Clusters of suicidal events among young people: Do clusters from one time period predict later clusters?

Accepted in Suicide and Life Threatening Behavior (Jan 2018)

LAY SAN TOO¹, PHD, JANE PIRKIS¹, PHD, ALLISON MILNER², PHD, JO ROBINSON³, PHD, and MATTHEW J. SPITTLAL¹, PHD

¹Centre for Mental Health, Melbourne School of Population and Global Health, University of Melbourne, Parkville Victoria 3052 Australia

²Centre for Health Equity, Melbourne School of Population and Global Health, University of Melbourne, Parkville Victoria 3052 Australia

³Orygen, The National Centre of Excellence in Youth Mental Health, Parkville, VIC 3052, Australia

Address correspondence to
Associate Professor Matthew Spittal
Centre for Mental Health,
Melbourne School of Population and Global Health,
University of Melbourne,
Parkville Victoria 3052 Australia
E-mail: m.spittal@unimelb.edu.au
Tel: +61 3 903 58230
ABSTRACT

Objective

We sought to compare clusters of suicidal events between two different time periods and examine the extent to which earlier clusters predict later clusters.

Method

We included data on suicides and suicide attempts from New South Wales between July 2001 and June 2012 and Western Australia between January 2000 and December 2011. Suicide attempts included admissions to hospital for deliberate self-harm and suicides were deaths due to deliberate self-harm. We combined data on suicides and suicide attempts and grouped them into two equal time periods. We detected clusters in each period using Poisson discrete scan statistics. We estimated the predictive values of earlier clusters on later clusters.

Results

The results showed that clusters from earlier time period had a moderate power (36%) in predicting later clusters. During later time period, some additional cluster areas (14%) were found and some earlier cluster areas subsided (64%).

Conclusions

Historical clusters predict 36% of the subsequent clusters, which is probably not sufficient for targeting interventions. Our study highlights the need for other strategies to detect emerging clusters, for example up to date data.

Keywords

Suicide, suicide attempt, cluster, prediction
INTRODUCTION

A suicide cluster is typically defined by a greater number of suicides (and/or suicide attempts) occurring in space and time than would be expected given the underlying size of the population and the rate in that population (Niedzwiedz, Haw, Hawton, & Platt, 2014; O’Carroll, Mercy, & Steward, 1988). Although suicide clusters are relatively rare (Niedzwiedz, Haw, Hawton, & Platt, 2014), their actual or perceived presence results in substantial concern in the communities where they occur. This is because when suicide clusters occur they can have significant adverse consequences for communities, including the risk of subsequent suicides through imitation, complicated grief responses, and ongoing trauma (Gould, Wallenstein, & Davidson, 1989; Haw, Hawton, Niedzwiedz, & Platt, 2013). Early identification of suicide clusters is therefore important to allow prevention efforts to be initiated in cluster areas to potentially avert these consequences.

Previous research on suicide clusters has examined the extent at which clusters exist and the risk factors contributing to a suicide occurring in a cluster. For example, Cheung, Spittal, Williamson, Tung, and Pirkis (2013) examined spatial-temporal suicide clusters occurring in Australia between 2004 and 2008. They identified 15 suicide clusters which accounted for 2.4% of all suicides. They also found that the risk of a suicide occurring in a cluster was high for suicides by people from Indigenous background and people who lived in remote areas (Cheung, Spittal, Williamson, Tung, & Pirkis, 2014). Jones et al. (2013) examined spatial-temporal suicide clusters in Wales for 2000-2009 and reported that less than 1% of suicides in people aged 15-34 years occurred as part of a cluster. In a 2016 study, we found that suicides among young people were more likely to occur in a cluster (five clusters, representing 5.6% of youth suicides) compared to suicides by adults (seven clusters, representing 2.3% of adult suicides) during a three-year period (Robinson, Too, Pirkis, & Spittal, 2016). We also found that for both young people and adults,
suicide clusters occurred more frequently among people of Indigenous origin than among non-Indigenous people. In another study, we found ten suicide clusters among unemployed persons, representing 1.7% of all unemployed suicides (Milner, Too, & Spittal, 2017).

