Journal of Adolescent Health 59 (2016) 668–673

ELSEVIER

JOURNAL OF ADOLESCENT HEALTH

www.jahonline.org

Original article

Data Mining of Web-Based Documents on Social Networking Sites That Included Suicide-Related Words Among Korean Adolescents

Juyoung Song, Ph.D. [a], Tae Min Song, Ph.D. [b], Dong-Chul Seo, Ph.D. [c,*], and Jae Hyun Jin, M.S. [b]

[a] Department of Administration of Justice, Pennsylvania State University, Schuylkill Haven, Pennsylvania
[b] Statistics and Information Department, Korea Institute for Health and Social Affairs, Sejong City, South Korea
[c] Department of Applied Health Science, Indiana University School of Public Health, Bloomington, Indiana

Article history: Received March 11, 2016; Accepted July 28, 2016
Keywords: Social big data; Suicide; Data mining; Adolescent health

A B S T R A C T

Purpose: To investigate online search activity of suicide-related words in South Korean adolescents through data mining of social media Web sites as the suicide rate in South Korea is one of the highest in the world.
Methods: Out of more than 2.35 billion posts for 2 years from January 1, 2011 to December 31, 2012 on 163 social media Web sites in South Korea, 99,693 suicide-related documents were retrieved by Crawler and analyzed using text mining and opinion mining. These data were further combined with monthly employment rate, monthly rental prices index, monthly youth suicide rate, and monthly number of reported bully victims to fit multilevel models as well as structural equation models.
Results: The link from grade pressure to suicide risk showed the largest standardized path coefficient (beta = .357, $p < .001$) in structural models and a significant random effect ($p < .01$) in multilevel models. Depression was a partial mediator between suicide risk and grade pressure, low body image, victims of bullying, and concerns about disease. The largest total effect was observed in the grade pressure to depression to suicide risk. The multilevel models indicate about 27% of the variance in the daily suicide-related word search activity is explained by month-to-month variations. A lower employment rate, a higher rental prices index, and more bullying were associated with an increased suicide-related word search activity.
Conclusions: Academic pressure appears to be the biggest contributor to Korean adolescents' suicide risk. Real-time suicide-related word search activity monitoring and response system needs to be developed.

IMPLICATIONS AND CONTRIBUTION

This study investigated search activity of suicide-related words by data mining 163 social media Web sites frequented by South Korean adolescents over a two-year period. The analysis was linked to data on employment rate, rental prices, and bullying victimization, finding that academic pressure was the greatest contributor to suicide risk in this population.

Suicide is a global leading cause of death among adolescents and young adults. Yet, these populations are difficult to study using traditional public health monitoring. Many youth are not included in surveys or clinical screening because they are not covered by existing sampling frames or they forgo medical care. Herein, we pilot a novel method for assessing the risks of suicide by monitoring the search queries on Korean social media Web pages that are commonly used with adolescents and young adults.

The suicide rate in South Korea was 29.1 for every 100,000 people in 2012, which was the highest among the 34 countries that comprise the Organization for Economic Cooperation and Development (OECD) [1]. In 2013, suicide was the leading cause of death among adolescents and young adults ages up to 39 years

Conflicts of Interest: There are no actual or potential conflicts of interest involved with this study.

* Address correspondence to: Dong-Chul Seo, Ph.D., Department of Applied Health Science, Indiana University School of Public Health, Suite 116, 1025, E. 7th Street, Bloomington, IN 47405.
E-mail address: seo@indiana.edu (D.-C. Seo).

1054-139X/
http://dx.doi.org/10.1016/j.jadohealth.2016.07.025

in South Korea [2]. Indeed, according to the 2013 Survey by the Korean Ministry of Health and Welfare, 30.9% of adolescents had experienced depression; 16.6% had thought about committing suicide; 5.7% had planned to commit suicide; and 4.1% had attempted to commit suicide [3].

