

Authors' Responses to the Reviewers' Comments – Revision 1

World Journal of Psychiatry - Manuscript ID: 75588

Title: Predicting South Korea Adolescents Vulnerable to Depressive Disorder Using Bayesian Nomogram: a Community-based Cross-sectional Study

We appreciate the editor, who gave opportunity to revise our work. Also, we would like to thank the reviewers for careful and thorough reading of this manuscript and for the thoughtful comments and constructive suggestions, which help to improve the quality of this manuscript. We carefully considered your comments as well as those offered by the reviewers. We agree with most of them, and the manuscript has been revised thoroughly according to the reviewer's advice. We hope that these revisions improve the paper such that you and the reviewers now deem it worthy of publication in World Journal of Psychiatry. Also, we hope our revision meet your approval. We next detail our responses to each reviewer's concerns and comments.

Our response follows (the reviewer's comments and editor's comments are in italics).

Editor

1) Thank you for submitting your manuscript to World Journal of Psychiatry, a peer-reviewed, online, and open access journal. We are pleased to inform you that one of the peer reviewers has completed his/her review of your manuscript.

We appreciate the editor, who gave opportunity to revise our work. We carefully considered your comments as well as those offered by the reviewers. We agree with most of them, and the manuscript has been

revised thoroughly according to the reviewer's advice. We hope that these revisions improve the paper such that you and the reviewers now deem it worthy of publication in World Journal of Psychiatry.

Reviewer #1

The study developed a nomogram based on a naïve Bayesian algorithm using epidemiological data that could represent adolescents in the South Korean community and presented baseline data for screening adolescent depressive disorder. The study is interesting and relevant however there are some minor grammatical errors throughout the manuscript. I suggest that the manuscript should be proofread by native English speaker to correct all the grammatical errors.

We appreciate the reviewer's detailed evaluations and positive comments. We appreciate the positive feedback. Also, we appreciate you very much for giving us the opportunity to revise our paper. We agree with most of them, and the manuscript has been revised thoroughly according to the reviewer's advice. Our changes have been marked in red font and highlight in the revised manuscript. We hope that these revisions improve the paper such that reviewers now deem it worthy of publication in World Journal of Psychiatry. Also, we hope our revision meet your approval.

We have received native-speaking corrections based on reviewer's suggestions. Here's what's changed:

Abstract

BACKGROUND

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

AIM

This study developed a nomogram based on a naïve Bayesian algorithm by using epidemiological data that could represent adolescents in the South Korean community and presented baseline data for screening adolescent depressive disorder in adolescents.

METHODS

Epidemiological This study analyzed data from 2,438 subjects who completed the Brief Symptom Inventory (BSI) questionnaire were used to develop. The a prediction model was developed using based on the a Bayesian nomogram for predicting depressive disorder in adolescents.

RESULTS

The results of this study showed that physical symptoms, aggression, social withdrawal, attention, satisfaction with school life, mean sleeping hours, and conversation time with parents were the influencing influential factors for on adolescent depressive disorder in adolescents. Among them, the physical symptoms was were the most influential factor in the prediction of depressive disorder.

CONCLUSION

it will be required to provide a Active intervention and attention to by periodically checking the emotional state of adolescents and offering, provide individual counseling, and enable 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

Core tip: The Early detection and prevention of the depressive disorder in adolescents in adolescence is an important social issue because depressive disorder in adolescence not only adversely affects interpersonal relationships and academic achievement but also increases the probability of other accompanying related other mental disorders illnesses such as panic disorder. In our study, we developed a nomogram for screening adolescent depressive disorder using epidemiological data. As a result of developing a nomogram for on 2,438 adolescents, physical-Physical symptoms, aggression, social withdrawal, attention, satisfaction with school life, mean sleeping hours, and conversation time with parents were the influencing factors for on adolescent depressive disorder in adolescents.

