DIGITAL HEALTH

Understanding and comparing risk factors and subtypes in South Korean adult and adolescent women's suicidal ideation or suicide attempt using survey and social media data

Donghun Kim[1] (iD) **, Ting Jiang[1], Ji Hyun Baek[2],
Sou Hyun Jang[3] and Yongjun Zhu[1]**

Abstract

Objective: This study aimed to investigate the similarities and differences in risk factors for suicide among adult and adolescent women in South Korea and identify subtypes of suicidal ideation or suicide attempt in each group.

Methods: Multifaceted data were collected and analyzed by linking survey and social media data. Interpretable machine learning models were constructed to predict suicide risk and major risk factors were extracted by investigating their feature importance. Additionally, subtypes of suicidal adult and adolescent women were identified and explained using risk factors.

Results: The risk factors for adult women were primarily related to mental disorders, while those for adolescent women were primarily related to interpersonal experiences and needs. Two subtypes of suicidal adult women were one with high psychiatric symptoms and mental disorders of them and/or their families and the other with excessive social media use and high online victimization. Two subtypes of suicidal adolescent women were one with high psychiatric symptoms, high ACEs, and high social connectedness, and the other with frequent social media use, high online sexual victimization, and high social assurance.

Conclusions: These findings enable a stratified and targeted understanding of suicide in women and help develop customized suicide prevention plans in South Korea.

Keywords

Feature analysis, suicide in women, risk factors, subtype analysis, South Korea

Submission date: 17 August 2023; Acceptance date: 29 April 2024

Introduction

Background

Suicide is a significant health problem among women. Although global mortality by suicide is three times higher in men than in women, the suicide attempt rate is three times higher in women than in men.[1] Additionally, women are more likely to have suicidal ideation than men[2] and are more vulnerable to mental disorders such as depression which can result in suicide.[3] This health issue is important to both adult and adolescent women. In South Korea, which has the highest suicide rate among the Organization for Economic Co-operation and Development (OECD) countries,[4] adult women thought about suicide twice as much as adult men and adolescent women attempted suicide twice as often as adolescent men.[5] According to Statistics Korea,[6,7] suicide is the

[1]Department of Library and Information Science, Yonsei University, Seoul, Republic of Korea
[2]Department of Psychiatry, Sungkyunkwan University School of Medicine, Samsung Medical Center, Seoul, Republic of Korea
[3]Department of Sociology, Korea University, Seoul, Republic of Korea

Corresponding author:
Yongjun Zhu, Department of Library and Information Science, Yonsei University, 50 Yonsei-ro Seodaemun-gu, Seoul, 03722, Republic of Korea.
Email: zhu@yonsei.ac.kr

leading cause of death among women aged from 10 to 40, and regarding the suicidal ideation, the prevalence was 16.1% among adolescent women, which is about twice the rate of adolescent men (9.5%). Thus, it is necessary to take dedicated governmental action to prevent suicide in women based on a broad and in-depth understanding of various factors related to suicide in South Korean women.

A comprehensive understanding of various risk factors associated with suicide is essential to prevent or reduce suicide.[8–10] This is because suicide is a heterogeneous phenomenon interacting with many factors related to psychiatry,[11] sociodemographics,[12] society,[13] and social media use.[14] Previous studies on suicide in women have reported major risk factors such as older age,[12] depression,[13] family conflicts,[13] and low social support.[15] These factors differ between adults and adolescents because the circumstances and conditions facing adults and adolescents are different (e.g. unemployment vs school problems).[16] For example, Kawashima[16] studied suicide attempters in Japan and found that risk factors for adolescent women suicide attempters were school problems and parent loss while those for their adult counterparts were a history of mental disorders. Similarly, Kolves and De Leo[17] studied suicides in Australia and reported that suicides in adolescents were highly impacted by family conflicts and school problems while those in adults were highly impacted by psychiatric disorders such as unipolar depression, bipolar disorder, and anxiety disorder. In addition to their findings on high-level differences between adult and adolescent suicides, considering social media use to explain the difference between the groups is highly important. With the increasing popularity of social media, excessive use of social media, various forms of online victimization on social media (e.g. cyberbullying and sexual harassment), expression of negative moods on social media, and talking about suicide on social media are increasingly being considered as risk factors for suicide.[18–20] Therefore, we need to broaden our understanding on suicide risk by exploring more fine-grained differences using various factors including social media use.

As suicide can be triggered by many different factors and traits, prevention strategies should be customized by considering suicide subtypes.[21] Researchers have attempted to investigate suicide subtypes by focusing on differences among subtypes in factors related to demographics, psychosocial aspects, and behavioral characteristics. Pfeffer[21] identified two subtypes with or without assaultive behavior among adolescent psychiatric inpatients in the US Cross and colleagues[22] studied adolescents in the United States with a history of at least one suicide attempt and found six subtypes: internalizing, externalizing, emotional dysregulation, high functioning, narcissistic, and immature. Schaffer and colleagues[23] studied Canadian suicide attempters with bipolar disorder and identified five subtypes with differences in factors such as substance abuse, history of suicide attempts, living circumstances, financial stressors, legal stressors, and interpersonal stressors. Most studies focused on suicide attempters without distinction in genders

and age groups by only considering high-level factors such as employment, mental disorders, and demographics as criteria for subtyping. Furthermore, social media data have become a valuable source of understanding suicidal behavior as Gunn[24] illustrated that people freely express their thoughts and emotions on social media posts due to its anonymity.

Therefore, to have a better understanding of subtypes of women thinking and/or attempting suicide, we should consider fine-grained data collected from both survey and social media while making a clear distinction between adults and adolescents.

Objectives

This study aimed to compare risk factors for suicide between adult and adolescent women in South Korea and identify subtypes of suicidal ideation or attempt in each group. To the best of our knowledge, no study has been conducted with a similar objective, even though there has been a rise in the occurrence of suicidal ideation and suicide attempt among this population. To achieve this goal, multifaceted data were analyzed by linking survey and social media data to investigate suicide in women from various perspectives. Therefore, this study targets women who actively use social media. Survey data included 270 questions classified into 7 categories: demographics, objective psychiatric symptoms and self-reported mental disorders, internet use, online/offline social networks, social media use, interpersonal experiences and needs, and subjective wellbeing. Social media data included five variables including the number of posts and the mean sentiment score about posts. Using these data, interpretable machine learning models were built to predict suicidal ideation or suicide attempt as well as identify and compare major risk factors for suicide in the two groups. Additionally, cluster analyses were conducted using the features which were had been identified as important in predicting adult and adolescent suicide, to identify the subtypes of suicidal adult and adolescent women as well as consider their unique characteristics. The following research questions were developed, and we emphasize that the targets of our study are women who use social media.

> RQ1: What are the similarities and differences in risk factors for suicide between adult and adolescent women in South Korea?

> RQ2: What are the subtypes of suicidal adult women in South Korea?

> RQ3: What are the subtypes of suicidal adolescent women in South Korea?

Methods

The study consisted of three parts: data collection and preprocessing, comparing major risk factors for suicide

between adult and adolescent women, and understanding subtypes of suicidal ideation or suicide attempt in adult and adolescent women. The overall process is shown in Figure 1 and each of the steps is discussed in detail in the following sections.

Participants and procedures

The study protocol was approved by the Institutional Review Board (IRB) at Sungkyunkwan University in South Korea. The study participants were adolescent women aged 13 to 18 years and adult women aged 20 to 39 years who had been using Twitter or Instagram and posted at least 10 times within the past year. In data collection, we did not consider any constraints except for gender (i.e. only woman), age (13–40) and whether participants' social media use. In South Korea, the criteria of distinguishing between adults and adolescents have recently been changed from calendar year to birthday basis and this led to situations where individuals aged 19 are sometimes classified as adults or adolescents depending on different criteria. After consulting with the online survey company that collected the data for this study, we have chosen to exclude individuals aged 19 in our study. We will discuss this in *Limitations* section. We collected survey and social media data from June 2020 to July 2020 for 500 adolescent women and from September 2020 to October 2020 for 500 adult women in South Korea through an online research company, *Macromill Embrain*. The company contacted randomly selected people in the panel to participate in our survey. If they agreed to participate in the study, they were required to read and agree to a consent statement. Throughout the data collection period, the company distributed surveys daily and provided us with the collected data every day until reaching a total of 1000 respondents. While participants were responding to the survey questionnaires, with their consent, we requested them to log into their Twitter or Instagram accounts and crawled their posts created in the past year. All personal information were modified into nonidentifiable formats. Among 1000 participants, 25 participants (20 adults and 5 adolescents) were removed because they had posted fewer than 10 posts in the past year. The final data included 480 adults and 495 adolescents. Their demographic information is available in Appendix 1. Of the 975 participants, 221 (96 adults and 125 adolescents) reported having suicidal ideation and/or suicide attempt within the past 2 weeks.

Measures

The survey questionnaires included 270 questions related to 7 higher-order constructs and 14 lower-order constructs as shown in Table 1. A detailed description of the variables is available in Appendix 2.

The survey data were linked to social media data with variables shown in Table 2.