In 2012, Larkin and Beautrais conducted a particularly novel study in this area, looking at the phenomenon of ‘echo clusters’ in New Zealand. These clusters occur in the same location as an initial cluster but are separated in time. They detected nine locations where 10 ‘echo clusters’ occurred; these clusters occurred on an average of 7.6 years after the initial cluster.

One limitation of previous studies is that they have focused almost entirely on identifying clusters of suicides, rather than using the broader definition that also includes suicide attempts. Suicide attempts are frequently mentioned in community concerns about suicide clusters (Robbins & Conroy, 1983) and more generally, are strongly associated with risk of suicide (Hawton & Harriss, 2007; Turecki & Brent, 2016). In our previous study, we identified clusters of suicides and clusters of suicide attempts separately (Too, Pirkis, Milner, & Spittal, 2016). We found there were more suicide attempt clusters (n = 7) than suicide clusters (n = 2) and these suicide clusters were located within a larger but later suicide attempt cluster. We also found that those suicide attempts that occurred in areas of low socioeconomic status, areas with a high proportion of Indigenous people, or areas with a high proportion of people who had moved in the last year were more likely to be part of a cluster. This suggests that both clusters of suicides and clusters of suicide attempts should be considered in study on cluster detection for appropriate responses and preventive purposes.

To address this gap in knowledge, we included both suicides and suicide attempts and sought to identify clusters among people aged ≤25 years. We were particularly interested in ‘echo clusters’ so we considered two time periods and sought to determine whether clusters detected in
the first period predicted clusters in the subsequent period. We restricted our analysis to people aged \( \leq 25 \) years because this group is at greatest risk of being in a cluster, as indicated above. To identify clusters, we used the scan statistic, which is a statistical method for detecting a greater than expected number of events in specific spatial areas and particular time periods. This method has been commonly used to detect clusters of communicable diseases (D'Aignaux et al., 2002; Shea, Kammerer, Winston, Navin, & Horsburgh, 2014) and suicides (Gould, Kleinman, Lake, Forman, & Midle, 2014; Jones et al., 2013; Qi, Hu, Page, & Tong, 2012).

**METHOD**

*Study sample and data*

We used data on suicides and suicide attempts from New South Wales (NSW) and Western Australia (WA), two states in Australia with well-established linkage systems. Suicides were defined as any deaths where the primary cause of death was intentional self-harm [ICD-10 codes X60-X84 (World Health Organisation, 2016)]. Suicide attempts were defined as any admission to hospital where intentional self-harm was coded as an external cause of injury (i.e., the same ICD-10 codes), irrespective of whether or not there was intent to die.

The data linkage agency in each state provided us with individual-level data on suicides from each state’s death registry and on hospital admissions for suicide attempts from either the Admitted Patients Data Collection (in NSW) or the Hospital Morbidity Data Collection (in WA). From these four datasets we were able to extract a core set of variables. These were date of death/admission, age at the time of death/admission, sex, method of attempt, ICD-10 code for cause of death/injury, and geographical area of usual residence. Unfortunately, the two states used different systems for coding geographic information; NSW used statistical area level 2 (SA2),
whereas WA used the older statistical local area (SLA). SA2 is an Australian Statistical Geography Standard (ASGS) defined area while SLA is an Australian Standard Geographical Classification (ASGC) defined area (Australian Bureau of Statistics, 2015). The ASGS replaced the ASGC since July 2011 and has been used for the 2011 Census. In general, SA2s have an average population of about 10,000 persons and SLAs have an average population of about 16,000 persons (Australian Bureau of Statistics, 2012). Our analysis was conducted separately in each state, minimizing the potential for this difference to influence the results.

**Inclusion and exclusion criteria**

Based on the availability of suicide and suicide attempt data in each state, we used NSW data from 1 July 2001 to 30 June 2012 and WA data from 1 January 2000 to 31 December 2011. We included only observations from people aged 15 to 24 years at the time of their suicide or suicide attempt. Figure 1 shows the number of suicides and suicide attempts that we identified from each state, the number of observations that were excluded (and the reason for these exclusions), and the final number of observations that we used for analysis. We grouped the datasets into two equal time periods based on day, month and year of event. These were 1 July 2001 to 31 December 2006 (period one) and 1 January 2007 to 30 June 2012 (period two) in NSW and 1 January 2000 to 31 December 2005 (period one) and 1 January 2006 to 31 December 2011 (period two) in WA.

