Check for updates

OPEN ACCESS

EDITED BY Giuseppe Plazzi, University of Modena and Reggio Emilia, Italy

REVIEWED BY Valentina Baldini, University of Bologna, Italy Luca Altieri, Civil Hospital of Brescia, Italy

*CORRESPONDENCE Lufeng Chen Chenlufenga@163.com

[†]These authors have contributed equally to this work

RECEIVED 01 March 2025 ACCEPTED 02 April 2025 PUBLISHED 25 April 2025

CITATION

Yao L, Chen Q, Yang K, Zheng Z, Chen Z, Wang D, Xia Y, Chen D and Chen L (2025) Novel insight into prediction model for sleep quality among college students: a LASSOderived sleep evaluation. *Front. Psychiatry* 16:1585732. doi: 10.3389/fpsyt.2025.1585732

COPYRIGHT

© 2025 Yao, Chen, Yang, Zheng, Chen, Wang, Xia, Chen and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Novel insight into prediction model for sleep quality among college students: a LASSOderived sleep evaluation

Ling Yao^{1,2†}, Qingquan Chen^{3†}, Kang Yang^{4†}, Zhihua Zheng⁴, Zhihan Chen⁴, Danna Wang⁴, Yining Xia⁴, Dingquan Chen⁴ and Lufeng Chen^{1*}

¹The Second Affiliated Hospital of Fujian Medical University, Quanzhou, Fujian, China, ²The Graduate School of Fujian Medical University, Fuzhou, Fujian, China, ³The School of Public Health, Fujian Medical University, Fuzhou, Fujian, China, ⁴Fujian Medical University, Fuzhou, Fujian, China

Background: Sleep disturbance has become a significant concern among college students, as it can lead to various mental and physical disorders. This study aims to provide a fresh perspective by developing and validating a predictive model for sleep quality among college students.

Methods: Data from 20,645 college students in Fujian Province, China, collected between 5th April and 16th April 2022, were analyzed. Participants completed the Pittsburgh Sleep Quality Index (PSQI) scale, a self-designed general data questionnaire, and a sleep quality influencing factor questionnaire. Multinomial logistic regression, LASSO regression, and Boruta feature selection methods were utilized to select relevant variables. The data were then divided into a training-testing set (70%) and an independent validation set (30%) using stratified sampling. Six machine learning techniques, including artificial neural network (ANN), decision tree, gradient-boosting tree, k-nearest neighbor, naïve Bayes, and random forest, were developed and validated. Finally, an online sleep evaluation website was established based on the best-fitting prediction model.

Results: The mean global PSQI score was 6.02 ± 3.112 , with a sleep disturbance prevalence of 28.9% (defined as a global PSQI score > 7). The LASSO regression model identified eight predictors: age, specialty, respiratory history, coffee consumption, staying up late, prolonged online activity, sudden changes, and impatient closed-loop management. Among the evaluated models, the ANN demonstrated superior performance with an area under the receiver operating characteristic curve (AUC) of 0.713 (95% CI: 0.696–0.730), accuracy of 0.669 (95% CI: 0.669–0.669), sensitivity of 0.682 (95% CI: 0.699–0.665), specificity of 0.637 (95% CI: 0.665–0.610). Decision curve analysis and clinical impact analysis further confirmed the model's clinical utility.

Conclusions: This study developed a prediction model for sleep disturbance among college students using a LASSO regression and ANN, incorporating eight predictors. The model can serve as an intuitive and practical tool for predicting sleep quality and supporting effective management and healthcare on college campuses.

KEYWORDS

sleep quality, college students, machine learning, LASSO regression, PSQI, ANN, prediction model

Introduction

Sleep quality among college students

Sleep disturbances represent a major health issue that encompasses a wide range of sleep complaints, such as difficulty initiating sleep (DIS), difficulty maintaining sleep (DMS) (1), earlymorning awakening (EMA) (2), non-restorative sleep (NRS), and poor sleep quality (3, 4). In addition, poor sleep quality is associated with certain medical conditions (e.g., fibromyalgia (5), arthritis/ rheumatism (6, 7), heart disease, and cancer (8, 9).

The characteristics of sleep among college students are different from the general public. College students report an average of 7-7.5 hours of sleep per night, which is 1-1.5 hours fewer than their selfreported ideal of 8.5 hours per night, according to studies that suggest they suffer from chronic sleep deprivation (10). The most common sleep disorders seen in college students are inadequate sleep hygiene (ISH), delayed sleep phase disorder (DSPD), and insomnia (11). Often, university students who are transitioning from adolescence to adulthood experience numerous challenges, such as having to adapt to new social situations, leaving home, and coping with high academic and social pressures and erratic living schedules, all of which could increase the risk of sleep disturbances (12). A meta analysis calculated 14 studies (n=22,297) and the pooled sleep disturbance prevalence of college students is 33% (95% CI: 22-44%) (13). The delivery of behavioral sleep medicine is particularly relevant for the college student population, as the early intervention on their sleep problems might prevent lifelong consequences.

Previous studies

Previous studies found varied rates of sleep disturbances (2) in this population, and a large proportion of them focused on specific variables that related to later adverse outcomes. For instance, chronic diseases, stressful issues (e.g., anxiety, lockdown, graduation) (14), sleep attitudes (15), lifestyles (e.g., exercise, midday rest) (16, 17), electronic device use (18), diet, and consumption of alcohol, cigarettes, drugs, and coffee were found to relate to sleep disturbances (19).

Furthermore, only a few studies have used these variables to predict sleep quality among college students. Nonetheless, despite studies that predicted sleep quality for medical staff (20), elderly patients (21), infants (22), adolescents (23), and children (24), a few models set college students as target patients, which provided a completely new insight on sleep quality predictions for campus students.

In a previous study conducted on the same population (25), 11 variables related to sleep quality were used as parameters for prediction models. Among these variables, residence had a significant impact on the sleep quality of college students. However, it was found that the place of residence could pose a challenge for migrating the model to other locations, affecting the simplicity and generalizability of the model.

LASSO algorithm and application

The ordinary least squares method is often used by researchers to explore the relationships between variables in previous studies. However, when many predictors are included, it leads to modeldata overfit and multi-collinearity.

The LASSO algorithm was first proposed by Robert Tibshirani in 1997 and is known as the least absolute shrinkage and selection operator (26). It can obtain a more refined model by constructing a penalty function that aids in compressing some coefficients while setting other coefficients to zero. Thus, it retains the advantage of subset shrinkage and is a biased estimator when dealing with data with complex covariance. Research has shown that the LASSO method is capable of constructing a more compressed model that offers greater prediction accuracy compared to other existing methods (27, 28).

