Available online at www.sciencedirect.com

Public Health

journal homepage: www.elsevier.com/puhe

Original Research

Google and suicides: what can we learn about the use of internet to prevent suicides?

V. Chandler

Saint Mary's University, Economics Department, 923, Robie Street, B3H 3C3 Halifax, Canada

ARTICLE INFO

Article history:
Received 7 April 2017
Received in revised form
29 August 2017
Accepted 23 October 2017

Keywords:
Suicide
Online behaviour
Suicide prevention

ABSTRACT

Objectives: This article studies the statistical relationship between the search propensity of suicide-related terms on Google and the number of suicides.
Study design: Suicide mortality data from all American states from January 2006 to November 2014 ($n = 5372$) and data on Google search intensity for the same period was collected.
Methods: Regression analysis with dynamic components was performed to determine the relationship between search intensity and the number of suicides.
Results: First, this article finds a positive simultaneous correlation between search intensity and the number of suicides. The magnitude of this relationship has grown from 2006 to 2014 suggesting an increased reliance on the internet for suicide-related information. Second, search propensity is a significant predictor for the number of suicides for youth and for males.
Conclusions: Suicide prevention websites should therefore be designed knowing that at-risk individuals in both groups are probably more prone to look for suicide-related information online.

Introduction

In 2014, suicide was the ninth leading cause of death for the American population and the second most important for youth (aged 20–24 years). Overall, 42,773 people committed suicide in 2014 of whom 3245 were aged between 20 and 24 years.[1]

With people spending on average more than 20 h online per week,[2] the internet represents a potentially important outreach tool to prevent suicides. It is a particularly well-suited platform to discuss such taboo topics, because it provides total anonymity.[3] In spite of the growing importance of this media and its potential role to prevent suicides, little is known about the relationship between online searches and the number of suicides.

First, the number of suicides could be uncorrelated to search propensity, because people searching for suicide terms have no intention to commit suicide[a] or because suicidal people take action without consulting the internet. In both cases, resources devoted to an online presence should be channelled towards other outreach strategies.

Second, the number of suicides could be negatively correlated to search propensity. Suicidal people or their relatives could use the internet to look for help. If websites like those appearing when googling 'I want to kill myself' do indeed prevent suicides, the propensity to Google suicide-related terms would be negatively correlated with the number of suicides. Such a negative coefficient would be consistent with the fact that suicide prevention organizations are successfully preventing suicides.

E-mail address: vincent.chandler@smu.ca.
[a] People could simply be curious or may search for information concerning the suicides of celebrities.
https://doi.org/10.1016/j.puhe.2017.10.016
0033-3506/

Third, if these websites are unable to prevent at-risk individuals looking for help from committing suicide, the relationship between search propensity and the number of suicides would be positive. In such a case, websites designed to prevent suicides may need to reconsider their outreach strategy. Moreover, such a positive correlation could be explained by the fact that at-risk people tend to consult pro-suicide websites following their internet search[4] or could be due to friends or relatives searching for suicide-related terms just following a suicide as a coping mechanism.[b]

Results from previous studies provide no consensus on this relationship partially due to their small samples. Some studies find no significant relationship,[c] whereas others do find a positive relationship.[d]

This article contributes to this literature by considering a large sample consisting of monthly data for 50 states during the period stretching from January 2006 to November 2014 (5388 observations[e]). This sample is more than 50 times larger than the samples used in previous studies, and therefore enables a more thorough analysis. Moreover, this article studies a later period which has not been considered in previous work. With the increasing use of the internet, more up-to-date studies are necessary. Finally, this article is the first to consider the suicides of different groups to make a first attempt at understanding who is performing such searches.

Overall, there is a positive relationship between the contemporaneous search intensity and the number of suicides: a one percentage point increase in search propensity is related to a yearly increase of 54 suicides in the United States. The magnitude of this relationship is particularly strong after 2010 possibly explaining the non-results of previous studies conducted before 2010. Moreover, overall search propensity is a statistically significant predictor for the number of youth (aged 16–35 years) suicides and for the number of male suicides.

