

Original Paper

Association Between School-Related Google Trends Search Volume and Suicides Among Children and Adolescents in Japan During 2016-2020: Retrospective Observational Study With a Time-Series Analysis

Takahiro Arai^{1,2,3}, MSHM; Hiroe Tsubaki⁴, PhD; Ayako Wakano⁵, PhD; Yasuyuki Shimizu³, BA

¹School of Management and Information Sciences, Tama University, Tokyo, Japan

²Graduate School of Health Management, Keio University, Kanagawa, Japan

³Japan Suicide Countermeasures Promotion Center, Tokyo, Japan

⁴The Institute of Statistical Mathematics, Tokyo, Japan

⁵Department of Economics, School of Political Science and Economics, Tokai University, Kanagawa, Japan

Corresponding Author:

Takahiro Arai, MSHM

School of Management and Information Sciences

Tama University

4-1-1 Hijirigaoka

Tama City

Tokyo, 206-0022

Japan

Phone: 81 42 337 7149

Email: arai.t@keio.jp

Abstract

Background: Suicide is the leading cause of death among children and adolescents in Japan. Internet search volume may be useful in detecting suicide risk. However, few studies have shown an association between suicides attempted by children and adolescents and their internet search volume.

Objective: This study aimed to examine the relationship between the number of suicides and the volume of school-related internet searches to identify the search terms that could serve as the leading indicators of suicide prevention among children and adolescents.

Methods: We used data on weekly suicides attempted by elementary, middle, and high school students in Japan from 2016 to 2020, provided by the National Police Agency. Internet search volume was weekly data for 20 school-related terms obtained from Google Trends. Granger causality and cross-correlation analysis were performed to estimate the temporal back-and-forth and lag between suicide deaths and search volume for the related terms.

Results: The search queries “I do not want to go to school” and “study” showed Granger causality with suicide incidences. The cross-correlation analysis showed significant positive correlations in the range of -2 to 2 for “I do not want to go to school” (highest value at time lag 0, $r=0.28$), and -1 to 2 for “study” (highest value at time lag -1 , $r=0.18$), indicating that the search volume increased as the number of suicides increased. Furthermore, during the COVID-19 pandemic period (January-December 2020), the search trend for “I do not want to go to school,” unlike “study,” was highly associated with suicide frequency.

Conclusions: Monitoring the volume of internet searches for “I do not want to go to school” could be useful for the early detection of suicide risk among children and adolescents and for optimizing web-based helpline displays.

(*J Med Internet Res* 2024;26:e51710) doi: [10.2196/51710](https://doi.org/10.2196/51710)

KEYWORDS

adolescent; children; COVID-19; Google Trends; internet; Japan; monitoring; suicide; surveillance; time series analysis

Introduction

In Japan, suicide is the leading cause of death among people aged 10-19 years [1]. Suicide among children and adolescents has continued to increase since 2016, with data for 2020 showing the suicide rate (per 100,000 population) for primary, secondary, and high school students totaled 4.0, with 0.2 for primary school students, 4.5 for secondary school students, and 11.0 for high school students. One of the key reasons for suicide is school-related problems [2]. Therefore, suicide prevention among children and adolescents is primarily conducted in schools. However, school-based suicide prevention programs have exhibited limited effectiveness in reducing suicide attempts and rates [3,4]. In recent years, suicide prevention using the internet has received greater attention to complement the effectiveness of school-based suicide prevention programs directed toward the young population. As the internet has become a prevailing medium for young people, suicide prevention has been examined using internet spaces rather than mental health care in schools and gatekeeping using the phone [5]. Recent studies have analyzed the association between the volume of suicide-related terms derived from internet search queries and suicide [6-11].

The field of public health has explored the development of monitoring methodologies using internet searches. During the COVID-19 pandemic, there have been suggestions that Google Trends could be used to aid disease surveillance systems [12-14]; their complementary function has also been explored in the context of suicide prevention [15,16].

The Japanese government focused on the relationship between internet search queries and suicide and proposed a suicide prevention strategy using internet search mechanisms in 2017 [17]. However, the empirical findings were quite limited. Sueki [18] reported a significant positive correlation between the number of suicides and the search term “utsu” (depression). Hagihara et al [19] further reported a strong association between suicide rates and the search term “hydrogen sulfide,” a means of attempting suicide. Both studies covered the period before the COVID-19 pandemic, and search terms were related to mental illness and crisis interventions, such as the means of attempting suicide. Furthermore, during the COVID-19 pandemic, Taira et al [20] found that keywords such as “no money” and “divorce” could predict the number of suicides, discovering a new linguistic marker by targeting search terms related to social factors rather than words directly related to specific disease names or suicide.

Internet search terms may provide clues for identifying suicide crisis pathways and be useful for optimizing web-based suicide prevention [21]. However, the research investigating an association between suicide deaths among children and adolescents and internet search queries is scant. Therefore, new linguistic markers that have the potential to detect suicide risk among children and adolescents need to be explored.

Our study aims to identify the search terms that could serve as the leading indicators of suicide among children and adolescents. For this purpose, we investigated the relationship between the number of suicides and internet search volumes. We further

examined how these trends changed in the social context during the COVID-19 pandemic. In this study, children and adolescents are defined based on their school attendance at elementary, junior high, and high schools rather than by age. This definition, as used in the school education system, is considered a prominent factor contributing to the occurrence of suicide among children and adolescents in Japan [2].

Several analytical challenges regarding the usefulness of search queries in suicide prevention remain unresolved. In other words, the coarseness of the monthly data on the number of suicides is problematic. To tackle this weakness, we propose that using weekly data can change the analysis results [22]. Our study has 2 unique features. First, we used weekly data on the number of suicides and internet search volumes because monthly analysis may not observe changes in highly fluid social conditions. Second, we used the number of suicides attempted by elementary, middle, and high school students rather than the total number of suicides. The Japanese suicide surveillance system includes occupational status.

Methods

Suicide Data

The National Police Agency provided data on the number of suicides. These were weekly data processed by the Japan Suicide Countermeasures Promotion Center under the supervision of the Ministry of Health, Labor, and Welfare after receiving confirmation from the National Police Agency. The data collection period was 5 years, from 2016 to 2020, and the analysis included elementary, middle, and high school students. The variables present in the National Police Agency data are demographic variables, gender (binary), age (integer), occupational status (51 categories including elementary, middle, and high school students), living together (3 levels—yes, no, and unknown), municipality of residence (character), suicide municipality (character), and other or not having a history of suicide attempts (3 levels—yes, no, and unknown). Selecting participants solely based on age would not adequately exclude the individuals who dropped out or did not attend high school after graduating from their junior high school. Therefore, this study used occupational status categories to accurately identify and include elementary, junior high, and high school students.

