LI, Yan and TENG, Luyan. Predicting Mental Health Among Adolescents with Risk Behaviors Based on Machine Learning. Pedagogy and Psychology of Sport. 2025;28:67521. eISSN 2450-6605.
https://doi.org/10.12775/PPS.2025.28.67521
https://apcz.umk.pl/PPS/article/view/67521

The authors declare that there is no conflict of interests regarding the publication of this paper.
Received: 16.11.2025. Revised: 19.12.2025. Accepted: 19.12.2025. Published: 19.12.2025.

Predicting Mental Health Among Adolescents with Risk Behaviors Based on Machine Learning

Yan Li[1], Luyan Teng[2,*]

[1] Department of Psychology, Faculty of Medicine, University of Helsinki, Helsinki, Finland, Email: yan.z.li@helsinki.fi https://orcid.org/0000-0002-2977-4945

[2] College of International Education, Sichuan International Studies University, China, Email: luyan.teng@outlook.com https://orcid.org/0000-0001-7673-3217

*Corresponding Author

Abstract

Objectives

While mental health is known to predict risk behaviors, less is understood about how specific risk behaviors contribute to mental health outcomes, particularly across genders. This study used machine learning to examine the predictive relationships between various risk behaviors and adolescent mental health, and to explore gender differences in these patterns.

Methods

We analyzed data from the nationally representative Chinese "Database of Youth Health," including 8,670 high school students surveyed in 2020. A gradient-boosted decision tree model (XGBoost) was used to predict mental health, measured by the Symptom Checklist-90 (SCL-90), based on 22 risk behaviors. SHAP (Shapley Additive ExPlanations) values were calculated to interpret individual feature contributions.

Results

The model showed good performance (RMSE = 0.49, MAE = 0.35). Frequent dizziness during sports, lack of seat belt use, and alcohol consumption were identified as significant risk factors.

Gender differences emerged: earlier age of first smoking was more strongly associated with poorer mental health among girls, while exercise frequency was a stronger protective factor for boys.

Conclusion

These findings underscore the need for gender-sensitive mental health interventions that address both physical and behavioral risk factors, and demonstrate the utility of machine learning in identifying nuanced predictors of adolescent mental health.

Keywords: risk behaviors, mental health; adolescents, machine learning; gender differences

Introduction

Adolescent risk behaviors encompass activities such as drug and alcohol abuse, reckless driving, and other dangerous behaviors (Sullivan et al., 2010). Mental health problems such as anxiety, depression, and emotional dysregulation in adolescents can predict increased engagement in risky behavior (Deng et al., 2024; Jones et al., 2011; Kessler et al., 2005). However, limited studies focusing on how risk behaviors predict mental health outcomes, although studies found they strongly correlated with each other (Brooks et al., 2002). During adolescence, males are more prone to exhibit externalizing disorders such as substance abuse, which even lead to suicide death (Miranda-Mendizabal et al., 2019). In contrast, females are more likely to develop internalizing mental health issues such as depression (Rosenfield & Mouzon, 2013). However, we know little about how specific risk behavior affect the mental health of male and female differently.

Machine learning (ML) offers an advanced approach over traditional statistical methods, excelling in handling high-dimensional and complex datasets and revealing non-linear relationships (Abdolali & Gillis, 2021; Aghaabbasi & Chalermpong, 2023). This capability enhances predictive accuracy and uncovers deeper data patterns often missed by conventional approaches. Therefore, the current study aims to build an explainable predictive model to identify how specific risk behaviors predict mental health outcomes among adolescents in China by using the ML data analysis methods. Through this investigation, we seek to provide valuable insights into the interplay between risk behaviors and mental health, identify gender-specific patterns, and contribute to the development of targeted interventions for improving adolescent well-being. We intend to address the following key research questions (RQ):

RQ1: What are the most significant high-risk behaviors that predict adolescents' mental health problems in China?

RQ2: Is there any gender difference in the relationship between risk behaviors and mental health problems among Chinese adolescents?

Method

Participants

This study utilized data from the openly available Chinese nationally representative dataset "Database of Youth Health" (DYH), including 8,670 high school students (boy = 4144, girl = 4526) surveyed in 2020, aged 15 to 18 years (Shengfa et al., 2022). The detailed background information is presented in Table 1.

