



COMPUTER SCIENCE

NOVEL-RESULT

# Identifying spatial and temporal suicide clusters in a Californian county

Anders K. Waalen<sup>1,2,\*</sup> , Seraphim Telep<sup>1</sup> and Rimal Bera<sup>1</sup>

<sup>1</sup>Department of Psychiatry, UC Irvine School of Medicine, Irvine, California, USA, and <sup>2</sup>Zucker School of Medicine, Hofstra University, Port Jefferson, New York, USA

\*Corresponding author: Email: [awaalen@northwell.edu](mailto:awaalen@northwell.edu)

(Received 31 January 2022; Revised 08 January 2023; Accepted 25 January 2023)

## Abstract

Barriers to suicide cluster detection and monitoring include requiring advanced software and statistical knowledge. We tested face validity of a simple method using readily accessible household software, Excel 3D Maps, to identify suicide clusters in this county, years 2014–2019. For spatial and temporal clusters, respectively, we defined meaningful thresholds of suicide density as 1.39/km<sup>2</sup> and 33.9/yearly quarter, defined as the 95th percentile of normal logarithmic and normal scale distributions of suicide density per area in each ZIP Code Tabulated Area and 24 yearly quarters from all years. We generated heat maps showing suicide densities per 2.5 km viewing diameter. We generated a one-dimensional temporal map of 3-month meaningful cluster(s). We identified 21 total population spatial clusters and one temporal cluster. For greater accessibility, we propose an alternative method to traditional scan statistics using Excel 3D Maps potentially broadly advantageous in detecting, monitoring, and intervening at suicide clusters.

**Key words:** cluster; epidemiology; hotspot; monitor; suicide

## Background

A suicide cluster is defined by the Center for Disease Control (CDC) as “a group of suicides or suicide attempts, or both, that occur closer in time and space than would normally be expected in a given community” (O’Carroll et al., 1988). Unlike suicide hotspots, which are usually community-identified at public or iconic sites, clusters may span a wider geographic area and go unnoticed (be hidden) to the general population (Too et al., 2017).

Literature has proposed three theoretical mechanisms of clusters. The first explanation is the psychology concept of suicide contagion, which functions through social learning theory (Bandura, 1977). The second is that people in the same area may share high-risk characteristics for suicide, such as depressed patients in an inpatient psychiatric facility (Too et al., 2019). Third, Durkheim’s classical suicide theory offers a sociological explanation hinging on social integration and regulation concepts (Durkheim, 1897).

A report from the Orange County Health Care Agency (OCHCA) found that suicides were predominantly by people middle-aged, White, male, and using firearms. The report found suicide density by resident population to be greatest in coastal ZIP codes (Sanchez et al., 2019; Walker, 2019). It did not, however, examine suicide density by land mass *area*, which is arguably more clinically relevant when it comes to fair and equitable municipal-specific allocation of funds toward effective (Pirkis et al., 2015) but often costly (Kerkhof, 2003) prevention interventions (e.g., barriers) in high-risk areas such as suicide

hotspots. The OCHCA study may appear to suggest greater risk in affluent locations where there are fewer people and suicides, despite the vast majority occurring further inland in more resource-poor areas. While this type of analysis may be useful for addressing risk associated with specific *types* of people, such as White suburban males, it does nothing to address risk associated with specific *locations* and may lend itself to community bias (e.g., racial and classist). For example, this approach may lead to unfair allocation of additional prevention resources to predominantly White, resource-rich communities, abandoning locations of greater suicide numbers and risk. Analysis of suicide density by population, although fairly standard in epidemiology, fails to identify physical locations which by themselves attract suicides (e.g., through notoriety of well-known hotspots, access to means like jumping, or simply increased traffic of a downtown area) and as such, leaves out the clinical and practical relevance of suicide clusters and their effective interventions.

To the best of our knowledge, there has been only one study of clusters in Orange County (OC), assessing multiple hotspots on a university campus to date (Waaalen et al., 2019). However, the ZIP codes wherein that study's clusters lie are considered by the OCHCA study to be of the least suicide-dense regions by population. Therefore, it is likely that other hidden clusters exist in more suicide-dense OC regions. It is worth identifying possible hotspots, for the OCHCA failed to identify any actions at hotspots, or suicide clusters, in its report of suicide interventions (Sanchez et al., 2019).

The CDC's definition for suicide clusters offers no gold standard method for scanning for hidden clusters. A growing body of research has created multiple statistically advanced methods for detecting spatial, temporal, and spatiotemporal clusters, which require specialized software and/or trained statisticians (Cheung et al., 2013; Kulldorff, 1997; Tango & Takahashi, 2005). Each method is imperfect in its design and, most importantly, is not easily applied by the average community stakeholders who the CDC ultimately holds responsible for suicide prevention (Hill et al., 2020; O'Carroll et al., 1988). Furthermore, these methods' emphasis on statistical significance requires comparisons with unrelated, outside populations, lacks sensitivity, and may not apply to community-identified clusters (Boyce, 2011; O'Carroll & Mercy, 1990). This presents a major barrier to ongoing, real-time cluster detection required for necessary targeted interventions in the midst of our growing national suicide epidemic. "There is no reason to believe that all suicide clusters will inevitably come to the attention of health authorities through passive identification by groups in the community"... and... "early intervention might provide the best chance for minimizing the hypothesized contagious effect of suicide" (O'Carroll & Mercy, 1990). As such, the community is in urgent need of an accessible, operational definition and efficient modality to detect and monitor clusters in ongoing, real time.

The goal of this study is threefold: (a) test via face validity, a simple method using readily accessible household software to identify suicide clusters; (b) identify spatial and temporal suicide clusters in OC, and (c) examine possible mechanisms for these clusters, including suicide contagion (via relationship between temporal and spatial clusters), shared characteristics, and danger of location.