INTRODUCTION

Depressive disorder is causes an emotional disorder response which that can makes a person experience feel depression, a sense of failure, a sense of loss, and/or a sense of worthlessness as a result of a negative perception of himself or 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 two weeks in the past year [1]. It has been reported that South Koreans experience depressive disorder most frequently during the adolescence period compared to other stages of life [2]. A national survey using of South Korean adolescents [2] reported that one out of four male students and one out of three female students 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 is a period of involves the most experiencing physical, mental, and social changes that occur the most and developing physical, mental, and social elements the most in a human lifetime [5]. The depressive disorder experienced in adolescence Adolescents experiencing depressive disorder are highly likely to be expressed as self-torturing behaviors and/or express delinquent behaviors, and aggressive behaviors [3]. Moreover, if a depressive disorder is not identified and managed early, it may progress to become a chronic illness or with depression is more likely to recur during a person's lifetime even after recovering from depression [6]. The Early detection and prevention of the depressive disorder in adolescence is an important social issue because depressive disorder in adolescence not only adversely affects interpersonal relationships and academic achievement but also increases the probability of accompanying developing other mental disorders illnesses such as panic disorder [7].

Although South Korea has developed and carried out evidence-based interventions and prevention programs to ~~prevent/mitigate~~ depressive disorder ~~and maintain mental health~~ in adolescents ⁽²⁾, the number ~~of adolescents with depressive disorder in South Korea~~ has increased every year for the past ~~ten-10~~ years ⁽²⁾. Consequently, it is necessary to identify ~~complex-the~~ influential factors ~~affecting-causing~~ depression and develop a predictive model with high accuracy that can ~~predict-identify~~ groups highly vulnerable to depressive disorder ~~based on it for identifying adolescents with depressive order~~ as soon as possible. ⁽³⁾

Recently, ~~a-the~~ naïve Bayesian nomogram has been used ~~in recent years~~ as a method for predicting ~~high-risk~~ groups ~~of at high risk of developing~~ diseases ⁽⁹⁾. ~~One of the advantages of this method is that The naïve Bayesian nomogram~~ presents the risk probability ~~due according to the~~ multiple risk factors of a disease visually so that clinicians can easily understand ~~the results, which is an advantage of this method.~~ ⁽¹⁰⁾. ~~In This~~ ~~this~~ study, ~~developed~~ a nomogram based on a naïve Bayesian algorithm using epidemiological data ~~that could represent on~~ adolescents in ~~the~~ South Korean ~~was community-developed~~ and ~~presented~~ baseline data for screening ~~adolescent~~ depressive disorder ~~in adolescents is presented.~~ ⁽⁴⁾

METHODOLOGYS

Data source

This is a secondary data analysis study using ~~the~~ raw data ~~of from~~ the 2019 Korean Children Youth Panel Study (KCYPs) surveyed 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). ⁽⁴⁾

The survey method ~~of for the~~ KCYPs ~~was specifically~~ presented in Cho et al. (2018) ⁽¹¹⁾. Briefly ~~speaking~~, the KCYPs sampled 7th-grade students attending 162 middle schools across South Korea using ~~the a~~ stratified multi-stage cluster sampling method. ~~This study selected target s~~ Schools ~~were selected~~ according to the probability proportional to ~~the~~ size sampling method for 27 clusters; ~~which were extracted by~~ ~~acrossing~~ 16 metropolitan cities ~~and provinces and city sizes (i.e., metropolitan cities,~~ small and medium-sized cities, and rural areas). After checking the information on the number of 7th-grade classes and the number of students in each class ~~for at~~ each school, samples were extracted by randomly selecting ~~sample~~ classes. The KCYPs collected data using ~~the 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 This the present study,~~ we analyzed 2,438 subjects after excluding 152 cases with missing values in the depressive disorder screening ~~test part~~ among 2,590 people who completed the KCYPs ~~questionnaire~~ in 2019. ⁽⁴⁾

4

Measurements

Depression, the outcome variable, was defined by using ~~ten~~ 10 items for measuring depression in the Brief Symptom Inventory (BSI) (1984) ^[13], which was ~~developed~~ adapted for the South Korean population by standardizing the Symptom Checklist-90-Revision (SCL-90-R) ~~developed by Derogatis (1973)~~ ^[12]. The BSI is a self-reporting test ~~and with~~ each item ~~is-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. (2015) ^[14], the threshold for depression in this study was 24 points, corresponding to 1 standard deviation (less than ~~the~~ 16th percentile). Choi ~~& and~~ Shin (2016) ^[15] reported that ~~the~~ Cronbach's α , ~~(a measurement of reliability), for~~ the BSI was 0.904, ~~and (0.882 in Cronbach's α - the present study) was .882 in this study.~~

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 ~~presented~~ provided in Table 1.