The LASSO method has been utilized to provide clinical information for performing early identification of S-COVID-19-P on admissions in fever clinics with a 100% recall score, and its model has been deployed as an online triage tool (29). In addition, it performed well in selecting features to improve the mortality predictions of hospitalized patients with COVID-19 with electronic health records (30) and predict the outcomes of SARS-CoV-2 pneumonia patients based on laboratory findings (31).

In comparison to a previous study (26), which utilized a multivariate unconditional LR analysis to determine predictors, LASSO is a tool that selects fewer parameters and guarantees greater prediction accuracy, making it advantageous in applications. The LASSO method can be used as a suitable alternative machine learning technique for exploring the key predictors that affect sleep quality among college students.

Aims

The aim of this study was to investigate the most common potential risk factors associated with poor sleep quality and further develop and validate a LASSO-derived prediction model to measure the risks of poor sleep quality among university students.

This study hypothesized that the significant variables associated with poor sleep quality could be identified and used to create an easy-to-operate website that would accurately and individually evaluate the probability of suffering from poor sleep, especially among university students.

We hope that this website will be an intuitive and practical tool for sleep quality predictions that will support early prevention in colleges, enhance more personalized and precise medical aids in hospitals, and assist in allocating appropriate health resources for governments and societies.

Methods

Aim and design

The objective of this study was to explore prevalent risk factors linked to impaired sleep quality among university students and subsequently develop and validate a LASSO-derived prediction model to assess the risk of poor sleep quality. Figure 1. shows the study design and model workflow of this study. And we ensured the security of the data.

Setting and participants

An internet-based cross-sectional sleep quality survey was conducted for 33 universities in Fujian Province. We collected data from 23,572 full-time undergraduate or graduate students (with an age range of 17–35 years) in Fujian Province who underwent an internet-based cross-sectional sleep quality survey between April 5 and 16, 2022. Full-time undergraduate and graduate students who delayed enrollment due to the epidemic, lived outside of a student residence, or had significant sleep or mental disorders were excluded from the study.

Scales and questionnaires

The Pittsburgh Sleep Quality Index (PSQI) is a self-rated questionnaire which assesses sleep quality and disturbances over a 1-month time interval, It is one of the most extensively used and useful tools for assessing sleep disorders. Its clinical and clinemetric qualities point to potential use in mental clinical practice and research of college students (32). Higher sleep scores on the PSQI scale equate to poorer sleep quality.

A self-designed general data questionnaire and a sleep quality influencing factor questionnaire include age, gender, residence,



specialty (medical-related majors, science and engineering, or liberal arts), grade (graduating or non-Graduating), Body Mass Index (BMI), with or without respiratory history, the frequency of coffee consumption, staying up late, spending long hours online, suffering sudden changes, fearing infection of COVID-19, feeling impatient with closed-loop management (Supplementary Text S1).

Definition of the results

In this study, a PSQI score of 7 was used as the cut-off point for sample grouping, with poor sleep being defined as a PSQI of > 7 and good sleep quality being defined as a PSQI of \leq 7.

Statistical analysis

The qualitative data were expressed as numbers and percentages and compared using a chi-square test or a Fisher's exact test. A pvalue of < 0.05 in the univariate analysis and a p value of < 0.01 in the multivariate analysis were considered statistically significant. The questionnaires included in the study contained no missing data.

Predictor selection

We used three algorithms to select predictors in the dataset. First, predictors with a *p* value of < 0.10 from the univariate analysis were entered into a multivariate logistic regression. Second, a LASSO algorithm was used for a 10-fold cross-validation to select potential predictors with non-zero coefficients. Third, the Boruta feature selection was used to identify key categorical variables. The performance of these methods was assessed against the following: Nagelkerke R^2 (larger values are better), root-mean-square error (RMSE; lower values are better). Finally, the optimal predictor selection algorithm was determined based on Occam's razor.

Prediction model development and validation

Six prediction models were built using artificial neural network (ANN), decision tree (DT), gradient-boosting tree (GBT), k-nearest neighbor (K-nn), naïve Bayes (NB), and random forest (RF). The incorporated data were divided into a training-testing set (70%) and an independent validation set (30%) using stratified sampling. To avoid overfitting and promote the models, we used a 10-fold cross-validation for the training-testing set and referenced the best models to the independent validation set. We evaluated the model's performance by calculating the area under the receiver operating curve (AUROC) for the six models in the independent validation set, and, in addition, we calculated the accuracy, sensitivity,

specificity, precision, F1-score, and KAPPA to further evaluate the model's performance. In this study, a calibration curve analysis was performed to assess the agreement by the slope of the calibration curve (an ideal value of 1), intercept, and Brier score (an ideal value of 0; a value of >0.3 indicates poor calibration).

A decision curve analysis was performed by quantifying the net clinical benefit at different threshold probabilities, and a clinical impact curve analysis was performed by quantifying the costbenefit ratio at different threshold probabilities to determine the clinical usefulness of the prediction model.

All machine learning models were developed and validated using R, version 4.2.1.

Clinical applications

Prediction models have traditionally been assessed using sensitivity and specificity statistics, but these results are silent on if using the model in clinical practice would be advantageous or disadvantageous.

In this study, we utilized calibration curve analysis, decision curve analysis, clinical impact curve analysis, and net clinical benefit to compare the clinical practice performance of six models.

The calibration curve assesses the agreement between predicted probabilities and actual observations. The baseline is typically an ideal 45-degree diagonal line, representing perfect calibration where predicted probabilities equal observed probabilities. The decision curve analysis curve evaluates the clinical utility of a prediction model across different probability thresholds. The baseline represents the net benefit without the model, avoiding any benefit or harm from predictions. The clinical impact curve assesses the impact of a model's predictions on patient management across different thresholds. Net clinical benefit is useful for determining whether basing clinical decisions on a model would do more good than harm. This is in contrast to traditional measures such as sensitivity, specificity, or area under the curve, which are statistical abstractions not directly informative about clinical value. Estimating net clinical benefit makes possible to clarify the basis for therapeutic decisions on an individual and collective level.

Net clinical benefit is defined as (33):

Net benefit =
$$\frac{\text{True positives}}{n} - \frac{\text{False positives}}{n} \left(\frac{P}{1-p}\right)$$

n is the total number of patients in the study and p is the threshold probability.

In addition, we also developed an easy-to-operate website to put model into practice. In reality, many college students refrain from visiting the hospital for a sleep quality assessment due to the inconvenience of attending an appointment. Thus, we provided a simple application for predicting the sleep quality of college students. College students who consider themselves in need of a simple screening for poor sleep quality can access the online website we created and enter the appropriate predictors into the website, which will generate predictions in real time to assist participants in making medical decisions.