Internet Search Data

Search queries used weekly data (7-day interval from Sunday to Monday) from January 3, 2016, to December 26, 2020, using Google Trends. The Google Trends search volumes during this period were normalized between 0 and 100. According to Google Trends, the normalization method is described as “search results are normalized to the time and location of a query by the following process: Each data point is divided by the total searches for the geography and time range. The resulting numbers are then scaled on a scale of 0 to 100 based on a topic’s proportion to all searches on all topics” [23]. Therefore, it only shows a relative value rather than a pure search frequency. Relative search values (RSV) are calculated as follows.

For the search volume of a word “*i*” in a certain week “*t*,” in a certain region “*j*,” let $Q_{max_{i,j}}$ be the maximum search volume when it was the maximum in all periods.

$$Q_{max_{i,j}} = \max\{search_{i,j,t}\} \tag{1}$$

The amount of retrieval of the search word “*i*” in a given week “*t*” is relative to the maximum value.

$$RSV_{i,j,t} = \left(\frac{search_{i,j,t}}{Q_{max_{i,j}}}\right) * 100 \tag{2}$$

The search terms used in this study were selected based on a list of the causes of student suicide [2]. We excluded “suicide” and “I want to die,” which have been used in many studies, because they are likely to be searched by many people unrelated to suicide and are not necessarily representative of suicide risks among children and adolescents [24]. In addition, we integrated words related to school resumption because the school resumption timing is considered to cause strong psychological stress among young people [25].

Based on the above information, 20 keywords were selected for this study, such as “school,” “school reopening,” “I do not want to go to school,” “school hard,” “school anxiety,” “When is school?” “When does school start?” “school start,” “school bullying,” “school friends,” “school grades,” “bullying,” “friendships,” “study,” “anxiety about studying,” “I do not know how to study,” “mid-term or term-end exam,” “entrance examination,” “future career,” and “grades.”

Statistical Analysis

We used the time-series analysis approach used in previous studies [26]. Referring to previous studies, we constructed a concept map to vividly illustrate time-series analysis (Figure 1).

First, a unit root test (an augmented Dickey-Fuller [ADF] test) was performed on the collected time-series data (n=260 weeks) using the R package (R Core Team) “tseries” [27] to examine whether the sequence was stationary or not. Stationarity is an important assumption in many statistical analyses, including Granger causality analysis. The null hypothesis of the ADF test (H0) assumes the existence of a unit root and indicates nonstationarity, while the alternative hypothesis (H1) indicates stationarity. With the significance level set at .05, a *P* value below it justifies the rejection of H0 and affirms the stationarity of the data. Next, the vector autoregressive (VAR) model was estimated using the R package “vars” [28]. It incorporates the optimal lag variables, whose order was determined by Akaike’s information criterion (AIC), and the Granger causality was verified. Granger causality analysis is a statistical method used to determine if one time-series variable can predict another time-series variable. To assess the significance of the relationship, we performed the *F* test. If the *P* value was below the significance level (*P*<.05), we rejected the null hypothesis

and concluded that there is evidence of Granger causality. However, if the *P* value was above the significance level, we failed to reject the null hypothesis, indicating that there was no evidence of Granger causality. Furthermore, to adjust for the influence of the school calendar, we incorporated dummy variables for weeks including April 8, September 1, and January 8, which are anticipated to have spikes in suicide risk as they mark the end of spring, summer, and winter breaks, respectively [25]. We corrected the number of suicides using these dummy variables in our Granger causality testing model (the school calendar model). We tested the Granger causality test using this adjusted number of suicides.

The Google Trends model for the Granger causality test is as follows.

$$Y_t = \alpha_0 + \sum_{i=1}^p \alpha_{1i} Y_{t-i} + \sum_{i=1}^p \alpha_{2i} X_{t-i} + \epsilon_{Y,t} \tag{3}$$

$$X_t = \beta_0 + \sum_{i=1}^p \beta_{1i} X_{t-i} + \sum_{i=1}^p \beta_{2i} Y_{t-i} + \epsilon_{X,t} \tag{4}$$

Y_t is the number of suicides at time *t*. X_t is the causal variable (the number of searches for the keyword on Google) at time *t*.

Adjustments to the number of suicides in the school calendar model are as follows:

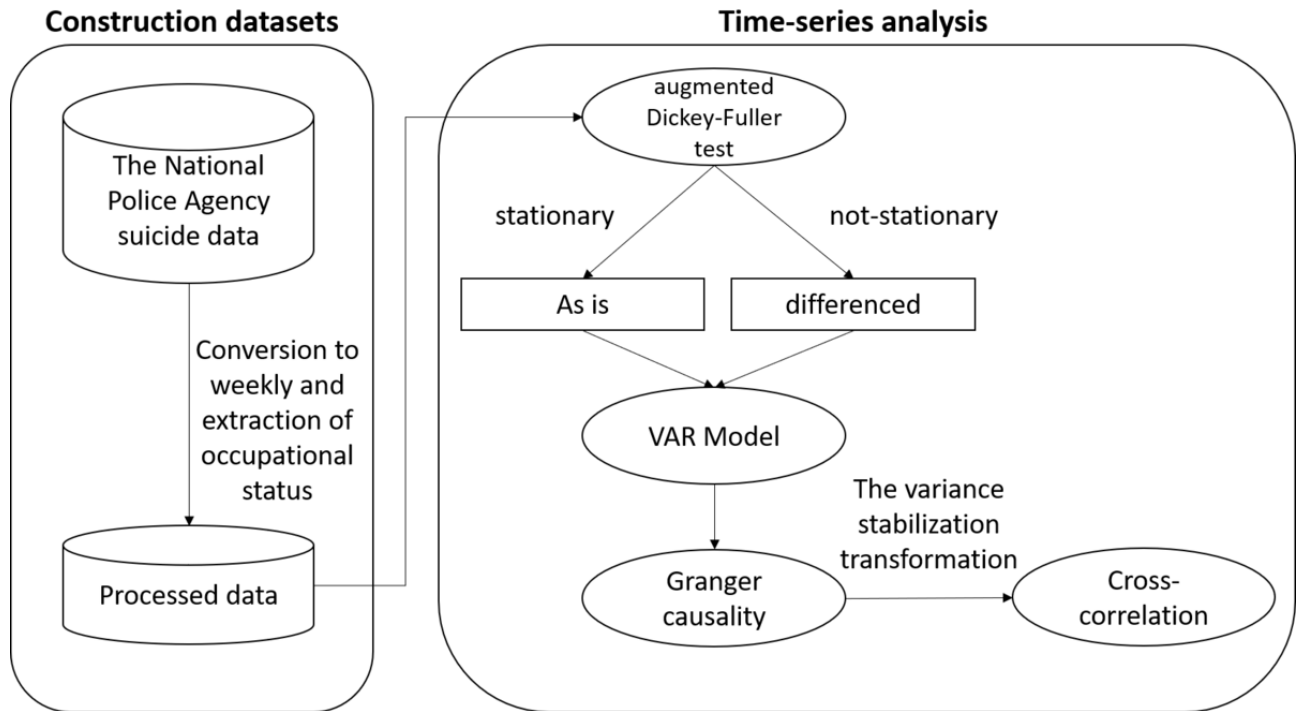
The Poisson regression model takes the following form:

$$\log(\lambda_t) = \gamma_0 + \gamma_1 D_{spring,t} + \gamma_2 D_{summer,t} + \gamma_3 D_{winter,t} \tag{5}$$