4

(Table 1 insertion)

4

Social withdrawal was measured by using five items from the Behavior Problem Scale for Children and Adolescence (BPSCA) developed by Kim ~~and~~ Kim (1998) ^[16] after excluding items overlapping with other sub-domains. Each item was measured on a 4-point scale, ~~and with~~ the total score ~~ranged-ranging~~ from 5 to 20 points. A higher score indicates more severe social withdrawal. Choi ~~& and~~ Shin ^[15] reported that ~~the~~ Cronbach's α (reliability) ~~offor~~ the tool was 0.850, ~~(and Cronbach's α was .894 in the presentis study).~~

4

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

4

Attention problems were measured ~~with-by using~~ 7 items in the EPS ^[17]. Each item was measured on a 4-point scale ~~with~~, ~~and~~ the total score ~~ranges-ranging~~ from 7 to 28 points. A higher score indicates more

severe attention problems. Kim and Song (2014) [18] reported that the Cronbach's α of the tool was 0.791, and Cronbach's α was (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. Physical symptoms These were measured with by using 8 items in the EPS [17]. Each item was measured on a 4-point scale, and with the total score ranges ranging from 8 to 32 points. A higher score indicates more severe physical symptoms. Choi et al., (2017) [19] reported that the Cronbach's α of the tool was 0.87, and Cronbach's α was (0.858 in the present study).

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

The A nomogram is used to visually present complex functions or calculations visually [20,21]. In particular, it is used as a method to visually present the diagnosis, recurrence, and survival prediction of a disease [20,21]. The nomogram It is expressed as a graphically (Figure 1), where 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 the attribute value becomes the individual score of the point displayed at the top.

(Fig 1 insertion)

This study We used the a naïve Bayesian classifier as an the algorithm to develop the nomogram. A The the naïve Bayesian 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 to of each other. When the attributes are assumed to be independent, the posterior probability $P(c|X)$ can be expressed as Equation (1), where the posterior probability indicates the probability that an object $(X = (a_1, a_2, \dots, a_m))$ belongs to class C can be expressed as follows:

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

When where c is the target class of a the nomogram. However, and if c_e is a class other than c , and $P(c_e|X)$ represents the probability that object X does not belong to class c [9], then:

The odds ratio (Pdds OR) for these two probabilities can be expressed as Equation (2), calculated as:

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

This study We presented-calculated the general accuracy, precision, recall, F1 score, the area under the curve (AUC), and calibration plot using leave-one-out cross-validation (LOOCV) of the developed Bayesian algorithm-based nomogram for-to validating-validate the-its predictive performance-of-the developed Bayesian algorithm-based nomogram.

Precision is defined as the proportion of what the model of classifies-classifications that as-are true that is actually being true.

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

Recall is defined as the ratio of the number of what the model predicts-predictions to-bethat are true out of what is 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 has-a-limit-in-to overcoming-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-the bias.

The F1-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 is derived-for-a-good-means that the performance of the model is good (that is, i.e., the the larger the Recall value compared to is larger than the Fallfall-out). All analyses were performed by using Python version 3.10.0 (<https://www.python.org/downloads/>; (accessed on 28 November 2021).

RESULTS

The General-general characteristics of the subjects

~~The total subjects were~~Of the 2,438 ~~students~~ subjects. ~~The~~ the majority of subjects were male (54.1%) ~~in a family with~~ two siblings including the subject (61%). ~~32.0% of them had~~ 30 minutes or more ~~and but~~ less than 1 hour of mean conversation time with ~~their~~ parents during weekdays (~~32.0%~~), ~~37.2% of them were not neither~~ satisfied ~~nor~~ dissatisfied with last semester's school performance (~~37.2%~~), ~~44.2% were~~ satisfied with last semester's school life (~~44.2%~~), and ~~40.4% slept for an average of~~ 8 hours of ~~mean sleeping hours per day~~ during weekdays (~~40.4%~~) (Table 2). Their aggression, attention problems, physical symptoms, and social withdrawal ~~of subjects~~ are presented in Figs. 2, 3, 4, and 5, ~~respectively~~.