Results

Participant characteristics

From the included participants, the mean global PSQI score was 6.02 ± 3.112 and 14,673 had good sleep quality (71.1%) and 5,972 had poor sleep quality (28.9%). Table 1 demonstrates a statistical analysis of the effect of different factors on sleep quality among college students. The distribution of the participants in this study covered the entire province of Fujian in China, as illustrated in Figure 2.

Prediction feature selection

Univariate and multivariate ordered logistic regressions were used to assess the variables associated with sleep quality among college students (Table 1). The multivariate analysis identified 11 candidate predictors.

Figure 3A shows the results for the 13 variables included in the LASSO algorithm. When the λ value was increased to 0.016 (one standard error of the minimum value of λ), only eight candidate predictors were retained in the model, which were presumably the most influential predictors of sleep quality among college students (Figure 3B).

The Boruta feature selection was used to identify key categorical variables, i.e., to statistically compare the importance of the feature variables that were actually present in the data with those that were randomly added. Finally, 12 candidate predictors were identified as important variables (Figure 4).

LASSO can compress the coefficients of some unimportant features to zero, thereby selecting the most relevant features. Boruta is a feature selection method based on random forests, which generates a large number of candidate features and may retain some redundant or unimportant features, leading to a higher model complexity. In our study, LASSO selected 8 of the most important predictor variables through 10-fold cross-validation, and these variables demonstrated high predictive accuracy (AUROC of 0.713) in subsequent models such as artificial neural networks (ANN). Although Boruta also identified 12 important features, the features selected by LASSO performed better in terms of model prediction performance. Compared with Boruta, the LASSO regularization algorithm was identified as the optimal predictor selection algorithm based on Occam's razor.

LASSO introduces sparsity through L1 regularization and performs automatic feature selection. Compared with other regularization methods, unlike Ridge regression (L2 regularization), which shrinks all coefficients but does not exclude any features, LASSO tends to produce sparse solutions by completely eliminating irrelevant features. While Elastic Net combines L1 and L2 regularization, it is computationally more complex in highdimensional settings. LASSO strikes an optimal balance between computational efficiency and sparsity, making it a preferred choice for this study.

The final predictors included in the prediction model were as follows: age, specialty, respiratory history, coffee consumption, staying up late, long hours online, sudden changes, and impatient closed-loop management. Table 2 represents the performance of three feature selection methods. The OR and 95% CI values of the included predictors are shown in Figure 5.

Development and validation of a sleep quality prediction model for college students

Finally, the eight predictors were integrated into the sleep quality risk prediction model for college students (Figure 6A). In the training-testing set, the AUROC values of ANN, DT, GBT, Knn, NB, and RF were 0.700 (95% CI: 0.691-0.708), 0.634 (95% CI: 0.624-0.643), 0.688 (95% CI: 0.679-0.696), 0.602 (95% CI: 0.593-0.611), 0.692 (95% CI: 0.684-0.701), and 0.694 (95% CI: 0.686-0.703), respectively (Figure 6B). In the independent validation set, the AUROC values of ANN, DT, GBT, K-nn, NB, and RF were 0.713 (95% CI: 0.696-0.730), 0.627 (95% CI: 0.610-0.644), 0.697 (95% CI: 0.679-0.714), 0.593 (95% CI: 0.575-0.612), 0.706 (95% CI: 0.689-0.723), and 0.706 (95% CI: 0.688-0.723), respectively (Figure 6C). Details on the model's performance are shown in Table 3. We plotted the predicted model and ideal calibration curves (Figure 7) and further evaluated the agreement in terms of calibration slope (an ideal value of 1) and Brier score (an ideal value of 0; a value >0.3 indicates poor calibration). Good calibration was observed for all six machine learning models (Figures 7A-F), with Brier scores of 0.182, 0.195, 0.193, 0.229, 0.208, and 0.185, respectively. However, the respective calibration slopes deviated slightly as follows: 1.083, 0.896, 2.769, 0.376, 0.259, and 1.062. Details are shown in Table 4.

In order to determine the clinical usefulness of the models, a decision curve analysis and a clinical impact curve analysis were performed on the prediction models. The clinical decision curves (Figure 8) showed that when the clinical decisions were performed using the ANN, DT, GBT, K-nn, NB, and RF prediction models, the threshold probabilities of achieving a greater net benefit than the "no treatment" or "all treatment" scenarios were 0.89, 0.89, 0.88, 0.81, 0.82, and 0.88, respectively.

A clinical impact curve analysis (Figure 9) showed the clinical effectiveness of the six predictive models. The ANN, DT, GBT, K-nn, NB, and RF models were judged to be a high match between those with poor sleep quality and those with actual poor sleep quality when the threshold probabilities were greater than 75%, 70%, 75%, 65%, 70%, and 75%, respectively, confirming the high clinical efficiency of the prediction model.

Variables	n=20645	Sleep Quality		Univariate analysis		Multivariate analysis	p
		Good(≤7)	Poor(>7)	χ^2	р	OR [95% CI]	
Age							
<20	9227	6819 (46.5%)	2408 (40.3%)			1.00 [Reference]	< 0.001
≥20	11418	7854 (53.5%)	3564 (59.7%)	- 64.732	<0.001	1.31 [1.23, 1.40]	
Gender		1		1		1	
Male	6326	4747 (32.4%)	1579 (26.4%)			1.00 [Reference]	<0.001
Female	14319	9926 (67.6%)	4393 (73.6%)	69.801	<0.001	1.14 [1.06, 1.23]	
Residence		I		1			
Quanzhou	4959	3502 (23.9%)	1457 (24.4%)			1.00 [Reference]	
Fuzhou	3782	2519 (17.2%)	1263 (21.1%)	-		0.92 [0.83, 1.02]	0.109
Longyan	7066	5293 (36.1%)	1773 (29.7%)	-		0.71 [0.64, 0.78]	< 0.001
Nanping	897	655 (4.5%)	242 (4.1%)	-		0.91 [0.77, 1.09]	0.313
Ningde	540	392 (2.7%)	148(2.5%)	114.574	< 0.001	0.83 [0.67, 1.03]	0.091
Putian	190	146(1.0%)	44(0.7%)	-		0.78 [0.54, 1.11]	0.172
Sanming	2133	1451 (9.9%)	682 (11.4%)	-		0.93 [0.82, 1.06]	0.293
Xiamen	437	291 (2.0%)	146 (2.4%)	-		0.93 [0.74, 1.16]	0.524
Zhangzhou	641	424 (2.9%)	217 (3.6%)	_		1.57 [1.29, 1.90]	< 0.001
Specialty							
Medical-related majors	4088	3120 (21.3%)	968 (16.2%)			1.00 [Reference]	
Science and engineering	7876	5780 (39.4%)	2096 (35.1%)	163.534	<0.001	1.42 [1.27, 1.58]	< 0.001
Liberal arts	8681	5773 (39.3%)	2908 (48.7%)	_		1.62 [1.46, 1.80]	< 0.001
Grade							
Graduating class	1459	998 (6.8%)	461 (7.7%)	5 442	0.021	1.00 [Reference]	0.628
Non-graduating class	19186	13675 (92.3%)	5511 (92.3%)	- 5.443	0.021	1.03 [0.91, 1.17]	
BMI							
<18.5	4804	3381(23.0%)	1423(23.8%)				
[18.5,24)	12513	8960(61.1%)	3553(59.5%)	5 127	0.1/2		
[24,28)	2303	1604(10.9%)	699(11.7%)	5.127	0.165		
≥28	1025	728(5.0%)	297(5.0%)	_			
Respiratory history							
No	15477	11342 (77.3%)	4135 (69.2%)	146.000		1.00 [Reference]	
Yes	5168	3331 (22.7%)	1837 (30.8%)	146.882	<0.001	1.35 [1.25, 1.45]	< 0.001
Coffee consumption	า						
No	10290	7884 (53.7%)	2406 (40.3%)			1.00 [Reference]	
Occasionally	8331	5652 (38.5%)	2679 (44.9%)	415 020	-0.001	1.18 [1.10, 1.27]	< 0.001
Often	1441	799(5.4%)	642 (10.8%)	413.938	<0.001	1.55 [1.37, 1.76]	<0.001
Almost everyday	583	338 (2.3%)	245(4.1%)			1.29 [1.07, 1.56]	0.007