## Materials and methods

### *Study design and data collection*

The university IRB exempted this study from full review. Our design consisted of a population-based retrospective study of completed suicides in Orange Country from 2014 to 2019 based on data from the OC Sheriff's Department Coroner Division. These data were Coroner-determined and confirmed. For each suicide, data obtained included death date and time; gender (listed by sex), age (which we grouped into 10-year cohorts), and race of individuals; attempt means/mode; event address and place (e.g., hotel, freeway, and beach); and death address and place. In certain cases (e.g., ocean drowning), data described locations in terms of nearest location. We identified 618 total suicides. We individually mapped, rather than aggregated, suicide locations using Microsoft Excel and its Excel 3D Maps system.

### *Spatial clusters*

From event locations, Excel 3D Maps generated heat maps showing suicide densities per 2.5 km viewing diameter region (3.9 km<sup>2</sup> area). Its spatial interface is Bing Maps. The diameter was measured with Google Maps' measuring tool with landmarks for reference. We arbitrarily defined a meaningful threshold of suicide density per area of 1.39/km<sup>2</sup> defined as the 95th percentile of a normal logarithmic scale distribution of suicide density per area in each ZIP Code Tabulated Area (ZCTA) cumulated from all years. Four outliers in this distribution were not included. We examined each data point for location accuracy. If there were any locations Excel 3D Maps could not recognize, we examined and manually altered the address in question for accurate plotting. We subsumed suicides of five Zip codes with no corresponding ZCTA under the surrounding Zip code. We removed three ZIP codes whose suicide attempts did not, but whose corresponding deaths did, occur in OC. We color-coded map densities on a scale ranging from blue (0 suicide events) to red ( $\geq 1.39/\text{km}^2$ ). We identified red regions as potential "hotspots" and examined the associated suicide victim (e.g., race and sex) and observed map (e.g., type of place and features) data to infer possible associations with suicide clusters.

### *Temporal clusters*

We defined a meaningful threshold of 33.9 suicides per yearly quarter as the 95th percentile of a normal distribution of suicide density per quarter using 24 three-month quarters between 2014 and 2019. We created a list of 70 potential "quarters" of three consecutive calendar months and identified the quarters exceeding the previously defined meaningful threshold of suicide density. We generated via Microsoft Excel a one-dimensional temporal map and labeled 3-month meaningful cluster(s). We then cross-referenced the suicides of these temporal clusters with those of spatial clusters to look for evidence of contagion at spatial clusters.

### *Statistics*

No formal cluster analysis was done. Clusters were described as above and identified manually by color on a map of OC. Associations with known geographic locations were identified manually.

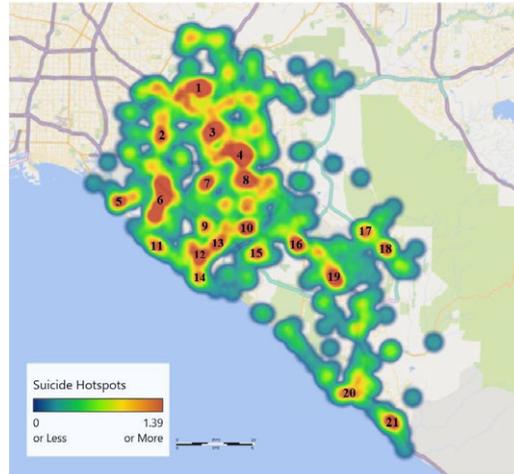
## **Results**

### *Data demographics*

Racial diversity ranged from 69% (425) White, 17% (103) Hispanic, 12% (72) Asian, and 3% (17) Black to 0% (1) unknown. Sex split into 83% (510) Male and 17% (108) Female. Age diversity ranged from 4% (24) 10–19, 20% (125) 20–29, 18% (112) 30–39, 15% (90) 40–49, 20% (121) 50–59, 15% (95) 60–69, 6% (34) 70–79, and 2% (12) 80–89 to 1% (5) 90–99. Mode ranged from 25% Asphyxia (157), 22% (139) Gunshot, 15% (91) Overdose, 14% (87) Fall, 7% (44) Vehicular-Traffic, 6% (38) Train, 3% (17) Cutting/Stabbing, 3% (21) Drowning, 1% (8) Fire, 1% (8) Blunt Force Trauma, 1% (4) Vehicular-Private Property, 0% (2) Other, and 0% (1) Religious Objections, to 0% (1) Broken Neck (via hanging).

### *Spatial clusters*

From all 618 total suicides, we identified 21 meaningful (greater than 1.39/km<sup>2</sup>) spatial clusters throughout OC for years 2014–2019, labeled as "All Suicide" (AS) clusters (Figure 1). By visual inspection, 31% of all suicides (192) appear to be in clusters. Of clusters (Table 1), 81% (17/21) included major transportation lines, 62% (13/21) included major tourist attractions or recreational parks, 48% (10/21) included educational centers or educational housing, 33% (7/21) included mobile home parks, 33% (7/21) included major bodies of water, 29% (6/21) included two-star hotels, 19% (4/21) contained



**Figure 1.** Meaningful spatial clusters of all suicides in Orange County (OC) years 2014–2019 (numbered). Out of 618 total suicides, we identified 21 meaningful spatial clusters (numbered here) throughout OC for the years 2014–2019. We arbitrarily defined a meaningful threshold of suicide density per area of  $1.39/\text{km}^2$  defined as the 95th percentile of a normal logarithmic scale distribution of suicide density per area in each ZIP Code Tabulated Area cumulated from all years. We color-coded map densities on a scale ranging from blue (0 suicide events) to red ( $\geq 1.39/\text{km}^2$ ). We identified red regions as potential “hotspots.”

medical centers, 19% (4/21) included downtown areas, 10% (2/21) included cemeteries, 5% (1/21) included an airport, and 5% (1/21) included LGBT Centers.

No age cohort produced meaningful spatial clusters (Figure 2). Hispanic (Figure 2a) and White (Figure 2b) were the only races with meaningful spatial clusters. The one Hispanic cluster coincided with AS cluster 8 and a Male Sex cluster. This was located in a Latino-predominant city and included its downtown area, two rail lines, the city zoo, the OC LGBT Pride office, the OC LGBT Center, and a two-star hotel. It was sandwiched between two high schools.