[Figure 1]
Population and geographical data

We obtained population estimates among people aged 15 to 24 years for each SA2 in NSW from the 2011 Australian Bureau of Statistics (ABS) Census data. Similarly, we obtained population estimates for each SLA in WA from the 2006 Census data. Ideally, 2006 data would have been used in the analyses for both states because 2006 was the study midpoint, but we used 2011 data for the NSW analyses because SA2 classification was not available in the 2006 Census data. We also obtained area data on socioeconomic status (measured using the Index of Relative Socioeconomic Disadvantage) in NSW and WA from the corresponding Census data. The Index of Relative Socio-economic Disadvantage summarises the socioeconomic status of an area based on a range of information from the Census [e.g., the proportion of people in the area with poor English language skills, low-skilled occupations, separated/divorced, disabled or living in a household with low household equivalised income (Australian Bureau of Statistics, 2013)].

We computed geographical coordinates of population weighted centroids using the mean centroid algorithm in ArcGIS 10.3.1. In both states this was done using the smallest available geographic unit – in NSW this was mesh blocks and in WA this was collection districts, which are an amalgamation of mesh blocks. We calculated the centroids for each SA2/SLA using estimates of the population size and the digital boundaries of mesh blocks/collection districts.

Statistical analysis

We combined suicides and suicide attempts into a single variable, referred to as suicidal events. We calculated the annual rates of suicidal events in each state in the two time periods using the
population estimates in 2011 (NSW) and 2006 (WA) as the denominators. We profiled suicidal events in terms of sex and method used.

To identify spatial-temporal clusters of high relative rates of suicidal events, we performed Poisson discrete scan statistics using SaTScan version 9.4.1 (Kulldorff, 2015). We performed separate analyses in each state and in each time period. We did this by entering into SaTScan the number of suicidal events per day in each SA2/SLA, the relative socioeconomic disadvantage of each SA2/SLA, the population estimates of each SA2/SLA and the centroids of each SA2/SLA. We set the time window from a minimum of 1 day to a maximum of 90 days and the maximum spatial window at commonly used 10% of the population at risk (Jones et al., 2013). The shape of the spatial scan window was fixed as circular. These settings mean that a set of cylinders was used to scan the space-time region where the base of the cylinder defines the area of the potential cluster and its height defines the time interval of the cluster. Because suicides and suicide clusters tend to occur in areas of low socioeconomic status (Exeter & Boyle, 2007; Too, Pirkis, Milner, & Spittal, 2016; Rehkopf & Buka, 2006), we treated area-level socioeconomic status as a confounder and controlled for this in our cluster detection analyses. We used Monte Carlo stimulation to assess the significance level of any potential clustered that we detected (Kulldorff, 1997). Clusters were included if their p value was lower than 0.10, signifying the presence of a possible cluster. We created maps using ArcGIS to show the locations of clusters in the two time periods.

The task of detecting clusters which include suicide attempts is potentially more complicated than identifying clusters of suicides only. This is because, unlike death, a single person can have multiple suicide attempts and potentially die by suicide. We needed to be sure that any identified clusters comprised suicidal events by different people, not multiple events by the same person. We addressed this using a two-step process. In the first step, we used all included
data to identify a set of candidate clusters. In the second step, we examined each candidate cluster to identify any person who contributed multiple observations to the cluster, and then retained one randomly-selected observation within the cluster for that person. We then re-identified the clusters multiple times, based on the maximum number of observations from any person within a cluster. Each time we randomly selected a different observation for that person. Our final set of clusters was therefore those clusters that were consistently identified after the random removal of repeated observations [see the details of this procedure from our previous study (Too, Pirkis, Milner, & Spittal, 2016)].

Following detection of clusters, we assessed the predictive ability of an area in period one having a cluster (yes or no) on the occurrence of a cluster in that area in period two (yes or no). We assessed sensitivity, specificity, positive predictive values (PPVs) and negative predictive values (NPVs) treating period two data as the gold standard and period one data as the test. This analysis was conducted using Stata 14.2.