Many studies showed that emotional imbalance such as depression is associated with suicidal risk [4–6]. Researchers found that 4%–15% of the variance in such depression can be accounted for by stress [7,8]. As a result, the stress-vulnerability model was suggested in an attempt to have a better understanding of the relationship between stress and depression [9]. This theory posits that those who have innate or acquired vulnerability to stress tend to develop emotional imbalance such as mental disease or depression when exposed to various stressors [10]. Alloy et al. advocate the cognitive diathesis-stress theories of depression, adding that stressors interact with social and psychological factors, which can lead to development of depression, addiction to alcohol use, or even extreme choices such as suicide [11]. In addition to this individualistic approach linking stress and mental health or suicide, macro-level social approach has been suggested. Agnew [12] claimed in his macro-level strain theory that financial deficiency, family breakdown, child abuse, and social disunion help build up aversive stimuli and that the affected individuals' anger exacerbates the level of tension and negative impact in the community, which could trigger emotional imbalance and even suicidal ideation among the affected individuals.

In terms of explanations on adolescent suicidal behavior, the cognitive-behavioral model of adolescent suicidal behavior indicates that certain risk factors serve as a trigger which, in turn, leads to suicidal ideation and attempts through the interactions between cognitive distortions, dysregulation, and maladaptive behaviors [13]. The interpersonal psychological theory of attempted and completed suicide adds that suicidal behavior is caused by acquired capability of committing suicide, feeling burdensome, and thwarted belongingness [14]. As for Korean adolescents who commit suicide, academic pressure to perform well on tests, peer support, and bullying are often regarded as major contributors to suicidal ideation and attempts [15–17]. Other contributors include financial hardship [18,19], low body image [20], and disease [21,22].

It is important to distinguish suicidal ideation or planning from suicidal attempts and behavior. Researchers argue that suicidal ideation and planning almost always precede completed suicides [23]. Therefore, it is critically important to identify individuals with suicidal ideation or who are planning to commit suicide beforehand and intervene at the right time. As almost all the Korean adolescents have access to the Internet and many of them tend to share their personal stories and even emotional events with others on social networking sites, social media Web sites may provide a great opportunity for monitoring suicidal ideation and making timely interventions against suicide in adolescents.

The proportion of households with access to the Internet in South Korea was 97% in 2011, the highest among the 34 OECD countries [24] and the proportion of South Korean adolescents who used social networking service more than once a day was 59.2% in 2014 [25]. The exponential growth of mobile Internet and explosive social networking service use provides a new opportunity for research as those create a wealth of digital data [26]. This study was conducted to investigate online search activity of suicide-related words in South Korean adolescents through data mining of social media Web sites commonly used in South Korea. The results of this study may be utilized to help establish an online automatic monitoring system (e.g., a robot or artificial intelligence) that can monitor and capture online search activity of suicide-related words and provide on a real-time basis ballooned texts of helpful information to suicidal individuals.

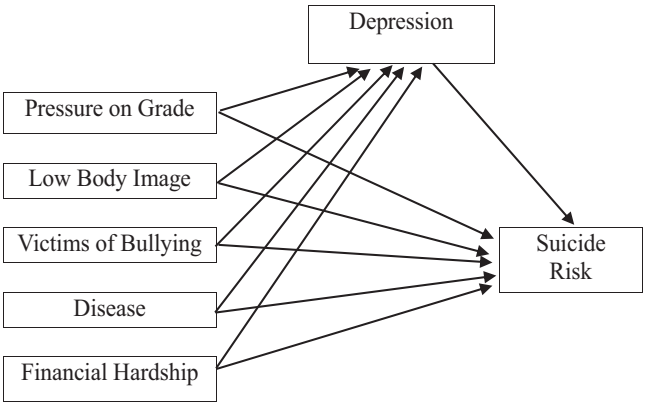

Figure 1. Conceptualization of analysis models.

Methods

Analytical framework and hypotheses

Based on prior research findings in the literature, the investigators identified six major potential suicide risk factors among Korean adolescents [16,18–20,27]. They were depression [27], pressure on grade [16], low body image [20], being a victim of bullying [18], having a disease with serious physical symptoms [18], and household financial hardship [19]. As shown in Figure 1, depression was hypothesized as both an exogenous predictor for suicide risk and a mediator between other predictors and suicide risk. Besides this structural model at the individual level, given the socioecological influence on suicide risk especially among adolescents [28], this study conceptualized a multilevel model where these individual potential suicide risk factors are daily characteristics nested under monthly macro or aggregate characteristics, such as monthly employment rate, monthly rental prices index, monthly youth suicide rate, and monthly number of reported bully victims. To avoid multicollinearity, the individual-level bully variable was excluded from the multilevel model. It was hypothesized that (1) each individual suicide risk factor (i.e., grade pressure, low body image, victims of bullying, concerns about disease, concerns about financial hardship, and depression) affects Korean adolescents' suicide-related words search activity and that (2) such online search is influenced by the aforementioned monthly macro or aggregate characteristics.