White race (15) and Male sex (17) (Figure 2c) produced the highest number of meaningful clusters. White clusters coincided with AS clusters #1, #3, #4–6, #10–13, and #16–21. All White clusters but #11, #13, and #20 coincide with Male clusters. Male clusters coincided with AS clusters #1–8, #10, #12, #14–19, #21, and the Hispanic cluster. All Male clusters but #2, #7, #8, #14, and #15 coincided with White clusters.

Female sex did not produce any meaningful clusters. Asphyxia was the only mode to produce a meaningful cluster (Figure 2d). This coincided with the lower (southern) portion of AS cluster 6, which was dumbbell-shaped. Hanging constituted 87% (136/157) of asphyxia suicides. Other types included suffocation and carbon monoxide poisoning. While suicide by train did not produce meaningful two-dimensional clusters, suicides clustered in a clearly identifiable linear form along two major rail lines, particularly at their junction passing by eight medical centers and a nationally recognized theme park (Figure 3). These corresponded with AS clusters #1, #3, #4, #8, #20, and #21. All but #8 coincide with White clusters. All but #20 coincide with Male clusters.

Of note, most of cluster 1’s suicides were near the rail line, which cut across the cluster center, or one of the two freeways. More than half of cluster 2 suicides centered around the two-star hotel and one of the mobile home parks. cluster 3 included a nationally recognized theme park. cluster 7 included eight mobile home parks. cluster 15 predominantly consisted of a major university and its campus housing. cluster 20 contained five major tourist attractions or recreational parks (Table 1).

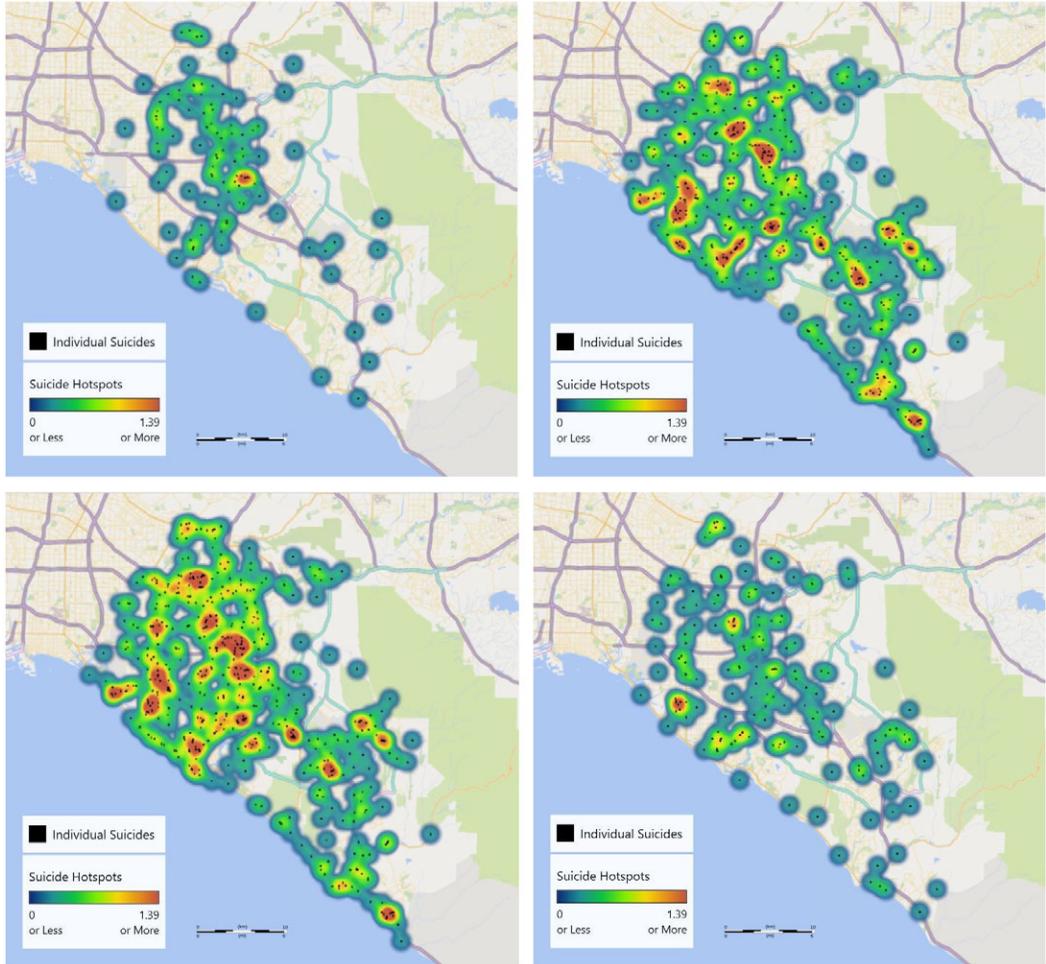
### Temporal clusters

Of 70 possible 3-month quarters, we identified one temporal cluster (Figure 4). This was July to September 2017. With 36 suicides, it yielded an average of 12 per month. Of those, 89% (32) were male, 63% (23) were White, 17% (6) were Hispanic, 11% (4) were Asian, and 8% (3) were Black. Of modes, 28%

**Table 1.** Frequency of features per cluster

Cluster	Medical center	Education center/ educational housing	Major transportation line	Mobile home park	Two-star hotel	Major tourist attraction/ recreational park	Downtown area	Convention/ major business center	Cemetery	LGBT center	Major body of water	Airport
1		1	2	1			1					
2		1	2	3	1							
3			2			1		1				
4	2		4			2						
5			1			1					1	
6-Upper		1	2	1								
6-Lower/Asphyxia		1		1					1			
7			2	8	1							
8/Hispanic		2	2		1	1	1			2		
9		2			2				1			
10		1	1					1				1
11			1			1	1				1	
12		1	1	2		1	1					
13		1	1	1	1	2						
14	1					2					2	
15		2										
16	2	3	2									
17			1			1						
18			1			1					1	
19	1		1			1						
20			3	1		5					1	
21			3		3	1					1	

Note. Depicted here are the numbers of features observed in each AS cluster. We defined education centers as high schools, colleges, or universities, educational housing as university campus housing, and major transportation lines as rail lines, freeways, highways, and major streets. Of these clusters, 81% (17/21) included major transportation lines, 62% (13/21) included major tourist attractions or recreational parks, 48% (10/21) included educational centers or educational housing, 33% (7/21) included mobile home parks, 33% (7/21) included major bodies of water, 29% (6/21) included two-star hotels, 19% (4/21) contained medical centers, 19% (4/21) included downtown areas, 10% (2/21) included cemeteries, 5% (1/21) included an airport, and 5% (1/21) included LGBT Centers.