Table 1

Characteristics	Boys (n = 4144) *M*	*SD*	Girls (n = 4526) *M*	*SD*	Gender differences
Mental health issues	1.51	0.60	1.58	0.59	$t(8668) = -5.31, p <.001$
Risk behaviors					
Seat belt use frequency	3.68	1.44	3.77	1.36	$t(8668) = -3.01, p < .01$
Rides with drunk driver	1.23	0.74	1.14	0.58	$t(8668) = 6.38, p < .001$
Age first smoking	1.76	1.64	1.32	1.13	$t(8668) = 14.19, p < .001$
Smoking days	1.34	1.12	1.15	0.75	$t(8668) = 9.16, p < .001$
Daily cigarettes	1.30	0.98	1.15	0.72	$t(8668) = 8.18, p < .001$
Alcohol days (lifetime)	1.97	1.83	1.50	1.39	$t(8668) = 13.38, p < .001$
Age first drink	2.31	2.10	1.72	1.69	$t(8668) = 14.43, p < .001$
Alcohol days (last 30 days)	1.59	1.47	1.28	1.06	$t(8668) = 11.01, p < .001$
Exercise days	3.46	2.30	2.98	2.14	$t(8668) = 10.20, p < .001$
Daily TV hours (weekdays)	1.69	1.40	1.47	1.14	$t(8668) = 8.08, p < .001$
Daily non-study screen time (weekdays)	2.18	1.65	1.84	1.35	$t(8668) = 10.48, p < .001$
Weekly PE days	2.77	1.21	2.70	1.12	$t(8668) = 2.95, p < .001$
Sports classes attended (12 months)	1.52	0.81	1.39	0.71	$t(8668) = 7.64, p < .001$
Duration of extracurricular sports	1.59	1.14	1.41	0.99	$t(8668) = 7.82, p < .001$
Father exercise	3.28	1.51	3.41	1.49	$t(8668) = -4.03, p < .001$

Mother exercise	3.23	1.49	3.35	1.46	$t(8668) = -3.73$, $p < .001$
Family exercise	2.30	1.38	2.23	1.32	$t(8668) = 2.31$, $p = .021$
Dizziness frequency during sport	1.59	1.08	1.57	1.05	$t(8668) = 0.75$, $p = 0.448$
Last Dental visit	2.63	1.49	2.55	1.51	$t(8668) = 2.48$, $p = .013$
	N	%	N	%	
School bullying victim (D = yes)	451	10.88	355	7.84	χ^2 (1,8670) = 23.35, p <.001
Cyberbullying victim (D = yes)	412	9.94	329	7.27	χ^2 (1,8670) = 19.43, p <.001
Tried smoking (D = yes)	667	16.10	335	7.40	χ^2 (1,8670) = 159.12, p <.001

Measurements

Risk behaviors

Risk behaviors were assessed based on the 2017 State and Local Youth Risk Behavior Survey and included 22 factors (Shengfa et al., 2022). These behaviors encompassed a range of activities during school, such as school bullying, cyberbullying, smoking, insufficient physical activity, and more. School bullying, cyberbullying, and tried smoking ever were measured with dichotomous (*yes/no*) responses, while other items, such as seatbelt use frequency, alcohol days (lifetime), daily non-study screen time (weekdays), and dizziness frequency during sport, were measured on continuous scales (see Supplementary 1 for details).

Mental health issues

Adolescents' mental health issue was measured using the Symptom Checklist 90 (SCL-90) (Derogatis & Cleary, 1977). The SCL-90 was comprised of 90 items, quantifying psychopathology in terms of nine primary symptom constructs: somatization, obsessive-compulsive symptoms, interpersonal sensitivity, depression, anxiety, hostility, terror, paranoia, and psychosis. Response options range from 1 (*not at all*) to 5 (*extremely*).

Covariates

We included background variables that have been shown to correlate with adolecents' mental health as covariates, covering both individual and family conditions. The individual information included *residential areas*, *having art class*, *boarding status*, *school activity participation*, *sports game participation*, *academic financial support.* Family conditions covered *parental education*, family *economic status*, *parental conflicts*, *father drunk*, *parental academic expectation*, *having computer & internet*, *having siblings*, *having study desk*, *book number* at home.

Data processing

The Extreme Gradient Boosting (XGBoost) is an advanced machine learning algorithm that combines several learning applications to produce higher prediction accuracy than any of the individual learning applications (Bentéjac et al., 2021). It has gained prominence in the performance and efficiency in handling large datasets and its robustness against overfitting. We divided the dataset into training (80%) and testing (20%) subsets to ensure that the XGBoost model generalizes effectively. For hyperparameter tuning, we employed random search combined with 5-fold cross-validation. We used root mean-squared error (RMSE) and mean absolute error (MAE) to assess the predictive performance of the XGBoost model.

In addition, Shapley Additive ExPlanations values (SHAP) were computed to interpret the model (Ekanayake et al., 2022), providing a detailed understanding of how each risk behavior predicted mental health. To explore the gender differences, we conducted the model for boys and girls separately.