To calculate *mean positive* and *mean negative sentiment scores of posts* for each participant, we firstly calculated a sentiment score for each post using the Python library KnuSentiLex.[25] It was trained using the bidirectional long
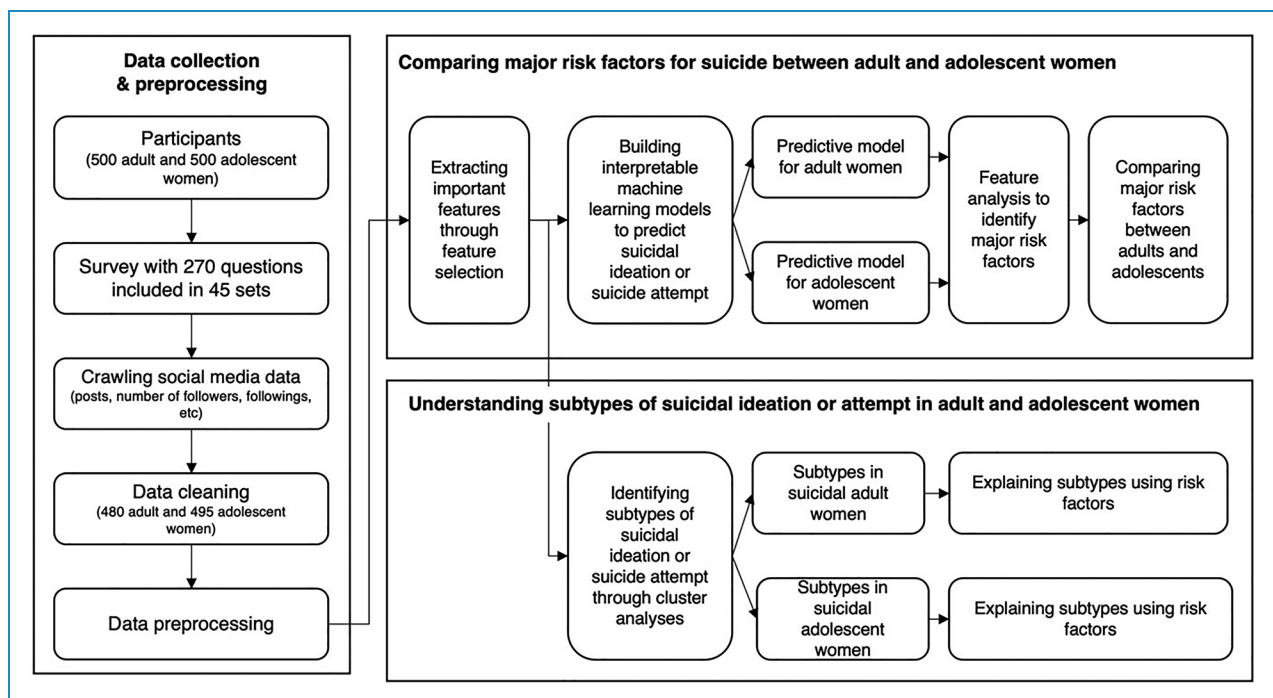
Figure 1. The process of data collection, preprocessing, and analysis.

Table 1. Description of survey data categories.

Higher-order construct	Lower-order construct	Description
Psychiatric symptoms measured by brief scales and self-reported mental disorders	Own psychiatric symptoms and mental disorders	Measures own depression symptoms, anxiety disorder symptoms, PTSD symptoms, alcohol use disorder symptoms, psychosis symptoms, obsessive-compulsive disorder, schizophrenia, and suicidal ideation or suicide attempt
	Family history of mental disorders	Measures family history of depression or bipolar disorder, alcohol use disorder or drug overdose, anxiety or panic disorders, obsessive-compulsive disorder, schizophrenia, and self-harm or suicide attempt
Internet use	Internet use	Measures how often participants accessed social media, online shopping platforms, adult websites, and others in the past week
Online/offline social networks	Online/offline communication and social networking	Measures how long participants usually spent in face-to-face talks, text messages, social media, and others
	Social support on social media	Measures the extent to which participants feel they are receiving help from others on social media
Social media use	Basic characteristics of social media use	Measures general use patterns of social media such as frequency of social media use, number of friends on social media, and social media platform in use
	Positive self-presentation	Measures how often participants selectively posted on social media to show their positive aspects
	Narcissistic self-presentation	Measures how often participants posted on social media to show their personal information to positively express their self-images
	Distress disclosure	Measures how often participants posted on social media to disclose their negative feelings and thoughts
	Feedback	Measures how often participants sent or received positive or negative feedback on social media in the past month
	Negative experiences on social media	Measures how often participants experienced negative things on social media in the past year
Interpersonal experiences and needs	Interpersonal adverse experiences	Measures how often participants experienced interpersonal adverse things such as neglect, sexual abuse, being touched or forced to touch someone else in their childhood, last year, or lifetime
	Need to belong	Measures participants' emotional distance between themselves and others, their tendency to rely on others, and their fear of being evaluated negatively by others
Subjective wellbeing	Satisfaction with life	Measures participants' general sense of satisfaction with their lives

(continued)

Table 1. Continued.

Higher-order construct	Lower-order construct	Description
Demographics	Demographics	Measures participants' age, region, living arrangements, number of family members living with, average monthly household income, and economic status

PTSD: posttraumatic stress disorder.

Table 2. Description of social media variables.

Variable	Description
User ID	Participants' Twitter or Instagram ID
Post content	Content of posts written in the past year
Number of posts	Number of participants' posts in the past year
Number of followers	The number of participants' followers on Twitter or Instagram
Number of followings	The number of participants' followings on Twitter or Instagram
Mean positive sentiment score about posts	Average sentiment score of participants' positive posts in the past year
Mean negative sentiment score about posts	Average sentiment score of participants' negative posts in the past year

short-term memory (Bi-LSTM) on a lexicon of over 300,000 sentiment words. It tokenizes a post and calculates each token's sentiment score which ranges from −2 (very negative) to +2 (very positive) and reports the sum of each token's score as the final score of the post. Given a participant's all posts, after calculating each post's sentiment score, we classified them into positive or negative sentiment groups. We calculated mean scores of the posts in each of the two groups and obtained each participant's positive and negative sentiment scores.

Statistical analyses

Interpretable machine learning. To identify and compare major risk factors for suicide between adult and adolescent women, we used eXtreme Gradient Boosting (XGBoost),[26] which is a gradient boosting method with regularization and has been used by many studies[27,28] to build high-performance prediction models for tabular data. Before training our models, we oversampled our data using the Synthetic Minority Oversampling Techniques (SMOTE) function from a Python library Imblearn[29] to tackle the unbalanced sample problem. To optimize performance, models were trained several times through an iterative process of feature selection using a scikit-learn library SelectFromModel[30] in a stratified 10-fold cross-validation setting. After building final models with the best performance, we visualized top risk factors and their impacts on predicting suicide using the SHAP summary plot.[31]

Cluster analysis. To understand subtypes of suicidal ideation or suicide attempt in adult and adolescent women, we used the Gaussian Mixture Model (GMM),[30] which is a probabilistic model that assumes data were generated from the finite mixture distributions.[32] To achieve interpretable results, we only considered important features obtained through a feature selection processes during building machine learning models. By doing so, we aimed to deliver more distinguishable and easy-to-interpret cluster results. To identify the optimal number of clusters, we built cluster models with the number of clusters ranging from two to ten and calculated the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), entropy, and Bootstrap Likelihood Ratio Test (BLRT) scores for each model. Models with lower AIC and BIC scores, higher entropy scores, and significant p values in BLRT are considered as better fitting models.[33,34] We then consulted two psychiatrists for selecting the optimal number of clusters. In this study, the psychiatrists were not asked to label individual participant because there are more than 60 features to consider when labeling and this is a laborious task which may result in incorrect results. Instead, we showed them nine candidates with the number of clusters ranges from 2 to 10 and asked them to select the number of clusters (2, 3, etc.) that best shows each cluster's distinct characteristics. The nine candidates were generated by using GMM. For example, if they think each of three clusters in the candidate with three clusters has distinct characteristics, they may choose three as the optimal number of clusters. This manual evaluation may favor smaller number of clusters due to interpretability and therefore we also considered four metrics of automated methods.

Study design and procedures

Comparing major risk factors for suicide between adult and adolescent women. First, respective interpretable machine learning models were trained to predict suicide risk for adult and adolescent women. In addition to make predictions, interpretable machine learning models describe reasoning processes of making predictions and explain relative importance of features used in predictions.[35] Second, with the predictive models, major factors were identified by exploring feature importance which indicates a feature's relative contribution to predicting suicide risk. Specifically, factors were identified that increase the probability of having suicidal ideation or attempting suicide (i.e. risk factors). Third, the two sets of major risk factors (one for each group) were compared to understand similarities and differences.

Understanding subtypes of suicidal ideation or suicide attempt in adult and adolescent women. Subtypes of suicidal ideation or suicide attempt were identified through cluster analyses. A cluster analysis detects homogeneous subgroups in a sample according to a measure of similarities among data points in the sample.[36] In this study, similarities among individuals in a group (either suicidal adult or adolescent women) were computed from the variables used in predictive models. After identifying subtypes, a two-tailed Student's *t*-test was used to explain the unique characteristics of each subtype.

Results

Predictive models of suicide ideation or attempt

After the feature selection process, 65 features for the adult women's model and 51 features for the adolescent women's model were extracted. The performance of the best-performing models is shown in Table 3.

The results of the models showed an F1-score of 0.92 and AUC of 0.92 in the adults' model and an F1-score of 0.89 and AUC of 0.89 in the adolescents' model.