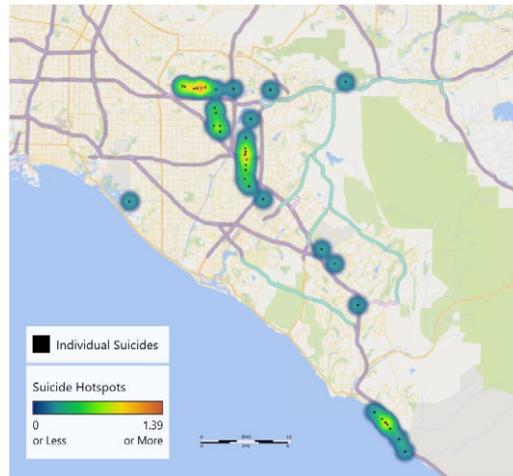


**Figure 2.** Meaningful spatial clusters of suicides by shared characteristics in Orange County (OC), years 2014–2019. Using our meaningful threshold of suicide density per area,  $1.39 \text{ km}^{-2}$ , meaningful clusters (depicted in red) of suicides with the following shared characteristics are identified. A) Hispanics in OC, years 2014–2019: one Hispanic cluster was identified. This coincided with AS cluster 8 and a Male Sex cluster. It was located in a Latino-predominant city and included its downtown area, two rail lines, the city zoo, the OC LGBT Pride office, the OC LGBT Center, and a two-star hotel. It was sandwiched between two high schools. B) Whites in OC, years 2014–2019: 15 White clusters were identified which coincided with AS clusters #1, #3, #4–6, #10–13, and #16–21. All White clusters but #11, #13, and #20 coincide with Male clusters. C) Males in OC, years 2014–2019: 17 Male clusters were identified, which coincided with AS clusters #1–8, #10, #12, #14–19, #21, and the Hispanic cluster. All Male clusters but #2, #7, #8, #14, and #15 coincided with White clusters. Female sex did not produce any meaningful clusters. D) Asphyxia in OC, years 2014–2019: Asphyxia was the only mode to produce a meaningful cluster. This coincided with the lower (southern) portion of AS cluster 6, which was dumbbell-shaped. Hanging constituted 87% (136/157) of asphyxia suicides. Other types included suffocation and carbon monoxide poisoning.

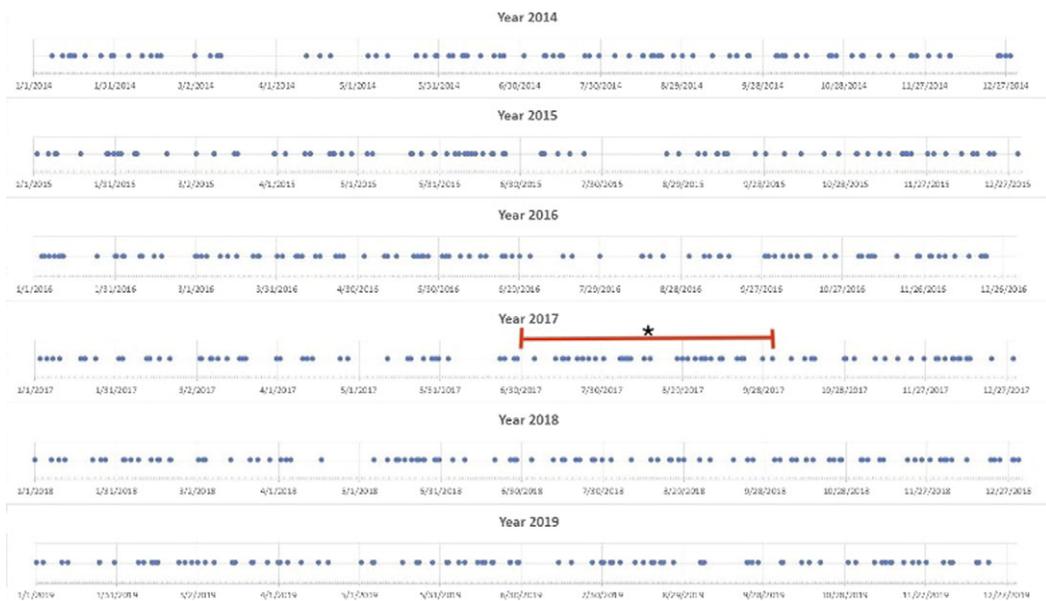
(10) were Asphyxia, 6% (2) were Drowning, 17% (6) were Fall, 17% (6) were Gunshot, 14% (5) were Vehicular-Traffic, 14% (5) were Overdose, 3% (1) were Train, and 3% (1) were Cutting/Stabbing. As these were evenly spaced across OC, we could not identify any temporospatial clusters. Numbers per city ranged from one (13 cities) to five (1 city). Two of similar age but distinct race and mode occurred at opposite sides of a freeway intersection in spatial cluster 13.

## Discussion

Through this study, we achieved a simple, novel method of identifying clinically meaningful clusters using the commonly available household computer program, Excel and its 3D Maps. To the best of our



**Figure 3.** Non-“meaningful” spatial suicide clusters of interest of suicides by train in Orange County years 2014–2019. While suicide by train did not produce meaningful two-dimensional clusters, suicides clustered in clearly identifiable linear form along two major rail lines, particularly at their junction passing by eight medical centers and a nationally recognized theme park. These corresponded with AS clusters #1, #3, #4, #8, #20, and #21. All but #8 coincide with White clusters. All but #20 coincide with Male clusters.