Results

Descriptive statistics

As shown in Table 1, significant gender differences were observed in most risk behaviors, with boys generally engaging in a higher frequency of such behaviors compared to girls. However, girls demonstrated more severe mental health issues than boys.

Table 1

Description for study variable

	Boys (n = 4144)		Girls (n = 4526)		
Characteristics	*M*	*SD*	*M*	*SD*	Gender differences
Seat belt use frequency	3.68	1.44	3.77	1.36	$t(8668) = -3.01, p < .01$
Rides with drunk driver	1.23	0.74	1.14	0.58	$t(8668) = 6.38, p < .001$
Age first smoking	1.76	1.64	1.32	1.13	$t(8668) = 14.19, p < .001$
Smoking days	1.34	1.12	1.15	0.75	$t(8668) = 9.16, p < .001$
Daily cigarettes	1.30	0.98	1.15	0.72	$t(8668) = 8.18, p < .001$
Alcohol days (lifetime)	1.97	1.83	1.50	1.39	$t(8668) = 13.38, p < .001$
Age first drink	2.31	2.10	1.72	1.69	$t(8668) = 14.43, p < .001$
Alcohol days (last 30 days)	1.59	1.47	1.28	1.06	$t(8668) = 11.01, p < .001$
Exercise days	3.46	2.30	2.98	2.14	$t(8668) = 10.20, p < .001$

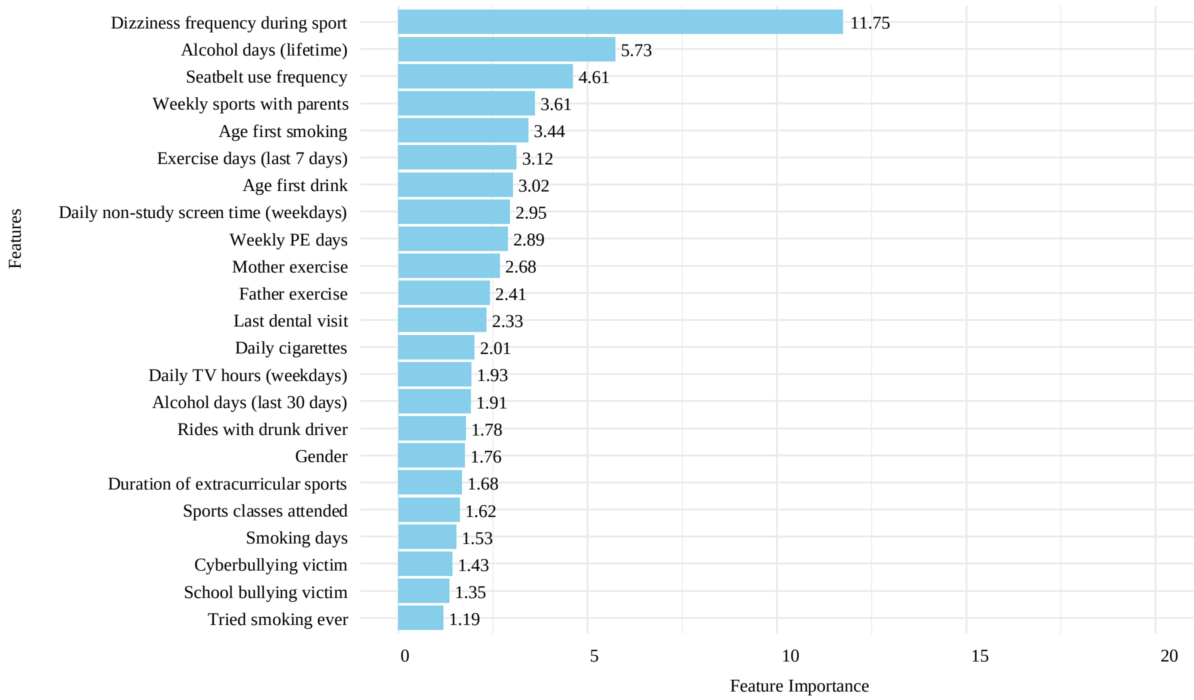
Daily TV hours (weekdays)	1.69	1.40	1.47	1.14	$t(8668) = 8.08, p < .001$
Daily non-study screen time (weekdays)	2.18	1.65	1.84	1.35	$t(8668) = 10.48, p < .001$
Weekly PE days	2.77	1.21	2.70	1.12	$t(8668) = 2.95, p < .001$
Sports classes attended (12 months)	1.52	0.81	1.39	0.71	$t(8668) = 7.64, p < .001$
Duration of extracurricular sports	1.59	1.14	1.41	0.99	$t(8668) = 7.82, p < .001$
Father exercise	3.28	1.51	3.41	1.49	$t(8668) = -4.03, p < .001$
Mother exercise	3.23	1.49	3.35	1.46	$t(8668) = -3.73, p < .001$
Family exercise	2.30	1.38	2.23	1.32	$t(8668) = 2.31, p = .021$
Dizziness frequency during sport	1.59	1.08	1.57	1.05	$t(8668) = 0.75, p = 0.448$
Last Dental visit	2.63	1.49	2.55	1.51	$t(8668) = 2.48, p = .013$
Mental health issues	1.51	0.60	1.58	0.59	$t(8668) = -5.31, p <.001$
	N	%	N	%	
School bullying victim (D = yes)	451	10.88	355	7.84	χ^2 (1,8670) = 23.35, p <.001
Cyberbullying victim (D = yes)	412	9.94	329	7.27	χ^2 (1,8670) = 19.43, p <.001
Tried smoking (D = yes)	667	16.10	335	7.40	χ^2 (1,8670) = 159.12, p <.001