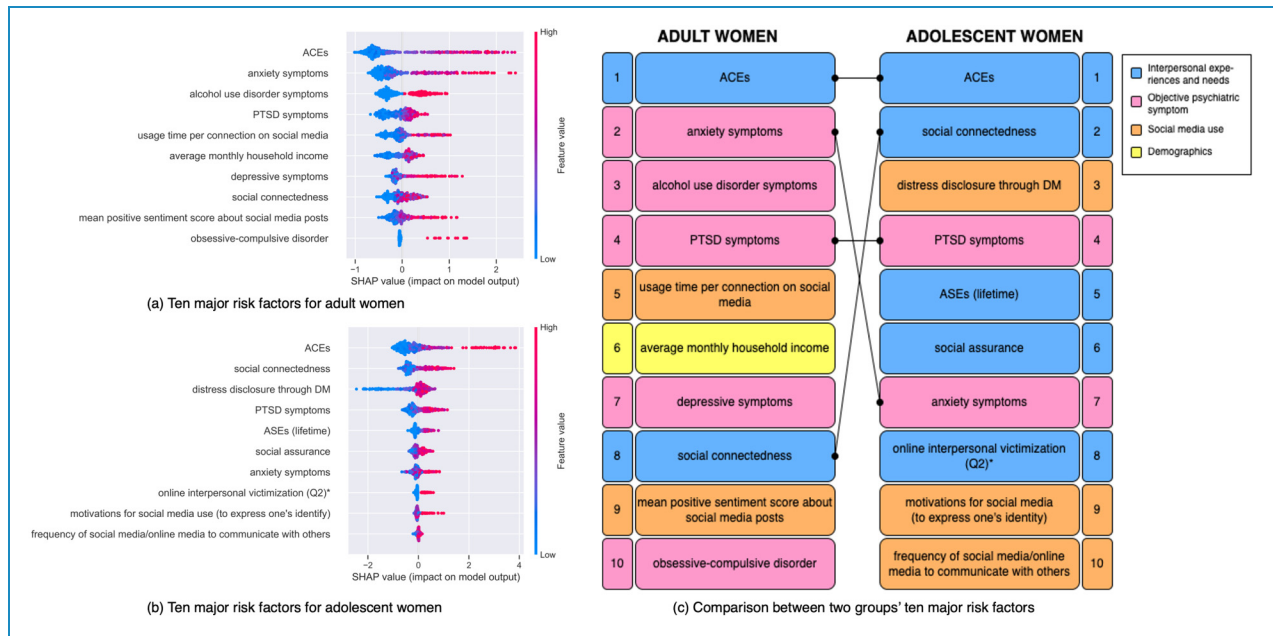### Comparing major risk factors for suicide between adult and adolescent women

Figure 2(a) and (b) show the 10 major risk factors of models for adult women and adolescent women, ranked by their

Table 3. Performance of the two predictive models.

Model	Accuracy	Precision	Recall	F1	Area under curve
Adult women	0.92	0.92	0.91	0.92	0.92
Adolescent women	0.89	0.90	0.89	0.89	0.89

relative importance. Figure 2(c) shows the comparison between the two groups' major risk factors. In Figure 2(a) and (b), given a feature, each dot denotes an individual case, and its horizontal position shows the individual's SHAP value which indicates its impact on the predictive model's output. A value of zero denotes no impact on suicide prediction, a positive SHAP value indicates its positive impact on increasing the probability of having suicidal ideation or attempt, and a negative SHAP value indicates the opposite. For instance, if someone has a positive SHAP value in *adverse childhood experiences (ACEs)* indicating abuse or other potentially damaging childhood experiences,[37] her *ACEs* then has an impact on increasing the probability of her having suicidal ideation or suicide attempt. The dot color represents feature values, with blue color representing low values and red color representing high values. Therefore, if most of a feature's red dots (i.e. high values) are associated with SHAP values higher than zero, this feature is considered a risk factor for suicide.

Five findings are highlighted as follows. First, as shown in Figure 2(a) and (c), half of the 10 major risk factors for adult women were related to mental health (i.e. *anxiety symptoms,*[38] *alcohol use disorder symptoms,*[39] *post-traumatic stress disorder (PTSD) symptoms,*[40] *depressive symptoms,*[38] and *obsessive-compulsive disorder*). Second, compared to adult women, as shown in Figure 2(b) and (c), half of the 10 major risk factors for adolescent women were related to interpersonal experiences and needs. The most important factor was *ACEs,* followed by *social connectedness* indicating an emotional distance between self and others,[41] *ASEs in lifetime* indicating sexual abuse including being touched without consent,[42] *social assurance* indicating the tendency to rely on others to feel a sense of belonging,[43] and *online interpersonal victimization (Q2)* indicating experience of having been asked about very private and sexual information online despite being unwilling.[44] Third, as shown in Figure 2(c), the most important factor in both groups was *ACEs* and the fourth was *PTSD symptoms.* Fourth, although both *anxiety symptoms* and *social connectedness* were identified as major risk factors for adult and adolescent women, these factors had different importance within the two groups. *Anxiety symptoms* ranked second in adults and seventh in adolescents. *Social connectedness* ranked second in adolescents and eighth in adults. Fifth, there were different risk factors related to social media use among adult and adolescent women. Adult women had two factors related to social media use (i.e. *usage time per connection on social media* and *mean positive sentiment score about social media posts).* These were ranked as the fifth and ninth risk factors. Adolescent women had three additional factors including *distress disclosure through direct message (DM),*[45] *motivations for social media (to express one's identity),*[46] and *frequency of social media/online media to communicate with others.* These were ranked as the third, ninth, and tenth risk factors.

Figure 2. Major risk factors for suicide in adult and adolescent women. *Online interpersonal victimization (Q2)** denotes the experience of having been asked about very private and sexual information online despite being unwilling.

Understanding subtypes of suicidal ideation or suicide attempt

Subtypes of suicidal adult ($n = 96$) and adolescent women ($n = 125$) were identified separately using cluster analyses.

Subtypes of suicidal adult women

After generating models with different number of clusters, we found that the model with two clusters had the lowest BIC and the highest entropy (Table 4) and excelled in two out of the four metrics. In addition to the results of the automated methods, we asked two psychiatrists to review the nine candidates with different number of clusters. They agreed that two is the optimal number of clusters which shows interpretable and distinguishable results. Synthesizing the results of the automated methods and psychiatrists' domain expertise, we selected two as the optimal number of clusters.

Table 5 shows each subtype's representative factors and their values. Significant differences were observed in four of the seven high-level categories.

As the results suggest, subtype 1 was characterized by severe psychiatric symptoms and mental disorders. They had more problems with both their own mental health and family histories of mental disorders, and family history of mental disorders has been shown to significantly increase an individual's own suicide risk.[47,48] Subtype 2 tended to spend more time on social media than subtype 1. Specifically, subtype 2 spent more than 50 min per connection on social media whereas subtype 1 spent less than 30 min. Additionally, among the various

Table 4. Results of four evaluation metrics for models with different number of clusters (suicidal adult women).

# of clusters	AIC	BIC	Entropy	BLRT
2	−12,175.393	−838.410	0.989	>.05
3	−17,315.978	−309.220	0.898	>.05
4	−15,249.323	7427.208	0.815	>.05
5	−21,841.651	6504.654	0.888	>.05
6	−20,754.287	13,261.792	0.883	>.05
7	−22,231.874	17,453.979	0.928	>.05
8	−19,302.275	26,053.351	0.915	>.05
9	−16,486.235	34,539.165	0.915	>.05
10	−14,988.371	41,706.803	0.934	>.05

AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion; BLRT: Bootstrap Likelihood Ratio Test.

motivations for using social media, the degree of using social media to escape from school or work was significantly higher in subtype 2. They also experienced a significantly higher level of *having been threatened or embarrassed by posting or sending messages about you for others to see* than subtype 1.

Table 5. Characteristics of subtypes in suicidal adult women.

Category	Factor	Total suicidal adults ($n = 96$)	Subtype 1 ($n = 54$)	Subtype 2 ($n = 42$)	p value
Psychiatric symptoms and mental disorders	Alcohol use disorder symptoms	0.53	0.69	0.33	.000
	Psychosis symptoms	0.22	0.31	0.10	.006
	Family history of anxiety or panic disorders	0.14	0.22	0.02	.002
	Family history of depression or bipolar disorders	0.33	0.52	0.10	.000
	Family history of self-harm or suicide attempts	0.09	0.15	0.02	.025
Online/offline networks	Ratio of social media in average daily communication	0.28	0.21	0.36	.000
Social media use	Usage time per connection on social media	0.17	0.11	0.25	.002
	Motivations for social media use (to escape from school or work)	0.30	0.22	0.40	.006
Interpersonal experiences and needs	Experiences having been threatened or embarrassed by posting or sending messages about you for others to see	0.15	0.06	0.26	.008

Table 6. Results of four evaluation metrics for models with different number of clusters (suicidal adolescent women).

Number of clusters	AIC	BIC	Entropy	BLRT
2	−3993.300	3798.705	0.896	<.001
3	−4108.797	7580.624	0.709	<.001
4	−10,097.776	5489.061	0.863	>.05
5	−13,507.312	5976.941	0.857	>.05
6	−17,985.984	5395.686	0.881	>.05
7	−17,916.132	9362.954	0.864	>.05
8	−13,092.318	18,084.185	0.798	>.05
9	−13,026.954	22,046.964	0.804	>.05
10	−18,360.386	20,610.949	0.859	>.05

AIC: Akaike Information Criterion; BIC: Bayesian Information Criterion; BLRT: Bootstrap Likelihood Ratio Test.

Subtypes of suicidal adolescent women

After generating models with different number of clusters, we found that the model with two clusters had the lowest BIC, the highest entropy, and a significant p value in

BLRT (Table 6) and excelled in three out of the four metrics. Again, two psychiatrists reviewed the results and agreed that two was the optimal number of clusters. Therefore, we selected two as the number of clusters for suicide in suicidal adolescent women.

Table 7 shows the two subtypes' characteristics, including the differences among five categories within the subtypes.