λ_t is the expected number of suicides at week *t*. $D_{spring,t}$, $D_{summer,t}$, and $D_{winter,t}$ are dummy variables at the end of the spring, summer, and winter breaks, respectively. The actual number of suicides data is assumed to follow a Poisson distribution. By subtracting the predictions made by the Poisson regression model from the measured data, we generated suicide count data adjusted for the effects of the school calendar.

After performing the Granger causality test, as other past studies did, we also calculated the cross-correlation coefficients based on the same method as in previous studies [22]. Correlation coefficients with the time lag data were estimated, and a significance test was performed. We assumed that the variance was equivariant when the correlation coefficient was obtained. In this case, the square root transformation was used for the variance stabilization transformation owing to the data frequency. Finally, search terms and suicide frequencies that showed significant differences in the number of suicides were plotted against the number of weeks in 2020 to determine social conditions during the COVID-19 pandemic. All statistical analyses were performed using R (version 4.1.0; The R Foundation), and the significance level was set to 5%.

Figure 1. Concept map expressing the methodology of time series analysis. VAR: vector autoregressive.



Ethical Considerations

Our study was approved by the Research Ethics Review Committee of Japan Suicide Countermeasures Promotion Center on September 16, 2021 (R03-01). Patient consent was waived because this study was a secondary analysis of existing data. The data were made available for research use by processing them on a weekly basis to prevent personal identification.

Results

Distribution of the Number of Suicides

The total number of suicides among elementary, middle, and high school students, from January 3, 2016, to December 26, 2020, was 1925 (mean 7.40, SD 3.42). The 52 weeks were divided into 5 periods. The summary statistics for each period were as follows: period 1 (January 3, 2016, to December 31, 2016) was 318 (mean 6.12, SD 2.94), period 2 (January 1, 2017, to December 30, 2017) was 352 (mean 6.77, SD 2.95), period 3 (December 31, 2017, to December 29, 2018) was 366 (mean

7.04, SD 2.81), period 4 (December 30, 2018, to December 28, 2019) was 394 (mean 7.58, SD 3.14), and period 5 (December 29, 2019, to December 26, 2019) was 495 (mean 9.52, SD 4.18). During our study period (260 total weeks), there were 0 weeks with no suicide deaths among elementary, middle, or high school students.

Evaluation Outcomes

The study variables were compiled and are presented in Table 1. They consist of 11 school-related terms and 9 terms integrated with school words, generating a total of 20 variables.

Unit root tests (ADF) were performed on the original series of suicide variables, all of which exhibited stationary behavior (Table 2). The test results are available upon request from the authors. Using the AIC, a VAR model was selected for Granger causality analysis. The results showed a statistically significant Granger causality between the search terms “I do not want to go to school” ($P=.006$) and “study” ($P=.02$) and the number of suicides (Table 3). The other search terms did not show Granger causality in either direction for the number of suicides.

Table 1. The Google Trends search terms selected as the subjects of analysis.

English equivalent search terms	Original Japanese search terms
Search terms combined with school	
school	学校 (gakkou)
school reopening	学校 再開 (gakkou saikai)
I do not want to go to school	学校 行きたくない (gakkou ikitakunai)
school hard	学校 辛い (gakkou tsurai)
school anxiety	学校 不安 (gakkou fuan)
When is school?	学校 いつ (gakkou itsu)
When does school start?	学校 いつから (gakkou itsukara)
school start	学校 始まる (gakkou hajimaru)
school bullying	学校 いじめ (gakkou ijime)
School-related search terms	
school friends	学校 友人 (gakkou yuujin)
school grades	学校 成績 (gakkou seiseki)
bullying	いじめ (ijime)
friendships	友人関係 (yuujinkankei)
study	勉強 (benkyou)
anxiety about studying	勉強 不安 (benkyou fuan)
I do not know how to study	勉強 わからない (benkyou wakaranai)
mid-term or term-end exam	試験 (shiken)
entrance examination	受験 (juken)
future career	進路 (shinro)
grades	成績 (seiseki)

Table 2. Results of augmented Dickey-Fuller test on time-series data.

Variables	ADF ^a test statistics	P value
the number of suicides	-4.61	<.001
school	-3.91	.01
school reopening	-3.91	.01
I do not want to go to school	-6.72	<.001
school hard	-5.96	<.001
school anxiety	-4.89	<.001
When is school?	-3.49	.04
When does school start?	-4.21	<.001
school start	-4.08	<.001
school bullying	-5.22	<.001
school friends	-5.38	<.001
school grades	-15.67	<.001
bullying	-4.94	<.001
friendship	-6.15	<.001
study	-5.77	<.001
anxiety about studying	-5.45	<.001
I do not know how to study	-6.37	<.001
mid-term or term-end exam	-5.84	<.001
entrance examination	-4.67	<.001
future career	-4.53	<.001
grades	-3.69	.03

^aADF: augmented Dickey-Fuller.

Table 3. Temporal pre-post relationship by Granger causality (Google Trends model).

