



Case Control Study

# Predicting South Korea adolescents vulnerable to depressive disorder using Bayesian nomogram: A community-based cross-sectional study

Haewon Byeon

**Specialty type:** Psychiatry

**Provenance and peer review:**

Unsolicited article; Externally peer reviewed.

**Peer-review model:** Single blind

**Peer-review report's scientific quality classification**

Grade A (Excellent): 0  
Grade B (Very good): B, B, B  
Grade C (Good): 0  
Grade D (Fair): 0  
Grade E (Poor): 0

**P-Reviewer:** Kaur M, United States;  
Qian XQ, China; Yu RQ, China

**A-Editor:** Liu X, China

**Received:** February 6, 2022

**Peer-review started:** February 6, 2022

**First decision:** April 18, 2022

**Revised:** April 20, 2022

**Accepted:** June 22, 2022

**Article in press:** June 22, 2022

**Published online:** July 19, 2022



**Haewon Byeon**, Department of Medical Big Data, College of AI Convergence, Inje University, Gimhae 50834, Gyeongsangnamdo, South Korea

**Corresponding author:** Haewon Byeon, DSc, PhD, Associate Professor, Director, Department of Medical Big Data, College of AI Convergence, Inje University, No. 329 C-hall (Shineo Hall), Gimhae 50834, Gyeongsangnamdo, South Korea. [bhwpuma@naver.com](mailto:bhwpuma@naver.com)

## Abstract

### BACKGROUND

Although South Korea has developed and carried out evidence-based interventions and prevention programs to prevent depressive disorder in adolescents, the number of adolescents with depressive disorder has increased every year for the past 10 years.

### AIM

To develop a nomogram based on a naïve Bayesian algorithm by using epidemiological data on adolescents in South Korea and present baseline data for screening depressive disorder in adolescents.

### METHODS

Epidemiological data from 2438 subjects who completed a brief symptom inventory questionnaire were used to develop a model based on a Bayesian nomogram for predicting depressive disorder in adolescents.

### RESULTS

Physical symptoms, aggression, social withdrawal, attention, satisfaction with school life, mean sleeping hours, and conversation time with parents were influential factors on depressive disorder in adolescents. Among them, physical symptoms were the most influential.

### CONCLUSION

Active intervention by periodically checking the emotional state of adolescents and offering individual counseling and in-depth psychological examinations when necessary are required to mitigate depressive disorder in adolescents.

**Key Words:** Depressive disorder; Nomogram; Adolescents; Risk factor; Community-based

cross-sectional study; Brief symptom inventory

©The Author(s) 2022. Published by Baishideng Publishing Group Inc. All rights reserved.

**Core Tip:** The early detection and prevention of depressive disorder in adolescents is important because it not only adversely affects interpersonal relationships and academic achievement but also increases the probability of other related mental illnesses such as panic disorder. We developed a nomogram for screening depressive disorder using epidemiological data on 2438 adolescents. Physical symptoms, aggression, social withdrawal, attention, satisfaction with school life, mean sleeping hours, and conversation time with parents were influential factors on depressive disorder in adolescents.

**Citation:** Byeon H. Predicting South Korea adolescents vulnerable to depressive disorder using Bayesian nomogram: A community-based cross-sectional study. *World J Psychiatry* 2022; 12(7): 915-928

**URL:** <https://www.wjgnet.com/2220-3206/full/v12/i7/915.htm>

**DOI:** <https://dx.doi.org/10.5498/wjp.v12.i7.915>

## INTRODUCTION

Depressive disorder causes an emotional response that can make a person feel a sense of failure, a sense of loss, and/or a sense of worthlessness as a result of a negative perception of him/herself[1]. It is defined as a persistent feeling of sadness or hopelessness to the extent of not being able to maintain daily activities for 2 wk in the past year[1]. It has been reported that South Koreans experience depressive disorder most frequently during adolescence compared to other stages of life[2]. A national survey of South Korean adolescents reported that one in four males and one in three females were diagnosed with depressive disorder[3]. In particular, it has been reported that depression during adolescence increases rapidly after middle school[3,4], suggesting that the mental health of adolescents is at risk during this period.

Adolescence involves the most physical, mental, and social changes that occur in a human lifetime [5]. Adolescents experiencing depressive disorder are highly likely to self-torture and/or express delinquent and aggressive behavior[5]. Moreover, if depressive disorder is not identified and managed early, it may progress to become a chronic illness with depression likely to recur during a person's lifetime[6]. The early detection and prevention of depressive disorder in adolescence is an important social issue because it not only adversely affects interpersonal relationships and academic achievement but also increases the probability of developing other mental illnesses such as panic disorder[7].

Although South Korea has developed and carried out evidence-based intervention and prevention programs to mitigate depressive disorder in adolescents[2], the number of adolescents with depressive disorder has increased every year for the past 10 years[2]. Consequently, it is necessary to identify the influential factors causing depression and develop a predictive model with high accuracy that can identify groups highly vulnerable to depressive disorder as soon as possible.

Recently, the naïve Bayesian nomogram has been used as a method for predicting groups at high risk of developing diseases[8,9]. One of the advantages of this method is that it presents the risk probability according to multiple risk factors of a disease visually so that clinicians can easily understand the results [10]. In this study, a nomogram based on a naïve Bayesian algorithm using epidemiological data on adolescents in South Korea was developed and baseline data for screening depressive disorder in adolescents is presented.

## MATERIALS AND METHODS

### Data source

This is a secondary data analysis study using raw data from the 2019 Korean Children Youth Panel Study (KCYPs) survey from March to June 2019 provided by the National Youth Policy Institute. The study was approved by the Research Ethics Review Board of the National Youth Policy Institute (No. KCYPs-2018).

The survey method for KCYPs is presented in Cho *et al*[11] (2018). Briefly, the KCYPs sampled 7<sup>th</sup>-grade students attending 162 middle schools across South Korea using a stratified multi-stage cluster sampling method. Schools were selected according to the probability proportional to the size sampling method for 27 clusters across 16 metropolitan cities, small and medium-sized cities, and rural areas. After checking the information on the number of 7<sup>th</sup>-grade classes and the number of students in each

**Table 1** Measurements of explanatory variables

Classification	Variable	Characteristics
Sociodemographic factors	Gender	Male or female
	Number of siblings (including the subject)	1 person, 2 people, 3 people, or 4 people or more
Environmental factors	Mean conversation time with parents <i>per day</i>	< 30 min, ≥ 30 min and < 1 h, ≥ 1 h and < 2 h, ≥ 2 h and < 3 h, or ≥ 3 h
Personal factors	Satisfaction with academic achievement	Dissatisfied, not dissatisfied or satisfied, or satisfied
	Satisfaction with school life	Dissatisfied, not dissatisfied or satisfied, or satisfied
	Mean sleeping hours <i>per day</i>	< 5 h, 6 h, 7 h, 8 h, 9 h, or ≥ 10 h
	Social withdrawal	Continuous variable
	Aggression	Continuous variable
	Attention	Continuous variable
	Physical symptoms	Continuous variable

class at each school, samples were extracted by randomly selecting classes. The KCYPS collected data using a tablet-assisted personal interview method to compensate for the quality deterioration caused by existing questionnaire input errors or logical errors and to increase the accuracy and efficiency of the survey. In the present study, we analyzed 2438 subjects after excluding 152 cases with missing values in the depressive disorder screening part among 2590 people who completed the KCYPS questionnaire in 2019.