Subtype 1 had severe psychiatric symptoms and high *general problematic internet use (GPIU)* which indicates problematic use of the internet that can cause psychological or social difficulties.[49] They used social media habitually to spend time without specific purposes. They had issues related to interpersonal experiences and needs such as *ACEs*. They felt a great sense of emotional distance from others (i.e. high *social connectedness*), which may increase their suicide risk. Different from the results in adult women where the two subtypes were characterized by having either mental disorders or interpersonal problems, adolescent women in subtype 1 had both problems. While people in subtype 1 used social media without specific purposes, subtype 2 used social media for managing social relationships, including maintaining relationships with friends and/or organizing gatherings. They also had more negative online sexual experiences than subtype 1. It is noteworthy that subtype 2 had different interpersonal needs than subtype 1. Subtype 1 had high *social connectedness* whereas subtype 2 had high *social assurance*.

Figure 3 visualizes the identified subtypes in three-dimensional spaces using principal component analysis

Table 7. Characteristics of subtypes in suicidal adolescent women.

Category	Factor	Total suicidal adolescents (n = 125)	Subtype 1 (n = 78)	Subtype 2 (n = 47)	p value
Psychiatric symptoms and mental disorders	Anxiety symptoms	0.29	0.33	0.23	.009
	PTSD symptoms	0.34	0.37	0.29	.033
Internet use	General problematic internet use	0.49	0.52	0.44	.020
Online/offline social networks	Frequency of social media/online media use to communicate with others	0.73	0.68	0.80	.007
Social media use	Frequency of checking feedback from others on social media	0.13	0.15	0.08	.007
	Period of social media usage	0.40	0.44	0.34	.018
	Motivations for social media use (to keep relationships with friends)	0.49	0.42	0.60	.001
	Motivations for social media use (to spend time)	0.70	0.75	0.62	.003
	Motivations for social media use (to use habitually)	0.66	0.70	0.59	.028
	Motivations for social media use (to organize gatherings)	0.61	0.57	0.60	.040
Interpersonal experiences and needs	Adverse childhood experiences	0.19	0.23	0.14	.013
	Experiences having been asked to engage in unwanted sexual talk online	0.38	0.29	0.53	.010
	Experiences having been threatened or embarrassed by posting or sending messages about you for others to see	0.17	0.22	0.09	.036
	Social connectedness	0.38	0.44	0.29	.001
	Social assurance	0.56	0.50	0.65	.000

PTSD: posttraumatic stress disorder.

(PCA)[50] and summarizes the characteristics. As shown in the figure, the suicidal adult and adolescent women were noticeably divided into two separate distinct subtypes.

Discussion

In contrast to prior research that examined a limited range of potential factors related to suicide risk in women of all age groups in Western countries, this study analyzed risk factors specifically for adult and adolescent women in South Korea. It further explored the similarities and differences between the two groups using multifaceted data organized into 7 higher-order and 14 lower-order constructs. In addition, previous studies on suicide subtypes did not

investigate women and primarily used general factors such as mental disorders to explain the subtypes.[21,22,51] To fill this gap, the subtypes of suicidal adult and adolescent women were examined separately using a comprehensive set of factors. In the following, we provide a detailed discussion of our research questions.

> RQ1: What are the similarities and differences in risk factors for suicide between adult and adolescent women in South Korea?

To answer RQ1, we initially built interpretable machine learning models to predict suicide risk for adult and adolescent women separately. We then identified major risk
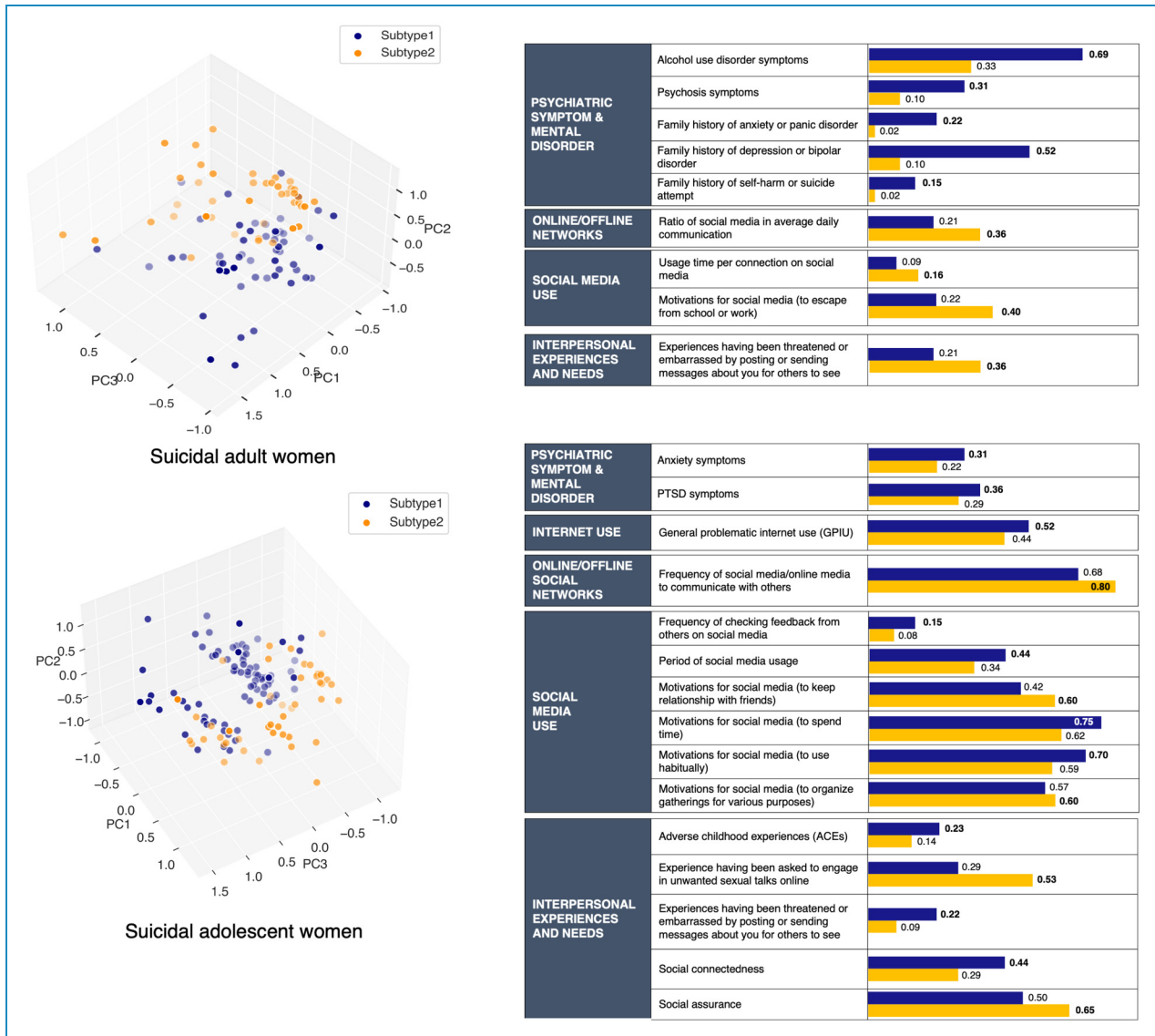
Figure 3. Summarization of subtypes of suicidal adult and adolescent women.

factors using the concept of feature importance which indicates a feature's relative contribution to predicting suicide risk. Finally, we compared the identified major risk factors for adult women with those for adolescent women to explore similarities and differences. In common, *ACEs* were the most important risk factor for suicidal ideation and attempt in both adult and adolescent women. These results illustrate that adverse experiences women undergo during childhood may have long-term effects even after they grow up and their impacts are so severe that they may cause the person to think about or attempt suicide.[52,53] Although both groups had *ACEs* as the most important factor, the extent to which they contributed to suicide risk was different. The highest SHAP value in adult women was approximately 2.5 but in adolescent women, it was 4. Thus, the negative impact of *ACEs* on

adolescents may be much higher given that they experienced it relatively recently and are less capable of handling negative events.[54,55] Government officials should consider *ACEs* a major risk factor for suicide in women and develop relevant prevention strategies. Although *anxiety symptoms* and *social connectedness* were identified as major risk factors for both adult and adolescent women, their importance differed. *Anxiety symptoms* were more important to adult women while *social connectedness* was more important to adolescent women. Studies reported that for adolescents who undergo social and physical changes, relationships with peers or parents are an important issue that is significantly related to emotional problems (e.g. loneliness or social isolation.)[56] Thus, *social connectedness* may have a greater impact on suicidal ideation or suicide attempt among adolescent women. Based on this

result, we can infer that adolescent women are more vulnerable to social and interpersonal problems than adult women. Different high-risk social media use behaviors were also identified between the two groups. Excessive social media use was an important factor for suicide in adult women, whereas the use of social media to communicate with others was an important risk factor in adolescent women. Previous studies explained that excessive use of social media may cause chronic sleep loss and low satisfaction with life, which have a negative influence on mental health.[57,58] For adolescents, Koutamanis[59] illustrated that adolescents who engaged in social media to communicate with others were more likely to receive negative feedback, which could lead to unhealthy mental status. Particularly, previous studies explained that engaging in risky online behaviors, such as self-promotion[60] and posting narcissistic content[61] during communication on social media, may result in receiving negative feedback. Therefore, adolescent women may need to be aware of potential negative consequences when interacting with others on social media.

RQ2: What are the subtypes of suicidal adult women in South Korea?