Sample

The sample for this study was Web-based individual documents (i.e., buzz) that contained suicide-related words among adolescents, which were collected from 163 social media Web sites in South Korea (i.e., Twitter, four Internet blogs, two online cafes, 13 message boards, and 143 online news sites and other social network services) from January 1, 2011 when Twitter initiated Korean language service to December 31, 2012.

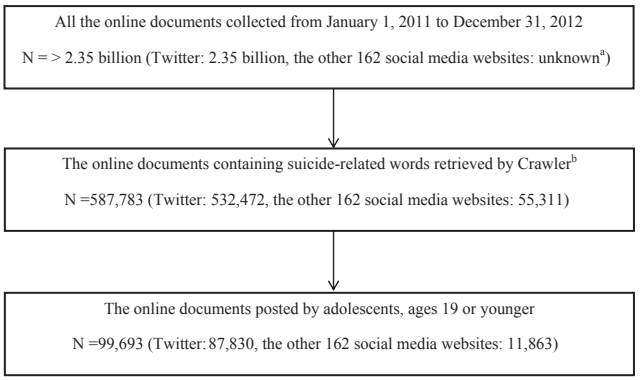

Figure 2. Flowchart of sample selection. [a]Whereas the entire number of online posts during the study period was available for Twitter due to the full crawl retrieval method, it was not for other social media Web sites due to the focused crawl retrieval method. [b]Suicide-related words included suicide, kill oneself, jump to one's death, drown oneself, hang oneself, and other forms of suicide as well as suffixes of these keywords in Korean. Noise was minimized by excluding the documents that contained such stop words as "score an own goal" and "research institute on committing suicide" because the first two syllables of these words are written and pronounced exactly the same as suicide in Korean.

As shown in Figure 2, out of more than 2.35 billion posts for 2 years from January 1, 2011 to December 31, 2012, a total of 587,783 contained suicide-related words, such as suicide, kill oneself, jump to one's death, drown oneself, hang oneself, and other forms of suicide. Noise was minimized by excluding the documents that contained such stop words as "score an own goal" and "research institute on committing suicide" because the first two syllables of these words are written and pronounced exactly the same as suicide in Korea. To limit the study sample to adolescents, only the documents that contained such descriptors as adolescents, secondary school students, or ages 19 or younger were selected and included in the sample. The final analysis sample was composed of 99,693 documents that were posted by adolescents, aged 19 or younger.

This sample section was carried out by a leading Korean telecommunications company that was contracted to conduct data collection and treatment. The telecommunications company employed a crawler (i.e., a set of programs that crisscross the Internet autonomously in order to retrieve Web-based content) [29] to collect documents from the selected channels; categorized and cleaned unstructured data through text mining; and converted the categorized unstructured data into binary representation in preparation for statistical analysis. The "21st Century Sejong Project" corpus developed by the National Institute of Korean Language was used as the lexicon for text mining. The Institutional Review Board of the authors' institutions approved the study protocol. No personally identifiable information was available in the collected data.

Measures

The dependent variable of the present study was the search activity of suicide-related words. Such dependent variable was conceptualized as suicide risk on structural and multilevel models in the present study. After suicide-related words (buzz) among adolescents were collected from the social media Web sites, they were transformed and coded into structured data through text mining and opinion mining as follows: 1 for expressions *approving of suicide* (e.g., "will commit suicide," "thinking of committing suicide," "suicide is a way of solving problems," and "it is not difficult to commit suicide") or *neutral to suicide* (those neither approving of suicide nor disapproving of suicide) and 0 for expressions *disapproving of suicide* (e.g., "committing suicide is sad," "committing suicide is bad," and "committing suicide is shocking"). Expression either *approving of suicide* or *neutral to suicide* was conceptualized as unprotected from suicide risk (i.e., indicative of suicide risk) and expression *disapproving of suicide* was conceptualized as protective from suicide risk.