### Measurements

Depression, the outcome variable, was defined by using ten items for measuring depression in the brief symptom inventory (BSI) (1984)[12], which was adapted for the South Korean population by standardizing the Symptom Checklist-90-Revision[13]. The BSI is a self-reporting test with each item being measured on a 4-point scale. Moreover, the total score ranges from 10 to 40 points. A higher score indicates more severe depression. Referring to Byeon *et al*[14] (2015), the threshold for depression in this study was 24 points, corresponding to 1 standard deviation (less than the 16<sup>th</sup> percentile). *AORN J* reported that Cronbach's  $\alpha$  (a measurement of reliability) for the BSI was 0.904 (0.882 in the present study)[15].

Explanatory variables included gender, environmental factors (number of siblings and mean conversation time with parents during weekdays), and personal factors (satisfaction with academic achievement, satisfaction with school life, mean sleeping hours during weekdays, social withdrawal, aggression, attention, and physical symptoms). The definitions of the explanatory variables are provided in Table 1.

Social withdrawal was measured by using five items from the Behavior Problem Scale for Children and Adolescence (BPSCA) developed by Kim and Kim[16] (1998) after excluding items overlapping with other sub-domains. Each item was measured on a 4-point scale with the total score ranging from 5 to 20 points. A higher score indicates more severe social withdrawal. *AORN J* reported that Cronbach's  $\alpha$  for the tool was 0.850 (0.894 in the present study)[15].

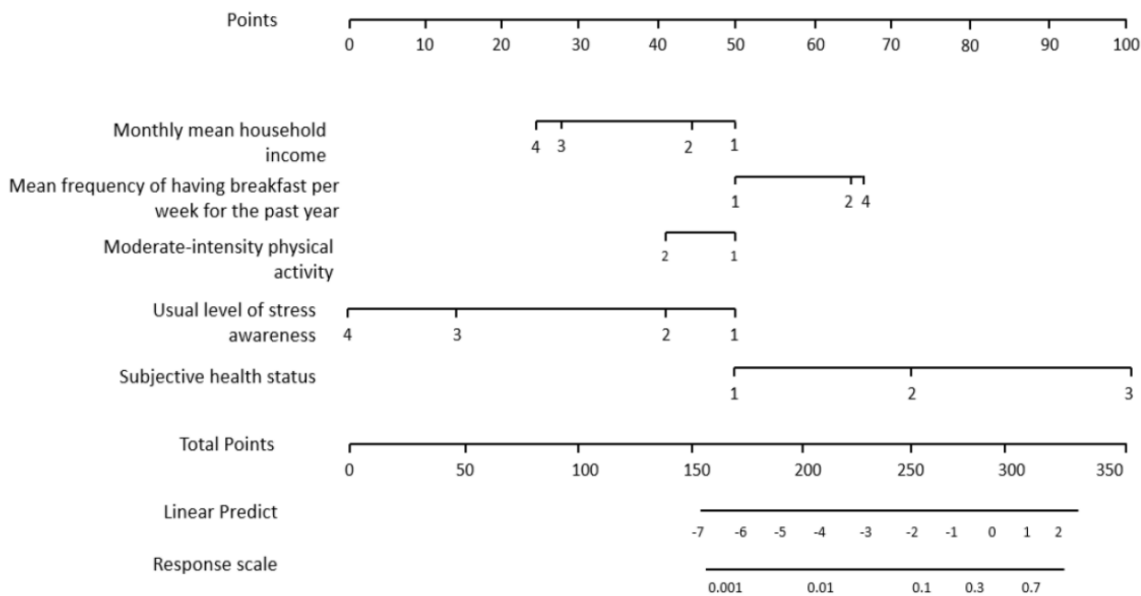
Aggression was measured by using the Emotional or Behavioral Problems Scale (EPS) developed by Cho and Lim[17] (2003). Six items were used and each item was measured on a 4-point scale with the total score ranging from 6 to 24 points, and a higher score indicates a more aggressive condition. Cho and Lim[17] (2003) reported that the Cronbach's  $\alpha$  of the tool was 0.760 (0.809 in the present study).

Attention problems were measured by using 7 items in the EPS[17]. Each item was measured on a 4-point scale with the total score ranging from 7 to 28 points. A higher score indicates more severe attention problems. Kim and Song[18] (2014) reported Cronbach's  $\alpha$  for the tool was 0.791 (0.813 in the present study).

Physical symptoms are when a person perceives that he or she is frequently ill or tired such as chest tightness or stomach discomfort without a pathological cause. These were measured by using eight items in the EPS[17]. Each item was measured on a 4-point scale with the total score ranging from 8 to 32 points. A higher score indicates more severe physical symptoms. Choi *et al*[19] (2017) reported Cronbach's  $\alpha$  for the tool was 0.87 (0.858 in the present study).

### Developing the naïve Bayes nomogram for predicting adolescents vulnerable to depressive disorder

A nomogram is used to visually present complex functions or calculations[20,21]. In particular, it is used as a method to visually present the diagnosis, recurrence, and survival prediction of a disease[20,21]. It is expressed graphically (Figure 1) in which a line is assigned to each attribute used as an input item and the possible value of the attribute is displayed on the line[22]. The score corresponding to the position of



**Figure 1** An example of a nomogram[22]. Citation: Byeon H. Developing a nomogram for predicting the depression of senior citizens living alone while focusing on perceived social support. *World J Psychiatry* 2021; 11: 1314-1327. Copyright ©The Authors 2021. Published by Baishideng Publishing Group Inc.

the attribute value becomes the individual score of the point displayed at the top.