+

(Table 2 insertion)⁺

(Fig 2 insertion)⁺

(Fig 3 insertion)⁺

(Fig 4 insertion)⁺

(Fig 5 insertion)⁺

The General-general sample characteristics by and the prevalence of depressive disorder⁺

~~The general characteristics of the subjects (n=2,438) are presented by the prevalence of depressive disorder (Table 3).~~ The results of ~~chi~~Chi-square tests showed ~~significant differences (p <0.05) that between the groups with and without depressive disorder and the group without depressive disorder were significantly (p<0.05) different in~~ according to 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).⁺

+

(Table 3 insertion)⁺

+

Correlations between the variables⁺

Correlation analysis ~~results was performed to identify the correlation~~ between the major variables used in this study, ~~and the results~~ are presented in Figure 6. Depressive disorder was significantly (~~p<0.05~~) and positively correlated with attention, aggression, social withdrawal, and physical symptoms (~~p <0.05~~).⁺

+

(Fig 6 insertion)⁺

+

~~Development of aP-model for predicting~~ the adolescent groups of adolescents vulnerable to depressive disorder by using ~~the~~ Bayesian nomograms⁺

Figure 7 shows the naive-Bayesian nomogram for predicting the South Korean the adolescent groups vulnerable to depressive disorder. This study We developed a nomogram by selecting comprising only seven variables with high importance. It was confirmed that physical symptoms, aggression, social withdrawal, attention, satisfaction with school life, mean sleeping hours, and conversation time with parents were the major influence-influential factors of associated with adolescent depression in adolescents. the physical symptoms was comprised the most influential factor for predicting adolescent depression in this a-high-risk groups. This study We predicted the depression risk of South Korean adolescents by using the developed nomogram (Figure 7). The results predicted that adolescents The high-risk group comprised those who received 15.5 points in-for physical symptoms (EPS test), 11.5 points in-for aggression (EPS test), 10.5 points in-for social withdrawal (BPSCA test), and 17.5 points in for attention (EPS test) and were dissatisfied with current-their school life, slept 10 hours or more per day on average, and talked with parents less than 30 minutes would have (84% of developing depression risk).⁴⁾

↓
(Figure 7 insertion)⁴⁾

The predictive performance of the developed nomogram for predicting the adolescent groups highly vulnerable to depressive disorder was validated by using AUC, F-1 score, accuracy, and a calibration plot. The results of the LOOCV evaluation Leave One Out Cross Validation showed that the model had an AUC was of 0.90 (Figure 8), an F-1 score was of 0.86, general accuracy of was 0.85, precision was of 0.88, and recall was of 0.86. Adolescents with and without depressive disorder and adolescents without depressive disorder were compared by using a calibration plot (Figure 9) and chi-square tests based on the predicted probability and observed probability, between which. There there was no significant difference between predicted probability and observed probability (P = 0.683).⁴⁾

↓
(Figure 8 insertion)⁴⁾

(Figure 9 insertion)⁴⁾

DISCUSSION

This study was conducted to present baseline data for preventing depressive disorder in adolescents by identifying multiple influential risk factors influencing the depressive disorder of adolescents. The results of this study revealed that physical symptoms, aggression, social withdrawal, attention, satisfaction with school life, mean sleeping hours, and conversation time with parents were significant predictors. Among them, physical symptoms had the greatest influence on the depressive disorder of adolescents. The outcomes of Numerous- numerous previous studies [23,24,25,26] explored-on variables associated with adolescent-depression in adolescents and reported that identify peer relationships, the home environment, and the school environment were-as significant risk factors [23,24,25,26], which supported the results-findings of this-the present study.^[1]