The independent variables included expressions on grade pressure, low body image, bully victims, concerns about disease, and concerns about financial hardship. The search activity of depression-related words served as a potential mediator. Grade pressure included such keywords as school grade, exam, college entrance, study, academic stress, scholastic aptitude test, college entrance examination, midterm exam, and final exam. Low body image included such keywords as appearance, body weight, body shape, chubby, obese, overweight, shape or width of face, and ugly. Concerns about disease included such keywords as disease, sick, ill, disabled, terminal disease, cancer, and hospitalization. Concerns about financial hardship included such keywords as bankrupt, broke, unemployed, fired, food stamps, poverty, and economic problem. Depression-related words included depression, manic-depressive illness, bipolar disorder, blues, and antidepressant. When each of these keywords was present in the retrieved online documents, it was coded as 1. Otherwise, it was coded as 0.

As for the multilevel model, information on the monthly employment rate, monthly rental prices index, and monthly youth suicide rate was collected from the Bank of Korea (i.e., the Korean Federal Reserve Bank) as level-2 variables. Additionally, the information on the monthly amount of phone consultations for adolescent victims of bullying ("monthly number of bully victims" hereafter) was collected as a level-2 variable from the Korean Youth Help Line Service (the phone number of 1,388 without any area code).

Analysis

Descriptive properties of the distribution of all the major variables were examined and appropriately transformed (i.e., log transformation) when the variables departed substantially from normality in terms of skewness and kurtosis. As for the structural model, the model fit was tested using incremental fit indices, such as normed fit index, comparative fit index, and Tucker-Lewis index as well as absolute fit indices, such as χ^2 and RMSEA (root mean square error of approximation). An incremental fit index greater than .90 is generally considered as an acceptable fit to the data [30]. RMSEA is a goodness of fit index that was developed to compensate for the drawback of χ^2 that is a function of sample size [31]. An RMSEA value less than .05 indicates a *very good* fit; a value between .05 and .08 a *good* fit; and a value greater than .10 a *poor* fit [32]. Sobel test was used to examine significance of indirect effect [33]. To test a potential mediating effect of depression, we used a method suggested by Heir et al. [34]. Parameter estimation was conducted using the restricted maximum likelihood method [35]. We used SPSS 22 (IBM Corporation, Armonk, NY) for descriptive statistics and AMOS 22.0 (IBM) for structural equation model analysis.

Table 1
Parameter estimates of the structural model

Log-transformed variables	B	Beta	SE	CR	*p*
Grade→depression	.139	.117	.048	2.873	.004
Appearance→depression	.261	.102	.104	2.518	.012
Bullying→depression	−.115	−.080	.054	−2.145	.032
Disease→depression	.162	.094	.074	2.198	.028
Finance→depression	−.047	−.023	.080	−.583	.560
Depression→suicide risk	.096	.138	.020	4.762	<.001
Grade→suicide risk	.296	.357	.027	11.124	<.001
Finance→suicide risk	.148	.106	.044	3.390	<.001
Disease→suicide risk	.200	.166	.040	4.944	<.001
Appearance→suicide risk	.266	.149	.057	4.664	<.001
Bullying→suicide risk	.104	.103	.029	3.536	<.001

B: unstandardized path coefficient, beta: standardized path coefficient, Grade: grade pressure, Appearance: low body image, Bullying: victims of bullying, Disease: concerns about disease, Finance: concerns about financial hardship, Suicide Risk: daily amount of online searches of suicide-related words.
CR = critical ratio; SE = standard error.

As for the multilevel model, we used HLM 7.0 (Scientific Software International, Inc., Skokie, IL) for 2-level analysis. Null model (model 1) was first constructed to examine the extent to which daily search activity of suicide-related words varied from month to month. Random coefficient regression model (model 2) was constructed by adding to model 1 the five individual potential suicide risk factors, such as grade pressure, low body image, concerns about disease, concerns about financial hardship, and depression. Each of these variables was allowed to vary by month. Then, level-2 variables, such as monthly employment rate, monthly rental prices index, monthly youth suicide rate, and monthly number of bully victims, were added to model 2 as predictors of daily search activity of suicide-related words. The monthly employment rate and the rental prices index were grand-mean centered. At this final model, level-1 variables with nonsignificant random effects at model 2 were not allowed to vary by month.