To answer RQ2, we performed cluster analyses to identify subtypes among suicidal adult women and summarized their characteristics. During cluster analyses, we utilized 65 important factors obtained from the best-performing prediction model we had built when addressing RQ1. Results of a cluster analysis using all the features may less distinguishable, hard to interpret and provide condensed information because relative importance of features are not considered, and every feature is treated as equal. To avoid this, we performed cluster analyses in answering RQ2 and RQ3 by utilizing the results of RQ1, therefore RQ1 is a necessary step toward RQ2 and RQ3. As a result, we identified two subtypes among suicidal adult women. The first subtype was characterized by severe psychiatric symptoms and mental disorders among them and/or their families. The second subtype used social media excessively and suffered from online victimization. Thus, in addition to national efforts to promote mental health, a safe online and social media environment may be important. Protection against exposure to online victimization may be extremely important to successful suicide prevention.

RQ3: What are the subtypes of suicidal adolescent women in South Korea?

To answer RQ3, we performed cluster analyses to identify subtypes among suicidal adolescent women. Again, we utilized 51 important factors obtained from the best-performing prediction model we had built when addressing RQ1. As a result, we identified two subtypes of suicidal adolescent women. The first subtype had severe psychiatric

symptoms, more adverse experiences in childhood, and a large emotional distance from others. The second subtype used social media frequently, had more experiences of online sexual victimization, and was highly dependent on others. Mitchell et al.[62] reported that adolescents who were asked to engage in unwanted sexual talk online, including requests for offline contact, had high levels of distress after the incidences. Furthermore, previous studies reported that people who was highly dependent on others may use social media to fill the deficiencies in face-to-face interaction with others, and it may reduce social participation in real life and lead to unhealthy mental status.[63,64] Thus, adolescents in subtype 2 may have felt a lack of social relationships in real life and used social media to solve this problem. Instead, they may have suffered more from negative sexual incidents and eventually thought about or attempted suicide. Therefore, promoting face-to-face interactions with parents and peers, developing offline social skills, and reducing adolescents' desire to rely on media[65] may be good suicide prevention strategies.

Limitations

This study has a few limitations. First, although this study identified major risk factors using a wide range of potential factors, it did not consider physical disorders (e.g. neoplasm) or pregnancy[66–68] which have been considered important factors associated with suicide in women. Thus, further research needs to include these factors for a more detailed understanding of suicide in women. Second, owing to limitations involved in data collection, this study only considered women aged between 13 and 39 years. The World Health Organization[69] reported that although the suicide rate is highest among women in this age range, it is also high among those older than 40. Third, we utilized simple metrics such as the number of posts or sentiment score derived from social media data for predicting suicide. In future studies, it would be beneficial to analyze the social media content such as linguistic patterns of suicidal users or main themes in their posts. Fourth, as the study participants were individuals who had used social media, the results may be different from those who had not used social media. Among the top risk factors for suicide in adult and adolescent women, five factors were related to social media. Therefore, the study results may have limited generalizability to individuals who do not use social media. Fifth, this study distinguished adults from adolescents based on the age of 19. That is, women over 19 were classified as adults, while those under 19 were categorized as adolescents. In addition to this criterion, some studies have examined individuals aged 18 to 25 as a distinct group (i.e. transition-aged youth or young adults) based on the assumption that they have common characteristics. Our future study will take this perspective into account when categorizing groups and understanding

differences. Finally, this study used a self-reported questionnaire to identify participants' gender and our data may include women whose self-reported gender identity is different form their biological sex. Because the study participants were randomly selected by a renowned survey company with rich experience, we believe that the proportion of gender minority in our sample may not largely different from the proportion in South Korea population. Therefore, it is difficult to attribute the high suicide rate in this study to the influence of gender minority population, but we see the limited assessment of different gender identities is a limitation.

Conclusions

In this study, by using multifaceted and linked data from surveys and social media, along with interpretable machine learning approaches, we investigated (1) major risk factors for suicide among South Korean adult women; (2) major risk factors for suicide among South Korean adolescent women; (3) similarities and differences in risk factors between adult and adolescent women; (4) subtypes of the subtypes of suicidal adult women in South Korea; and (5) subtypes of suicidal adolescent women in South Korea.

We summarize our contributions as follows. First, whereas previous studies primarily used a single type of data, we used multifaceted data by linking surveys and social media. Therefore, we investigated suicide risk among South Korean adult and adolescent women from various perspectives. Second, our study identified major risk factors for suicide separately for adult and adolescent women and compared the results enabling a stratified and targeted understanding. Third, we identified and explained subtypes in suicidal adult and adolescent women separately and further segmentalized the subjects of study. All of these mentioned contributions are further related to developing customized suicide prevention plans and programs.

ORCID iD: Donghun Kim   https://orcid.org/0000-0001-5441-1532

References

1. Meleiro AdS, Correa H. Suicide and suicidality in women. In: Rennó J, Valadares G, Cantilino A, et al. (eds) *Women's mental health: a clinical and evidence-based guide*. Cham: Springer, 2020, pp.221–225.
2. Tasnim R, Islam MdS, Sujan MdSH, et al. Suicidal ideation among Bangladeshi university students early during the COVID-19 pandemic: prevalence estimates and correlates. *Child Youth Serv Rev* 2020; 119: 105703–105703.
3. Wild P, Bovio N, Canu IG, et al. Part-time work and other occupational risk factors for suicide among working women in the Swiss National Cohort. *Int Arch Occ Env Hea* 2021; 94: 981–990.
4. Ministry of Health and Welfare of South Korea. *2021 White Paper on Suicide Prevention*. Report for the Ministry of Health and Welfare of South Korea, July 2021.
5. StatisticsKorea. Statistics Korea: Causes of death statistics in 2020, http://www.kostat.go.kr/portal/korea/kor_nw/1/6/2/index.board?bmode=read&bSeq=&aSeq=403046&pageNo=1&rowNum=10&navCount=10&currPg=&searchInfo=&sTarget=title&sTxt= (2021).
6. Statistics Korea. *Quality of Life Indicators in Korea* 2022, https://kostat.go.kr/board.es?mid=a90106000000&bid=12316&act=view&list_no=423794 (December 1, 2023, accessed December 1, 2023).
7. Statistics Korea. *Children and Youth Well-being in Korea* 2022, https://kostat.go.kr/board.es?mid=a90106000000&bid=12316&act=view&list_no=425844 (December 1, 2023, accessed December 1, 2023).
8. Seeman N, Reilly DK, Fogler S, et al. Suicide risk factors in U.S. college students: Perceptions differ in men and women. *Suicidol Online* 2017; 8: 20–26.
9. Sharma B, Nam EW, Kim HY, et al. Factors associated with suicidal ideation and suicide attempt among school-going urban adolescents in Peru. *Int J Environ Res Pu* 2015; 12: 14842–14856.
10. Kim D, Quan L, Seo M, et al. Interpretable machine learning-based approaches for understanding suicide risk and protective factors among South Korean females using survey and social media data. *Suicide Life Threat Behav* 2023; 53: 484–498.
11. Lee J, Bang YS, Min S, et al. Characteristics of adolescents who visit the emergency department following suicide attempts: comparison study between adolescents and adults. *BMC Psychiatry* 2019; 19: 231.

12. Rochat TJ, Bland RM, Tomlinson RM, et al. Suicide ideation, depression and HIV among pregnant women in rural South Africa. *Health* 2013; 05: 650–661.

13. Chin YR and Choi K. Suicide attempts and associated factors in male and female Korean adolescents a population-based cross-sectional survey. *Community Mental Health J* 2015; 51: 862–866.

14. Swedo EA, Beauregard JL, Fijter Sd, et al. Associations between social media and suicidal behaviors during a youth suicide cluster in Ohio. *J Adolescent Health* 2021; 68: 308–316.

15. Miller AB, Esposito-Smythers C and Leichtweis RN. Role of social support in adolescent suicidal ideation and suicide attempts. *J Adolescent Health* 2015; 56: 286–292.

16. Kawashima Y, Ito T, Narishige R, et al. The characteristics of serious suicide attempters in Japanese adolescents- comparison study between adolescents and adults. *BMC Psychiatry* 2012; 12: 191.

17. Kolves K and De Leo D. Child, adolescent and young adult suicides: a comparison based on the Queensland suicide registry. *J Child Adolescent Behav* 2015; 3: 1–5.

18. Young R, Subramanian R, Miles S, et al. Social representation of cyberbullying and adolescent suicide: a mixed-method analysis of news stories. *Health Commun* 2017; 32: 1082–1092.

19. Benítez-Hidalgo V, Henares-Montiel J, Ruiz-Pérez I, et al. Cyber sexual harassment against women and impact on health. A cross-sectional study in a representative population sample. *J Public Health* 2023; fdad182.

20. Bryan CJ, Butner JE, Sinclair S, et al. Predictors of emerging suicide death among military personnel on social media networks. *Suicide LifeThreat Behav* 2018; 48: 413–430.

21. Pfeffer CR, Newcorn J, Kaplan G, et al. Subtypes of suicidal and assaultive behaviors in adolescent psychiatric inpatients: a research note. *J Child Psychol Psychiatry* 1989; 30: 151–163.

22. Cross D, Westen D and Bradley B. Personality subtypes of adolescents who attempt suicide. *J Nerv Mental Dis* 2011; 199: 750–756.

23. Schaffer A, Sinyor M, Reis C, et al. Suicide in bipolar disorder: characteristics and subgroups. *Bipolar Disord* 2014; 16: 732–740.

24. Gunn JF and Lester D. Twitter postings and suicide: an analysis of the postings of a fatal suicide in the 24 hours prior to death. *Suicidologi* 17, Epub ahead of print 2015. DOI: 10.5617/suicidologi.2173.