**Figure 4.** Meaningful temporal clusters in Orange County (OC) years 2014–2019. Using a meaningful threshold temporal density of 33.9 suicides per yearly quarter, we identified one 3-month-long temporal cluster. This was July to September 2017. As these suicides appear to have been evenly spaced across OC, we could not identify any temporospatial clusters.

knowledge, we are the first to demonstrate the utility and face validity of this readily available system for identifying meaningful suicide clusters, which has the potential to allow communities to more easily recognize and address suicide hotspots. We identified 21 AS “All Suicides” spatial clusters and clusters of White, Hispanic, Male, and Asphyxia categories. White, Hispanic, Male, and Asphyxia cases were more common and thus more likely to yield meaningful thresholds of suicide area densities. All these subgroup

clusters coincided within the AS clusters, giving validity to their meaning as suicide hotspots. Upon examination of the entirety of OC, only one temporal suicide cluster was found.

Three possible explanations for clusters exist: (a) suicide contagion (which may be supported by temporal clusters), (b) shared characteristics of people completing suicide within clusters, and (c) characteristics of the area that predispose to suicide.

Our temporal cluster did not localize to any part of OC. Internal or external forces may have affected all of OC (e.g., seasonal [Partonen et al., 2004; Shapiro, 2019], social [Bridge et al., 2020], political [The Trevor Project, 2016; The Trevor Project et al., 2017], and random chance). Given our temporal cluster displayed diversity in its locations, sex, race, age, and mode, we found no evidence of internal suicide contagion.

Shared characteristics in our spatial clusters provided face validity to our method and support the idea of shared characteristics among people and areas predisposing to clusters. These included high schools, colleges, and universities. Out of OC's 28 colleges (*Orange County, 2020b*) and universities and nine high schools (U.S. News, 2020), we found 43% (16) in clusters. Surprisingly, in attempting to relate this to the generally higher risk of suicide among teens and young adults, we found relatively few suicides in the ages of 10–19 though stable frequency distribution through the ages of 20–69. These numbers actually match the rates among different age groups nationwide in 2018 (AFSP, 2020), the year of many of these suicides. Due to low death rates among teenagers in this population, the high school proximity to a cluster may have posed as less of a suicide risk factor compared to college/university proximity. Perhaps these results are due to underreporting of teen deaths or protective factors instituted within these particular high schools, such as suicide prevention campaigns, although this is purely speculation. Despite the association between adolescents and young adults with clusters (Gould & Lake, 2013), we could identify no clusters by any age groups. This may be due to our viewing diameter or differences in data collection. cluster 15 appeared to underrepresent the suicide number compared with that reported by its campus police department over the same time period (Waaalen et al., 2019). Age groups among those aged 20–59 did appear to track along areas of higher suicide density, possibly because they constituted the majority of suicides.

Other common characteristics included mobile home parks, possibly due to socioeconomic stressors, although this is purely speculative. Preexisting addiction problems (Romney & Hua, 1995) which can increase suicidality through substance use could have led to increased socioeconomic stressors and increased likelihood of living in a mobile home park. While we found only 7% of OC's 245 mobile home parks (*Orange County, 2020a*) in clusters, not all mobile home parks may be clustered in such a way as to predispose to suicide clusters. The relevance of mobile home park presence may lie in their proximity to one another, as was seen in cluster 7, whose standout feature was eight mobile home parks. Similar explanations can be had for two-star hotels. Medical centers were common, likely due to higher presence of psychiatric patients. LGBT Pride Office and OC LGBT Center were present in cluster 8, the Hispanic cluster. LGBT population concentration around these service sites may explain higher suicide numbers, due to higher risk (The Trevor Project, 2021). Cemeteries may prime the suicidal person on the notions of death, dying, and glorification of death through acts of remembrance and paying of respects done at cemeteries. Freeways and rail lines may contribute to increased suicide risk by area both through increased people traffic as well as via increased availability of vehicular and train collision suicide modes. Similarly other area characteristics of high people traffic, such as convention centers, theme parks, beaches, tourist attractions, and airports, may predispose to higher suicide numbers.

We found examining suicide risk by area instead of by population gave very different results than OCHCA's study (Sanchez et al., 2019). Unlike their study which found risk per population greater in the affluent ZIP codes along the coast, ours found greater risk per *area* in areas further inland, with lower economic resources (*Orange County Register, 2008*), and of high population density (Brinkhoff, 2021). Similar to theirs which found greater risk among Whites and Males, ours found spatial clusters presenting the greatest risk for Whites and Males. However, the Hispanic cluster, cluster 8, represented a never-before identified area showing risk for Hispanic people as well. This can be explained by a higher population density. Perhaps it also represents a higher risk of being Male, Hispanic, and LGBT, given the

Hispanic cluster's LGBT community centers and coincidence with the Male cluster. Unfortunately, LGBT identity is not routinely recorded in county data, but it would be interesting to compare risk of such identity intersection in places such as Los Angeles, which has several LGBT Centers and Hispanic-predominant neighborhoods. This method provides a more pragmatic way to focus preventive resources into culturally specific areas with more suicide.

Furthermore, there were differences between the OCHCA study and ours, possibly related to differences in data collection. We used free, readily accessible data from the OC Coroner's office, because bureaucratic and financial barriers impeded our use of the same data sources as OCHCA's study, the State of California's Department of Public Health Death Statistical Master File and Vital Records Business Intelligence System. We also included an additional year, 2019, in our study. While OCHCA's study found that suicides were primarily conducted via firearms, this study found most suicides were carried out through Asphyxia, which is not explained by the additional study year, 2019. Spatial cluster 6 presented the greatest risk for asphyxia. The vast majority of these suicides were by hanging, specifically. These may coincide within a hidden temporal cluster not identified via our parameters, thus representing contagion. Lastly, while the prior study found greatest suicide numbers among the middle-aged, this study found equal or greater numbers from those in their 20s (20%) than from those in their 40s (18%) or 50s (20%).