Null hypothesis	Optimal VAR ^a lag length	F test (df)	P value
“school” does not Granger-cause the number of suicides	7	1.10 (7, 476)	.36
The number of suicides does not Granger-cause “school”	7	0.63 (7, 476)	.74
“school reopening” does not Granger-cause the number of suicides	9	0.79 (9, 464)	.63
The number of suicides does not Granger-cause “school reopening”	9	1.32 (9, 464)	.22
“I do not want to go to school” does not Granger-cause the number of suicides	1	7.55 (1, 512)	.006
The number of suicides does not Granger-cause “I do not want to go to school”	1	3.79 (1, 512)	.05
“school hard” does not Granger-cause the number of suicides	1	0.01 (1, 512)	.94
the number of suicides does not Granger-cause “school hard”	1	1.51 (1, 512)	.22
“school anxiety” does not Granger-cause the number of suicides	1	0.94 (1, 512)	.33
the number of suicides does not Granger-cause “school anxiety”	1	0.39 (1, 512)	.53
“When is school?” does not Granger-cause the number of suicides	3	0.19 (3, 500)	.90
the number of suicides does not Granger-cause “When is school?”	3	0.21 (3, 500)	.89
“When does school start?” does not Granger-cause the number of suicides	12	0.93 (12, 446)	.51
the number of suicides does not Granger-cause “When does school start?”	12	0.56 (12, 446)	.88
“school start” does not Granger-cause the number of suicides	3	0.69 (3, 500)	.56
the number of suicides does not Granger-cause “School start”	3	0.74 (3, 500)	.53
“school bullying” does not Granger-cause the number of suicides	1	2.80 (1, 512)	.09
the number of suicides does not Granger-cause “school bullying”	1	0.13 (1, 512)	.72
“school friends” does not Granger-cause the number of suicides	1	0.11 (1, 512)	.74
the number of suicides does not Granger-cause “school friends”	1	1.42 (1, 512)	.23
“school grades” does not Granger-cause the number of suicides	2	0.32 (2, 506)	.73
The number of suicides does not Granger-cause “school grades”	2	1.72 (2, 506)	.18
“bullying” does not Granger-cause the number of suicides	2	1.85 (2, 506)	.16
the number of suicides does not Granger-cause “bullying”	2	0.75 (2, 506)	.47
“friendship” does not Granger-cause the number of suicides	1	0.02 (1, 512)	.89
the number of suicides does not Granger-cause “friendship”	1	0.07 (1, 512)	.79
“study” does not Granger-cause the number of suicides	5	2.63 (5, 488)	.02
the number of suicides does not Granger-cause “study”	5	0.54 (5, 488)	.74
“anxiety about studying” does not Granger-cause the number of suicides	4	0.39 (4, 494)	.82
the number of suicides does not Granger-cause “anxiety about studying”	4	0.40 (4, 494)	.81
“I do not know how to study” does not Granger-cause the number of suicides	1	0.08 (1, 512)	.77
the number of suicides does not Granger-cause “I do not know how to study”	1	0.26 (1, 512)	.61
“mid-term or term-end exam” does not Granger-cause the number of suicides	3	0.27 (3, 500)	.85
the number of suicides does not Granger-cause “mid-term or term-end exam”	3	0.39 (3, 500)	.76
“entrance examination” does not Granger-cause the number of suicides	5	0.86 (5, 488)	.51
the number of suicides does not Granger-cause “entrance examination”	5	0.58 (5, 488)	.71
“future career” does not Granger-cause the number of suicides	4	0.74 (4, 494)	.56
the number of suicides does not Granger-cause “future career”	4	0.89 (4, 494)	.47
“grades” does not Granger-cause the number of suicides	2	2.11 (2, 506)	.12
the number of suicides does not Granger-cause “grades”	2	2.07 (2, 506)	.13

^aVAR: vector autoregressive.

Robustness of the Result

For the search words that were found to be significant in the above analysis, we conducted the robustness check by generating a dataset that excludes the effect of the school calendar. The reason for this check is to test whether our search keywords remain statically significant after controlling for the effect of the school calendar. If it remains so, it infers that our search

keywords have a powerful predicting power for the number of suicides among elementary, middle, and high school students.

The results of the school calendar model rejected the null hypotheses “I do not want to go to school” does not Granger-cause the number of suicides and “study” does not Granger-cause the number of suicides both as in the Google Trends model (Table 4).

Table 4. Temporal pre-post relationship by Granger causality (school calendar model in the search terms “I do not want to go to school” and “study”).

Null hypothesis	Optimal VAR ^a lag length	F statistics (df)	P value
“I do not want to go to school” does not Granger-cause the number of suicides	1	10.91 (1, 512)	.001
The number of suicides does not Granger-cause “I do not want to go to school”	1	2.74 (1, 512)	.10
“study” does not Granger-cause the number of suicides	2	4.01 (2, 506)	.02
the number of suicides does not Granger-cause “study”	2	1.43 (2, 506)	.24

^aVAR: vector autoregressive.

Correlation Analysis

Subsequently, variance stabilization transformations were applied to the phrases shown by Granger causality. A cross-correlation analysis, which revealed that the time-series of suicides lagged by 2 weeks concerning both “I do not want to go to school” and “study,” was performed (Table 5).

In this study, the search queries “I do not want to go to school” and “study” showed Granger causality with suicide incidences. The results of the cross-correlation analysis showed significant positive correlations in the range of -2 to 2 for “I do not want to go to school” (highest value at time lag 0, $r=0.28$) and -1 to 2 for “study” (highest value at time lag -1, $r=0.18$), indicating that the search volume increased as the number of suicides increased.

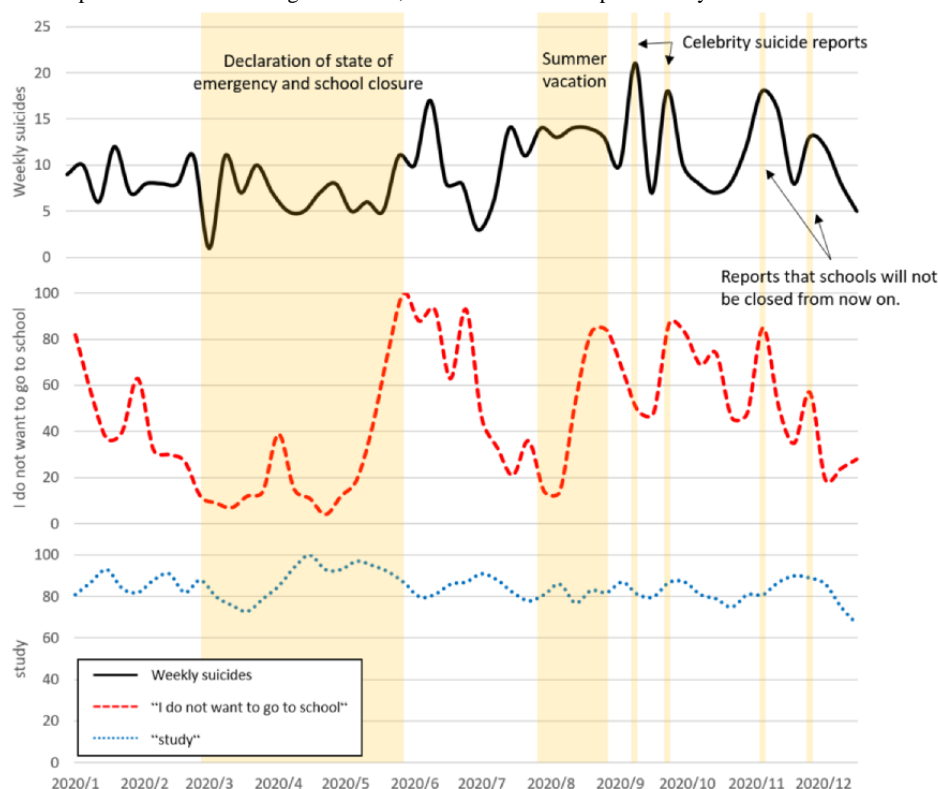
Furthermore, in 2020, during the COVID-19 pandemic, the time-series trend for “I do not want to go to school” resembled that of the number of suicides. Figure 2 shows the social situation in 2020 and the time-series of the number of suicides and search terms. Both variables decreased from March to May, when schools were closed nationwide, and increased from May to June, when the state of emergency was lifted. This relationship persisted even after the summer vacation and increased from September to October. In addition, 2 more sharp increases were observed after November. Conversely, the search volume for “study” increased from March to May and remained high thereafter, showing no significant change compared with “I do not want to go to school.”