25. On B, Park S and Na C. KnuSentiLex, https://github.com/park1200656/KnuSentiLex (2023, accessed February 1, 2023).

26. Chen T and Guestrin C. XGBoost: a scalable tree boosting system. *In: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, San Francisco, USA, 13 August–17 August 2016, pp.785–794. New York: Association for Computing Machinery.

27. Budholiya K, Shrivastava SK and Sharma V. An optimized XGBoost based diagnostic system for effective prediction of heart disease. *J King Saud Univ – Comput Inf Sci* 2022; 34: 4514–4523.

28. Islam M, Alam J, Maniruzzaman M, et al. Predicting the risk of hypertension using machine learning algorithms: a cross sectional study in Ethiopia. *PLoS ONE* 2023; 18: e0289613.

29. Lemaitre G, Nogueira F and Aridas CK. Imbalanced-learn: a Python toolbox to tackle the curse of imbalanced datasets in machine learning. *J Mach Learn Res* 2017; 18: 1–5.

30. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res* 2011; 12: 2825–2830.

31. Lundberg S. SHAP, https://shap.readthedocs.io/en/latest/example_notebooks/api_examples/plots/waterfall.html (2022, accessed January 1, 2022).

32. Nguyen T, Chen G and Chacon L. An adaptive EM accelerator for unsupervised learning of Gaussian mixture models. *arXiv*. Epub ahead of print 2020. DOI: 10.48550/arxiv.2009.12703.

33. Crosby RD, Wonderlich SA, Engel SG, et al. Daily mood patterns and bulimic behaviors in the natural environment. *Behav Res Ther* 2009; 47: 181–188.

34. Henning JS, Dusza SW, Wang SQ, et al. The CASH (color, architecture, symmetry, and homogeneity) algorithm for dermoscopy. *J Am Acad Dermatol* 2006; 56: 45–52.

35. Rudin C, Chen C, Chen Z, et al. Interpretable machine learning: fundamental principles and 10 grand challenges. *Stat Surv* 2021; 16: 1–85.

36. Garcia-Rudolph A, Garcia-Molina A, Opisso E, et al. Personalized web-based cognitive rehabilitation treatments for patients with traumatic brain injury: cluster analysis. *JMIR Med Inform* 2020; 8: e16077.

37. Felitti VJ, Anda RF, Nordenberg D, et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults: the adverse childhood experiences (ACE) study. *Am J Prev Med* 2019; 56: 774–786.

38. Chae S, Shin J, Ji D, et al. A brief depression and anxiety rating scale for sexual violence victims. *Psychiatry Investig* 2021; 18: 172–178.

39. Lee S. Health and policy tasks for Koreans. *Health and Welfare Forum* 1997; 8: 5–14.

40. Chung H, Kim J-W, Kwon J, et al. Development of brief post-traumatic stress disorder rating scale for sexual violence victims. *Psychiatry Investig* 2019; 16: 868–871.

41. Lee RM and Robbins SB. Measuring belongingness: the social connectedness and the social assurance scales. *J Couns Psychol* 1995; 42: 232–241.

42. Tucker W, Garn JV and Yang W. Relationship between adverse childhood experiences and gambling in Nevada. *Am J Health Behav* 2021; 45: 246–255.

43. Lee-Won RJ, Herzog L and Park SG. Hooked on Facebook: the role of social anxiety and need for social assurance in problematic use of Facebook. *Cyberpsychol Behav Soc Netw* 2015; 18: 567–574.

44. Ybarra ML, Mitchell KJ, Finkelhor D, et al. Internet prevention messages: targeting the right online behaviors. *Arch Pediatr Adolesc Med* 2007; 161: 138–145.

45. Kahn JH and Hessling RM. Measuring the tendency to conceal versus disclose psychological distress. *J Soc Clin Psychol* 2001; 20: 41–65.

46. Lai C-H. Motivations, usage, and perceived social networks within and beyond social media. *J Comput-Mediat Comm* 2019; 24: 126–145.

47. Qin P, Agerbo E and Mortensen PB. Suicide risk in relation to family history of completed suicide and psychiatric disorders: a nested case-control study based on longitudinal registers. *Lancet* 2002; 360: 1126–1130.

48. Sørensen HJ, Mortensen EL, Wang AG, et al. Suicide and mental illness in parents and risk of suicide in offspring. *Soc Psych Psych Epid* 2009; 44: 748–751.

49. Beard KW and Wolf EM. Modification in the proposed diagnostic criteria for internet addiction. *CyberPsychology Behav* 2001; 4: 377–383.

50. Granato D, Santos JS, Escher GB, et al. Use of principal component analysis (PCA) and hierarchical cluster analysis (HCA) for multivariate association between bioactive compounds and functional properties in foods: a critical perspective. *Trends Food Sci Tech* 2018; 72: 83–90.

51. Ortigo KM, Westen D and Bradley B. Personality subtypes of suicidal adults. *J Nerv Mental Dis* 2009; 197: 687–694.

52. Almuneef M, ElChoueiry N, Saleheen HN, et al. Gender-based disparities in the impact of adverse childhood experiences on adult health: findings from a national study in the Kingdom of Saudi Arabia. *Int J Equity Health* 2017; 16: 90.

53. Boullier M and Blair M. Adverse childhood experiences. *Paediatr Child Health (Oxford)* 2018; 28: 132–137.

54. Estévez A, Jáuregui P, Sánchez-Marcos I, et al. Attachment and emotion regulation in substance addictions and behavioral addictions. *J Behav Addict* 2017; 6: 534–544.

55. Guo J, Fu M, Liu D, et al. Is the psychological impact of exposure to COVID-19 stronger in adolescents with pre-pandemic maltreatment experiences? A survey of rural Chinese adolescents. *Child Abus Negl* 2020; 110: 104667–104667.

56. Helsen M, Vollebergh W and Meeus W. Social support from parents and friends and emotional problems in adolescence. *J Youth Adolescence* 2000; 29: 319–335.

57. Khalaf AM, Alubied AA, Khalaf AM, et al. The impact of social media on the mental health of adolescents and young adults: a systematic review. *Cureus* 2023; 15: e42990.

58. Marttila E, Koivula A and Räsänen P. Does excessive social media use decrease subjective well-being? A longitudinal analysis of the relationship between problematic use, loneliness and life satisfaction. *Telemat Inform* 2021; 59: 101556.

59. Koutamanis M, Vossen HGM and Valkenburg PM. Adolescents' comments in social media: why do adolescents receive negative feedback and who is most at risk? *Comput Hum Behav* 2015; 53: 486–494.

60. Nault KA and Yap AJ. The self-promotion boost: positive consequences for observers of high-rank self-promoters. *J Pers Soc Psychol* 2024; 126: 175–212.

61. große Deters F, Mehl MR and Eid M. Narcissistic power poster? On the relationship between narcissism and status updating activity on Facebook. *J Res Pers* 2014; 53: 165–174.

62. Mitchell KJ, Finkelhor D and Wolak J. Risk factors for and impact of online sexual solicitation of youth. *JAMA* 2001; 285: 3011–3014.

63. Seo M, Kim J and David P. Always connected or always distracted? ADHD symptoms and social assurance explain problematic use of mobile phone and multicommunicating. *J Comput Mediat Commun* 2015; 20: 667–681.

64. Kraut R, Kiesler S, Boneva B, et al. Internet paradox revisited. *J Soc Issues* 2002; 58: 49–74.

65. Kim J-H. Psychological issues and problematic use of smartphone: ADHD's moderating role in the associations among loneliness, need for social assurance, need for immediate connection, and problematic use of smartphone. *Comput Hum Behav* 2018; 80: 390–398.

66. Gradus JL, Rosellini AJ, Horváth-Puhó E, et al. Prediction of sex-specific suicide risk using machine learning and single-payer health care registry data from Denmark. *JAMA Psychiat* 2020; 77: 25–34.

67. Gressier F, Guillard V, Cazas O, et al. Risk factors for suicide attempt in pregnancy and the post-partum period in women with serious mental illnesses. *J Psychiatr Res* 2017; 84: 284–291.

68. Karmakar C, Luo W, Tran T, et al. Predicting risk of suicide attempt using history of physical illnesses from electronic medical records. *JMIR Mental Health* 2016; 3: e5475.

69. World Health Organization. *Suicide worldwide in 2019: global health estimates.* Report for the World Health Organization, 2021.

Appendix 1

Table A1. Demographic characteristics of participants.

Category	Value	Adults (n:480)	Adolescents (n:495)
Age	Mean	28.1	16.6
Most used social media platform	Facebook	13 (2.7%)	93 (18.8%)
	Instagram	387 (80.6%)	229 (46.3%)
	Twitter	78 (16.3%)	165 (33.3%)
	Others	2 (0.4%)	8 (1.6%)
Suicidal ideation or suicide attempt	Suicidal ideation	80 (16.7%)	99 (20.0%)
	Suicide attempt	58 (12.1%)	83 (16.8%)
Average monthly household income	Lower than 880 USD (1000 K KRW)	8 (1.7%)	12 (2.4%)
	880 USD (1000 K KRW)-1760 USD (2000 K KRW)	47 (9.8%)	35 (7.1%)
	1760 USD (2000 K KRW)-2640 USD (3000 K KRW)	100 (20.8%)	73 (14.7%)
	2640 USD (3000 K KRW)-3520 USD (4000 K KRW)	55 (11.6%)	86 (17.4%)
	3520 USD (4000 K KRW)-4400 USD (5000 K KRW)	61 (12.7%)	86 (17.4%)
	4400 USD (5000 K KRW)-5280 USD (6000 K KRW)	46 (9.6%)	57 (11.5%)
	5280 USD (6000 K KRW)-6160 USD (7000 K KRW)	25 (5.2%)	47 (9.5%)
	6160 USD (7000 K KRW)-7040 USD (8000 K KRW)	47 (9.8%)	32 (6.5%)
	7040 USD (8000 K KRW)-7920 USD (9000 K KRW)	18 (3.8%)	17 (3.4%)
	7920 USD (9000 K KRW)-8800 USD (10,000 K KRW)	17 (3.5%)	12 (2.4%)
	More than 8800 USD (10,000 K KRW)	54 (11.3%)	34 (6.9%)
	No income	2 (0.4%)	4 (0.8%)
Region	Seoul	112 (23.3%)	101 (20.4%)
	Busan	46 (9.6%)	21 (4.2%)
	Daegu	46 (9.6%)	20 (4.0%)
	Incheon	20 (4.2%)	43 (8.7%)
	Gwangju	33 (6.9%)	26 (5.3%)
	Daejeon	25 (5.2%)	24 (4.8%)
	Ulsan	9 (1.9%)	9 (1.8%)
	Gyeonggi-do	75 (15.6%)	133 (26.9%)

Table A1. Continued.