The rest of the article is divided in the following way. I first present the data. I then describe some the methodological challenges. Finally, I show the results and discuss briefly their implications.

Methods

Data

Data from Google searches is provided by Google Trends. This website supplies the propensity to search ratios by state for any keyword or combination thereof. These ratios are relative to the period with the highest search ratio for a given state. For example, assume that the word 'suicide' was googled 1000 times in period 1 in state 1. Moreover, suppose it was entered 2000 times in the search engine in period 2 in state 1. In such a case, Google Trends would report a ratio of 0.5 for period 1 and a ratio of 1 for period 2 assuming that period 2 had the highest search propensity per month for this state.

To protect confidentiality, Google Trends does not report results when the absolute number of searches for a certain keyword for a given unit is below a certain threshold.[f] When entering keywords, it is therefore important to cast a wide net to obtain results for all states and thus avoid missing data.

During the month of December 2016, data were collected on the search propensity of the following expression: suicide + "kill myself" + "kill maself" + "kill me" + "want to die" + "wanna die" + suicidal − squad for the period 2004 to 2014. The quotation sign indicate that a certain expression should be included precisely in this order, '+' means 'or' and '−' means 'not'. The term 'squad' was removed to avoid any confusion with the 2016 film 'Suicide Squad'.

Google Trends data are available from January 2004 to December 2016. I restrict the data to the interval between 2006 and 2014 for three reasons. First, observations in 2004 would not have all lags. Second, Google was not as widespread in the earlier part of the interval as it was in the later part of the period. Third, suicide data from the Centers for Disease Control and Prevention (CDC) is not available after 2014.

Data on suicides were collected from the CDC. CDC Wonder provides data on the number of deaths per state per month for intentional self-harm (X60-X84) until December 2014. To protect confidentiality, no data below 10 (i.e. <10 suicides) are reported by CDC. The dependent variable is therefore censored on the left. This issue represents a challenge in states with small populations like Wyoming, Vermont and North Dakota and in cases when I consider suicides from sub-samples based on age or gender.

To make data on suicides comparable with the output from Google Trends, I transform the suicide data into ratios relative to the month with the highest number of suicides per population for a given state. I therefore divide the ratio of suicides by state population of a given month by the highest ratio of suicides by state population by month of this state: $\frac{\text{Suicide}_{it}}{\text{Population}_{it}} / \frac{\text{Suicide}_{ik}}{\text{Population}_{ik}}$ where 'k' is the month in which state 'i' had the highest suicide ratio of the 2006 to 2014 period.

[b] Depending on the time unit of a study, searches following a suicide would be included in the suicide month. For example, if a suicide occurs on July 5th 2010, it will count towards the suicide intensity for July 2010 in a given state. If friends and relatives in this state search for suicide-related terms from July 6th to July 31st 2010, those searches will count towards the search intensity for July 2010 in a given state.

[c] Sueki,[5] for example, finds no correlation between Google searches for suicide using monthly Japanese data from 2004 to 2009 (72 observations). Moreover, Page, Change and Gunnell[6] find no evidence using 96 data points from February 2004 to February 2011 in Australia. Finally, Bruckner, McClure and Kim[7] find no correlation for England and Wales using 84 observations.

[d] Gunn and Lester[8] find a statistically significant contemporaneous correlation between suicide-related searches and the number of suicides in a study considering 50 American states in 2009. Similarly, McCarthy[9] uses weekly data for the whole United States between 2004 and 2009 and finds a relationship between searches and suicides only for youth suicide. Finally, Yang et al.[10] find that 'major depression' and 'divorce' were the search terms most highly correlated with the number of suicides in Taiwan with a sample of 60 observations.

[e] This number corresponds to the maximum number of observations. Depending on the sample chosen, some observations are censored. When using the total number of suicides per state per month, for example, 16 observations are censored. When considering youth suicides, 1400 observations are censored.

[f] Google does not disclose this threshold.