Feature importance and interpretation

The full sample model produced an RMSE of 0.49, and an MAE of 0.35. As shown in Figure 1, dizziness frequency during sports emerged as the strongest predictor of mental health issues in adolescents. Additionally, seat belt use and alcohol consumption days were significant predictors, with SHAP gains exceeding 4. Other notable factors included weekly sports participation with parents, exercise frequency in the past seven days, and the age of first exposure to smoking and drinking, all of which demonstrated gains exceeding 3. The model, including covariates, further confirmed the significant roles of these factors (see Supplementary 1).

Figure 1

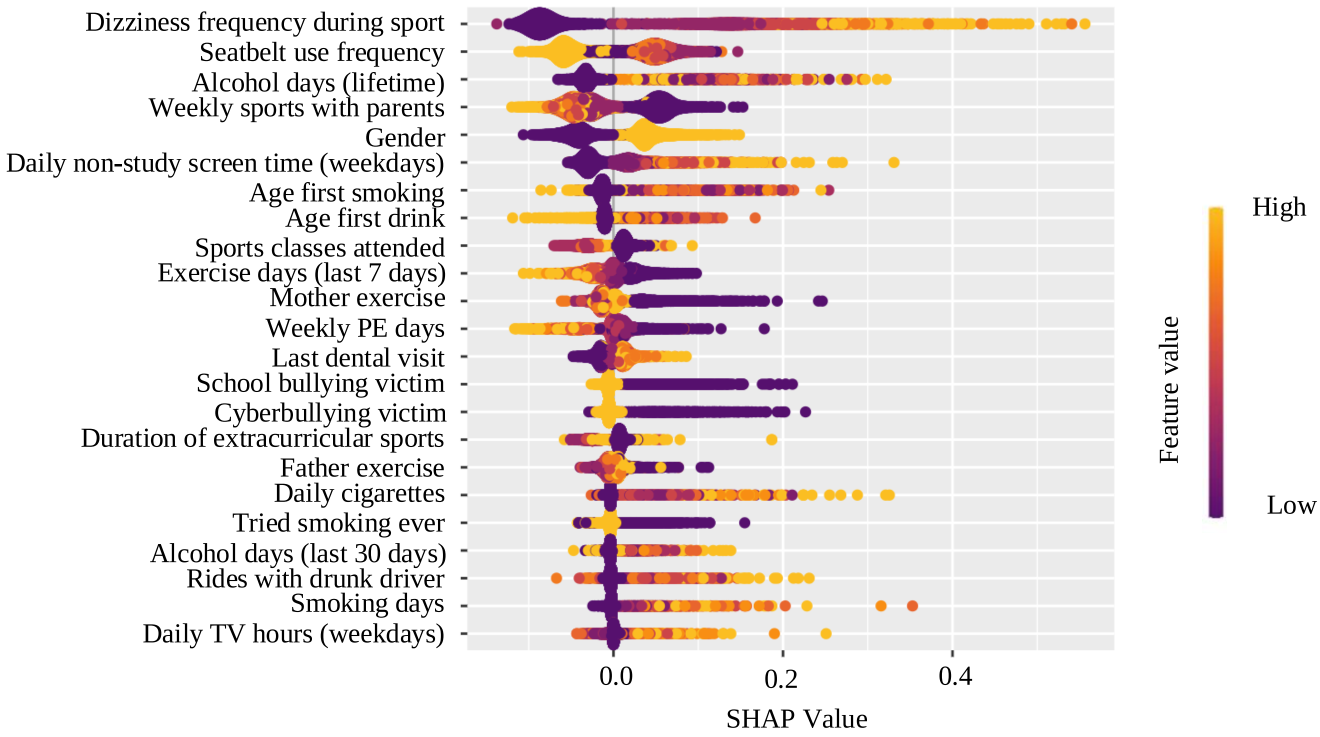
Importance of risk behaviors for adolescents' mental health issues



Specifically, more frequent dizziness during sports is associated with more severe mental health problems, as indicated in Figure 2. Increased seat belt use is linked to fewer mental health issues, whereas more alcohol consumption days is associated with greater mental health challenges. Additionally, a lack of sports participation, fewer sports activities with parents, and an earlier age of first exposure to smoking and drinking are associated with higher levels of mental health problems.