Results

Structural model of suicide risk

A total of 23.1% in 2011 and 22.2% in 2012 of online expressions in the present study were indicative of suicide risk. These results are just 3%–4% different from the statistics published by the Korean Youth Risk Behavioral Surveillance System administered by the Korean Centers for Disease Control and Prevention where 19.6% of adolescents showed suicidal ideation in 2011 and 18.3% in 2012 [36]. All the computed goodness of fit indices on the structural model of suicide risk indicated a good fit to the data: $\chi^2 = 7.57$ ($df = 5$, $p = .182$), normed fit index = .992, comparative fit index = .997, Tucker-Lewis index = .984, and RMSEA = .027. Individual unstandardized and standardized path coefficients along with critical ratios are shown in Table 2. All the paths were significant at the .05 level and showed the expected directionality except the link from concerns about financial hardship to depression. Interestingly, the link from grade pressure to suicide risk showed the largest standardized path coefficient. As shown in Tables 1 and 2, depression was a partial mediator between suicide risk and grade pressure, low body image, victims of bullying, and concerns about disease. The largest total effect was observed in the grade pressure to depression to suicide risk. Sensitivity analysis was conducted with several reduced samples where less-than-obvious keywords were dropped. The results did not show substantial changes in terms of directionality, magnitudes, and statistical significance of path coefficients.

Table 2
Mediation effect of depression on suicide risk

Path	Total effect	Direct effect	Indirect effect[a]
Grade→depression→suicide risk	.541**	.514**	.027**
Disease→depression→suicide risk	.446**	.417**	.029**
Appearance→depression→suicide risk	.390**	.359**	.031**
Bullying→depression→suicide risk	.224**	.232**	−.008

Grade: grade pressure, Appearance: low body image, Bullying: victims of bullying, Disease: concerns about disease, Suicide Risk: daily search activity of suicide-related words. The coefficients are standardized path coefficients. The variable "concerns about financial hardship" is not shown here because it did not have a significant association with depression (see Table 2).
**$p < .01$.
[a] Based on Sobel test.

Multilevel model of suicide risk

As is shown in Table 3, there was a significant month-to-month variation in the outcome variable ($\chi^2 = 262.12$, $p < .01$). The intraclass correlation coefficient was .269 (.07/[.07 + .19]), which means that about 27% of the variance in the daily search activity of suicide-related words is explained by month-to-month variations in the search. This justifies the use of multilevel analytical approach to this investigation.

Table 3
Multilevel analysis of daily amount of search activity of suicide-related words

Model	Model 1		Model 2		Model 3	
Parameter						
Fixed effect	Coef.	SE	Coef.	SE	Coef.	SE
Level 1						
Intercept	.62	.05***	.14	.05**	−.97	.40
Grade			.22	.05**	.17	.06**
Depression			.16	.03**	.21	.04**
Finance			.10	.04*	.09	.04*
Disease			.17	.04**	.17	.04**
Appearance			.29	.05**	.27	.06**
Level 2						
MER					−.00	.00**
MRPI					.04	.01**
MYSR					.05	.09
MVB					.95	.38*
Random effect	Variance	χ^2	Variance	χ^2	Variance	χ^2
Level 1 residual	.19		.17		.12	
Level 2 residual	.07	262.12**	.03	47.82**	.01	24.71
Grade			.05	94.03**	.04	83.00**
Depression			.01	36.01**	.01	36.64*
Finance			.01	20.67		
Disease			.01	31.50		
Appearance			.01	22.85		
ICC	.27		.20		.08	
Deviance	935.18		620.38		636.01	

Grade: pressure on grade, Appearance: low body image, Disease: concerns about disease, Finance: concerns about financial hardship.
**$p < .01$ and *$p < .05$.
Coef. = coefficient; ICC = intraclass correlation coefficient; MER = monthly employment rate; MRPI = monthly rental prices index; MVB = monthly victims of bullying; MYSR = monthly youth suicide rate; SE = standard error.