Table 5. Cross-correlation with the number of suicides at lags -3 to +3 after variance stabilization.

Search terms	-3	-2	-1	0	1	2	3
“I do not want to go to school”	0.04	0.09 ^a	0.21 ^a	0.28 ^a	0.26 ^a	0.17 ^a	0.06
“study”	0.05	0.04	0.18 ^a	0.17 ^a	0.13 ^a	0.12 ^a	0.09

^aStatistically significant values at $P<.05$.

Figure 2. Number of suicides and search queries during the COVID-19 pandemic period. Highlights indicate social conditions. The top line plots weekly suicide, the middle line plots “I do not want to go to school,” and the bottom line plots “study.”



Discussion

Principal Findings

This study examines the association between the number of suicides among Japanese children and adolescents and the volume of school-related internet searches. We used Granger causality analysis to show that specific school-related internet search volumes may precede the number of suicides among Japanese elementary, middle, and high school students over time. In the Granger causality test, “I do not want to go to school” and “study” were statistically significant in both the Google Trends model and the school calendar model. Internet search terms that were significantly and positively correlated with the number of suicides were “I do not want to go to school” and “study.” Regarding the time lag, a negative direction starting from 0 indicates the periods after the increase in search volumes, whereas a positive direction indicates the periods before the increase in search volumes. Therefore, to investigate whether the number of suicides increases after an increase in search volumes, the negative lag direction should be primarily considered. In this study, lags were calculated weekly, whereas previous studies used monthly aggregates, where a lag of approximately 2-3 months was observed [16,18,22]. However, the lag in this study was less than 2 weeks, indicating the high accuracy of this leading indicator. The results in Table 4 imply that a lag in either direction indicates a short-term association between the search terms and suicide. The positive correlation between the number of suicides and the search volumes suggests that search terms may be a leading indicator of suicide. Furthermore, the volume of searches for “I do not want to go to school” showed a time-series trend similar to the number of

suicides among elementary, middle, and high school students, even in the social context of the 2020 COVID-19 pandemic. To the best of our knowledge, our study is the first to investigate the association between the weekly number of elementary, middle, and high school student suicides and the volume of internet searches for school-related terms in Japan. The discovery of new linguistic markers associated with suicide among elementary, middle, and high school students may contribute to research and practice related to suicide prevention. As our analysis is exploratory, aimed at generating hypotheses for future research, we did not apply multiple comparison corrections [29]. Future studies will need to apply more stringent statistical controls when testing these initial hypotheses to validate the findings.

In 2020, during the COVID-19 pandemic, similar time-series trends were revealed between the search volume “I do not want to go to school” and the number of suicides. In Japan, schools were closed nationwide on March 2, 2020, and were reopened when the state of emergency was lifted on May 25, 2020. During this period, a sharp increase was noted in both the number of suicides and search volumes, which could be attributed to the increasing tension over the school’s reopening. This trend could also be observed during the summer vacation from August to September. The increase in the number of suicides in September cannot necessarily be explained solely by the psychological burden of the resumption of school, since reports of celebrity suicides occurred. However, the number of searches for “I do not want to go to school” gradually began to increase before school resumed in September. This phenomenon may be attributed to an increase in the number of searches considered as suicide risk owing to the timing of the end of the long

vacation. Therefore, this result may have reinforced the “broken promise effect” theory [30], which suggests that suicide rates tend to increase after a long break from school.

Other negative school-related terms searched during the same period did not yield significant Granger causality or correlation coefficients. This suggested the robustness of the “I do not want to go to school” term being the high-risk keyword for suicide among elementary, middle, and high school students.

The 2 peaks in search volumes in November 2020 might have been influenced by press conferences conducted by the Ministry of Education, Culture, Sports, Science, and Technology on November 10 and November 27, which stated that entrance examinations and the nationwide closure of all schools would not be canceled or requested [31,32]. This statement may have recalled the psychological stress children and adolescents feel about schools. Therefore, government statements related to schools should be made with caution.

This study’s findings are consistent with those of previous studies. A meta-analysis and systematic review indicated that school absenteeism and dropout rates are associated with an increased risk of self-harm and suicide attempts [33,34], suggesting that educational factors may influence suicide. The search term “I do not want to go to school” in this study may be precisely a proxy variable for measuring educational factors such as expectations and the stress of attending school. The results imply that suicide is closely related to the educational system, especially the opening and closing of schools. In other words, it is possible that school opening and closing are positively and negatively correlated with suicide risk, respectively. However, caution must be exercised while interpreting these results. This is because it has been reported that school ties may be a powerful protective factor against youth suicide [35].

They should enable children who do not want to go to school to continue their learning at home or in other settings by providing them with extensive support. For example, the government could provide financial support to encourage participation in alternative education programs. This support could include subsidizing the cost of accessing online learning platforms or providing the necessary funding to attend free schools. Schools should also adopt more flexible scheduling to motivate students to learn. This allows students to progress at their own pace and helps them to balance their academic and personal interests and activities. In addition, mental health support in schools and communities needs to be strengthened to help children for whom school is a distressing experience. The findings of this study contribute to a better understanding of support strategies for children who do not want to go to school and to proposals for promoting comprehensive change in the education system.

In a study in South Korea on suicide-related internet search terms among children and adolescents, the search volume for “dropout” showed a significant positive correlation [36], indicating that school-related search terms were associated with the number of suicides, which is consistent with our findings. However, our study differs in that we used weekly data rather than daily data, which would be affected by each day of the

week, Monday through Sunday. Furthermore, our study fills a gap in previous studies because it shows the potential to be used as a risk factor not only in normal times but also in emergencies such as school closures associated with the COVID-19 pandemic.

This study’s strength is exhibited in the process of selecting search terms through an exploratory approach. The internet search terms selected in this study were chosen through an exploratory review of official sources on suicide in Japan. The selection process did not follow a rigorous and systematic methodology. However, relying solely on a systematic list of search terms based on previous research does not necessarily uncover search terms that are effective for suicide prevention. One study hypothesized that internet searches related to past orientation would be positively related to suicide rates because people with regret and depression are more likely to dwell on the past [37]. In another study, from a clinical psychology perspective, search terms associated with absolutist thinking, such as “completely” and “totally,” were used [38]. These emerging linguistic markers may not have been identified through standardized and broad-ranging methodologies. Although we could have selected specific “words” related to suicide, it may have been difficult to select the sentence “I do not want to go to school.”