TABLE 1 Analysis of factors affecting sleep quality among college students in this survey.

(Continued)

TABLE 1 Continued

Variables	n=20645	5 Sleep Quality Univariate analysis		nalysis	Multivariate analysis	р		
		Good(≤7)	Poor(>7)	χ ²	р	OR [95% CI]		
Staying up late		'						
Not matched	10951	8519 (58.1%)	2432 (40.7%)			1.00 [Reference]		
Sometimes matched	7738	5195 (35.4%)	2543 (42.6%)	700 ((1) 0.00		1.35 [1.26, 1.45]	< 0.001	
Often matched	1442	746 (5.1%)	696 (11.7%)	/80.661	<0.001	1.93 [1.71, 2.18]	<0.001	
Always matched	514	213(1.5%)	301(5.0%)			2.24 [1.84, 2.73]	<0.001	
Long hours online								
Not matched	5892	4838(33.0%)	1054 (17.6%)			1.00 [Reference]		
Sometimes matched	8915	6483 (44.2%)	2432 (40.7%)	086 207	<0.001	1.29 [1.18, 1.41]	<0.001	
Often matched	4005	2457 (16.7%)	1548 (25.9%)	980.397	<0.001	1.83 [1.66, 2.03]	<0.001	
Always matched	1833	895 (6.1%)	938 (15.7%)			2.67 [2.36, 3.02]	<0.001	
Sudden changes								
No	19792	14207 (96.8%)	5585 (93.5%)	117.000	<0.001	1.00 [Reference]		
Yes	853	466 (3.2%)	387 (6.5%)	- 117.000 <0.0		1.89 [1.63, 2.20]	< 0.001	
Fears of infection								
Not matched	8038	5922 (40.4%)	2116 (35.4%)			1.00 [Reference]		
Sometimes matched	10274	7302 (49.8%)	2972 (49.8%)	120.260	<0.001	0.98 [0.92, 1.06]	0.641	
Often matched	1673	1055 (7.2%)	618 (10.3%)	120.200	<0.001	1.19 [1.05, 1.34]	0.005	
Always matched	660	394 (2.7%)	266(4.5%)			1.25 [1.04, 1.49]	0.015	
Impatient closed-lo	Impatient closed-loop management							
Not matched	7663	6250 (42.6%)	1413 (23.7%)			1.00 [Reference]		
Sometimes matched	8944	6252 (42.6%)	2692 (45.1%)	1027 292	<0.001	1.59 [1.47, 1.72]	<0.001	
Often matched	2252	1293 (8.8%)	959 (16.1%)	1037.303	<0.001	2.40 [2.15, 2.67]	<0.001	
Always matched	1786	878(6.0%)	908 (15.2%)			3.06 [2.72, 3.45]	<0.001	

OR, odds ratio; CI, confidence interval; BMI, Body Mass Index.

Clinical applications

Clinically, this model can provide actionable insights for assessing sleep-related risks and informing intervention measures. By inputting participant-specific information-such as age, specialty, respiratory history, coffee consumption, late-night habits, prolonged online activity, sudden life changes, and impatient closed-loop management-the model generates a risk score for poor sleep quality ranging from 0% to 100%. Using a threshold of 50%, the model offers tailored recommendations. If the risk score is ≥50%, further diagnostic evaluation and targeted treatment are advised. If the risk score is <50%, lifestyle modifications and regular follow-ups are recommended. To enhance accessibility and practical application, we have developed a user-friendly website (cosleep.angelong.cn) that integrates the ANN model. This platform allows both participants and clinicians to input relevant predictors and obtain immediate risk assessments along with actionable advice.

Discussion

Main findings

The current study is a valuable addition to the field as it created and verified a prediction model for estimating sleep quality. This model is easily accessible due to the utilization of the following eight readily available variables: age, specialty, respiratory history, coffee consumption, stay up, long hours online, sudden changes, and impatient closed-loop management.

During the COVID-19 pandemic, the contagious nature and uncertainty surrounding the virus have heightened fear of infection, leading to increased psychological stress and anxiety, which in turn adversely affected their sleep patterns. The implementation of closed-loop management significantly restricted students' mobility, limiting their ability to move freely on campus, participate in extracurricular activities, or engage in leisure activities as they previously could. Such restrictions contributed to



feelings of depression and irritability among students. Prolonged exposure to these emotional states has been shown to negatively impact sleep quality. In a study conducted by Kwon, Mihyoung et al., data analysis revealed that fear of COVID-19 is a significant factor influencing sleep quality, with a strong positive correlation observed between COVID-19-related fear and declines in sleep quality. These findings align closely with our conclusions (34).

Additionally, our results showed good predictive ability of our fitted models (i.e., cutoff, AUROC, accuracy, sensitivity, specificity,

precision, F1-score, and KAPPA values of 0.710, 0.713, 0.669, 0.682, 0.637, 0.822, 0.745, and 0.284, respectively). In addition, the Brier score was 0.182. The calibration curves showed good agreement between the predictions and the observations. The decision curve analysis demonstrated that the model could achieve a net benefit. The clinical impact curve confirmed the high clinical efficiency of the prediction model.