As hypothesized, each individual suicide risk factor significantly affected Korean adolescents' online search of suicide-related words ($p < .05$ for both model 2 and model 3). However, the hypothesis that such online search would be influenced by monthly employment rate, rental prices index, youth suicide rate, and number of bully victims was partially supported. A lower employment rate, a higher rental prices index, and more victims of bullying ($p < .05$), but not monthly youth suicide rate, were associated with increased online search of suicide-related words. In terms of random effect of individual variables, only the effect of grade pressure ($p < .01$) and depression ($p < .05$) on the search activity showed a significant monthly variation that may reflect seasonal effect.

Discussion

This study was the first attempt that investigated online search activity of suicide-related words in South Korean adolescents through data mining of social media Web sites for a 2-year period, which was based on the stress-vulnerability model [9] and Agnew's macro-level strain theory [12]. A noteworthy finding of the present study was that among the six significant individual suicide risk factors, such as grade pressure, low body image, victims of bullying, concerns about disease, concerns about financial hardship, and depression, grade pressure was the most important predictor for Korean adolescents' suicide risk in terms of effect size. Not only for the bivariate relation between individual risk factors and suicide risk but also for the depression-mediated relation, was the largest effect observed in grade pressure. This finding is alarming although previous studies [15,16] have alluded to this possibility. Indeed, many Korean adolescents report academic pressure to perform well on school tests and the nationwide scholastic assessment tests as one of the biggest intolerable stressors [16,21,37]. Given that all Korean high school students have only one chance to take the nationwide scholastic assessment tests at a set date in a given year, not surprisingly, buzz on grade pressure reached its peak around the test date and the random effect (i.e., monthly variation) of grade pressure on suicide risk was significant. In order to reduce the highest suicide rate of South Korea among the 34 OECD countries [1], the Korean society and government may want to explore and institutionalize ways to relieve students' academic pressure.

The findings of the present study affirm that adolescents do express their stress, depressive emotions, and even suicidal thoughts in cyber space and communicate these emotions and thoughts with other Internet users. Besides, they implicitly or explicitly reveal what drives them to think about committing suicide. Therefore, it might save a lot of lives if real-time online suicide-related word search activity monitoring and response system are in place where social media Web sites are constantly monitored against suicide-related words by robots and instant messages tailored to each suicidal risk factor are provided. For example, if the robot detects a phrase of thinking about committing suicide on a social media Web site, it might float a ballooned text informing a telephone number on which they can talk to someone who can provide counseling and support. The present study should be replicated to other settings and populations in order to accumulate more evidence base and should be expanded to develop such an online suicidal ideation response system.

It is well known that depression is an important suicide risk factor [11]. A finding of note in the present study was that depression was not only an exogenous predictor for suicide risk with seasonal effect but also a partial mediator between other predictors of suicide risk and the outcome variable (i.e., search activity of suicide-related words). Specifically, the present study indicates that grade pressure, low body image, victims of bullying, and concerns about disease make Korean adolescents feel depressed which, in turn, contributes to their search activity of suicide-related words, although these individual risk factors are also directly associated with such search activity. It is interesting that western literature [38] reports victims of bullying as a major risk factor for adolescent depression and suicidality, whereas the link from victims of bullying to depression to suicide risk in the present study showed the smallest total effect among the four mediation effects of depression. The largest mediation effect of depression was found in the link from grade pressure to depression to suicide risk. This indicates that cultural differences may exist in the adolescent suicidal risk factors. Future research would be desirable that investigates cultural differences in the relative magnitude of suicidal risk factors in adolescents.

Another finding of the present study that deserves mention is that monthly employment rate and monthly rental prices index were significant predictors for search activity of suicide-related words. This implies that macroeconomic hardship that might threaten adolescents' present or future disposable income may contribute to their suicidal risk as suggested by Agnew's macro-level strain theory [12]. Agnew [12] argued that community-level characteristics may add to macro-level strain that may make individuals vulnerable to negative consequences above and beyond individual proneness to such negativity. Unlike monthly employment rate and monthly rental prices index, monthly youth suicide rate failed to predict search activity of suicide-related words. This finding might appear counterintuitive, but a few speculative explanations can be made. First, online search activity of suicide-related words might not coincide with completed suicides. Rather, such online search very likely precedes the act of committing suicide, and the aggregated number of completed suicides in a given month might not reflect search activity of suicide-related words in that specific month. Second, even if the time distance between the search activity and suicidal attempt is narrow for completed suicides, not all those who searched suicide-related words would attempt committing suicide or lead to completed suicides. Thus, time discrepancy between the search activity and monthly youth suicide rate may occur. Finally, all the completed suicides are not likely to be necessarily accompanied by search activity of suicide-related words. More research is warranted to investigate the impact of macro-level characteristics on suicidal ideation above and beyond individual characteristics.