Category	Value	Adults (n:480)	Adolescents (n:495)
	Gangwon-do	5 (1.0%)	11 (2.2%)
	Chungcheongbuk-do	15 (3.1%)	12 (2.4%)
	Chungcheongnam-do	16 (3.3%)	16 (3.2%)
	Jeollabuk-do	14 (2.9%)	12 (2.4%)
	Jeollanam-do	11 (2.3%)	12 (2.4%)
	Gyeongsangbuk-do	20 (4.2%)	10 (2.0%)
	Gyeongsangnam-do	20 (4.2%)	37 (7.5%)
	Jeju-do	11 (2.3%)	3 (0.6%)
	Sejong	2 (0.4%)	5 (1.0%)

Appendix 2

Psychiatric symptoms measured by brief scales and self-reported mental disorders

Own psychiatric symptoms and mental disorders

[1] *Posttraumatic stress disorder (PTSD) symptoms* were measured using a brief PDS-5 rating scale (i.e., a 5-point Likert scale) for adults and a brief CPSS-V rating scale (5-point Likert scale) for adolescents developed by Chung et al.[1] consisting of nine items for adults and seven items for adolescents including (1) "Have you had recurrent or intrusive distressing thoughts or recollections about the assault?" and (2) "Have you had unwanted and upsetting thoughts or images about the assault that came into your mind?"

[2] *Depressive symptoms* were measured by a brief CES-D rating scale (4-point Likert scale) for adults and a brief CES-DC rating scale (4-point Likert scale) for adolescents developed by Chae et al.[2] consisting of four items for adults and three for adolescents including (1) "I was bothered by things that usually don't bother me," (2) "I had trouble focusing on what I was doing," and (3) "My sleep was restless."

[3] *Anxiety symptoms* were measured by a brief BAI rating scale (i.e. 4-point Likert scale) developed by Chae et al.[2] consisting of four items for adults and five items for adolescents including (1) fear of the worst happening, (2) scared, and (3) terrified or afraid.

[4] *Alcohol use disorder symptoms* were measured by the CAGE rating scale[3] which is a binary scale. We utilized one item ("Have you ever felt bad or guilty about your drinking?").

[5] *Suicide ideation or attempt* was measured by a C-SSRS rating scale (a binary scale) developed by Posner et al.[4] We used two items including (1) "Have you started to work out

or worked out the details on how to kill yourself? Do you intend to carry out this plan?" and (2) "Have you taken any steps toward making a suicide attempt or are preparing to kill yourself (such as collecting pills, giving valuables away, or writing a suicide note)?"

[6] *Self-reported mental disorders* were measured whether the participant had several mental disorders such as *obsessive-compulsive disorder* and *schizophrenia*.

Family history of mental disorders

[1] Family history of depression or bipolar disorder/alcohol use disorder or drug overdose/anxiety or panic disorder/obsessive compulsive disorder/schizophrenia/self-harm, or suicide attempt was measured by whether participants' family had histories related to mental disorders including depression, bipolar disorder, and alcohol use disorder using binary scales which were self-reported.

Internet use

[1] *Frequency of Internet use* was measured by questionnaire developed by the Korea Press Foundation.[5] It consists of 10 items, including how many times participants partake in (1) social media, (2) online shopping, and (3) online gaming in a week.

[2] *General problematic internet use (GPIU)* refers to the use of the Internet by participants that creates psychological, social, school, or work difficulties in participants' lives.[6] It was measured by GPIU2 (5-point Likert scale), which was developed by Caplan.[7] It consists of 15 items,

including (1) "I prefer online social interactions over face-to-face communication," (2) "I have difficulty controlling the amount of time I spend online," and (3) "My Internet use has created problems in my life."

[3] *Online behavior* was measured by the questionnaire (7-point Likert scale) developed by Ybarra et al.[8] It consists of 10 items, including (1) "Have you ever made rude or nasty comments online?", (2) "Have you ever visited an adult website on purpose?", and (3) "Have you ever talked about sex with someone known only online?"

Online/offline social networks

[1] *Online/offline communication and social networking* was measured through the questionnaire (5-point Likert scale), which was developed by referring to Kim's research[9] and consists of seven items, including (1) "How much time do you spend in face-to-face communication or interactions with others?", (2) "How much time do you spend using social media?", and (3) "How much time do you spend communicating via text messages or instant messages?"

[2] *Social support on social media* refers to the feeling that you are getting help to others[10] on social media. We used ISEL-12 (7-point Likert scale), which was developed by Cohen and Hoberman.[11] It consists of eight items, including (1) "When I need suggestions for how to deal with a personal problem, I know there is someone on social media I can turn to," (2) "I have someone on social media who takes pride in my accomplishments," and (3) "There is someone on social media with whom I can share my most private worries and fears."

Social media use
Basic characteristics of social media use

[1] *Basic characteristics of social media use* evaluates general use patterns on social media. It consists of 10 items, including (1) the type of social media platform used (Twitter, Facebook, Instagram, etc.), (2) the frequency of social media use, and (3) the number of friends on social media.

[2] *Motivations for social media use* was measured by Lai's questionnaire[12] which uses a 5-point Likert scale consisting of 11 items, including social media use (1) to maintain relationships with existing friends, (2) to express personal identity, and (3) to avoid being left out of friends' conversations.

Positive self-presentation

[1] *Positive self-presentation (PSP)* implies that participants selectively show positive aspects of themselves through social media.[13] We used Kim's questionnaire [13] based on a 7-point Likert scale consisting of seven items, including (1) "I post photos that only show the happy side of me" and (2) "I avoid writing about negative things that happen to me when I update my status."

[2] *Appraisal of PSP* reflects participants' self-appraisals when they express *PSP* on social media. We used Arbel's questionnaire[14] based on a 7-point Likert scale to measure this concept (i.e., "How does it make you feel when you express *PSP* on social media?").

Narcissistic self-presentation

[1] *Narcissistic self-presentation (NSP)* implies that participants share personal information to positively express their image on social media.[15] NSP (*exhibitionism*) refers to disclosing one's daily life to the public to seek or attract attention from others.[16] It was measured by Hollenbaugh's questionnaire[17] based on a 7-point Likert scale and consists of five items, including social media use (1) for attention, (2) to make yourself appear cool among your peers, and (3) to gain fame or notoriety.

[2] *NSP (ingratiation)* refers to strategies used so that one is liked by others[18] on social media. It was measured by Dominick's questionnaire[18] based on a 7-point Likert scale and consists of four items, including social media use (1) to receive ingratiation, (2) to present a good image to others, and (3) to express intimacy to others to obtain favors.

[3] *NSP (competence)* refers to being perceived as skilled and qualified[18] on social media. It was measured by Dominick's questionnaire[18] based on a 7-point Likert scale and consists of four items, including social media use (1) to post what you are good at or confident about, (2) to post that you have done a good job when doing something that others find difficult, and (3) to post about the importance of an affair that you finished successfully.

[4] *Appraisal of NSP* involves participants' self-appraisals when they express *NSP* on social media. We used Arbel's questionnaire[14] based on a 7-point Likert scale to measure this concept (i.e. "How does it make you feel when you express *NSP* on social media?").

Distress disclosure

[1] *Distress disclosure* refers to participants' active disclosure of their negative thoughts or emotions[19] on social media. We converted a general version of Kahn's questionnaire[19] for social media using a 7-point Likert scale that consisted of 13 items including (1) "When I feel upset, I usually post on social media," (2) "If I have a bad day, I like to talk about it on social media at the end of the day," and (3) "I typically don't discuss things that upset me on social media."

[2] *Appraisal of distress disclosure* involves participants' self-appraisal when they perform *distress disclosure* on social media. We used Arbel's questionnaire[14] based on a 7-point Likert scale to measure this concept ("How did it make you feel when you expressed *distress disclosure* on social media?").

[3] *Distress disclosure through Direct Message (DM)* is the same as *distress disclosure* but is expressed through DM.

[4] *Appraisal of distress disclosure through DM* is the same as *appraisal of distress disclosure* but is expressed through DM.

Feedback

[1] *Positive and negative feedback* refers to the frequency of sending or receiving feedback. We used Lai's questionnaire[12] and Miura's questionnaire[20] which are 5-point Likert scales. Nine items were used in the measurement, including (1) "In the past 30 days, how often have you sent positive feedback to your friends on social media?", (2) "In the past 30 days, how often have you sent negative feedback to your friends on social media?", and (3) "In the past 30 days, how often have you received negative feedback from your friends on social media?"