To evaluate whether we could gain a sufficient sample size to draw conclusions, we performed a *post hoc* sample size calculation based on



FIGURE 3

Feature selection using least absolute shrinkage and selection operator (LASSO) regression. (A) A LASSO-based ordinal logistic model with a 10-fold cross-validation based on the minimum mean squared error (MSE) was employed to find the optimal parameter (λ). The MSE vs. log (λ) is shown in the plot. The two dotted vertical lines indicate the optimal λ values based on the criterion of minimal MSE (λ min) or the criterion of one standard error of the minimum (λ 1se). (B) A LASSO coefficient profile of all feature variables against the log (λ) sequence. The blue and red dotted vertical lines represent the log(λ min) and log(λ 1se), respectively.



an online interactive tool (https://riskcalc.org/samplesize/). In the final model with eight predictors, we used the C-statistic in conjunction with the expected incidence to approximate the Cox–Snell R-squared and found that the poor sleep quality incidence was 28.9% for all participants. A minimal sample of 316 participants and a minimum of 11.42 events per predictor parameter were required. Thus, the actual sample of 20,645 patients in this study likely provided sufficient power to ensure the reliability of our results.

Strengths

To the best of our knowledge, this is the first prediction model derived from a LASSO algorithm for sleep prediction aimed at college students. Our team has made a significant breakthrough by creating an intuitive website (cosleep.angelong.cn) that allows students and administrators to effectively monitor their sleep quality in comparison to a previous study (26). This innovative platform will provide valuable insights into sleep patterns and ultimately improve overall wellness. It could boost more precise, data-driven, individualized risk estimations and promote better healthcare resource allocations. A novel insight into this application is shown in Figure 10.

This study is superior to a previous study (26) because here we updated the algorithm to select the eight most significant variables using LASSO and developed and validated a model with additional evaluation metrics (i.e., cutoff, F1-score, Brier score, decision curve analysis, and impact curve analysis).

Compared with other studies, Kim B.J. et al. used OSA and obesity as predictors to assess sleep quality using logistics, which lacked certain clinical predictive efficacy and generalizability (35).

TABLE 2	Performance	of	three	feature	selection	methods.

Feature Selection Method	No. of Feature Variables	AUROC (95% CI)	Accuracy (95% CI)	Nagelkerke R ²	RMSE	BIC
Univariate and multivariate stepwise regression	11	0.704(0.697-0.712)	0.670(0.670-0.670)	0.091	0.428	22705.47
LASSO regression	8	0.700(0.693-0.708)	0.676(0.676-0.676)	0.100	0.430	-34694.65
Boruta and stepwise regression	12	0.704(0.697-0.712)	0.678(0.678-0.678)	0.091	0.428	22714.36

AUROC, the area under the receiver operating curve; RMSE, root-mean-square error; BIC, Bayesian information criterion.

Age < 20 ≥20 Specialty	6819	2409		
< 20 ≥20 Specialty	6819	2409		
≥20 Specialty		2408	•	1.00 (1.00 to 1.00)
Specialty	7854	3564	i i i i i i i i i i i i i i i i i i i	1.28 (1.20 to 1.36)
Medical-related majors	3120	968	•	1.00 (1.00 to 1.00)
Science and engineering	5780	2096	H B -1	1.12 (1.02 to 1.23)
Liberal arts	5773	2908	H e H	1.36 (1.25 to 1.49)
Respiratory history				
No	11342	4135	•	1.00 (1.00 to 1.00)
Yes	3331	1837	H⊕t	1.34 (1.24 to 1.43)
Coffee consumption				
No	7884	2406	÷	1.00 (1.00 to 1.00)
Occasionally	5652	2679	I O I	1.21 (1.13 to 1.29)
Often	799	642	⊢●- 1	1.62 (1.43 to 1.83)
Almost everyday	338	245		1.34 (1.11 to 1.61)
Stay up			1	
Not matched	8519	2432	•	1.00 (1.00 to 1.00)
Sometimes matched	5195	2543	H e t	1.37 (1.28 to 1.47)
Often matched	746	696		1.99 (1.76 to 2.24)
Always matched	213	301	⊢	2.31 (1.90 to 2.82)
Long hours online				
Not matched	4838	1054		1.00 (1.00 to 1.00)
Sometimes matched	6483	2432	- + e +	1.32 (1.21 to 1.44)
Often matched	2457	1548	.	1.89 (1.71 to 2.09)
Always matched	895	938	⊢	→ 2.78 (2.46 to 3.14)
Sudden changes				
No	14207	5585		1.00 (1.00 to 1.00)
Yes	466	387	⊢ ●1	1.90 (1.64 to 2.21)
Impatient closed-loop manager	nent			
Not matched	6250	1413	l l	1.00 (1.00 to 1.00)
Sometimes matched	6252	2692	H e H	1.60 (1.48 to 1.73)
Often matched	1293	959		2.40 (2.16 to 2.67)
Always matched	878	908		• 3.06 (2.72 to 3.43)

FIGURE 5

Forest plot of ORs for 8 predictors included in the prediction model. The black dots and horizontal lines correspond to the OR and 95% CI values ORs with 95% CIs are shown on the right-hand side of the figure.

Lang, C. et al. studied sleep quality in adolescents by combining subjective and objective approaches using physical activity as a predictor (36), and Qing Hai Gong et al. used the dietary behaviors of adolescents as predictors. However, the above indicators are difficult to obtain and record and do not have advantages in largescale predictions (37).

The preliminary identification of variables for the final modeling using a multinomial logistic regression, LASSO regression, and Boruta feature selection was less dependent on the researcher's intuition. The use of machine learning (i.e., ANN, DT, GBT, K-nn, NB, and RF models) was also possible because large-scale representative data was used. In this study, the LASSO algorithm was used to select eight easy-to-obtain predictors and combined with machine learning algorithms to build a prediction model. Additionally, a *post hoc* sample size analysis based on an interactive online website showed good clinical prediction efficacy and generalizability of the model.