*Sensitivity analyses*

We repeated the analyses using a different time window (a minimum of one month to a maximum of 12 months) as there is no consensus regarding how long a cluster usually lasts. We also repeated the analyses using a different time frame, considering there might be carry-over clustering effects (i.e. a cluster may crossover the end of period one into the beginning of period two). To do this, we excluded the three-month data in between the two periods. This meant we included NSW data from 1 July 2001 to 15 November 2006 (period one) and 16 February 2007 to 30 June 2012 (period two) and WA data from 1 January 2000 to 15 November 2005 (period one) and 16 February 2006
to 31 December 2011 (period two). Finally, we undertook analyses including suicides only using windows of three and 12 months.

RESULTS

Characteristics of suicidal events

During the study period there were 718 (2.6%) suicides and 27,286 (97.4%) suicide attempts in people aged 15 to 24 years in NSW, and 328 (3.3%) suicides and 9,550 (96.7%) suicide attempts in WA. The annual rate of suicidal events (suicides and suicide attempts, per 100,000 persons) was 285.7 in NSW and 299.2 in WA (Table 1). Within these states, approximately 70% of suicidal events involved females. Poisoning by drugs was the most frequently used method (70% in NSW, 67% in WA). This was followed by cutting or piercing (21% in NSW, 22% in WA).

[Table 1]

Cluster detection

In NSW, during period one, we initially identified ten candidate spatial-temporal clusters of high relative risk. However, many individuals (56/265, 21%) had multiple suicidal events within a cluster (up to 13 suicidal events per person per cluster). We thus randomly selected one suicidal event per person within a cluster and repeated the cluster detection analysis 13 times, each time randomly selecting a different observation. As a result, we found six clusters consistently identified across the analyses (Table 2). These clusters contained 3% of all suicides and suicide attempts (432/14,370), with the cluster size ranging from 21 to 116 suicidal events. During period two, we initially detected 14 candidate spatial-temporal clusters with up to 15 suicidal events per person
per cluster. After repeating the analysis 15 times, each time randomly selecting just one of the repeated suicidal events, we detected ten clusters. We found 2.0% (269/13,381) of suicides and suicide attempts occurred within these ten clusters. The cluster sizes ranged from 6 to 52 events.

In WA, after undertaking analysis to identify a candidate set of clusters and repeating the analysis to account for repeated suicidal events, we identified three clusters in period one and no clusters in period two. The clusters in period one consisted of 1.8% (57/3,198) of all suicides and suicide attempts and ranged in size from 9 to 28 events.

[Table 2]

Location of clusters

Figure 2 shows the location of the clusters we detected. We have separated out those areas where a cluster appeared in period one only, period two only, and in both periods. In NSW, some cluster areas persisted, some cluster areas subsided, and some additional cluster areas were found in period two. We could not make this comparison in WA because no clusters were detected in period two.

[Figure 2]

Diagnostic properties of an earlier cluster predicting a later cluster

The PPV (the proportion of areas in period two where there was a cluster given a cluster in the same area in period one) was 36%, meaning that 64% of areas identified in period one were false positives (Table 3). The NPV (the proportion of areas in period two there was no cluster given no cluster in the same area in period one) was 86%. Sensitivity (among those areas where there was a cluster in period two, the proportion of areas that were a cluster in period one) was 59% and
specificity was 70% (among those areas where there was not a cluster in period two, the proportion of areas that were not a cluster in period one). Thus, using the presence or absence of a cluster as an indicator of a cluster in a later period was successful 36% of the time, and this level is probably not sufficient to be used reliably for ruling in areas of risk. We also found 14% of new cluster areas in period two and 64% of the earlier cluster areas subsided in the later period.

[Table 3]

Findings from the sensitivity analyses

In the analyses with time window of one month to twelve months, we consistently identified 10 clusters in period one (10.2%, 1,439/14,049 events) and five clusters (5.9%, 777/13,174 events) in period two in NSW while identified six clusters in period one (12.7%, 393/3,095 events) and four clusters (2.9%, 191/6,552 events) in period two in WA. We found that the PPV was 41%, the NPV was 87%, sensitivity was 68%, and specificity was 69%.