Online gatekeeping reduces suicidal ideation [39]. However, “Suicide Prevention Results” are not currently implemented for the search term “I do not want to go to school” in the Japanese Google search system. It is believed that whether a helpline is displayed depends on the language and search terms used [40,41]. Our study has the potential to improve the suicide prevention results for children and adolescents and enable effective referrals to counseling services.

Suicide surveillance in Japan has been delayed for more than 6 months [42]. Monitoring high-risk suicide keywords will help improve such delays and aid in rapid administrative support and mental health planning at the national and municipal levels. Furthermore, monitoring and evaluating surveillance systems involves not only promptness but also cost-effectiveness factors [43]. The Google Trends data used in this study can be obtained in real time and free of charge and could complement existing surveillance systems. This study’s results may provide useful information for many stakeholders, such as schools, internet service providers, governments, and municipalities.

Limitations

This study has several limitations. First, the Google Trends dataset may be susceptible to noise, as search volumes can be influenced by external factors such as social conditions and media coverage, independent of the searcher’s characteristics. In addition, Google Trends has been criticized for its inability to capture user demographics such as age and gender [7]. Our study assumes that the use of school-related search terms may limit the demographics of searchers to younger age groups. Indeed, future experimentation is needed to determine the exact demographic distribution of individuals who perform school-related searches.

Second, an increase in search volumes for specific keywords does not always correspond to an increase in the number of suicides. Some search terms are accompanied by online helpline advertisements in the browser, which may reduce suicide risk by guiding individuals to seek professional help. Consequently, the correlation between internet queries and the number of suicides may be nullified.

Third, the data used in this study only observes search behavior. Recent forms of social media, such as Twitter (rebranded as X) and Instagram (Meta), offer the distinct feature of allowing individuals to express their thoughts and situations, which diverges from searching behavior [24]. Consequently, these new media forms are currently being investigated regarding adolescent suicides [44]. In the future, it will be necessary to not only continually observe search terms through Google Trends but also scrutinize new media that enable the expression of psychological states.

Finally, while this study is concentrated on suicide among elementary, junior high, and high school students in Japan, yielding findings specific to this demographic within the Japanese context, the global availability of Google Trends and

suicide data presents an opportunity for broadening the research scope. Therefore, future studies are planned to explore the applicability of these findings across different regions and cultures, aiming to understand the universality and limitations of our conclusions.

Conclusions

This study identifies an association between the volume of internet searches for school-related terms such as “I do not want to go to school” and “study,” and the number of suicides among children and adolescents in Japan. By analyzing the search volume for school-related terms, which have not received much attention in previous studies, the study suggests the possibility of improving online helpline displays for children and adolescents. The volume of searches for school-related terms may also be useful in monitoring suicide risk among children and adolescents, not only during normal times but also during emergencies such as school closures. However, because internet search volume includes external factors unrelated to suicidal characteristics, it may not be easy to generalize the findings and incorporate them into national and municipal suicide prevention strategies. Further empirical studies will be needed in the future.

Acknowledgments

We thank all the members of the Japan Suicide Countermeasures Promotion Center for their warm guidance and encouragement throughout this research, as well as Professor Keita Yamauchi of Keio University for his important advice. This work was supported by JSPS KAKENHI (grants JP21H04403 and JP24K23761). The funding sources had no role in the study design, data collection, data analysis, data interpretation, or preparation of the manuscript.

Data Availability

Data supporting this study's results require permission from the Japan Center for Suicide Prevention. Requests to access these datasets should be directed to TA, the corresponding author.

Authors' Contributions

TA contributed to the conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing original draft, writing-review and editing, visualization, supervision, and project administration. HT performed the formal analysis and writing-review and editing. AW assisted the interpretation of the findings, formal analysis, and writing-review and editing. YS performed the supervision.

Conflicts of Interest

None declared.

References

1. Ministry of Health, Labour and Welfare. Annual Report on Suicide Prevention. 2020. URL: https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/hukushi_kaigo/seikatsuhogo/jisatsu/jisatsuhakusyo2022.html [accessed 2024-09-14]
2. Ministry of Education, Culture, Sports, Science and Technology. Conference of Research Collaborators on Suicide Prevention for Students. 2021. URL: https://www.mext.go.jp/content/20210216-mxt_jidou01-000012837_009.pdf [accessed 2024-09-14]
3. Cusimano MD, Sameem M. The effectiveness of middle and high school-based suicide prevention programmes for adolescents: a systematic review. *Inj Prev*. 2011;17(1):43-49. [doi: [10.1136/ip.2009.025502](https://doi.org/10.1136/ip.2009.025502)] [Medline: [21059602](https://pubmed.ncbi.nlm.nih.gov/21059602/)]
4. Katz C, Bolton S, Katz LY, Isaak C, Tilston-Jones T, Sareen J, et al. Swampy Cree Suicide Prevention Team. A systematic review of school-based suicide prevention programs. *Depress Anxiety*. 2013;30(10):1030-1045. [doi: [10.1002/da.22114](https://doi.org/10.1002/da.22114)] [Medline: [23650186](https://pubmed.ncbi.nlm.nih.gov/23650186/)]
5. Sueki H, Ito J. Suicide prevention through online gatekeeping using search advertising techniques: a feasibility study. *Crisis*. 2015;36(4):267-273. [doi: [10.1027/0227-5910/a000322](https://doi.org/10.1027/0227-5910/a000322)] [Medline: [26440623](https://pubmed.ncbi.nlm.nih.gov/26440623/)]
6. Barros JM, Melia R, Francis K, Bogue J, O'Sullivan M, Young K, et al. The validity of google trends search volumes for behavioral forecasting of national suicide rates in Ireland. *Int J Environ Res Public Health*. 2019;16(17):3201. [FREE Full text] [doi: [10.3390/ijerph16173201](https://doi.org/10.3390/ijerph16173201)] [Medline: [31480718](https://pubmed.ncbi.nlm.nih.gov/31480718/)]