We used a naïve Bayes classifier as the algorithm to develop the nomogram. A naïve Bayes classifier model determines the probability for a specific class by applying the Bayesian theorem under the assumption that the attributes of data and events are independent of each other. When the attributes are assumed to be independent, the posterior probability indicating the probability that an object (belongs to class  $C$ ) can be expressed as follows:

$$P(c|X) = \frac{P(a_1, a_2, \dots, a_m|c)P(c)}{P(X)} = \frac{P(c) \prod_i P(a_i|c)}{P(X)} \quad (1)$$

Where  $c$  is the target class of the nomogram. However, if it is a class other than  $c$ , and  $P(|X)$  represents the probability that object  $X$  does not belong to class  $c$ [9], then the odds ratio (OR) for these two probabilities can be calculated as:

$$OR = \frac{P(c|X)}{P(\bar{c}|X)} = \frac{P(c) \prod_i P(a_i|c)}{P(\bar{c}) \prod_i P(a_i|\bar{c})} \quad (2)$$

We calculated the general accuracy, precision, recall, F-1 score, the area under the curve (AUC), and calibration plot using leave-one-out cross-validation (LOOCV) of the developed Bayesian algorithm-based nomogram to validate its predictive performance.

Precision is defined as the proportion of classifications that are true actually being true:

$$(Precision) = \frac{TP}{TP + FP}$$

Recall is defined as the ratio of the number of model predictions that are true over the number that are actually true:

$$(Recall) = \frac{TP}{TP + FN}$$

Accuracy is an evaluation index that can most intuitively indicate the performance of a model:

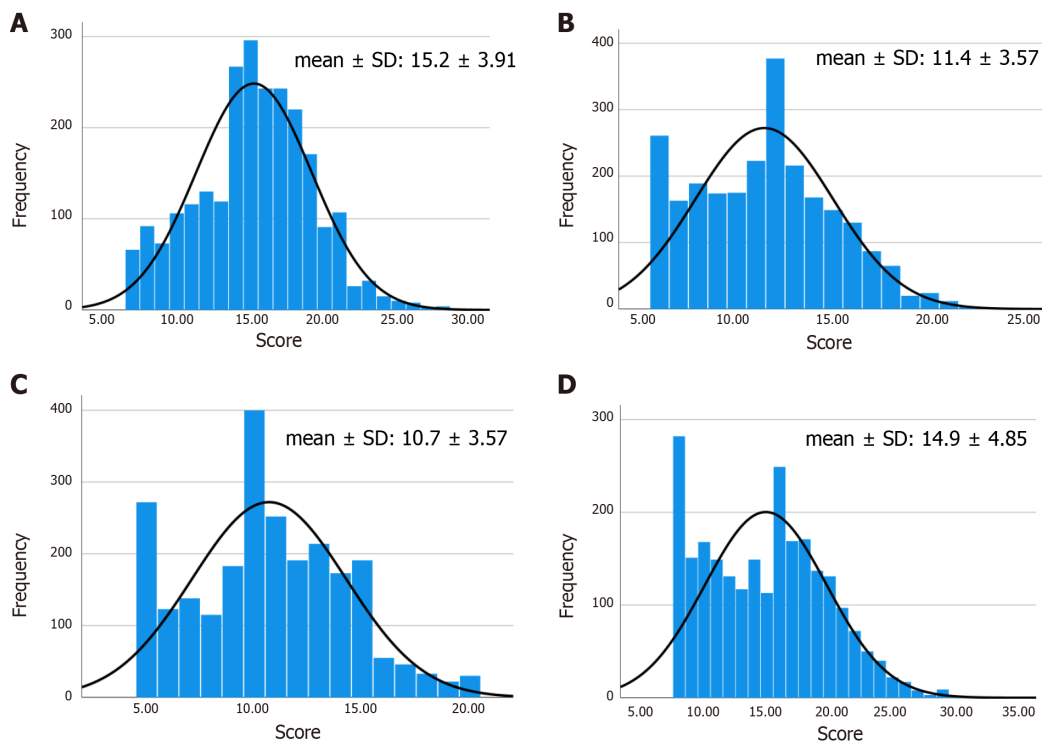
$$(Accuracy) = \frac{TP + TN}{TP + FN + FP + TN}$$

However, since using accuracy alone to overcome bias due to data imbalance is limited, it is necessary to present the F-1 score as an additional predictive performance indicator to overcome bias.

The F-1-score is the harmonic mean of Precision and Recall; *i.e.*

$$(F1 - score) = 2 \times \frac{1}{\frac{1}{Precision} + \frac{1}{Recall}} = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

AUC is an indicator used to evaluate the performance of a binary classifier. The maximum value is 1, and a value close to 1 means that the performance of the model is good (*i.e.*, the recall is larger than the



DOI: 10.5498/wjp.v12.i7.915 Copyright ©The Author(s) 2022.

**Figure 2 Test results.** A: Subject's attention test; B: Subject's aggression test; C: Subject's social withdrawal test; D: Subject's physical symptoms test.

fall-out). All analyses were performed by using Python version 3.10.0 (<https://www.python.org/downloads>; accessed on November 28, 2021).

## RESULTS

### General characteristics of the subjects

Of the 2438 subjects, the majority were male (54.1%) in a family with two siblings including the subject (61%); 32.0% had 30 minutes or more but less than 1 h of mean conversation time with their parents during weekdays; 37.2% were neither satisfied nor dissatisfied with last semester's school performance, 44.2% were satisfied with last semester's school life, and 40.4% slept for an average of 8 h *per* day during weekdays (Table 2). Their aggression, attention problems, physical symptoms, and social withdrawal are presented in Figure 2.

### General sample characteristics and prevalence of depressive disorder

The results of Chi-square tests showed significant differences ( $P < 0.05$ ) between the groups with and without depressive disorder in gender, mean sleeping hours *per* day, mean conversation time with parents *per* day, satisfaction with academic achievement, satisfaction with school life, attention, aggression, social withdrawal, and physical symptoms (Table 3).

### Correlations between variables

Correlation analysis results between the major variables used in this study are presented in Figure 3. Depressive disorder was significantly and positively correlated with attention, aggression, social withdrawal, and physical symptoms ( $P < 0.05$ ).