Figure 2.

SHAP values of risk behaviors for adolescents' mental health



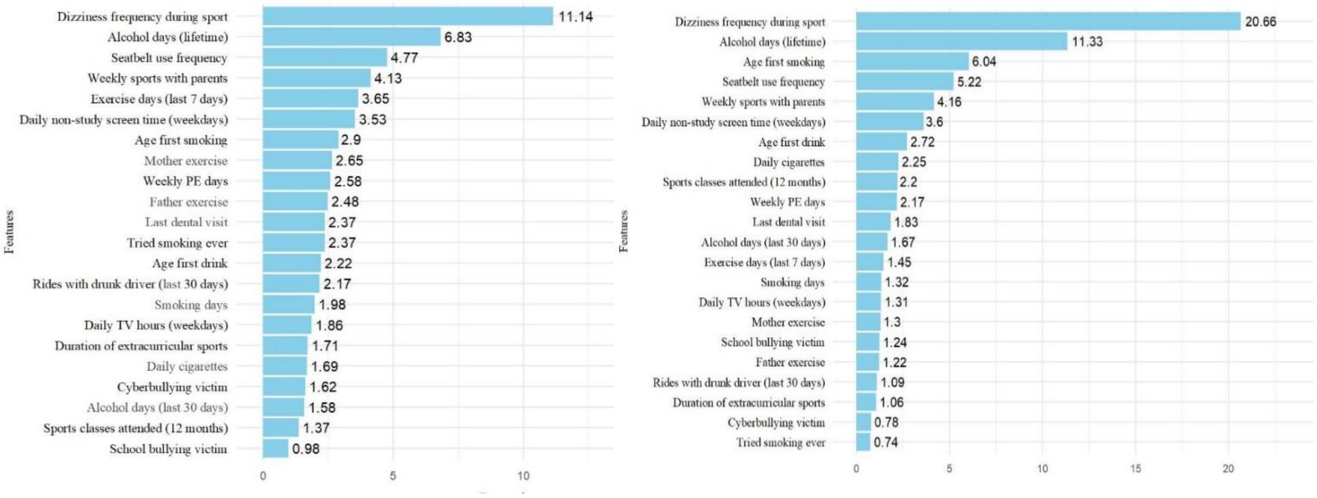
Note: This figure illustrates the SHAP values of the predictors, showing yellow dots for higher predictor values and purple dots for lower ones. The left side of the coordinate axis origin represents a negative impact, while the right side indicates a positive impact. In this research, the target variable is the adolescents' mental health issues, where higher scores reflect poorer mental health.

Gender difference

The gender-based analysis, as shown in Figure 3, indicated that risk behaviors play different roles in boys' and girls' mental health. While dizziness during sports and alcohol consumption days were the two strongest predictors for both genders, the association was more pronounced for girls. Furthermore, the age of first smoking was a stronger predictor for girls, whereas exercise frequency in the past seven days was a stronger predictor on boys' mental health outcomes.

Figure 3

Importance of risk behaviors for adolescents' mental health issues across gender



Discussion

The current study identified frequent dizziness during sports, lack of seat belt use, and alcohol consumption as significant predictors of adolescents' mental health. Moreover, gender differences emerged in the predictive patterns: an earlier age of first smoking was more strongly associated with mental health outcomes among girls, while exercise frequency was a more salient predictor for boys.

The present study identified several key behavioral and physiological predictors of adolescents' mental health. Notably, frequent dizziness during sports, lack of seat belt use, and alcohol consumption emerged as important risk factors. These findings align with previous literature emphasizing the role of physical well-being and risk-taking behaviors in shaping psychological outcomes during adolescence (Brooks et al., 2002; Patton et al., 2016). Dizziness during physical activity may reflect underlying health issues or poor physical fitness, which could exacerbate stress and emotional dysregulation. Similarly, the absence of seat belt use and alcohol consumption may indicate broader patterns of risky behaviors, which are often associated with impulsivity and lower self-regulation—factors linked to mental health problems (Steinberg, 2008). These findings underscore the importance of integrating physical health monitoring and behavioral risk assessments into school-based mental health screenings. Targeted interventions promoting bodily awareness, safety behaviors, and substance use prevention may serve as effective strategies for improving adolescents' psychological well-being.

