y

United States Census Bureau. (2021). *ZIP code tabulation areas (zctas). How zctas are created. Statistical data available for zctas*. United States Census Bureau.

https://www.census.gov/programs-surveys/geography/guidance/geo-areas/zctas.html

Van Orden, K. A., Witte, T. K., Cukrowicz, K. C., Braithwaite, S. R., Selby, E. A., & Joiner, T. E. (2010). The interpersonal theory of suicide. *Psychological Review*, *117*(2), 575–600. https://doi.org/10.1037/a0018697

Zahl, D. L., & Hawton, K. (2004). Repetition of deliberate self-harm and subsequent suicide risk: Long-term follow-up study of 11 583 patients. *British Journal of Psychiatry*, *185*(JULY), 70–75. https://doi.org/10.1192/bjp.185.1.70

Zhang, Z., Assunção, R., & Kulldorff, M. (2010). Spatial scan statistics adjusted for multiple clusters. *Journal of Probability and Statistics*, *2010*. https://doi.org/10.1155/2010/642379