7. Capron DW, Andel R, Voracek M, Till B, Niederkrotenthaler T, Bauer BW, et al. Time-series analyses of firearm-related Google searches and U.S. suicide rates 2004-2016. *Suicide Life Threat Behav*. 2021;51(3):554-563. [doi: [10.1111/sltb.12728](https://doi.org/10.1111/sltb.12728)] [Medline: [33426750](https://pubmed.ncbi.nlm.nih.gov/33426750/)]
8. Halford EA, Lake AM, Gould MS. Google searches for suicide and suicide risk factors in the early stages of the COVID-19 pandemic. *PLoS One*. 2020;15(7):e0236777. [FREE Full text] [doi: [10.1371/journal.pone.0236777](https://doi.org/10.1371/journal.pone.0236777)] [Medline: [32706835](https://pubmed.ncbi.nlm.nih.gov/32706835/)]
9. Jimenez A, Santed-Germán MA, Ramos V. Google searches and suicide rates in Spain, 2004-2013: correlation study. *JMIR Public Health Surveill*. 2020;6(2):e10919. [FREE Full text] [doi: [10.2196/10919](https://doi.org/10.2196/10919)] [Medline: [32281540](https://pubmed.ncbi.nlm.nih.gov/32281540/)]
10. Lee JY. Search trends preceding increases in suicide: a cross-correlation study of monthly Google search volume and suicide rate using transfer function models. *J Affect Disord*. 2020;262:155-164. [doi: [10.1016/j.jad.2019.11.014](https://doi.org/10.1016/j.jad.2019.11.014)] [Medline: [31733460](https://pubmed.ncbi.nlm.nih.gov/31733460/)]
11. Lopez-Agudo LA. The association between internet searches and suicide in Spain. *Psychiatry Res*. 2020;291:113215. [doi: [10.1016/j.psychres.2020.113215](https://doi.org/10.1016/j.psychres.2020.113215)] [Medline: [32559673](https://pubmed.ncbi.nlm.nih.gov/32559673/)]
12. Adem Doganer ZJZ. An infodemiological analysis of google trends in COVID-19 outbreak: predict case numbers and attitudes of different societies. *Journal of Database Management*. 2021;32(2):19. [doi: [10.4018/jdm.2021040101](https://doi.org/10.4018/jdm.2021040101)]
13. Mavragani A, Gkillas K. COVID-19 predictability in the United States using Google trends time series. *Sci Rep*. 2020;10(1):20693. [FREE Full text] [doi: [10.1038/s41598-020-77275-9](https://doi.org/10.1038/s41598-020-77275-9)] [Medline: [33244028](https://pubmed.ncbi.nlm.nih.gov/33244028/)]
14. Peng Y, Li C, Rong Y, Pang CP, Chen X, Chen H. Real-time prediction of the daily incidence of COVID-19 in 215 countries and territories using machine learning: model development and validation. *J Med Internet Res*. 2021;23(6):e24285. [FREE Full text] [doi: [10.2196/24285](https://doi.org/10.2196/24285)] [Medline: [34081607](https://pubmed.ncbi.nlm.nih.gov/34081607/)]
15. McCarthy MJ. Internet monitoring of suicide risk in the population. *J Affect Disord*. 2010;122(3):277-279. [FREE Full text] [doi: [10.1016/j.jad.2009.08.015](https://doi.org/10.1016/j.jad.2009.08.015)] [Medline: [19748681](https://pubmed.ncbi.nlm.nih.gov/19748681/)]
16. Yang AC, Tsai S, Huang NE, Peng C. Association of Internet search trends with suicide death in Taipei City, Taiwan, 2004-2009. *J Affect Disord*. 2011;132(1-2):179-184. [doi: [10.1016/j.jad.2011.01.019](https://doi.org/10.1016/j.jad.2011.01.019)] [Medline: [21371755](https://pubmed.ncbi.nlm.nih.gov/21371755/)]
17. Prime Minister's Office of Japan. The general principles of suicide prevention policy. URL: https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/hukushi_kaigo/seikatsuhogo/jisatsu/taikou_h290725.html [accessed 2024-09-14]
18. Sueki H. Does the volume of Internet searches using suicide-related search terms influence the suicide death rate: data from 2004 to 2009 in Japan. *Psychiatry Clin Neurosci*. 2011;65(4):392-394. [FREE Full text] [doi: [10.1111/j.1440-1819.2011.02216.x](https://doi.org/10.1111/j.1440-1819.2011.02216.x)] [Medline: [21569178](https://pubmed.ncbi.nlm.nih.gov/21569178/)]
19. Hagihara A, Miyazaki S, Abe T. Internet suicide searches and the incidence of suicide in young people in Japan. *Eur Arch Psychiatry Clin Neurosci*. 2012;262(1):39-46. [doi: [10.1007/s00406-011-0212-8](https://doi.org/10.1007/s00406-011-0212-8)] [Medline: [21505949](https://pubmed.ncbi.nlm.nih.gov/21505949/)]
20. Taira K, Hosokawa R, Itatani T, Fujita S. Predicting the number of suicides in Japan using Internet search queries: vector autoregression time series model. *JMIR Public Health Surveill*. 2021;7(12):e34016. [FREE Full text] [doi: [10.2196/34016](https://doi.org/10.2196/34016)] [Medline: [34823225](https://pubmed.ncbi.nlm.nih.gov/34823225/)]
21. Arendt F, Scherr S. Optimizing online suicide prevention: a search engine-based tailored approach. *Health Commun*. 2017;32(11):1403-1408. [doi: [10.1080/10410236.2016.1224451](https://doi.org/10.1080/10410236.2016.1224451)] [Medline: [27739876](https://pubmed.ncbi.nlm.nih.gov/27739876/)]
22. Tran US, Andel R, Niederkrotenthaler T, Till B, Ajdacic-Gross V, Voracek M. Low validity of Google trends for behavioral forecasting of national suicide rates. *PLoS One*. 2017;12(8):e0183149. [FREE Full text] [doi: [10.1371/journal.pone.0183149](https://doi.org/10.1371/journal.pone.0183149)] [Medline: [28813490](https://pubmed.ncbi.nlm.nih.gov/28813490/)]
23. FAQ about Google trends data. Google. 2023. URL: <https://support.google.com/trends/answer/4365533?hl=en> [accessed 2024-09-14]
24. Fahey RA, Boo J, Ueda M. Covariance in diurnal patterns of suicide-related expressions on Twitter and recorded suicide deaths. *Soc Sci Med*. 2020;253:112960. [FREE Full text] [doi: [10.1016/j.socscimed.2020.112960](https://doi.org/10.1016/j.socscimed.2020.112960)] [Medline: [32251933](https://pubmed.ncbi.nlm.nih.gov/32251933/)]
25. Matsubayashi T, Ueda M, Yoshikawa K. School and seasonality in youth suicide: evidence from Japan. *J Epidemiol Community Health*. 2016;70(11):1122-1127. [doi: [10.1136/jech-2016-207583](https://doi.org/10.1136/jech-2016-207583)] [Medline: [27225682](https://pubmed.ncbi.nlm.nih.gov/27225682/)]
26. Sato K, Mano T, Iwata A, Toda T. Need of care in interpreting Google trends-based COVID-19 infodemiological study results: potential risk of false-positivity. *BMC Med Res Methodol*. 2021;21(1):147. [FREE Full text] [doi: [10.1186/s12874-021-01338-2](https://doi.org/10.1186/s12874-021-01338-2)] [Medline: [34275447](https://pubmed.ncbi.nlm.nih.gov/34275447/)]
27. Trapletti. Time series analysis and computational finance. *Tseries*. 2019. URL: <https://cran.r-project.org/web/packages/tseries/tseries.pdf> [accessed 2024-09-14]
28. Pfaff B. VAR, SVAR and SVEC models: implementation within R package vars. *Journal of Statistical Software*. 2008;27(4):1-32. [FREE Full text]
29. Armitage P, Berry G, Matthews JNS. *Statistical Methods in Medical Research*. Malden, MA. Wiley-Blackwell; 2001.
30. Gabennesch H. When promises fail: a theory of temporal fluctuations in suicide. *Social Forces*. 1988;67(1):129-145.
31. Minister of Education, Culture, Sports, Science and Technology Koichi Hagiuda Press Conference Transcript. 2020. URL: https://www.mext.go.jp/b_menu/daijin/detail/mext_00108.html [accessed 2023-07-31]
32. Minister of Education, Culture, Sports, Science and Technology Koichi Hagiuda Press Conference Transcript. 2020. URL: https://www.mext.go.jp/b_menu/daijin/detail/mext_00108.html [accessed 2023-07-31]