From the perspective of the socioecological model, family, peer group, and school are three major environments-domains directly affecting the mental health of adolescents [27,28]. Moreover, Since the socioecological model considers that risk factors and protective factors are generated in these three domains [27,28]. Therefore, the viewpoint of the socioecological-socioecological model is a useful model to-for explaining the depressive disorder of adolescents, which is an-as the outcome of multiple risk factors. Nevertheless, there were limitations in previous studies [2,25] had limitations in-about explaining the relationship between multiple risk factors and depressive disorder [2,25]. First, they-researchers mainly used regression models as a-the method to-for exploring the risk factors for-associated with adolescent depressive disorder in adolescents. Although Presenting-calculating the odds-ratio ORs by using regression analysis is useful for identifying individual risk factors, but its is limited inability to identifying complex multiple risk factors is limited. Second, some previous studies [23,24] identified-only sociodemographic and environmental factors as risk factors for-associated with depressive disorder were identified in previous studies [23,25]. Indeed, There are still not enough studies that comprehensively analyzed-analysis of risk factors for-associated with adolescent-depressive disorder in adolescents by using individual factors such as the-difficulty in-with attention, and social withdrawal recognized-by adolescents, as well as environmental factors, has still not been sufficiently conducted. Third, as the regression model assumes the-normality and independence of-between the variables are assumed in regression analysis, it is difficult to draw accurate results-conclusions because the data on many diseases data, such as depressive disorder, are unbalanced, thereby-data violating the normality assumptions.^[1]

In summary, there are limitations in-when utilizing adolescent-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 *odd-ratio* ORs based on regression models. Thus, ⁴

~~This epidemiological study~~we identified ~~the-an~~ adolescent groups highly vulnerable to depressive disorders ~~by~~ using multiple risk factors based on ~~the-a~~ Bayesian nomogram to overcome ~~these~~ limitations of previous studies. ~~The-Our~~ results of this study 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 hours or more per day on average, and talked with ~~their~~ parents less than 30 minutes ~~would have a 84% of depression risk of 84%~~. Therefore, communities and schools ~~are needed to must~~ ~~continuously-continually~~ monitor the high-risk groups ~~for depressive disorder~~ for the early identification ~~and prevention~~ of depressive disorder ~~for-in~~ adolescents with these multiple risk factors ~~and the prevention of adolescent depressive disorder~~.⁴

Another important finding of ~~this-the~~ present study ~~was-is~~ that ~~the~~ physical symptoms ~~of-in~~ adolescents ~~was-comprised~~ the most influential risk factor in predicting depressive disorder. Ryu & Hong (2019) ^[29] also explored ~~the~~ factors affecting depressive disorder ~~using-in~~ 1,881 middle school students and confirmed that ~~the~~ physical symptoms of adolescents ~~was-comprised~~ the ~~major-main~~ risk factor influencing depressive disorder ~~the most~~. Choi et al., (2017) ^[49] also revealed that physical symptoms and depressive disorder had a positive correlation ~~after evaluating-in~~ fourth graders.⁴

~~Physical symptoms in adolescents, which refer to symptoms that a person perceives that he or she is frequently ill or tired such as chest tightness or stomach discomfort without a pathological cause. They are related to mental activities and the psychological state. These physical symptoms of adolescents are generally overlooked as an early symptom of depressive disorder because physical abnormality is they can-not be found in-by internal-medicine or neurological examination, or even if-when a physical abnormality is found, the symptoms are insignificant-insufficient to be diagnosed as a-for disease diagnosis. However, although the-depressive disorder of adolescents is basically not different-similar from-to the adult psychopathology, unlike in adults, clinical characteristics are characteristically-often accompanied by physical symptoms (e.g., fatigue, insomnia, muscle pain, and headache) and aggression [36]. In Particularly-particular]~~ Jung et al., (2004) ^[31] reported that depressed people ~~paid more attention to their internal status to~~ 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 ~~if-when~~ the physical symptoms ~~of adolescents~~ seem superficially

less severe. Consequently, ~~it will be needed for~~ the community and school ~~to must actively intervene and~~ pay attention to them ~~and actively intervene so that~~ by adolescents with physical symptoms can ~~periodically checking~~ their emotional state ~~of adolescents periodically and, as well as receive~~ providing individual counseling and in-depth psychological ~~testtesting~~.⁴