We propose this alternative method, for its accessibility and ease of use, to identify and monitor potentially overlooked meaningful clusters, at-risk features, and possible targets of intervention. Possible interventions specific to OC are mental health resources and suicide prevention campaigns that target cluster sites as well as their shared characteristics such as high schools, colleges, and universities; mobile home parks; areas of increased people traffic such as trains, highways, downtown areas, convention centers, beach and theme parks, tourist attractions, business complexes, airports, and educational campuses; and medical centers. We warn against the use of suicide by ZIP code population data to determine allocation of suicide prevention resources. We recommend focusing the allocation of said resources to the regions with the *greatest number* of suicides, rather than to the ZIP code populations with the *greatest risk* of suicide, in order to have the greatest impact on suicide prevention. One caveat would be that consideration must be made regarding cultural barriers in providing resources to any high-risk population.

### *Limitations and future work*

"The sensitivity with which attempted suicides are identified may be much greater during a community-identified cluster of completed suicides than otherwise.... Statistically verified clustering of suicide attempts [alone] tells us little about the statistical probability of an associated community-identified cluster of completed suicides." (O'Carroll & Mercy, 1990). Even statistically significant clusters alone do not prove contagion (Suicide Contagion: A Systematic Review of Definitions and Research Utility). While our identification of cluster 15, corresponding with a community-identified temporospatial hotspot studied elsewhere (Waaen et al., 2019), does provide face validity to our method, it would be inappropriate to rule out a cluster due to lack of statistical significance, particularly if the community has identified it first.

Further limitations of this study include observer bias affecting which shared characteristics of clusters were identified. Future studies should aim to account for such bias with a more structured approach. While the mode of Train did not show any hotspots, this may be due to its linear shape, which is not as easily captured by our circular model of capturing suicide density by area. Other methods of identifying clusters along linear pathways should be considered (Too et al., 2017). This study was limited by the magnification parameters at which we searched for clusters. The magnification we chose was somewhat arbitrary. Examination of spatial and temporal clusters at different viewing diameters may be of additional utility. For example, cluster 15's suicides did not show up in the temporal cluster we identified when looking on the scale of OC as a whole. By studying individual clusters such as cluster 15 on their

own, however, it may be possible to identify temporal clusters on preset time blocks ranging from weeks to months to years.

On first inspection, it seems strange that 31 and 7% of suicides were located in spatial and temporal clusters, respectively, defined to have a density threshold of the 95th percentile. This is both a limitation and advantage of the study method. Depending on viewing diameter, clusters may aggregate, merging into others and potentially leading to overrepresentation of cluster size. For future use, this may be easily corrected by altering the viewing diameter. At the same time, heat map technology transcends the limitations of other scan statistics, which often operate only in either circular/elliptical or linear form, allowing for irregular shapes and interesting connections between locations of relevance, such as circular centers and linear rail and roadways.

For simplicity, we used ZCTA distribution to define the 95th percentile threshold suicide density as opposed to a census block group. Locations came with ZIP codes, which made it easy to use ZCTA's. This presented its own room for error, however, as ZCTA's, like census block groups, come in varying shapes and sizes. Had an overlying grid system been as easily accessible to the average layman, we would have used it to ensure uniform areas from which to draw our distribution of densities.

Another study limitation is the Coroner's data supplied addresses (or nearest addresses) of events, not exact global positioning system coordinates. Only 8% appeared to have involved a vehicle, suggesting several near roads and railways occurred indoors. Upon closer inspection, a fair number appear to have been in parked automobiles and hotels. An important aim of a future study might be to examine more closely characteristics of exact locations along roads and railways possibly making them more dangerous for suicide.

With any statistical study, there is the potential for random error. For simplicity, we chose not to use an outside control population because every population is unique. Instead, we examined the same population over the course of several years. It is possible that certain clusters may appear randomly as part of a normal distribution without relationship to high-risk locations or time periods. Regardless, it is always important to monitor any clusters in ongoing real time to monitor for persistent or recurrent patterns (Waaalen et al., 2019), which is why a widely and readily accessible method such as ours is so important. In this study, we valued ease of use over specificity with the assumption that increased access would promote much needed long-term cluster monitoring.

## Conclusions

Through this proof of principle study, we achieved a simple, novel method of identifying suicide clusters using the commonly available household computer program, Excel. Shared characteristics in these spatial clusters provided face validity to our method and support the idea of shared characteristics among people and areas predisposing to clusters. We found that examining suicide risk by area instead of by population gave very different results than OCHCA's study. We propose this alternative method, for its accessibility and ease of use, to identify and monitor in real-time potentially overlooked meaningful geographic or temporal clusters, at-risk features, and possible targets of intervention.

**Acknowledgment.** We give special thanks to Hollis Krug, MD, for language and technical editing.

**Open peer review.** To view the open peer review materials for this article, please visit <http://doi.org/10.1017/exp.2023.2>.

**Data availability statement.** Given the sensitivity of information, data and analysis may be available on request only.