This research has limitations. First, analysis was conducted using group data instead of individual data. For this reason, caution is warranted in interpretation of the results at the individual level to avoid ecological fallacy [9]. Second, although all the computed goodness of fit indices of the structural model of suicide risk indicated a good fit to the data, there is a possibility that a better fitting structural model than the tested model could exist. Third, as with most correlational studies, the results of the present study cannot be utilized to ascertain the mechanism by which a number of risk factors affect adolescents' suicidal ideation or individual-level motivations of online search activity of suicide-related words. Fourth, there might be a number of

unmeasured confounding factors, especially spurious correlations. Finally, as the majority of online documents on the social networking sites do not leave information on age and sex, subgroup heterogeneity analysis by age and sex could not be conducted.

Despite these limitations, this study contributes to the literature as it is the first attempt to document and examine suicide risk factors using data mining of social media Web sites among adolescents in South Korea that shows one of the highest suicide rates in the world and at the same time one of the highest Internet access rates. This study was novel and innovative as it employed analysis of both nonstereotyped social media data and stereotyped data such as the Korean Youth Help Line Database in the investigation of adolescent suicide risk factors. The results of this study may be utilized to help establish an online automatic monitoring system (e.g., a robot or artificial intelligence) that can monitor and capture online search activity of suicide-related words and provide on a real-time basis ballooned texts of helpful information to suicidal individuals.

Acknowledgments

The funding source was not involved in any of the study design, data collection/analysis, interpretation of data, and writing of the report.

Funding Sources

This research was supported by a grant from the Korean Ministry of Science, ICT and Future Planning (MSIP), ICT R&D program (grant number: I0114-14-1016, Youth Crisis Forecasting and Real-time Response Technology Development by Big Data Analysis), the Creative Vitamin Project, and the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2015S1A3A2046566).