Gender differences also surfaced in the predictive patterns. For girls, the age at first smoking was more strongly associated with later mental health problems, which may reflect greater vulnerability to early substance use due to biological or psychosocial factors. Girls who begin smoking at an earlier age may be more exposed to peer pressure, internalizing symptoms, or maladaptive coping strategies (Thompson et al., 2015), which could compound their mental health challenges over time. In contrast, for boys, exercise frequency was a stronger predictor of mental health outcomes. This suggests that maintaining regular physical activity may serve as a protective factor against emotional or behavioral problems in male adolescents, consistent with prior research that the benefits of exercise for mood regulation and stress relief was stronger for boys (Halliday et al., 2019; Hands & Parker, 2016). These gender-specific findings highlight the need for differentiated mental health strategies that account for adolescents' gendered patterns of risk and protection. Preventive programs may benefit from prioritizing early smoking prevention among girls and promoting consistent physical activity among boys to enhance mental health outcomes.

Advantages and Limitations

This study has several notable advantages. First, this study utilized large-sample data from a representative survey in China, offering more statistical power than small-sample or regional data. Second, we included 22 risk behaviors such as school bullying, cyberbullying, smoking, insufficient physical activity, providing a comprehensive understanding of the importance of each factor. Third, we developed a ML model to examine the associations between risk behaviors and mental health. This model outperformed conventional statistical methods, such high dimensional data dealing, thereby deepening our understanding of the adverse effects of pregnancy and newborn risk factors.

However, there are some limitations to this study. First, due to limitations in data availability, some risk behaviors, such as, suicidality, delinquency or violence (Botsis, 2003), which may influence students' mental health, were not included in this study. Future study should incorporate a wider range of adolescents' risk behaviors, enhancing the predictive accuracy of future models. Second, the primary focus of the study is to evaluate the importance of risk behaviors and their associations with students' mental health, rather than to establish causal inferences. Future research could focus on establishing causal relationships between risk behaviors and mental health, which will help to confirm and extend the associations identified in this study.

Conclusion

This study contributes to the growing body of literature on adolescent mental health by identifying key behavioral and physiological predictors, including frequent dizziness during sports, lack of seat belt use, and alcohol consumption. These factors may reflect underlying health vulnerabilities and broader patterns of risk-taking behaviors that compromise psychological well-being. Importantly, gender differences emerged, suggesting that early smoking initiation poses a greater risk for girls, while regular physical activity serves as a protective factor for boys. These findings underscore the importance of incorporating gender-sensitive approaches into mental health prevention and intervention strategies. By targeting modifiable risk behaviors and promoting protective practices, particularly through school- and community-based programs, stakeholders can better support the psychological resilience and healthy development of adolescents.

Acknowledgment

The data used in this study were obtained from the Chinese Population Health Data Archive (PHDA) of the National Population and Health Science Data Center, specifically the "Adolescent Health Thematic Database" provided by Shandong University (Shandong University, 2021, https://doi.org/10.12213/11.A0031.202107.209.V1.0). We sincerely thank the data providers for their invaluable support and contribution to scientific research.

Funding

This study was supported by the European Commission under the Jean Monnet Module [Project No.101238708].

References

Abdolali, M., & Gillis, N. (2021). Beyond linear subspace clustering: A comparative study of nonlinear manifold clustering algorithms. *Computer Science Review*, *42*, 100435. https://doi.org/10.1016/j.cosrev.2021.100435

Aghaabbasi, M., & Chalermpong, S. (2023). Machine learning techniques for evaluating the nonlinear link between built-environment characteristics and travel behaviors: A systematic review. *Travel Behaviour and Society*, *33*, 100640. https://doi.org/10.1016/j.tbs.2023.100640

Bentéjac, C., Csörgő, A., & Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review*, *54*(3), 1937–1967. https://doi.org/10.1007/s10462-020-09896-5

Botsis, A. (2003). High risk behaviours in young adults: is there a common substrate? *Annals of General Hospital Psychiatry*, *2*(S1), S6. https://doi.org/10.1186/1475-2832-2-S1-S6

Brooks, T. L., Harris, S. K., Thrall, J. S., & Woods, E. R. (2002). Association of adolescent risk behaviors with mental health symptoms in high school students. *Journal of Adolescent Health*, *31*(3), 240–246. https://doi.org/10.1016/S1054-139X(02)00385-3