In the analyses that excluded three months in-between two periods, in NSW, six clusters were consistently detected in period one (3.1%, 432/14,031 events) and seven clusters (1.6%, 205/12,916 events) in period two. In WA, three clusters were consistently detected in period one (1.8%, 57/3,100 events) and two clusters (1.0%, 64/6,528 events) in period two. We found that the PPV was 29%, the NPV was 79%, sensitivity was 44%, and specificity was 66%. These findings were similar to the findings from the main analyses, indicating prediction of a cluster could not reliably rely on the presence or absence of an earlier cluster.

Finally, restricting the analysis to the detection of suicide clusters only with a three month temporal window, we identified one cluster in period one in NSW (3.0%, 12/397 events) and no clusters in period two. In WA, we identified no suicide clusters in either period. The PPV was
therefore 0% and the NPV 100%. Sensitivity and specificity were 0% and 92%, respectively. Essentially the same results were observed using a longer 12 month temporal window.

**DISCUSSION**

Our study examined whether clusters of suicides and suicide attempts among young people detected in an earlier time period predicted clusters in a later time period. We found that earlier clusters predicted 36% of the later clusters occurring in the same areas. Ours is an area-based measure of risk, and stands in contrast with other scales and instruments which typically measure risk using data gathered at the individual level. For this reason, it is difficult to draw firm conclusions about whether the values we observed represents a useful level of prediction. Our view is that it represents only moderate level of prediction, and that it is too low to be useful for directing resources toward prevention efforts. Such prediction is difficult probably because suicidal events are rare and clusters of these events are even rarer. The power of our prediction was probably weakened by including data from only two Australian states which meant that we were only able to include a small number of clusters in our analysis. Interestingly, our findings are not consistent with an earlier study which found nine locations where ten clusters recurred over a 18-year period (Larkin & Beautrais, 2012). When we restricted our analysis to the examination of suicide clusters only (as Larkin and Beautrais did, although their study used data from all ages whereas our study was restricted to young people) we did not detect any suicide clusters recurring in the subsequent period. These findings suggest that the inclusion of attempts in the definition of the cluster increases the reliability of prediction.

What mechanisms might explain why clusters reoccur in an area? We think there are at least two explanations, although the evidence base for this is thin. One is that people living in the same area share high-risk characteristics of suicide. One example of this from Australia is the
finding that Indigenous status is associated with an increased risk of a given suicide being in a cluster (Cheung, Spittal, Williamson, Tung, & Pirkis, 2014). If many people with high risk characteristics live close together (and continue to live close together) then this may explain why clusters re-occur in the same area. The alternate explanation is that areas themselves have characteristics that increase the risk of a cluster forming and re-forming; for example, poor access to mental health services, and low levels of social integration. If the high-risk characteristics of the area remain in place, the suicide cluster may be more likely to perpetuate. We suggest that future research is needed to better understand this, and the effects of individual-level risk factors piled on top of area-level risk factors.

We had hoped that we might be able to identify areas at risk of clusters of suicidal events on the basis of past clusters. This would have meant that we could reliably use relatively old data to identify areas at heightened risk. Because we were unable to do this, the key message is that more timely data are needed to identify the presence of clusters. Suicide data and even data on hospitalisations are subject to significant delays, so we may need to consider other novel, more timely data sources such as ambulance attendance data or police reports.

**Strengths and limitations**

Our study assessed the extent to which clusters of suicidal events detected in one period predicted clusters occurring in a later period. Our study has been strengthened by: (i) including both suicides and suicide attempts occurring among young people; and (ii) adjusting for socioeconomic status at the area level in analyses.

Our study was limited by several issues. First, we defined suicide attempt using the ICD codes that designate deliberate self-harm. This means we may have included individuals who
intended to harm themselves, but did not intend to die. The extent to which using this definition biases the results is unknown. Second, and relatedly, we relied on hospital admission data to define suicide attempts, rather than emergency department data (Bergen et al., 2012; Hawton et al., 2012; Kapur et al., 2015). This means that we may have excluded less serious attempts that did not require hospitalisation. Consequently, it is possible that we have missed additional clusters that may have been present, or that the clusters we did identify may in reality have been larger.