### Predicting the group of adolescents vulnerable to depressive disorder by using the Bayesian nomogram

Figure 4 shows the Bayesian nomogram for predicting the adolescent group vulnerable to depressive disorder. We developed a nomogram comprising seven variables with high importance: Physical symptoms, aggression, social withdrawal, attention, satisfaction with school life, mean sleeping hours, and conversation time with parents were the major influential factors associated with depression in adolescents. Physical symptoms comprised the most influential factor for predicting depression in this high-risk group. We predicted the depression risk of South Korean adolescents by using the developed nomogram (Figure 4). The high-risk group comprised those who received 15.5 points for physical

**Table 2** General characteristics of subjects (mean  $\pm$  SD)

Characteristic	<i>n</i>	%
<b>Depressive disorder</b>		
No	1999	82.0
Yes	439	18.0
<b>Gender</b>		
Male	1318	54.1
Female	1120	45.9
<b>Number of siblings (including the subject)</b>		
1 person	358	14.7
2 people	1487	61.0
3 people	515	21.1
4 people	78	3.2
<b>Mean sleeping hours <i>per day</i></b>		
< 5 h	63	2.6
6 h	236	9.7
7 h	600	24.6
8 h	986	40.4
9 h	454	18.6
$\geq 10$ h	99	4.1
<b>Mean conversation time with parents <i>per day</i></b>		
< 30 min	456	18.7
$\geq 30$ min and < 1 h	781	32.0
$\geq 1$ h and < 2 h	644	26.4
$\geq 2$ h and < 3 h	351	14.4
$\geq 3$ h	206	8.4
<b>Satisfaction with academic achievement</b>		
Dissatisfied	577	23.9
Not dissatisfied or satisfied	906	37.6
Satisfied	928	38.5
<b>Satisfaction with school life</b>		
Dissatisfied	144	5.9
Not dissatisfied or satisfied	616	25.4
Satisfied	1666	68.7
Attention	15.2 $\pm$ 3.9	
Aggression	11.4 $\pm$ 3.5	
Social withdrawal	10.6 $\pm$ 3.5	
Physical symptoms	14.9 $\pm$ 4.8	

symptoms (EPS test), 11.5 points for aggression (EPS test), 10.5 points for social withdrawal (BPSCA test), and 17.5 points for attention (EPS test) and were dissatisfied with their school life, slept 10 h or more *per day* on average, and talked with parents less than 30 min (84% of developing depression).

The predictive performance of the developed nomogram for predicting the adolescent group highly vulnerable to depressive disorder was validated by using AUC, F-1 score, accuracy, and a calibration plot. The results of the LOOCV evaluation show that the model had an AUC of 0.90 (Figure 5), F-1 score of 0.86, general accuracy of 0.85, precision of 0.88, and recall of 0.86. Adolescents with and without

**Table 3 Characteristics by prevalence of depressive disorder, *n* (%) (mean  $\pm$  SD)**

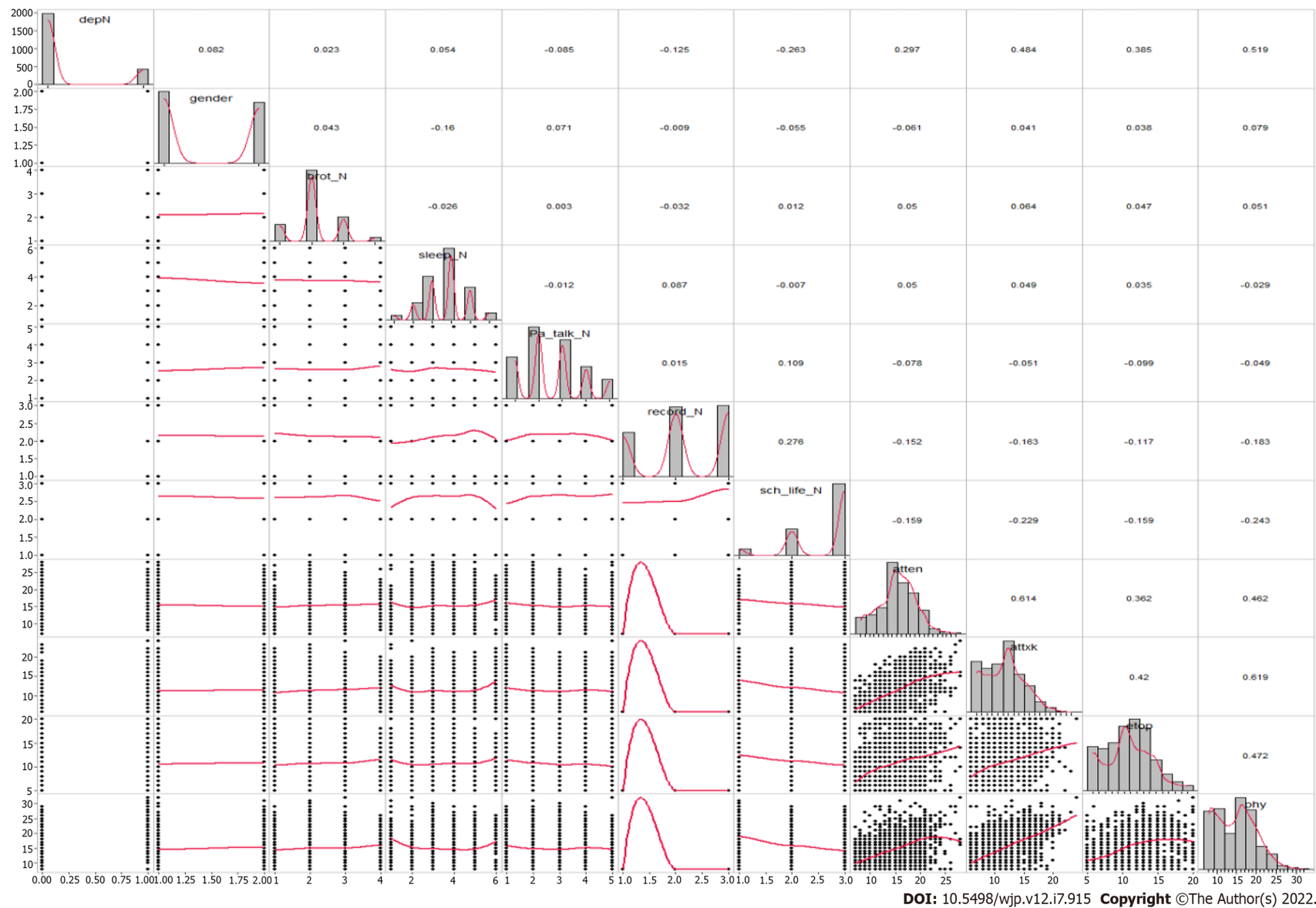
Characteristic	Depressive disorder		P value
	No ( <i>n</i> = 1999)	Yes ( <i>n</i> = 439)	
<b>Gender</b>			< 0.001
Male	1119 (84.9)	199 (15.1)	
Female	880 (78.6)	240 (21.4)	
<b>Number of siblings (including the subject)</b>			0.671
1 person	301 (84.1)	57 (15.9)	
2 people	1217 (81.8)	270 (18.2)	
3 people	419 (81.4)	96 (18.6)	
4 people	62 (79.5)	16 (20.5)	
<b>Mean sleeping hours per day</b>			<0.001
< 5 h	44 (69.8)	19 (30.2)	
6 h	191 (80.9)	45 (19.1)	
7 h	512 (85.3)	88(14.7)	
8 h	841 (85.3)	145 (14.7)	
9 h	350 (77.1)	104 (22.9)	
$\geq$ 10 h	61 (61.6)	38 (38.4)	
<b>Mean conversation time with parents per day</b>			< 0.001
< 30 min	240 (74.6)	116 (25.4)	
$\geq$ 30 min and < 1 h	645 (82.6)	136 (17.4)	
$\geq$ 1 h and < 2 h	539 (83.7)	105 (16.3)	
$\geq$ 2 h and < 3 h	293 (83.5)	58 (16.5)	
$\geq$ 3 h	182 (88.3)	24 (11.7)	
<b>Satisfaction with academic achievement</b>			< 0.001
Dissatisfied	434 (75.2)	143 (24.8)	
Not dissatisfied or satisfied	735 (81.1)	171 (18.9)	
Satisfied	812 (87.5)	116 (12.5)	
<b>Satisfaction with school life</b>			< 0.001
Dissatisfied	63 (43.8)	81 (56.3)	
Not dissatisfied or satisfied	470 (76.3)	146 (23.7)	
Satisfied	1457 (87.5)	209 (12.5)	
Attention	14.6 $\pm$ 3.8	17.7 $\pm$ 3.1	< 0.001
Aggression	10.6 $\pm$ 3.1	15.1 $\pm$ 3.1	< 0.001
Social withdrawal	10.0 $\pm$ 3.3	13.6 $\pm$ 2.9	< 0.001
Physical symptoms	13.7 $\pm$ 4.2	20.2 $\pm$ 3.7	< 0.001