Deng, S., Wang, F., Cai, Y., Wang, H., Wang, Z., Qian, Q., & Ding, W. (2024). *Prediction and Analysis of Multiple Causes of Mental Health Problems Based on Machine Learning* (pp. 150–160). https://doi.org/10.1007/978-3-031-57867-0_11

Derogatis, L. R., & Cleary, P. A. (1977). Confirmation of the dimensional structure of the scl-90: A study in construct validation. *Journal of Clinical Psychology*, *33*(4), 981–989. https://doi.org/10.1002/1097-4679(197710)33:4<981::AID-JCLP2270330412>3.0.CO;2-0

Ekanayake, I. U., Meddage, D. P. P., & Rathnayake, U. (2022). A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using Shapley additive explanations (SHAP). *Case Studies in Construction Materials*, *16*, e01059. https://doi.org/10.1016/j.cscm.2022.e01059

Halliday, A. J., Kern, M. L., & Turnbull, D. A. (2019). Can physical activity help explain the gender gap in adolescent mental health? A cross-sectional exploration. *Mental Health and Physical Activity*, *16*, 8–18. https://doi.org/10.1016/j.mhpa.2019.02.003

Hands, B., & Parker, H. (2016). Male and Female Differences in Health Benefits Derived from Physical Activity: Implications for Exercise Prescription. *Journal of Womens Health, Issues and Care*, *5*(4). https://doi.org/10.4172/2325-9795.1000238

Jones, A. C., Schinka, K. C., van Dulmen, M. H. M., Bossarte, R. M., & Swahn, M. H. (2011). Changes in Loneliness during Middle Childhood Predict Risk for Adolescent Suicidality Indirectly through Mental Health Problems. *Journal of Clinical Child & Adolescent Psychology*, *40*(6), 818–824. https://doi.org/10.1080/15374416.2011.614585

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime Prevalence and Age-of-Onset Distributions of DSM-IV Disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *62*(6), 593. https://doi.org/10.1001/archpsyc.62.6.593

Miranda-Mendizabal, A., Castellví, P., Parés-Badell, O., Alayo, I., Almenara, J., Alonso, I., Blasco, M. J., Cebrià, A., Gabilondo, A., Gili, M., Lagares, C., Piqueras, J. A., Rodríguez-Jiménez, T., Rodríguez-Marín, J., Roca, M., Soto-Sanz, V., Vilagut, G., & Alonso, J. (2019). Gender differences in suicidal behavior in adolescents and young adults: systematic review and meta-analysis of longitudinal studies. *International Journal of Public Health*, *64*(2), 265–283. https://doi.org/10.1007/s00038-018-1196-1

Patton, G. C., Sawyer, S. M., Santelli, J. S., Ross, D. A., Afifi, R., Allen, N. B., Arora, M., Azzopardi, P., Baldwin, W., Bonell, C., Kakuma, R., Kennedy, E., Mahon, J., McGovern, T., Mokdad, A. H., Patel, V., Petroni, S., Reavley, N., Taiwo, K., … Viner, R. M. (2016). Our future: a Lancet commission on adolescent health and wellbeing. *The Lancet*, *387*(10036), 2423–2478. https://doi.org/10.1016/S0140-6736(16)00579-1

Rosenfield, S., & Mouzon, D. (2013). *Gender and Mental Health* (pp. 277–296). https://doi.org/10.1007/978-94-007-4276-5_14

Shengfa, Z., Wei, L., Xiaosheng, D., Wenxin, C., Xiangren, Y., Wei, Z., Yue, Z., & Yuanzhi, Z. (2022). A Dataset on the Status Quo of Health and Health-Related Behaviors of Chinese Youth: A Longitudinal Large-Scale Survey in the Secondary School Students of Shandong Province. *Chinese Medical Sciences Journal*, *37*(1), 60. https://doi.org/10.24920/004051

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review*, *28*(1), 78–106. https://doi.org/10.1016/j.dr.2007.08.002

Sullivan, C. J., Childs, K. K., & O'Connell, D. (2010). Adolescent Risk Behavior Subgroups: An Empirical Assessment. *Journal of Youth and Adolescence*, *39*(5), 541–562. https://doi.org/10.1007/s10964-009-9445-5

Thompson, A. B., Tebes, J. K., & McKee, S. A. (2015). Gender differences in age of smoking initiation and its association with health. *Addiction Research & Theory*, *23*(5), 413–420. https://doi.org/10.3109/16066359.2015.1022159

Supplementary materials

Supplementary Table 1

Measurement for risk behaviours

No.	Variable Name	Scale	Description
1	Seatbelt use frequency	1-5: 1. Never 2. Rarely 3. Sometimes 4. Most of the time 5. Always	How often do you wear a seatbelt when riding in someone else's car?
2	Rides with drunk driver	1-5: 1. 0 times 2. 1 time 3. 2–3 times 4. 4–5 times 5. 6 or more times	In the past 30 days, how many times have you ridden in a car driven by someone under the influence of alcohol?
3	School bullying victim	1-2: 1. Yes 2. No	In the past 12 months, have you experienced bullying at school?
4	Cyberbullying victim	1-2: 1. Yes 2. No	In the past 12 months, have you experienced cyberbullying (e.g., through text messages, QQ, WeChat, Facebook, or other social media platforms)?
5	Tried smoking ever	1-2: 1. Yes 2. No	Have you ever tried smoking, even just one or two puffs?