The strength of the ~~presentis~~ study ~~was-is~~ that it identified ~~high-risk the groups for at high risk of~~ ~~developing the adolescent~~ depressive disorder based on multiple risk factors ~~by~~ using epidemiological data ~~that could represent~~ South Korean adolescents and ~~presented-provides~~ evidence for ~~the~~ early screening and management of depression ~~based on findings~~. ~~However, it does have some limitations. The~~ ~~the first limitation of this study was being~~ that there could be more potential variables for depressive disorder in addition to the explanatory variables ~~included-used~~ in this study because ~~this study we~~ analyzed secondary data. Second, the results ~~of this epidemiological study could-annot~~ be generalized for ~~all~~ high school students because ~~this study we~~ identified ~~a~~ high-risk groups for depressive disorder ~~using-in~~ seventh graders ~~only~~. Third, the variables used ~~in this study,~~ (including depressive disorder, ~~-~~) were measured based on a self-report questionnaire. ~~Thus, Future future~~ studies are needed to identify ~~high risk~~ 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 ~~of this study~~ were based on a cross-sectional approach, it is difficult to ~~interpret the results as adetermine~~ causal relationships. ~~Hence, It will be necessary to conduct~~ additional prospective cohort studies ~~should be conducted~~ to prove ~~the~~ causality between the depressive disorder high-risk group and depressive disorder, found in ~~this-the present~~ study.⁴

CONCLUSIONS

~~The results of this study~~ We showed that physical symptoms, aggression, social withdrawal, attention, satisfaction with school life, mean sleeping hours, and conversation time with parents were ~~the~~ ~~influencing-influential~~ factors ~~for associated with adolescent~~ depressive disorder ~~in adolescents~~. Among them, ~~the~~ physical symptoms ~~was-comprised~~ the most influential factor in the prediction of depressive disorder. Therefore, ~~it will be required to provide active intervention and attention to~~ periodically ~~checking on~~ the emotional state of adolescents ~~is required, along with,~~ providing individual counseling, and ~~enable-conducting~~ in-depth psychological examinations ~~when necessary. Moreover, It is necessary to~~ ~~conduct additional~~ longitudinal studies based on clinical depressive disorder data targeting depressive disorder ~~in the~~ high-risk groups confirmed in this study ~~should be conducted~~.⁴

Reviewer #2

This study was conducted to present baseline data for preventing depressive disorder in adolescents by identifying multiple risk factors influencing the depressive disorder of adolescents.

We appreciate the reviewer's detailed evaluations and positive comments. We appreciate the positive feedback. Also, we appreciate you very much for giving us the opportunity to revise our paper. We agree with most of them, and the manuscript has been revised thoroughly according to the reviewer's advice. Our changes have been marked in red font and highlight in the revised manuscript. We hope that these revisions improve the paper such that reviewers now deem it worthy of publication in World Journal of Psychiatry. Also, we hope our revision meet your approval.

1.) It is important to screen for and prevention of Adolescent Depression. But there are still two small problems: 1. Please include relevant information on statistics in the Methods section.

We would like to thank the reviewer for careful and thorough reading of this manuscript and for the thoughtful comments and constructive suggestions, which help to improve the quality of this manuscript. Also, we appreciate you very much for giving us the opportunity to revise our paper. We agree with the reviewers' comments. In response to a reviewer's suggestion, we've added the following text to the Methods section:

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 F1-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 fall-out).

2.) 2. Please explain what is the recall? Why is so high to 0.86?.

We appreciate the reviewer's detailed evaluations and positive comments. Also, we appreciate you very much for giving us the opportunity to revise our paper. Recall is defined as the ratio of the number of model predictions that are true over the number that are actually true. The reason that the recall was derived as 0.86 in this study is thought to be because this predictive model was derived with high precision and recall. In response to a reviewer's suggestion, we've added the following text to the Methods section:

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

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

We hope our revision meet your approval.

Reviewer #3

The study is well-designed and authors have explained the objective of the study with great clarity. The rationale of choosing the target participants of the study was explained well. The statistical methods used have also been described well. Various bodies of literature have emphasized complaints of physical symptoms as a presenting complaint for children and adolescent belonging to Asian and Hispanic ethnicities. The findings objectively highlight that physical symptoms can be a predictive factor for MDD in children and adolescents. Authors have done justice in noting the limitations of the study including use of self-report questionnaire. I personally enjoyed reading this study and found the findings to be of great clinical relevance.

We appreciate the reviewer's detailed evaluations and positive comments. Also, we appreciate the positive feedback. We appreciate you very much for giving us the opportunity to revise our paper. We agree with most of them, and the manuscript has been revised thoroughly according to the reviewer's advice. Our changes have been marked in red font and highlight in the revised manuscript. We hope that these revisions improve the paper such that reviewers now deem it worthy of publication in World Journal of Psychiatry. Also, we hope our revision meet your approval.