33. Castellví P, Miranda-Mendizábal A, Alayo I, Parés-Badell O, Almenara J, Alonso I, et al. Assessing the relationship between school failure and suicidal behavior in adolescents and young adults: a systematic review and meta-analysis of longitudinal studies. *School Mental Health*. 2020;12(3):429-441. [doi: [10.1007/s12310-020-09363-0](https://doi.org/10.1007/s12310-020-09363-0)]
34. Epstein S, Roberts E, Sedgwick R, Polling C, Finning K, Ford T, et al. School absenteeism as a risk factor for self-harm and suicidal ideation in children and adolescents: a systematic review and meta-analysis. *Eur Child Adolesc Psychiatry*. 2020;29(9):1175-1194. [FREE Full text] [doi: [10.1007/s00787-019-01327-3](https://doi.org/10.1007/s00787-019-01327-3)] [Medline: [30989389](https://pubmed.ncbi.nlm.nih.gov/30989389/)]
35. Tomek S, Burton S, Hooper LM, Bolland A, Bolland J. Suicidality in black American youth living in impoverished neighborhoods: is school connectedness a protective factor? *School Mental Health*. 2018;10(1):1-11. [doi: [10.1007/s12310-017-9241-4](https://doi.org/10.1007/s12310-017-9241-4)]
36. Choi WS, Han J, Hong HJ. Association between Internet searches related to suicide/self-harm and adolescent suicide death in South Korea in 2016-2020: secondary data analysis. *J Med Internet Res*. 2023;25:e46254. [FREE Full text] [doi: [10.2196/46254](https://doi.org/10.2196/46254)] [Medline: [37079349](https://pubmed.ncbi.nlm.nih.gov/37079349/)]
37. Lee D, Lee H, Choi M. Examining the relationship between past orientation and US suicide rates: an analysis using big data-driven Google search queries. *J Med Internet Res*. 2016;18(2):e35. [FREE Full text] [doi: [10.2196/jmir.4981](https://doi.org/10.2196/jmir.4981)] [Medline: [26868917](https://pubmed.ncbi.nlm.nih.gov/26868917/)]
38. Adam-Troian J, Arciszewski T. Absolutist words from search volume data predict state-level suicide rates in the United States. *Clinical Psychological Science*. 2020;8(4):788-793. [doi: [10.1177/2167702620916925](https://doi.org/10.1177/2167702620916925)]
39. Sueki H, Takahashi AT, Ito J. Changes in suicide ideation among users of online gatekeeping using search-based advertising. *Arch Suicide Res*. 2023;27(4):1339-1350. [doi: [10.1080/13811118.2022.2131491](https://doi.org/10.1080/13811118.2022.2131491)] [Medline: [36237123](https://pubmed.ncbi.nlm.nih.gov/36237123/)]
40. Scherr S, Haim M, Arendt F. Equal access to online information? Google's suicide-prevention disparities may amplify a global digital divide. *New Media & Society*. 2018;21(3):562-582. [doi: [10.1177/1461444818801010](https://doi.org/10.1177/1461444818801010)]
41. Arendt F, Haim M, Scherr S. Investigating Google's suicide-prevention efforts in celebrity suicides using agent-based testing: a cross-national study in four European countries. *Soc Sci Med*. 2020;262:112692. [doi: [10.1016/j.socscimed.2019.112692](https://doi.org/10.1016/j.socscimed.2019.112692)] [Medline: [32057542](https://pubmed.ncbi.nlm.nih.gov/32057542/)]
42. Baran A, Gerstner R, Ueda M, Gmitrowicz A. Implementing real-time data suicide surveillance systems. *Crisis*. 2021;42(5):321-327. [FREE Full text] [doi: [10.1027/0227-5910/a000829](https://doi.org/10.1027/0227-5910/a000829)] [Medline: [34528451](https://pubmed.ncbi.nlm.nih.gov/34528451/)]
43. Live life: an implementation guide for suicide prevention in countries. WHO. 2021. URL: <https://apps.who.int/iris/handle/10665/341726> [accessed 2023-07-31]
44. Niederkrotenthaler T, Stack S, Till B, Sinyor M, Pirkis J, Garcia D, et al. Association of increased youth suicides in the United states with the release of 13 reasons why. *JAMA Psychiatry*. 2019;76(9):933-940. [FREE Full text] [doi: [10.1001/jamapsychiatry.2019.0922](https://doi.org/10.1001/jamapsychiatry.2019.0922)] [Medline: [31141094](https://pubmed.ncbi.nlm.nih.gov/31141094/)]

Abbreviations

- AIC:** Akaike's information criterion
ADF test: augmented Dickey-Fuller test
RSV: relative search value
VAR: vector autoregressive

Edited by A Coristine; submitted 09.08.23; peer-reviewed by A Jimenez, H Kautz, A AL-Asadi; comments to author 24.02.24; revised version received 20.05.24; accepted 23.08.24; published 21.10.24

Please cite as:

Arai T, Tsubaki H, Wakano A, Shimizu Y
Association Between School-Related Google Trends Search Volume and Suicides Among Children and Adolescents in Japan During 2016-2020: Retrospective Observational Study With a Time-Series Analysis
J Med Internet Res 2024;26:e51710
URL: <https://www.jmir.org/2024/1/e51710>
doi: [10.2196/51710](https://doi.org/10.2196/51710)
PMID:

©Takahiro Arai, Hiroe Tsubaki, Ayako Wakano, Yasuyuki Shimizu. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 21.10.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.