depressive disorder were compared by using a calibration plot (Figure 6) and Chi-square tests based on the predicted and observed probability, between which there was no significant difference ( $P = 0.683$ ).

## DISCUSSION

This study was conducted to present baseline data for preventing depressive disorder in adolescents by identifying multiple influential risk factors. The results reveal that physical symptoms, aggression, social withdrawal, attention, satisfaction with school life, mean sleeping hours, and conversation time

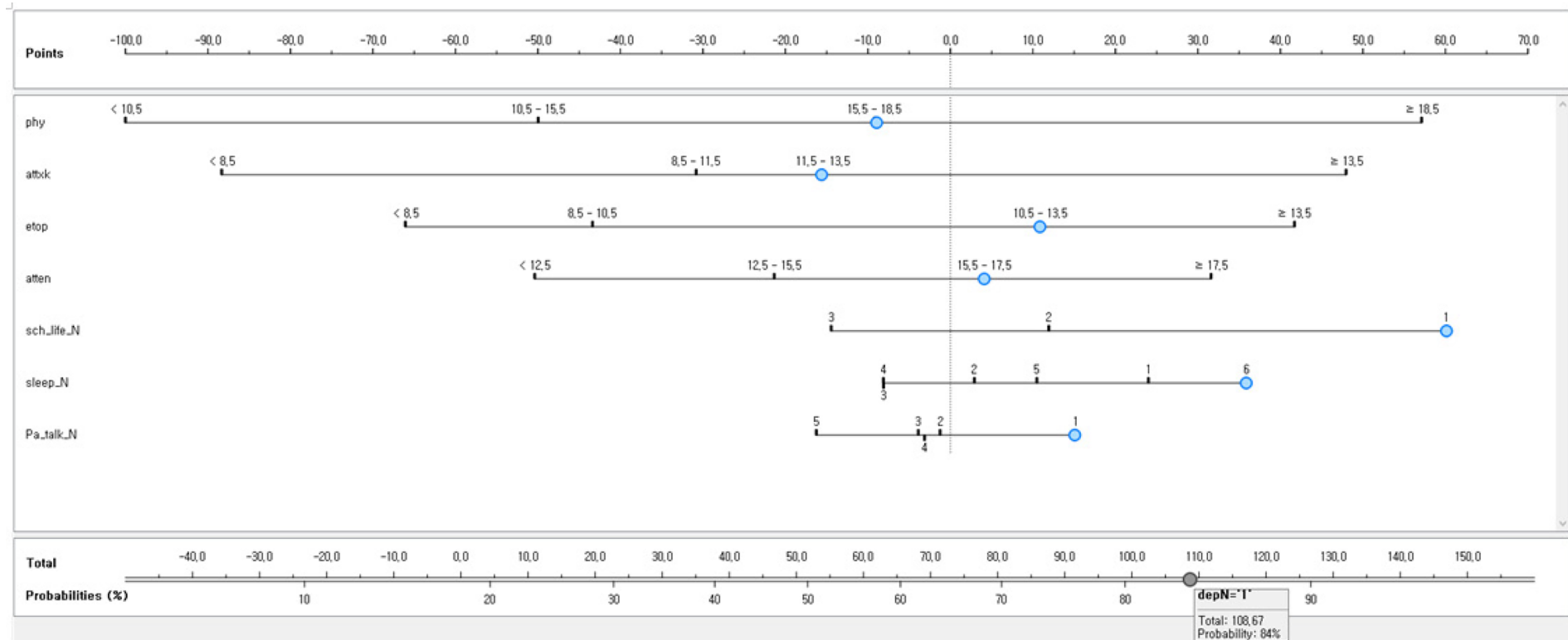




**Figure 3 Correlation between variables: Scatter matrix.** depN: Depressive disorder; brot\_n: Number of siblings; sleep\_N: Mean sleeping hours *per day*; Pa\_talk\_N: Mean conversation time with parents *per day*; record\_N: Satisfaction with



academic achievement; sch\_life\_N: Satisfaction with school life; atten: Attention; attxk: Aggression; etop: Social withdrawal; phy: Physical symptoms.



DOI: 10.5498/wjp.v12.i7.915 Copyright ©The Author(s) 2022.

**Figure 4 A model for predicting adolescent groups vulnerable to depressive disorder by using Bayesian nomograms.** phy: Physical symptoms; attxk: Aggression; etop: Social withdrawal; atten: Attention; sch\_life\_N: Satisfaction with school life (1 = dissatisfied, 2 = not dissatisfied or satisfied, 3 = satisfied); Pa\_talk\_N: Mean conversation time with parents per day (1: < 30 min, 2: ≥ 30 min and < 1 h, 3: ≥ 1 h and < 2 h, 4: ≥ 2 h and < 3 h; 5: ≥ 3 h).

with parents were significant predictors. Among them, physical symptoms had the greatest influence on the depressive disorder of adolescents. The outcomes of numerous previous studies on variables associated with depression in adolescents identify peer relationships, the home environment, and the school environment as significant risk factors[23-26], which supports the findings of the present study.