**Funding statement.** This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

**Conflict of interest.** The authors declare none.

## References

- AFSP. (2020). *Suicide statistics*. AFSP. <https://afsp.org/suicide-statistics/> (Retrieved 1 January 2021).
- Bandura, A.** (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, **84**, 191–215. <https://doi.org/10.1037/0033-295X.84.2.191>

- Boyce, N. (2011). Suicide clusters: The undiscovered country. *The Lancet*, **378**, 1452. [https://doi.org/10.1016/S0140-6736\(11\)61634-6](https://doi.org/10.1016/S0140-6736(11)61634-6)
- Bridge, J. A., Greenhouse, J. B., Ruch, D., Stevens, J., Ackerman, J., Sheftall, A. H., Horowitz, L. M., Kelleher, K. J., & Campo, J. V. (2020). Association between the release of Netflix's *13 Reasons Why* and suicide rates in the United States: An interrupted time series analysis. *Journal of the American Academy of Child and Adolescent Psychiatry*, **59**, 236–243. <https://doi.org/10.1016/j.jaac.2019.04.020>
- Brinkhoff, T. (2021). *Orange (County, California, USA)—Population statistics, charts, map and location*. City Population. [https://www.citypopulation.de/en/usa/california/06059\\_\\_orange/](https://www.citypopulation.de/en/usa/california/06059__orange/) (Retrieved 19 May 2021).
- Cheung, Y. T., Spittal, M. J., Williamson, M. K., Tung, S. J., & Pirkis, J. (2013). Application of scan statistics to detect suicide clusters in Australia. *PLoS One*, **8**, e54168. <https://doi.org/10.1371/journal.pone.0054168>
- Durkheim, E. (1897). *Suicide*. Routledge.
- Gould MS, Lake AM. (2013). *The contagion of suicidal behavior*. National Academies Press. <https://www.ncbi.nlm.nih.gov/books/NBK207262/> (Retrieved 19 May 2021).
- Hill, N. T. M., Too, L. S., Spittal, M. J., & Robinson, J. (2020). Understanding the characteristics and mechanisms underlying suicide clusters in Australian youth: A comparison of cluster detection methods. *Epidemiology and Psychiatric Sciences*, **29**, e151. <https://doi.org/10.1017/S2045796020000645>
- Kerkhof, A. (2003). Railway suicide: Who is responsible?. *Crisis*, **24**, 47–48.
- Kulldorff, M. (1997). A spatial scan statistic. *Communications in Statistics—Theory and Methods*, **26**, 1481–1496. <https://doi.org/10.1080/03610929708831995>
- O'Carroll, P. W., & Mercy, J. A. (1990). Responding to community-identified suicide clusters: Statistical verification of the cluster is not the primary issue. *American Journal of Epidemiology*, **132**, S196–S202. <https://doi.org/10.1093/oxfordjournals.aje.a115783>
- O'Carroll, P. W., Mercy, J. A., & Steward, J. A. (1988). CDC recommendations for a community plan for the prevention and containment of suicide clusters. *Morbidity and Mortality Weekly Report*, **37**, 1–12. <https://www.cdc.gov/mmwr/preview/mmwhtml/00001755.htm> (Retrieved 16 March 2020).
- Orange County. (2020a). *245 mobile home parks in Orange County, CA*. MHVillage. <https://www.mhvillage.com/parks/ca/orange-county> (Retrieved 1 January 2021).
- Orange County. (2020b). *California—Colleges & universities*. <https://www.ocgov.com/about/infooc/links/oc/colleges> (Retrieved 1 January 2021).
- Orange County Register. (2008). *Poorest ZIPs in O.C. and California*. <https://www.ocregister.com/2008/01/01/poorest-zips-in-oc-and-california/> (Retrieved 19 May 2021).
- Partonen, T., Haukka, J., Nevanlinna, H., & Lönnqvist, J. (2004). Analysis of the seasonal pattern in suicide. *Journal of Affective Disorders*, **81**, 133–139. [https://doi.org/10.1016/S0165-0327\(03\)00137-X](https://doi.org/10.1016/S0165-0327(03)00137-X)
- Pirkis, J., Too, L. S., Spittal, M. J., Kryszynska, K., Robinson, J., & Cheung, Y. T. (2015). Interventions to reduce suicides at suicide hotspots: A systematic review and meta-analysis. *The Lancet Psychiatry*, **2**, 994–1001. [https://doi.org/10.1016/S2215-0366\(15\)00266-7](https://doi.org/10.1016/S2215-0366(15)00266-7)
- Romney, L., & Hua, T. (1995). Methamphetamine labs leave growing toxic legacy: Crime: Drug “kitchens” in motels, rentals pose serious health threats, and cleanup costs outstrip resources. *Los Angeles Times*. <https://www.latimes.com/archives/la-xpm-1995-11-27-mn-7660-story.html> (Retrieved 2 January 2021).
- Sanchez, R. et al. (2019). *Suicide deaths in Orange County, California (2014–2018)*. Orange County Health Care Agency. [www.ochealthinfo.com/suicidereport](http://www.ochealthinfo.com/suicidereport). (Retrieved 15 March 2020).
- Shapiro, M. (2019). Suicide rates spike in spring, not winter, Dome. <https://www.hopkinsmedicine.org/news/articles/suicide-rates-spike-in-spring-not-winter> (Retrieved 19 May 2021).
- Tango, T., & Takahashi, K. (2005). A flexibly shaped spatial scan statistic for detecting clusters. *International Journal of Health Geographics*, **4**, 11. <https://doi.org/10.1186/1476-072X-4-11>
- The Trevor Project. (2016). *Crisis contacts from youth to The Trevor Project surge immediately following the election*. The Trevor Project. [https://www.thetrevorproject.org/trvr\\_press/crisis-contacts-from-youth-to-the-trevor-project-surge-immediately-following-the-election/](https://www.thetrevorproject.org/trvr_press/crisis-contacts-from-youth-to-the-trevor-project-surge-immediately-following-the-election/) (Retrieved 19 May 2021).
- The Trevor Project. (2021). *Preventing suicide: Facts about suicide*. The Trevor Project. <https://www.thetrevorproject.org/resources/preventing-suicide/facts-about-suicide/> (Retrieved 24 September 2018).
- The Trevor Project, Lunn, S. A., & Daniel, R. (2017). *Spike in crisis contacts related to anti-trans rhetoric*. The Trevor Project. [https://www.thetrevorproject.org/trvr\\_press/spike-in-crisis-contacts-related-to-anti-trans-rhetoric/](https://www.thetrevorproject.org/trvr_press/spike-in-crisis-contacts-related-to-anti-trans-rhetoric/) (Retrieved 19 May 2021).
- Too, L. S., Pirkis, J., Milner, A., Bugeja, L., & Spittal, M. J. (2017). Railway suicide clusters: How common are they and what predicts them? *Injury Prevention*, **23**, 328–333. <https://doi.org/10.1136/injuryprev-2016-042029>
- Too, L. S., Pirkis, J., Milner, A., Robinson, J., & Spittal, M. J. (2019). Clusters of suicidal events among young people: Do clusters from one time period predict later clusters?. *Suicide and Life-Threatening Behavior*, **49**, 561–571. <https://doi.org/10.1111/sltb.12460>