References

[1] Korean Ministry of Health and Welfare. OECD health data 2014. Seoul, Korea: Government Printing Office; 2014.
[2] Korean Administration of Statistics. Annual report on the causes of death statistics 2013. Seoul, Korea: Government Printing Office; 2014.
[3] Korean Ministry of Health and Welfare. The Korean Youth Risk Behavior Web-Based Survey. Seoul, Korea: Government Printing Office; 2013.
[4] Dempsey AG, Sulkowski ML, Nichols R, Storch EA. Differences between peer victimization in cyber and physical settings and associated psychosocial adjustment in early adolescence. Psychol Sch 2009;46:962–72.
[5] Butler D, Kift S, Campbell M. Cyber bullying in schools and the law: Is there an effective means of addressing the power imbalance. Elaw J 2009;16:84–114.
[6] Perren S, Dooley J, Shaw T, Cross D. Bullying in school and cyberspace: Associations with depressive symptoms in Swiss and Australian adolescents. Child Adolesc Psychiatry Ment Health 2010;4:1–10.
[7] Hammen C. Depression and cognitions about personal stressful life events. In: Alloy LB, ed. Cognitive Processes in Depression. New York, NY: The Guilford Press; 1988:77–108.
[8] Cohen S, Edwards JR. Personality characteristics as moderators of the relationship between stress and disorder. In: Meufeld RWJ, ed. Advanced in Investigation of Psychological Stress. New York, NY: Wiley; 1989:235–83.
[9] Song TM, Song JY, An JY, et al. Psychological and social factors affecting Internet search on suicide in Korea: A big data analysis of Google search trends. Yonsei Med J 2014;55:254–63.
[10] Lee HY, Ham BJ. Stress and mental illness. J Korean Med Assoc 2013;56:471–7.
[11] Abramson LY, Metalsky GI, Alloy LB. Hopelessness depression: A theory-based subtype of depression. Psychol Rev 1989;96:358–72.
[12] Agnew R. A general strain theory of community differences in crime rates. J Res Crime Delinq 1999;36:123–55.
[13] Kendall PC. Child and adolescent therapy: Cognitive-behavioral procedures. 3rd edition. New York, NY: The Guilford Press; 2005.
[14] Stellrecht NE, Gordon KH, Van Orden K, et al. Clinical applications of the interpersonal-psychological theory of attempted and completed suicide. J Clin Psychol 2006;62:211–22.
[15] Yoon MS, Lee HS. The relationship between depression, job preparing stress and suicidal ideation among college students: Moderating effect of problem drinking. Korean J Youth Stud 2012;19:109–37.
[16] Seo SJ, Jung MS. Effect of family flexibility on the idea of adolescent suicide: The senior year of high school boys. J Korean Contents Assoc 2013;13:262–74.
[17] Kim HJ, Noh JE. A causal analysis of suicidal impulse in the context of parents, friends, teachers and community support: Gender difference. Popul Assoc Korea 2011;34:135–62.
[18] Toprak S, Cetin I, Guven T, et al. Self-harm, suicidal ideation and suicide attempts among college. Psychiatry Res 2011;87:140–4.
[19] Kim SA. Effects of childhood stress, depression, and social support on middle school adolescent suicidal ideation. Korean J Fam Welfare 2009;14:5–27.
[20] Kim KM, Youm YS, Park YM. Impact of school violence on psychological well-being Korean students happiness and suicidal impulse. J Korean Contents Assoc 2013;13:236–47.
[21] King RA, Schwab-Stone M, Flisher AJ, et al. Psychosocial and risk behavior correlates of youth suicide attempts and suicidal ideation. J Am Acad Child Adolesc Psychaiatry 2001;40:837–46.
[22] Park E. The influencing factors on suicide attempt among adolescents in South Korea. J Korean Acad Nurs 2008;38:465–73.
[23] Hawton K, van Heeringen K. The international handbook of suicide and attempted suicide. New York: John Wiley & Sons; 2006.
[24] OECD. OECD factbook 2013: Economic, environmental and social statistics. Paris, France: OECD Publishing; 2013.
[25] Korean Ministry of Science, ICT and Future Planning. The 2014 survey on the Internet usage. Seoul, Korea: Government Printing Office; 2014.
[26] Song TM. Development of suicide search prediction model through analysis of social big data. Health Welfare Pol Forum 2013;202:74–86.
[27] Konick LC, Gutierrez PM. Testing a model of suicide ideation in college students. Suicide Life Threat Behav 2005;35:181–92.
[28] Hankin BL, Roberts J, Gotlib IH. Elevated self-standards and emotional distress during adolescence: Emotional specificity and gender differences. Cogn Ther Res 1997;21:663–79.
[29] Stassopoulou A, Dikaiakos MD. Crawler detection: A Bayesian approach. In: ICISP 2006 Proceedings of the International Conference on Internet Surveillance and Protection. Washington, DC: IEEE Computer Society; 2006:1–16.
[30] Hu LT, Bentler PM. Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equ Model 1999;6:1–55.
[31] Kline RB. Principles and practice of structural equation modeling. 3rd edition. New York, NY: Guilford Press; 2010.
[32] Bollen KA, Long JS. Testing structural equation models. Newbury Park, CA: Sage; 1993.
[33] Preacher KJ, Hayes AF. SPSS and SAS procedures for estimating indirect effects in simple mediation models. Behav Res Methods Intsrum Comput 2004;36:717–31.
[34] Hair JF Jr, Black WC, Babin BJ, et al. Multivariate data analysis. Upper Saddle River, NJ: Pearson Prentice Hall; 2006.
[35] Raudenbush SW, Bryk AS. Hirearchical linear Models: Applications and data analysis methods. Thousand Oaks, CA: Sage; 2002.
[36] Korea Centers for Disease Control and Prevention. The statistics of 2015 Korean Youth Risk Behavioral Surveillance Data. Seoul, Korea: Government Printing Office; 2015.
[37] Korean Administration of Statistics. Social indicator 2013. Seoul, Korea: Government Printing Office; 2014.
[38] Klomek AB, Marrocco F, Kleinman M, et al. Bullying, depression, and suicidality in adolescents. J Am Acad Child Adolesc Psychiatry 2007;46:40–9.