6	Age first smoking	1-7: 1. Never 2. ≤8 years 3. 9–10 years 4. 11–12 years 5. 13–14 years 6. 15–16 years 7. ≥17 years	How old were you when you first tried smoking, even just one or two puffs?
7	Smoking days	1-7: 1. 0 days 2. 1–2 days 3. 3–5 days 4. 6–9 days 5. 10–19 days 6. 20–29 days 7. 30 days	In the past 30 days, on how many days did you smoke?
8	Daily cigarettes	1-7: 1. None 2. <1 cigarette/day 3. 1 cigarette/day 4. 2–5/day 5. 6–10/day 6. 11–12/day 7. >20/day	In the past 30 days, on days when you smoked, how many cigarettes did you smoke per day?
9	Alcohol days (lifetime)	1-7: 1. 0 days 2. 1–2 days 3. 3–9 days 4. 10–19 days	In your lifetime, on how many days have you had at least one drink of alcohol?

		5. 20–39 days 6. 40–99 days 7. ≥100 days	
10	Age first drink	1-7: 1. Never 2. ≤8 years 3. 9–10 years 4. 11–12 years 5. 13–14 years 6. 15–16 years 7. ≥17 years	How old were you when you first drank a whole drink of alcohol, not just a sip?
11	Alcohol days (last 30 days)	1-7: 1. 0 days 2. 1–2 days 3. 3–5 days 4. 6–9 days 5. 10–19 days 6. 20–29 days 7. 30 days	In the past 30 days, on how many days did you have at least one drink of alcohol?
12	Exercise days (last 7 days)	1-8: 1. 0 days 2. 1 day 3. 2 days 4. 3 days 5. 4 days 6. 5 days 7. 6 days 8. 7 days	In the past 7 days, on how many days did you exercise for at least 30 minutes?

13	Daily TV hours (weekdays)	1-7: 1. None 2. <1 hour/day 3. 1 hour/day 4. 2 hours/day 5. 3 hours/day 6. 4 hours/day 7. ≥5 hours/day	During this school term, how many hours per day on average do you watch TV from Monday to Friday?
14	Daily non-study screen time (weekdays)	1-7: 1. None 2. <1 hour/day 3. 1 hour/day 4. 2 hours/day 5. 3 hours/day 6. 4 hours/day 7. ≥5 hours/day	During this school term, how many hours per day do you spend playing video games or using a computer for non-school activities?
15	Weekly PE days	1-6: 1. 0 days 2. 1 day 3. 2 days 4. 3 days 5. 4 days 6. 5 days	During school, how many days per week do you attend physical education classes?
16	Sports classes attended	1-4: 1. 0 2. 1 3. 2 4. 3 or more	In the past 12 months, how many extracurricular sports activities have you participated in?

17	Duration of extracurricular sports	1-5: 1. ≤3 months 2. 3–6 months 3. 7–12 months 4. 13–24 months 5. ≥25 months	How long did your extracurricular sports activities last?
18	Father exercise	1-5: 1. 0 times 2. 1 time 3. 2 times 4. 3 times 5. ≥4 times	How many times per week does your father exercise?
19	Mother exercise	1-5: 1. 0 times 2. 1 time 3. 2 times 4. 3 times 5. ≥4 times	How many times per week does your mother exercise?
20	Weekly sports with parents	1-5: 1. 0 times 2. 1 time 3. 2 times 4. 3 times 5. ≥4 times	How many times per week do you exercise with your parents?
21	Dizziness frequency during sport	1-5: 1. 0 times 2. 1 time 3. 2 times 4. 3 times	In the past 12 months, how many times have you experienced dizziness while engaging in sports?

		5. ≥4 times	
22	Last dental visit	1-5: 1. In the past 12 months 2. 1–2 years ago 3. Over 2 years ago 4. Never 5. Unsure	When was the last time you visited a dentist for a check-up, cleaning, or other dental care?

Supplementary Table 2

Importance of risk behaviours for adolescents' mental health issues including covariates