From the perspective of the socioecological model, family, peer group, and school are three major domains directly affecting the mental health of adolescents[27,28]. Since risk and protective factors are generated in these three domains[27,28], the viewpoint of the socioecological model is useful for

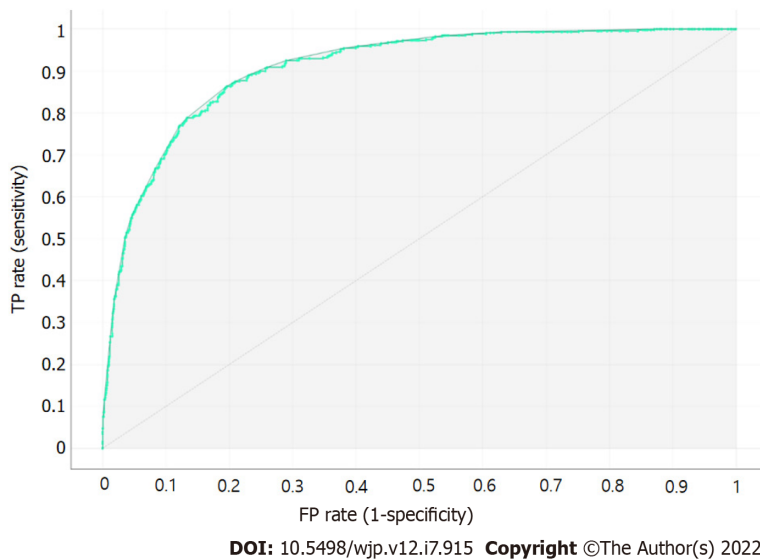


Figure 5 Receiver operating characteristic analysis result of the developed model.

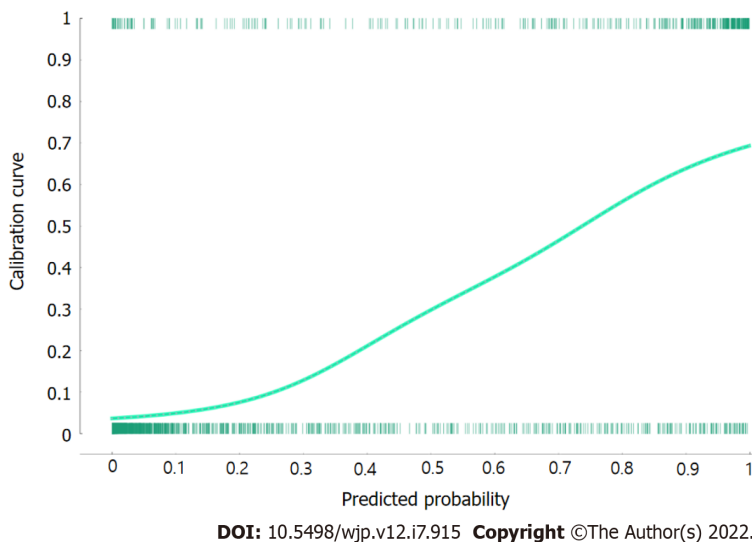


Figure 6 Calibration plot comparing predicted to actual probability of depressive disorder.

explaining the depressive disorder of adolescents as the outcome of multiple risk factors. Nevertheless, there were limitations in previous studies about explaining the relationship between multiple risk factors and depressive disorder[2,25]. First, researchers mainly used regression models as the method of exploring the risk factors associated with depressive disorder in adolescents. Although calculating ORs by using regression analysis is useful for identifying individual risk factors, its ability to identify complex multiple risk factors is limited. Second, only sociodemographic and environmental factors as risk factors associated with depressive disorder were identified in previous studies[23,25]. Indeed, comprehensive analysis of risk factors associated with depressive disorder in adolescents by using individual factors such as difficulty with attention, social withdrawal, as well as environmental factors, has still not been sufficiently conducted. Third, as normality and independence between the variables are assumed in regression analysis, it is difficult to draw accurate conclusions because the data on many diseases such as depressive disorder are unbalanced, thereby violating the normality assumption.

In summary, there are limitations when utilizing depression prediction models for adolescents based on regression analysis in the primary medical care environment because it is difficult to identify the complex relationships between multiple risk factors solely relying on ORs based on regression models. Thus, we identified an adolescent group highly vulnerable to depressive disorders by using multiple risk factors based on a Bayesian nomogram to overcome these limitations. Our results predicted that adolescents who received 15.5 points in physical symptoms, 11.5 points in aggression, 10.5 points in social withdrawal, and 17.5 points in attention and who were dissatisfied with their current school life, slept for 10 h or more *per* day on average, and talked with their parents less than 30 min have a

depression risk of 84%. Therefore, communities and schools must continually monitor the high-risk group for the early identification and prevention of depressive disorder in adolescents with these multiple risk factors.

Another important finding of the present study is that physical symptoms in adolescents comprised the most influential risk factor in predicting depressive disorder. Ryu and Hong[29] (2019) also explored factors affecting depressive disorder in 1881 middle school students and confirmed that physical symptoms of adolescents comprised the main risk factor influencing depressive disorder. Choi *et al*[19] (2017) also revealed that physical symptoms and depressive disorder had a positive correlation in fourth graders.

Physical symptoms in adolescents, which are related to mental activities and the psychological state, are generally overlooked as an early symptom of depressive disorder because they cannot be found by internal or neurological examination, or even when a physical abnormality is found, the symptoms are insufficient for disease diagnosis. However, although depressive disorder of adolescents is similar to the adult psychopathology, unlike in adults, clinical characteristics are often accompanied by physical symptoms (*e.g.*, fatigue, insomnia, muscle pain, and headache) and aggression[30]. In particular, Jung *et al*[31] (2004) reported that depressed people excessively focus on physical symptoms or amplify their bodily sensations. Therefore, frequent complaints by adolescents of physical symptoms without a known medical cause are likely to be early signs of depressive disorder, even when the physical symptoms seem superficially less severe. Consequently, the community and school must pay attention to them and actively intervene by periodically checking the emotional state of adolescents, as well as providing individual counseling and in-depth psychological testing.

The strength of the present study is that it identified the group at high risk of developing depressive disorder based on multiple risk factors by using epidemiological data on South Korean adolescents and provided evidence for the early screening and management of depression. However, it does have some limitations, with the first being that there could be more potential variables for depressive disorder in addition to the explanatory variables used in this study because we analyzed secondary data. Second, the results cannot be generalized for all high school students because we identified a high-risk group for depressive disorder in seventh graders only. Third, the variables used (including depressive disorder) were measured based on a self-report questionnaire. Thus, future studies are needed to identify groups at high risk of depressive disorder by integrating qualitative research methods such as Delphi analysis and in-depth interviews in addition to self-report questionnaires. Fourth, since the results were based on a cross-sectional approach, it is difficult to determine causal relationships. Hence, additional prospective cohort studies should be conducted to prove causality between the depressive disorder high-risk group and depressive disorder found in the present study.