Limitations

There were several limitations to this study. First, since a crosssectional design was used, the shortcomings associated with this design could not be avoided. Cross-sectional studies only reflect a situation at a particular point in time, but due to their shortcomings in causal co-occurrence research, they cannot determine the causal relationship between sleep quality and factors such as psychological conditions. Second, since these college students were not independently sampled, some bias might have been introduced in the sampling process. Third, the data were derived from selfassessments via online surveys, which inevitably introduced some instability in the results. Fourth, according to the adherence survey, fewer students were willing to undergo evening sleep quality monitoring (i.e., use wearable devices or smart bracelets) and post-follow-ups, making it challenging to conduct further cohort studies. This could also pose difficulties to conduct further assessment and analysis of their sleep quality at different times. Fifth, although our machine learning modeling appeared to have good predictive ability, the results are dependent on the data used in the development and validation stages. If possible, external validation of college students from other provinces in China should be performed to produce a better-trained prediction model. Sixth, the adaptability of the findings may face challenges when applied to diverse populations. The core characteristic variables in the model-such as age, major, coffee intake, and staying up late habits-primarily reflect the lifestyle and



FIGURE 6

Development and application of a model for predicting the risk of a PSQI score of > 7. The prediction model. (**B**, **C**) Receiving operating characteristic curves showing the performance of the model in predicting sleep quality among college students in the (**B**) training–testing set and (**C**) independent validation set.

TABLE 3 Model performance.

	Discrimination tests										
Algorithm	Cutoff	AUROC (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Precision (95% CI)	F1-score (95% CI)	KAPPA (95% CI)			
ANN	0.710	0.713 (0.696-0.730)	0.669 (0.669-0.669)	0.682 (0.699-0.665)	0.637 (0.665-0.610)	0.822 (0.837-0.807)	0.745 (0.729-0.762)	0.284 (0.313-0.255)			
DT	0.765	0.627 (0.610-0.644)	0.688 (0.688-0.688)	0.780 (0.795-0.765)	0.463 (0.491-0.435)	0.781 (0.796-0.766)	0.780 (0.765-0.795)	0.243 (0.274-0.211)			
GBT	0.718	0.697 (0.679-0.714)	0.617 (0.617-0.617)	0.576 (0.594-0.558)	0.719 (0.745-0.694)	0.835 (0.851-0.818)	0.682 (0.663-0.700)	0.241 (0.267-0.214)			
K-nn	0.590	0.593 (0.575-0.612)	0.659 (0.659-0.659)	0.778 (0.793-0.763)	0.367 (0.394-0.339)	0.751 (0.767-0.736)	0.764 (0.749-0.780)	0.149 (0.180-0.117)			
NB	0.802	0.706 (0.689-0.723)	0.669 (0.670-0.669)	0.680 (0.697-0.664)	0.642 (0.670-0.615)	0.824 (0.839-0.809)	0.745 (0.729-0.761)	0.286 (0.315-0.257)			
RF	0.718	0.706 (0.688-0.723)	0.652 (0.652-0.652)	0.649 (0.667-0.632)	0.658 (0.685-0.631)	0.824 (0.839-0.808)	0.726 (0.709-0.743)	0.267 (0.295-0.238)			

ANN, artificial neural network; DT, decision tree; GBT, gradient-boosting tree; K-nn, k-nearest neighbor; NB, naïve Bayes; RF, random forest.



environmental characteristics of college students in Fujian, China. However, factors influencing sleep patterns can vary significantly across different demographic groups (e.g., medical students vs. nonmedical students, adolescents with vs. without a family psychiatric history, graduate vs. undergraduate students) (19, 38, 39). Furthermore, certain variables included in the model—such as "impatience with closed-loop management"—are closely tied to the specific pandemic-related management practices of Chinese universities, which may not be applicable to other regions or time periods. Additionally, cultural differences in behaviors such as coffee consumption and staying up late habits (e.g., higher coffee intake habits in international students from Europe or America) could further limit the predictive validity of the model when applied to populations outside the study context.

Future directions

In terms of further research, we will explore the performance of this model in other regions and utilize wearable devices to track the sleep quality of college students continuously. We would not just estimate the risk factors for sleep disturbance but also transform them into later interventions. The published literature on early interventions among the university student population with no prior sleep-related pathologies is scarce, despite the fact that they are considered a high-risk group (40–43). Once we identify students with high risk of sleep disturbance, we can study different early interventions (e.g., pharmacological intervention, behavioral sleeppromoting intervention, or no intervention) and quantify the effects of different intervention types on the sleep characteristics of adolescents and emerging adults who do or do not have a sleep disorder by studying the outcomes, including changes in the sleep measure scores that represent at least one of the key sleep metrics such as total sleep time (TST), sleep efficiency (SE), wake after sleep onset (WASO), and sleep onset latency (SOL) measured with actigraphy/polysomnography (PSG), sleep stages (rapid eye movement sleep and non-rapid eye movement sleep [stages 1, 2, and 3], as evaluated with PSG only). Furthermore, we recommend that future studies incorporate accelerometers to objectively assess physical activity levels. This approach would significantly improve the validity and reliability of the findings by reducing reliance on

TABLE 4 Results of calibration curve analysis of six machine learning models for predicting sleep quality in college students.

	Calibration								
Algorithm	Brier score	Slope	Intercept						
ANN	0.182	1.083	-0.119						
DT	0.195	0.896	0.093						
GBT	0.193	2.769	-1.533						
K-nn	0.229	0.376	0.682						
NB	0.208	0.259	0.648						
RF	0.185	1.062	-0.046						

ANN, artificial neural network; DT, decision tree; GBT, gradient-boosting tree; K-nn, knearest neighbor; NB, naïve Bayes; RF, random forest.



self-reported data and providing more accurate measurements of both sleep and activity patterns (36).

Distress Scale (K6) (45), Kessler Psychological Distress Scale (K10) (46), Fatigue Scale 14 (FS-14) (47), Generalized Anxiety Disorder Scale (GAD-7) (48) and Eating Attitudes Test (EAT-26) (49) to identify other significant variables from different perspective.

Additionally, we can combine some relevant psychological scales, such as Beck Depression Inventory (BDI) (44), Kessler Psychological





While our research focused on college students in Fujian Province, China, we acknowledge that cultural, academic, and lifestyle differences across regions may impact the model's performance in other settings. To address this, we plan to conduct external validation using our an intuitive website (cosleep.angelong.cn) in other provinces in China and, if possible, from international cohorts. This will allow us to assess the model's robustness and adaptability across diverse cultural and educational contexts.

Implications

In our previous study, it was found that a student's residence significantly affected their quality of sleep. However, the previous study's limitations highlighted the need to confirm the effectiveness of residence as a predictor in the model (26). In the current study, the LASSO regression analysis did not identify residence as a significant variable. This implied that the residence factor might not have a decisive impact on the sleep quality of college students. Therefore, this supported the use and popularization of this model at a broader geographical range of universities.