Third, we may also have missed some clusters by removing 15% of suicides and 9% of suicide attempts with missing information on residential area. This is particularly relevant for WA (where 33% of suicides and 22% of suicide attempts were removed) and may explain a lower number of clusters detected within the state. Fourth, SaTScan only permits the detection of clusters that are circular shape by using a cylindrical scan, and not clusters that are non-circular or of irregular shape. However, our previous study identified irregular shaped clusters showed that they were 87% similar with circular clusters in terms of geographical location (Milner, Too, & Spittal, 2017). Lastly, as mentioned above, we included data only from NSW and WA. Future research should include data from a wider geographical scope if possible because this is likely to increase the power of cluster prediction.

**CONCLUSIONS**

Our study showed that the presence of clusters of suicide and suicide attempts among young people in a given area in one period does not strongly predict clusters occurring in the same area in a later period. This highlights that relatively up-to-date data is needed if cluster detection is to be used for identifying areas for intervention because community risk cannot be adequately predicted from earlier data.
ACKNOWLEDGEMENTS

We would like to thank Australian Rotary Health for funding this study. We would also like to thank the Department of Health Western Australia and NSW Ministry of Health for providing us with data on fatal and non-fatal suicides.

AUTHORS’ CONTRIBUTIONS

MJS, JP, and LST designed the study. LST and MJS collected data. LST performed the analyses and interpreted data with input from MJS. LST wrote the first draft of the manuscript. All authors revised the draft and contributed to the final version of manuscript.

CONFLICT OF INTERESTS

None declared.

FINANCIAL SUPPORT

This work was supported by Australian Rotary Health. JR was supported by a NHMRC Early Career Fellowship. JP was supported by a NHMRC Senior Research Fellowship.

ETHICS APPROVAL

Ethical approval was obtained from the Health Sciences Human Ethics Committee (the University of Melbourne), the Department of Health Western Australia Human Research Ethics Committee, and NSW Population and Health Services Research Ethics Committee.
REFERENCES


Table 1. Characteristics of suicidal events (suicides and suicide attempts) among people aged 15-24 years

<table>
<thead>
<tr>
<th></th>
<th>New South Wales (Jul 2001- Jun 2012)</th>
<th>Western Australia (Jan 2000-Dec 2011)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Rate (per 100 000 persons)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>28,004</td>
<td>285.7</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>9,079</td>
<td>32.4</td>
</tr>
<tr>
<td>Female</td>
<td>18,925</td>
<td>67.6</td>
</tr>
<tr>
<td><strong>Method used</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poisoning by drugs</td>
<td>19,579</td>
<td>69.9</td>
</tr>
<tr>
<td>Cutting/piercing</td>
<td>5,816</td>
<td>20.8</td>
</tr>
<tr>
<td>Hanging</td>
<td>1,189</td>
<td>4.3</td>
</tr>
<tr>
<td>Poisoning by motor</td>
<td>174</td>
<td>0.6</td>
</tr>
<tr>
<td>vehicle exhaust</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jumping from a height</td>
<td>225</td>
<td>0.8</td>
</tr>
<tr>
<td>Other</td>
<td>1,021</td>
<td>3.6</td>
</tr>
</tbody>
</table>
Table 2. Information on spatial-temporal clusters of suicidal events among people aged 15-24 years