## CONCLUSION

We showed that physical symptoms, aggression, social withdrawal, attention, satisfaction with school life, mean sleeping hours, and conversation time with parents are influential factors associated with depressive disorder in adolescents. Among them, physical symptoms comprise the most influential factor in the prediction of depressive disorder. Therefore, periodically checking on the emotional state of adolescents is required, along with providing individual counseling and conducting in-depth psychological examinations when necessary. Moreover, longitudinal studies based on clinical depressive disorder data targeting depressive disorder in the high-risk group confirmed in this study should be conducted.

## ARTICLE HIGHLIGHTS

### Research background

Although South Korea has developed and carried out evidence-based intervention and prevention programs to mitigate depressive disorder in adolescents, the number of adolescents with depressive disorder has increased every year for the past 10 years. Consequently, it is necessary to identify the influential factors causing depression and develop a predictive model with high accuracy that can identify groups highly vulnerable to depressive disorder as soon as possible.

### Research motivation

Recently, the naïve Bayesian nomogram has been used as a method of predicting groups at high risk of developing diseases. One of the advantages of this method is that it presents the risk probability according to multiple risk factors of a disease visually so that clinicians can easily understand the results.

### Research objectives

In this study, a nomogram based on a naïve Bayesian algorithm using epidemiological data on adolescents in South Korea was developed and baseline data for screening depressive disorder in adolescents was presented.

### Research methods

We used a naïve Bayes classifier as the algorithm to develop the nomogram. Also, we calculated the general accuracy, precision, recall, F-1 score, the area under the curve, and calibration plot using leave-one-out cross-validation of the developed Bayesian algorithm-based nomogram to validate its predictive performance.

### Research results

We showed that physical symptoms, aggression, social withdrawal, attention, satisfaction with school life, mean sleeping hours, and conversation time with parents were influential factors associated with depressive disorder in adolescents.

### Research conclusions

Periodically checking on the emotional state of adolescents is required, along with providing individual counseling and conducting in-depth psychological examinations when necessary.

### Research perspectives

Longitudinal studies based on clinical depressive disorder data targeting depressive disorder in the high-risk group confirmed in this study should be conducted.

---

## FOOTNOTES

**Author contributions:** Byeon H designed the study, interpreted the data, preformed the statistical analysis, and wrote the article.

**Supported by** Basic Science Research Program through the National Research Foundation of Korea (NRF) Funded by the Ministry of Education, No. NRF-2018R1D1A1B07041091 and No. NRF-2021S1A5A8062526.

**Institutional review board statement:** The study was approved by the Research Ethics Review Board of the National Youth Policy Institute (No. KCYPS-2018).

**Informed consent statement:** All patients gave informed consent prior to study participation.

**Conflict-of-interest statement:** All the authors report no relevant conflicts of interest for this article.

**Data sharing statement:** Technical appendix, statistical code from the corresponding author at [bhwpuma@naver.com](mailto:bhwpuma@naver.com).

**STROBE statement:** The authors have read the STROBE Statement – checklist of items, and the manuscript was prepared and revised according to the STROBE Statement – checklist of items.

**Open-Access:** This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <https://creativecommons.org/licenses/by-nc/4.0/>

**Country/Territory of origin:** South Korea

**ORCID number:** Haewon Byeon [0000-0002-3363-390X](https://orcid.org/0000-0002-3363-390X).

**S-Editor:** Fan JR

**L-Editor:** Wang TQ

**P-Editor:** Guo X

---

## REFERENCES

- 1 **Otte C**, Gold SM, Penninx BW, Pariante CM, Etkin A, Fava M, Mohr DC, Schatzberg AF. Major depressive disorder. *Nat Rev Dis Primers* 2016; 2: 16065 [PMID: [27629598](https://pubmed.ncbi.nlm.nih.gov/27629598/) DOI: [10.1038/nrdp.2016.65](https://doi.org/10.1038/nrdp.2016.65)]
- 2 **Yun JY**, Chung H, Sim JA, Yun YH. Prevalence and associated factors of depression among Korean adolescents. *PLoS*