Undiagnosed sleep problems can worsen the mental stress experienced by college students, potentially leading to long-term health consequences for both the individuals and the healthcare system (50). By examining the sleep quality of college students in the post-epidemic era, we can identify those at high risk for sleep problems and design targeted health promotion interventions that address modifiable factors. This study aimed to identify significant variables associated with poor sleep quality and use them to create an easy-to-use website. The website accurately evaluated the probability of suffering from poor sleep quality, particularly among university students. Improving the sleep quality of college students through early interventions can lead to increased awareness of sleep health and, ultimately, better wellbeing and academic performance. By predicting sleep quality and implementing interventions, we can promote universal education on sleep health among college students. Ultimately, optimizing sleep health will benefit the overall wellbeing and academic success of college students.

Conclusion

The prediction model, which incorporated eight predictors, was built using a LASSO regression and an ANN to estimate the probability of sleep disturbance among college students. Additionally, based on this model, we built an easy-to-operate website (cosleep.angelong.cn) for improved monitoring, which may be used as an intuitive and practical tool by both individuals and school management.

Data availability statement

The source code and the datasets used in this study are freely available at https://github.com/ChunmeiFan/ PSQI-Prediction.git.

Ethics statement

The studies involving humans were approved by the Medical Ethics Committee of the Second Affiliated Hospital of Fujian Medical University, Sleep Medicine Key Laboratory of University in Fujian, and the Sleep Disorder Medicine Center of the Second Affiliated Hospital of Fujian Medical University (IRB No. 2021-309). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

LY: Formal analysis, Validation, Writing – original draft, Writing – review & editing. QC: Formal analysis, Investigation, Project administration, Writing – original draft, Writing – review & editing. KY: Formal analysis, Software, Writing – original draft, Writing – review & editing. ZZ: Formal analysis, Validation, Writing – review & editing. ZC: Software, Supervision, Writing – review & editing. DW: Formal analysis, Investigation, Writing – review & editing. YX: Software, Validation, Writing – review & editing. DC: Supervision, Validation, Writing – review & editing. LC: Conceptualization, Funding acquisition, Methodology, Resources, Writing – review & editing.

References

1. Werner-Seidler A, O'Dea B, Shand F, Johnston L, Frayne A, Fogarty AS, et al. A smartphone app for adolescents with sleep disturbance: development of the sleep ninja. *JMIR Ment Health.* (2017) 4(3):e28. doi: 10.2196/mental.7614

2. Li L, Wang YY, Wang SB, Zhang L, Li L, Xu DD, et al. Prevalence of sleep disturbances in Chinese university students: a comprehensive meta-analysis. *J Sleep Res.* (2018) 27:e12648. doi: 10.1111/jsr.12648

3. Liu X, Chen H, Bo Q-G, Fan F, Jia C-X. Poor sleep quality and nightmares are associated with non-suicidal self-injury in adolescents. *Eur Child Adolesc Psychiatry*. (2016) 26:271–9. doi: 10.1007/s00787-016-0885-7

4. Morin CM, Drake CL, Harvey AG, Krystal AD, Manber R, Riemann D, et al. Insomnia disorder. *Nat Rev Dis Primers*. (2015) 1:15026. doi: 10.1038/nrdp.2015.26

 Baker S, McBeth J, Chew-Graham CA, Wilkie R. Musculoskeletal pain and co-morbid insomnia in adults; a population study of the prevalence and impact on restricted social participation. *BMC Family Practice*. (2017) 18(1):17. doi: 10.1186/s12875-017-0593-5

6. Vitiello MV, McCurry SM, Shortreed SM, Baker LD, Rybarczyk BD, Keefe FJ, et al. Short-term improvement in insomnia symptoms predicts long-term improvements in sleep, pain, and fatigue in older adults with comorbid osteoarthritis and insomnia. *Pain.* (2014) 155:1547–54. doi: 10.1016/j.pain.2014.04.032

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyt.2025. 1585732/full#supplementary-material

SUPPLEMENTARY TEXT S1

A Survey on the Sleep Quality of College Students during COVID-19 Epidemic in Fujian Province.

7. McBeth J, Dixon WG, Moore SM, Hellman B, James B, Kyle SD, et al. Sleep disturbance and quality of life in rheumatoid arthritis: prospective mHealth study. J Med Internet Res. (2022) 24:e32825. doi: 10.2196/32825

8. Galiano-Castillo N, Arroyo-Morales M, Ariza-Garcia A, Fernández-Lao C, Fernández-Fernández AJ, Cantarero-Villanueva I. Factors that explain the cancerrelated insomnia. *Breast J.* (2017) 23:387–94. doi: 10.1111/tbj.12759

9. Javaheri S, Redline S. Insomnia and risk of cardiovascular disease. Chest. (2017) 152:435-44. doi: 10.1016/j.chest.2017.01.026

10. Taylor DJ, Bramoweth AD. Patterns and consequences of inadequate sleep in college students: substance use and motor vehicle accidents. *J Adolesc Health.* (2010) 46:610–2. doi: 10.1016/j.jadohealth.2009.12.010

11. Kloss JD, Nash CO, Horsey SE, Taylor DJ. The delivery of behavioral sleep medicine to college students. *J Adolesc Health.* (2011) 48:553-61. doi: 10.1016/j.jadohealth.2010.09.023

12. Peltzer K, Pengpid S. Nocturnal sleep problems among university students from 26 countries. *Sleep Breath.* (2015) 19:499–508. doi: 10.1007/s11325-014-1036-3

13. Deng J, Zhou F, Hou W, Silver Z, Wong CY, Chang O, et al. The prevalence of depressive symptoms, anxiety symptoms and sleep disturbance in higher education

students during the COVID-19 pandemic: A systematic review and meta-analysis. *Psychiatry Res.* (2021) 301:113863. doi: 10.1016/j.psychres.2021.113863

14. Zhang L, Zheng H, Yi M, Zhang Y, Cai G, Li C, et al. Prediction of sleep quality among university students after analyzing lifestyles, sports habits, and mental health. *Front Psychiatry.* (2022) 13:927619. doi: 10.3389/fpsyt.2022.927619

15. Li J, Zhou K, Li X, Liu M, Dang S, Wang D, et al. Mediator effect of sleep hygiene practices on relationships between sleep quality and other sleep-related factors in Chinese mainland university students. *Behav Sleep Med.* (2016) 14:85–99. doi: 10.1080/15402002.2014.954116

16. Li W, Chen J, Li M, Smith AP, Fan J. The effect of exercise on academic fatigue and sleep quality among university students. *Front Psychol.* (2022) 13:1025280. doi: 10.3389/fpsyg.2022.1025280

17. Lane HY, Chang CJ, Huang CL, Chang YH. An Investigation into Smartphone Addiction with Personality and Sleep Quality among University Students. *Int J Environ Res Public Health*. (2021) 18:7588. doi: 10.3390/ijerph18147588

18. Mesquita G, Reimão R. Quality of sleep among university students: effects of nighttime computer and television use. *Arq Neuropsiquiatr.* (2010) 68:720–5. doi: 10.1590/s0004-282x2010000500009