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Start Date</th>
<th>End Date</th>
<th>Period (in days)</th>
<th>p value</th>
<th>Number of cases</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>New South Wales</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Period one</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>02/09/2005</td>
<td>28/11/2005</td>
<td>87</td>
<td>&lt;0.001</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>19/01/2005</td>
<td>18/04/2005</td>
<td>89</td>
<td>&lt;0.001</td>
<td>110</td>
</tr>
<tr>
<td>3</td>
<td>13/08/2006</td>
<td>30/10/2006</td>
<td>78</td>
<td>&lt;0.001</td>
<td>27</td>
</tr>
<tr>
<td>4</td>
<td>21/08/2004</td>
<td>16/11/2004</td>
<td>87</td>
<td>&lt;0.001</td>
<td>62</td>
</tr>
<tr>
<td>5</td>
<td>20/02/2005</td>
<td>20/05/2005</td>
<td>89</td>
<td>&lt;0.001</td>
<td>116</td>
</tr>
<tr>
<td>6</td>
<td>30/08/2005</td>
<td>25/11/2005</td>
<td>87</td>
<td>0.005</td>
<td>96</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>432</td>
</tr>
<tr>
<td>Period two</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>02/10/2007</td>
<td>29/12/2007</td>
<td></td>
<td>&lt;0.001</td>
<td>48</td>
</tr>
<tr>
<td>2</td>
<td>30/01/2008</td>
<td>27/04/2008</td>
<td></td>
<td>0.003</td>
<td>26</td>
</tr>
<tr>
<td>3</td>
<td>12/11/2011</td>
<td>05/02/2012</td>
<td></td>
<td>0.013</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>27/05/2011</td>
<td>01/06/2011</td>
<td></td>
<td>0.057</td>
<td>13</td>
</tr>
<tr>
<td>5</td>
<td>29/04/2007</td>
<td>02/07/2007</td>
<td></td>
<td>0.073</td>
<td>22</td>
</tr>
<tr>
<td>6</td>
<td>13/05/2012</td>
<td>28/06/2012</td>
<td></td>
<td>0.083</td>
<td>48</td>
</tr>
<tr>
<td>7</td>
<td>30/10/2008</td>
<td>25/01/2009</td>
<td></td>
<td>0.084</td>
<td>28</td>
</tr>
<tr>
<td>8</td>
<td>29/09/2009</td>
<td>24/11/2009</td>
<td></td>
<td>0.090</td>
<td>52</td>
</tr>
<tr>
<td>9</td>
<td>30/01/2007</td>
<td>08/03/2007</td>
<td></td>
<td>0.098</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>16/07/2009</td>
<td>05/09/2009</td>
<td></td>
<td>0.098</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>269</td>
</tr>
<tr>
<td><strong>Western Australia</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Period one</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>12/05/2004</td>
<td>8/07/2004</td>
<td>57</td>
<td>0.001</td>
<td>28</td>
</tr>
<tr>
<td>2</td>
<td>20/04/2001</td>
<td>5/07/2001</td>
<td>76</td>
<td>0.017</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>28/01/2004</td>
<td>4/03/2004</td>
<td>36</td>
<td>0.074</td>
<td>20</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>57</td>
</tr>
<tr>
<td>Period two</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 3. Prediction of clusters of suicidal events in period two based on clusters in period one

<table>
<thead>
<tr>
<th>Period 1 (test)</th>
<th>Period 2 (gold standard)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Areas with a cluster</td>
<td>Areas with no cluster</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>n</td>
<td></td>
</tr>
<tr>
<td>Areas with a cluster</td>
<td>88</td>
<td>157</td>
<td></td>
</tr>
<tr>
<td>Areas with no cluster</td>
<td>62</td>
<td>369</td>
<td></td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>36</td>
<td>30–42</td>
<td></td>
</tr>
<tr>
<td>Negative predictive value</td>
<td>86</td>
<td>82–89</td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>59</td>
<td>50–67</td>
<td></td>
</tr>
<tr>
<td>Specificity</td>
<td>70</td>
<td>66–74</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1. Selection of suicide and suicide attempts in NSW and WA for analyses

NSW (July 2001-June 2012)

1,117 suicide attempts being removed due to missing residential S.A.

759 suicide attempts being removed because these persons had more than one admission per day and only the last admission was included.

28,202 suicide attempts from Data Linkage

747 suicides from Data Linkage

37 suicide attempts being removed due to missing residential S.A.

27,328 suicide attempts and 718 suicides were combined in one dataset.

3 observations being removed due to their S.A. did not correspond to geographical coordinates data and 1 observation being removed due to zero population.

Final total number of observations: 29,024

Identification of clusters

WA (January 2000-December 2011)

1,763 suicide attempts being removed due to missing residential S.A.

1,064 suicide attempts being removed because these persons had more than one admission per day and only the last admission was included.

12,450 suicide attempts from Data Linkage

408 suicides from Data Linkage

126 suicide attempts being removed due to missing residential S.A.

9,500 suicide attempts and 328 suicides were combined in one dataset.

No suicide attempt being removed in the condition that when suicide attempt and suicide death of any person occurred on the same day, keep only suicide death.

No observation being removed due to their S.A. did not correspond to geographical coordinates data and no observation being removed due to zero population.

Final total number of observations: 9,878

Identification of clusters
Figure 2. Clusters of suicidal events in Western Australia and New South Wales in period one and period two