- U.S. News. (2020). *Orange County Department of Education public schools | California | Best High Schools*. U.S. News. <https://www.usnews.com/education/best-high-schools/california/districts/orange-county-department-of-education-104617> (Retrieved 1 January 2021).
- Waaen, A., Bera, K., & Bera, R.** (2020). Suicide hotspots, interventions, and future areas of work at a Californian university. *Death Studies*, **44**, 1–9. <https://doi.org/10.1080/07481187.2019.1595221>
- Walker, T.** (2019). *Suicide rate in Orange County reached highest point of the century last year*. *Orange County Register*. <https://www.ocregister.com/2019/11/05/suicide-rate-in-orange-county-reached-highest-point-of-the-century-last-year/> (Retrieved 15 March 2020).

# Peer Reviews

**Reviewing editor:** Prof. Emanuele Frontoni

University of Macerata, Information Engineering Department - DII, Macerata, Italy, 62100

Minor revisions requested.

doi:10.1017/exp.2023.2.pr1

## Review 1: Identifying Spatial and Temporal Suicide clusters in a Californian County

**Reviewer:** Prof. Marina Paolanti 

Università Politecnica delle Marche, Ingegneria dell'informazione, Ancona, Italy, 60121

Date of review: 30 August 2022

© The Author(s), 2023. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

**Conflict of interest statement.** Reviewer declares none.

### Comment

Comments to the Author: The paper deals with an extremely interesting topic and, as can be inferred from the analysis produced, presents itself as a first attempt to make a very uneven literature organic. In detail, the paper claims to present a real-time cluster detection required for necessary targeted interventions in the midst of our growing national suicide epidemic.

The problem addressed and the solution proposed look interesting, the paper is well structured but not well organized. It could be useful to provide a thoughtful discussion of how the paper serves as a foundation for future research with a dedicated section.

In particular, the introduction should give an overall review of the paper (including contributions, approach, results, advantages) while a "related work" section clearly outlines the perspective on this field with the discussion of all the references cited. In the same related work section, the authors can report other references in which these algorithms are applied. This could prevent a clear understanding of what should be the contributions of this paper, starting from its novelties.

The discussions should provide more explanations and particulars of the underlying meaning of this research, noting possible implications in other areas of study, and exploring possible improvements.

The English is correct, but more attention could be addressed to text errors and stylistic inaccuracies.

The section Conclusion should be expanded and improved by introducing more explanations and particulars of the underlying meaning of this research, noting possible implications in other areas of study, as well as results should be mainly discussed in relation to previous researches.

### Score Card

#### Presentation



Is the article written in clear and proper English? (30%)

3/5

Is the data presented in the most useful manner? (40%)

3/5

Does the paper cite relevant and related articles appropriately? (30%)

3/5

## Context



- Does the title suitably represent the article? (25%) 3/5
- Does the abstract correctly embody the content of the article? (25%) 2/5
- Does the introduction give appropriate context? (25%) 2/5
- Is the objective of the experiment clearly defined? (25%) 3/5

## Analysis



- Does the discussion adequately interpret the results presented? (40%) 2/5
- Is the conclusion consistent with the results and discussion? (40%) 3/5
- Are the limitations of the experiment as well as the contributions of the experiment clearly outlined? (20%) 2/5

## Review 2: Identifying Spatial and Temporal Suicide clusters in a Californian County

Reviewer: Dr. Riccardo Rosati 

Università Politecnica delle Marche Facoltà di Ingegneria, Information Engineering, via Breccie Bianche n.12, Ancona, Ancona, Italy, 60131

Date of review: 03 December 2022

© The Author(s), 2023. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

**Conflict of interest statement.** Reviewer declares none.

### Comment

Comments to the Author: In this paper, the authors investigate the suicide phenomenon in Californian country identifying spatial and temporal clusters. Clusters were detected using readily accessible software, i.e. Excel 3D Maps, processing a dataset related to the years 2014-2019.

The paper is well written and quite clear in all the parts, even if the organisation could be improved.

My major remarks are as follows:

- I suggest the authors strongly emphasise the major contributions of the work with respect to related literature.

- Why do the authors describe the employed dataset (i.e. Data Demographics) in Results section? In my opinion, it would be more appropriate to introduce the dataset earlier in the paper. An option might be to redefine the section “Methods” as “Materials & Methods” section and insert this description there.

- What is the rationale behind choosing a density threshold of the 95th percentile? Is it defined in state-of-the-art? Moreover, even if already implemented in the software used, the authors should provide more details about the cluster analysis performed in this study.

- It is not clear to me how the authors retrieved the features in Table 1 and how they aggregated this information. For instance, it seems that “major transportation line” is not included in the original dataset. How did they choose which feature to consider for the analysis?

- Section “Limitations” can be renamed “Limitations and Future work”, given the interesting insights the authors provide as future work to address the limitations of the present study.

### Score Card

#### Presentation



Is the article written in clear and proper English? (30%)

5/5

Is the data presented in the most useful manner? (40%)

3/5

Does the paper cite relevant and related articles appropriately? (30%)

4/5

#### Context



Does the title suitably represent the article? (25%)

5/5

Does the abstract correctly embody the content of the article? (25%)

4/5

Does the introduction give appropriate context? (25%)

4/5

Is the objective of the experiment clearly defined? (25%)

3/5

Analysis



Does the discussion adequately interpret the results presented? (40%)

---

5/5

Is the conclusion consistent with the results and discussion? (40%)

---

5/5

Are the limitations of the experiment as well as the contributions of the experiment clearly outlined? (20%)

---

5/5