Negative experiences on social media

[1] *Negative experience on social media* was measured by Rosenthal and Stern's questionnaires[21] based on a 4-point Likert scale consisting of five items, including (1) "Throughout your use of social media over the past year, have you ever encountered bullying or meanness?", (2) "Throughout your use of social media over the past year, have you ever encountered unwanted contacts?", and (3) "Throughout your use of social media over the past year, have you ever encountered misunderstandings?"

Interpersonal experiences and needs
Interpersonal adverse experiences

[1] *Adverse childhood experiences (ACEs)* refer to abuse or other potentially damaging childhood experiences.[22] We used Lee's questionnaire[23] based on a 4-point Likert scale. It consists of 20 items, including (1) "Have your parents or other adults in your family often or very often sworn at or put you down?", (2) "Have you ever lived with a problematic drinker or alcoholic?", and (3) "Has your mother (or stepmother) been repeatedly assaulted for at least a few minutes?"

[2] *Adverse sexual experiences (ASEs)* refer to sexual abuse such as being touched without consent, being forced to touch someone else, or being forced to have penetrative sex.[24] We used Koss's questionnaire[25] based on a 4-point Likert scale consisting of 16 items, including (1) "Someone stared at me in a sexual way or looked at the sexual parts of my body after I had asked them to stop," (2) "Someone sent me sexual or obscene materials such as pictures, jokes, or stories in the mail or over the Internet after I had asked them to stop," and (3) "Someone made sexual gestures at me, such as grabbing their crotch, pretending to masturbate, or imitating oral sex without my consent."

[3] *Online interpersonal victimization* refers to an unwanted sexual solicitation or harassment online.[8] It was measured by the Ybarra et al. questionnaire[8], which uses a binary scale. It consisted of seven items, including, (1) "In the past year, did you experience the development of a close relationship with someone you met online?", (2) "Has anyone in the past year been asking for sexual information about you on the Internet when you don't want to answer those questions? (e.g. what your body looks like or details regarding your sexual behaviors?)", and (3) "Has anyone in the past year been asking you for unwanted sexual activity through the Internet?"

Need to belong

[1] *Social connectedness* refers to an emotional distance between self and others.[26] We used Lee and Robbins' questionnaire[26] based on a 7-point Likert scale consisting of eight items, including (1) "I feel disconnected from the world around me," (2) "Even around people I know, I don't feel like I belong," and (3) "I don't think I belong with anyone or to any group."

[2] *Social assurance* refers to the tendency to rely on others to feel a sense of belonging.[27] We used Lee and Robbins' questionnaire[26] based on a 7-point Likert consisting of eight items, including (1) "I feel comfortable when someone continues to be with me," (2) "If I don't' have a buddy beside me, my life is incomplete," and (3) "I always try to stick with my friends."

[3] *Fear of negative evaluation* refers to the fear that others will evaluate one negatively in social situations.[28] We used Carleton's questionnaire[29] based on a 7-point Likert scale consisting of 11 items, including (1) "I worry about what kind of impression I am making on someone," (2) "I am concerned about other people's opinion of me," and (3) "If I know someone is judging me, it tends to bother me."

Subjective wellbeing

[1] *Satisfaction with life* refers to an individual's general sense of satisfaction with their life as a whole.[30] We used

Diener's questionnaire[31] based on a 7-point Likert scale consisting of five items, including (1) "I think that my life is close to my ideal," (2) "I think that the conditions of my life are very good," and (3) "I am satisfied with my life."

Demographics

[1] *Age*—We measured each participant's age.
[2] *Region*—We identified each participant's current region of residence in South Korea.
[3] *Most used social media platform*—We measured which social media platform each participant uses from Twitter, Facebook, Instagram, and others.
[4] *Living arrangements*—We measured who cohabits with each participant, including (1) mother, (2) father, and (3) grandfather.
[5] *Number of family members living with*—We measured the number of family members each participant was living with.
[6] *Average monthly household income*—We measured each participant's average monthly household income, which was divided into 12 categories including (1) lower than 880 USD, (2) 880–1760 USD, and (3) 1760–2640 USD.
[7] *Economic status*—We measured each participant's economic status based on five: categories: (1) upper, (2) upper middle, (3) middle, (4) middle low, and (5) low.

References for appendix 2

1. Chung H, Kim J-W, Kwon J, et al. Development of brief post-traumatic stress disorder rating scale for sexual violence victims. *Psychiatry Investig* 2019; 16: 868–871.
2. Chae S, Shin J, Ji D, et al. A brief depression and anxiety rating scale for sexual violence victims. *Psychiatry Investig* 2021; 18: 172–178.
3. Lee S. Korean health and policy challenges. *Health and Welfare Policy Forum* 1997; 5–14.
4. Posner K, Brown GK, Stanley B, et al. The Columbia–Suicide Severity Rating Scale: Initial validity and internal consistency findings from three multisite studies with adolescents and adults. *Am J Psychiatry* 2011; 168: 1266–1277.
5. Korea Press Foundation. *2016 Teenagers Media Use Survey*. April 1, 2016.
6. Beard KW, Wolf EM. Modification in the proposed diagnostic criteria for Internet addiction. *CyberPsychol Behav* 2001; 4: 377–383.
7. Caplan SE. Theory and measurement of generalized problematic Internet use: a two-step approach. *Comput Hum Behav* 2010; 26: 1089–1097.
8. Ybarra ML, Mitchell KJ, Finkelhor D, et al. Internet prevention messages: targeting the right online behaviors. *Arch Pediatr Adolesc Med* 2007; 161: 138–145.

9. Kim J-H, Seo M, David P. Alleviating depression only to become problematic mobile phone users: can face-to-face communication be the antidote? *Comput Hum Behav* 2015; 51: 440–447.
10. Kim DH, Lee H-K, Kim JW, et al. Reliability and validity of the Korean version of interpersonal support evaluation list-12 (ISEL-12). *J Korean Neuropsychiatr Assoc* 2012; 51: 416–421.
11. Cohen S, Hoberman HM. Positive events and social supports as buffers of life change stress 1. *J Appl Soc Psychol* 1983; 13: 99–125.
12. Lai C-H. Motivations, usage, and perceived social networks within and beyond social media. *J Comput-Mediat Commun* 2019; 24: 126–145.
13. Kim J, Lee J-ER. The Facebook paths to happiness: effects of the number of Facebook friends and self-presentation on subjective well-being. *Cyberpsychol Behav Soc Netw* 2011; 14: 359–364.
14. Arbel R, Perrone L, Margolin G. Adolescents' daily worries and risky behaviors: the buffering role of support seeking. *J Clin Child Adolesc Psychol* 2018; 47: 900–911.
15. Yu E, Kim H-C. Is she really happy? A dual-path model of narcissistic self-presentation outcomes for female facebook users. *Comput Hum Behav* 2020; 108: 106328.
16. Guo M, Liu R-D, Ding Y, et al. How are extraversion, exhibitionism, and gender associated with posting selfies on WeChat friends' circle in Chinese teenagers? *Pers Individ Differ* 2018; 127: 114–116.
17. Hollenbaugh EE, Ferris AL. Facebook self-disclosure: examining the role of traits, social cohesion, and motives. *Comput Hum Behav* 2014; 30: 50–58.
18. Dominick JR. Who do you think you are? Personal home pages and self-presentation on the World Wide Web. *Journal Mass Commun Q* 1999; 76: 646–658.
19. Kahn JH, Hessling RM. Measuring the tendency to conceal versus disclose psychological distress. *J Soc Clin Psychol* 2001; 20: 41–65.
20. Miura A, Yamashita K. Psychological and social influences on blog writing: an online survey of blog authors in Japan. *J ComputMediat Commun* 2007; 12: 1452–1471.
21. Rosenthal SR, Buka SL, Marshall BDL, et al. Negative experiences on Facebook and depressive symptoms among young adults. *J Adolesc Heal* 2016; 59: 510–516.
22. Felitti VJ, Anda RF, Nordenberg D, et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults the adverse childhood experiences (ACE) study. *Am J Prev Med* 1998; 14: 245–258.
23. Jooyoung L, Su-Jin L, Ji-Ae B, et al. Gender differences in effects of child adversity, impulsivity, and drug use on risky sexual behaviors of juvenile probationers. *Korean J Probat* 2017; 2: 109–138.
24. Tucker W, Garn JV, Yang W. Relationship between adverse childhood experiences and gambling in Nevada. *Am J Heal Behav* 2021; 45: 246–255.

25. Koss MP, Abbey A, Campbell R, et al. Revising the SES: a collaborative process to improve assessment of sexual aggression and victimization. *Psychol Women Q* 2007; 31: 357–370.

26. Lee RM, Robbins SB. Measuring belongingness: the social connectedness and the social assurance scales. *J Couns Psychol* 1995; 42: 232–241.

27. Lee-Won RJ, Herzog L, Park SG. Hooked on Facebook: the role of social anxiety and need for social assurance in problematic use of Facebook. *Cyberpsychol Behav Soc Netw* 2015; 18: 567–574.

28. Rapee RM, Heimberg RG. A cognitive-behavioral model of anxiety in social phobia. *Behav Res Ther* 1997; 35: 741–756.

29. Carleton RN, McCreary DR, Norton PJ, et al. Brief fear of negative evaluation scale—revised. *Dépress Anxiety* 2006; 23: 297–303.

30. Pavot W, Diener E. The Satisfaction With Life Scale and the emerging construct of life satisfaction. *J Posit Psychol* 2008; 3: 137–152.

31. Diener E, Emmons RA, Larsen RJ, et al. The Satisfaction With Life Scale. *J Pers Assess* 1985; 49: 71–75.