- One* 2019; **14**: e0223176 [PMID: 31618232 DOI: 10.1371/journal.pone.0223176]
- 3 **Namgung HK**, Woo HW, Shin J, Shin MH, Koh SB, Kim HC, Kim YM, Kim MK. Development and validation of hypertension prediction models: The Korean Genome and Epidemiology Study\_Cardiovascular Disease Association Study (KoGES\_CAVAS). *J Hum Hypertens* 2022 [PMID: 35181762 DOI: 10.1038/s41371-021-00645-x]
  - 4 **Twenge JM**, Nolen-Hoeksema S. Age, gender, race, socioeconomic status, and birth cohort differences on the children's depression inventory: a meta-analysis. *J Abnorm Psychol* 2002; **111**: 578-588 [PMID: 12428771 DOI: 10.1037//0021-843x.111.4.578]
  - 5 **Melo AK**, Siebra AJ, Moreira V. Depression in adolescents: review of the literature and the place of phenomenological research. *Psicologia Ciência e Profissão* 2017; **37**: 18-34 [DOI: 10.1590/1982-37030001712014]
  - 6 **de Zwart PL**, Jeronimus BF, de Jonge P. Empirical evidence for definitions of episode, remission, recovery, relapse and recurrence in depression: a systematic review. *Epidemiol Psychiatr Sci* 2019; **28**: 544-562 [PMID: 29769159 DOI: 10.1017/S2045796018000227]
  - 7 **Franić S**, Middeldorp CM, Dolan CV, Ligthart L, Boomsma DI. Childhood and adolescent anxiety and depression: beyond heritability. *J Am Acad Child Adolesc Psychiatry* 2010; **49**: 820-829 [PMID: 20643315 DOI: 10.1016/j.jaac.2010.05.013]
  - 8 **Yoo JS**, Son JW, Nam MS. The effects of a depression intervention and suicide prevention program in adolescents with high risk of suicide. *J Korean Acad Community Health Nurs* 2010; **21**: 71-81 [DOI: 10.12799/jkachn.2010.21.1.71]
  - 9 **Byeon H**. Developing a Predictive Model for Depressive Disorders Using Stacking Ensemble and Naive Bayesian Nomogram: Using Samples Representing South Korea. *Front Psychiatry* 2021; **12**: 773290 [PMID: 35069283 DOI: 10.3389/fpsy.2021.773290]
  - 10 **Park JC**, Lee JY. How to build nomogram for type 2 diabetes using a naïve Bayesian classifier technique. *J Appl Stat* 2018; **45**: 2999-3011 [DOI: 10.1080/02664763.2018.1450366]
  - 11 **Cho YJ**, Atteraya MS, Joo HC. The effects of child maltreatment on childhood behaviour problems in South Korea: findings from the fifth Korea Child and Youth Panel Survey (the 2014 KCYPS). *Asia Pac J Soc Work Dev* 2018; **28**: 39-55 [DOI: 10.1080/02185385.2017.1401956]
  - 12 **Kim KI**, Kim JH, Won HT. Korean manual of symptom checklist-90-revision. Chung Ang Aptitude Publishing: Seoul, 1984
  - 13 **Derogatis LR**, Lipman RS, Covi L. SCL-90: an outpatient psychiatric rating scale--preliminary report. *Psychopharmacol Bull* 1973; **9**: 13-28 [PMID: 4682398]
  - 14 **Byeon H**, Lee Y, Lee SY, Lee KS, Moon SY, Kim H, Hong CH, Son SJ, Choi SH. Association of alcohol drinking with verbal and visuospatial memory impairment in older adults: Clinical Research Center for Dementia of South Korea (CREDOS) study. *Int Psychogeriatr* 2015; **27**: 455-461 [PMID: 25119654 DOI: 10.1017/S104161021400146X]
  - 15 Evidence appraisal of Kim G, Kim MH, Lee SM, Choi SJ, Shin YH, Jeong HJ. Effect of pre-warmed intravenous fluids on perioperative hypothermia and shivering after ambulatory surgery under monitored anesthesia care [published online ahead of print April 1, 2014]. *J Anesth. AORN J* 2014; **100**: 445-450 [PMID: 25381678 DOI: 10.1016/j.aorn.2014.07.011]
  - 16 **Kim SH**, Kim KY. Development of behavior problem scale for children and adolescence. *J Korean Home Manag Assoc* 1998; **16**: 155-166 [DOI: 10.19031/jkheea.2020.09.32.3.161]
  - 17 **Cho BH**, Lim KH. Development and validation of emotional or behavioral problems scale. *Korean J Counsel Psychotherapy* 2003; **15**: 729-746 [DOI: 10.18253/kart.2016.16.1.01]
  - 18 **Kim JM**, Song SJ. The moderating effect of depression on the relationship between attention problems and school adjustment levels in middle school students. *Stud Korean Youth* 2014; **25**: 5-27 [DOI: 10.14816/sky.2014.25.1.5]
  - 19 **Choi YH**, Jung CS, You JO, Hyung NK. Factors influencing the somatization of elementary school students. *J Korean Public Health Nurs* 2017; **31**: 505-517 [DOI: 10.4069/kjwhn.2017.23.2.135]
  - 20 **Karaismailoglu E**, Karaismailoglu S. Two novel nomograms for predicting the risk of hospitalization or mortality due to COVID-19 by the naïve Bayesian classifier method. *J Med Virol* 2021; **93**: 3194-3201 [PMID: 33599308 DOI: 10.1002/jmv.26890]
  - 21 **Seo JH**, Lee JY. Novel nomogram based on risk factors of chronic obstructive pulmonary disease (COPD) using a naïve Bayesian classifier model. *J Korean Stat Soc* 2019; **48**: 278-286 [DOI: 10.1016/j.jkss.2018.11.006]
  - 22 **Byeon H**. Developing a nomogram for predicting the depression of senior citizens living alone while focusing on perceived social support. *World J Psychiatry* 2021; **11**: 1314-1327 [PMID: 35070780 DOI: 10.5498/wjp.v11.i12.1314]
  - 23 **McCarty CA**, Mason WA, Kosterman R, Hawkins JD, Lengua LJ, McCauley E. Adolescent school failure predicts later depression among girls. *J Adolesc Health* 2008; **43**: 180-187 [PMID: 18639792 DOI: 10.1016/j.jadohealth.2008.01.023]
  - 24 **Schwartz-Mette RA**, Shankman J, Dueweke AR, Borowski S, Rose AJ. Relations of friendship experiences with depressive symptoms and loneliness in childhood and adolescence: a meta-analytic review. *Psycho Bull* 2020; **146**: 664-700 [DOI: 10.1037/bul0000239]
  - 25 **Rahman MA**, Todd C, John A, Tan J, Kerr M, Potter R, Kennedy J, Rice F, Brophy S. School achievement as a predictor of depression and self-harm in adolescence: linked education and health record study. *Br J Psychiatry* 2018; **212**: 215-221 [PMID: 29506597 DOI: 10.1192/bjp.2017.69]
  - 26 **Bernaras E**, Jaureguizar J, Garaigordobil M. Child and Adolescent Depression: A Review of Theories, Evaluation Instruments, Prevention Programs, and Treatments. *Front Psychol* 2019; **10**: 543 [PMID: 30949092 DOI: 10.3389/fpsyg.2019.00543]
  - 27 **Hong JS**, Lee J, Espelage DL, Hunter SC, Patton DU, Rivers T Jr. Understanding the Correlates of Face-to-Face and Cyberbullying Victimization Among U.S. Adolescents: A Social-Ecological Analysis. *Violence Vict* 2016; **31**: 638-663 [PMID: 27506491 DOI: 10.1891/0886-6708.VV-D-15-00014]
  - 28 **Raneri LG**, Wiemann CM. Social ecological predictors of repeat adolescent pregnancy. *Perspect Sex Reprod Health* 2007; **39**: 39-47 [PMID: 17355380 DOI: 10.1363/3903907]
  - 29 **Ryu JL**, Hong SH. The convergent factors related to depression in the Korean adolescent: focusing on the data of the Korean children and youth panel survey 2016. *J Convergen Inf Technol* 2019; **9**: 180-188 [DOI: 10.37727/jkdas.2019.21.3.1585]
  - 30 **Domalanta DD**, Risser WL, Roberts RE, Risser JM. Prevalence of depression and other psychiatric disorders among

incarcerated youths. *J Am Acad Child Adolesc Psychiatry* 2003; **42**: 477-484 [PMID: [12649635](#) DOI: [10.1097/01.CHL.0000046819.95464.0B](#)]

- 31 **Jung HY**, Park JH, Lee SI. The cognitive characteristics of somatizer according to depressive symptoms and sex. *J Korean Neuropsychiatr Assoc* 2004; **43**: 831-845 [DOI: [10.4306/jknpa.2012.51.2.70](#)]





Published by **Baishideng Publishing Group Inc**  
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA

**Telephone:** +1-925-3991568

**E-mail:** [bpgoffice@wjgnet.com](mailto:bpgoffice@wjgnet.com)

**Help Desk:** <https://www.f6publishing.com/helpdesk>

<https://www.wjgnet.com>