19. Azad MC, Fraser K, Rumana N, Abdullah AF, Shahana N, Hanly PJ, et al. Sleep disturbances among medical students: a global perspective. *J Clin Sleep Med.* (2015) 11:69–74. doi: 10.5664/jcsm.4370

20. Li Y, Fang J, Zhou C. Work-related predictors of sleep quality in Chinese nurses: testing a path analysis model. *J Nurs Res.* (2019) 27:e44. doi: 10.1097/jnr.00000000000319

21. Yang CY, Chiou AF. Predictors of sleep quality in community-dwelling older adults in Northern Taiwan. J Nurs Res. (2012) 20:249–60. doi: 10.1097/jnr.0b013e3182736461

22. Grimes M, Camerota M, Propper CB. Neighborhood deprivation predicts infant sleep quality. *Sleep Health.* (2019) 5:148–51. doi: 10.1016/j.sleh.2018.11.001

23. Sathyanarayana A, Joty S, Fernandez-Luque L, Srivastava J, Elmagarmid A, Arora T, et al. Sleep quality prediction from wearable data using deep learning. *JMIR Mhealth Uhealth.* (2016) 4:e125. doi: 10.2196/mhealth.6562

24. Magee CA, Robinson L, Keane C. Sleep quality subtypes predict health-related quality of life in children. *Sleep Med.* (2017) 35:67–73. doi: 10.1016/j.sleep.2017.04.007

25. Zheng W, Chen Q, Yao L, Zhuang J, Huang J, Hu Y, et al. Prediction models for sleep quality among college students during the COVID-19 outbreak: cross-sectional study based on the internet new media. *J Med Internet Res.* (2023) 25:e45721. doi: 10.2196/45721

26. Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med.* (1997) 16:385–95. doi: 10.1002/(sici)1097-0258(19970228)16:4<385::aid-sim380>3.0.co;2-3

27. Sun K, Huang SH, Wong DS, Jang SS. Design and application of a variable selection method for multilayer perceptron neural network with LASSO. *IEEE Trans Neural Netw Learn Syst.* (2017) 28:1386–96. doi: 10.1109/TNNLS.2016.2542866

28. Schilling M, Rickmann L, Hutschenreuter G, Spreckelsen C. Reduction of platelet outdating and shortage by forecasting demand with statistical learning and deep neural networks: modeling study. *JMIR Med Inform.* (2022) 10:e29978. doi: 10.2196/29978

29. Feng C, Wang L, Chen X, Zhai Y, Zhu F, Chen H, et al. A novel artificial intelligence-assisted triage tool to aid in the diagnosis of suspected COVID-19 pneumonia cases in fever clinics. *Ann Transl Med.* (2021) 9:201. doi: 10.21037/atm-20-3073

30. Vaid A, Jaladanki SK, Xu J, Teng S, Kumar A, Lee S, et al. Federated learning of electronic health records to improve mortality prediction in hospitalized patients with COVID-19: machine learning approach. *JMIR Med Inform.* (2021) 9:e24207. doi: 10.2196/24207

31. Wu G, Zhou S, Wang Y, Lv W, Wang S, Wang T, et al. A prediction model of outcome of SARS-CoV-2 pneumonia based on laboratory findings. *Sci Rep.* (2020) 10:14042. doi: 10.1038/s41598-020-71114-7

32. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res.* (1989) 28:193–213. doi: 10.1016/0165-1781(89)90047-4

33. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ.* (2016) 352:i6. doi: 10.1136/bmj.i6

34. Kwon M, Oh J. Factors affecting sleep quality of college students during the coronavirus disease 2019 pandemic: A cross-sectional study. *Medicina (Kaunas).* (2023) 59:416. doi: 10.3390/medicina59020416

35. Kim BJ, Park KM. Which factors are the most important for predicting sleep quality in obstructive sleep apnea patients with obesity? *Eur Neurol.* (2019) 81:190–6. doi: 10.1159/000502003

36. Lang C, Brand S, Feldmeth AK, Holsboer-Trachsler E, Pühse U, Gerber M. Increased self-reported and objectively assessed physical activity predict sleep quality among adolescents. *Physiol Behav.* (2013) 120:46–53. doi: 10.1016/j.physbeh.2013.07.001

37. Gong QH, Li H, Zhang XH, Zhang T, Cui J, Xu GZ. Associations between sleep duration and physical activity and dietary behaviors in Chinese adolescents: results from the Youth Behavioral Risk Factor Surveys of 2015. *Sleep Med.* (2017) 37:168–73. doi: 10.1016/j.sleep.2017.06.024

38. Baldini V, Gnazzo M, Rapelli G, Marchi M, Pingani L, Ferrari S, et al. Association between sleep disturbances and suicidal behavior in adolescents: a systematic review and meta-analysis. *Front Psychiatry*. (2024) 15:1341686. doi: 10.3389/fpsyt.2024.1341686

39. Baldini V, Gnazzo M, Maragno M, Biagetti R, Stefanini C, Canulli F, et al. Suicidal risk among adolescent psychiatric inpatients: the role of insomnia, depression, and social-personal factors. *Eur Psychiatry*. (2025) 68:e42. doi: 10.1192/j.eurpsy.2025.29

40. Ruiz-Zaldibar C, Gal-Iglesias B, Azpeleta-Noriega C, Ruiz-López M, Pérez-Manchón D. The effect of a sleep intervention on sleep quality in nursing students: study protocol for a randomized controlled trial. *Int J Environ Res Public Health*. (2022) 19:13886. doi: 10.3390/ijerph192113886

41. Friedrich A, Schlarb AA. Let's talk about sleep: a systematic review of psychological interventions to improve sleep in college students. *J Sleep Res.* (2018) 27:4–22. doi: 10.1111/jsr.12568

42. Griggs S, Conley S, Batten J, Grey M. A systematic review and meta-analysis of behavioral sleep interventions for adolescents and emerging adults. *Sleep Med Rev.* (2020) 54:101356. doi: 10.1016/j.smrv.2020.101356

43. Goncalves A, Bernal C, Korchi K, Nogrette M, Deshayes M, Philippe AG, et al. Promoting physical activity among university students during the COVID-19 pandemic: protocol for a randomized controlled trial. *JMIR Res Protoc.* (2022) 11: e36429. doi: 10.2196/36429

44. Garbuio ALP, Carvalhal TAO, Tomcix MFR, Dos Reis IGM, Messias LHD. Sleep quality, latency, and sleepiness are positively correlated with depression symptoms of Brazilians facing the pandemic-associated stressors of COVID-19. *Med (Baltimore).* (2022) 101:e28185. doi: 10.1097/MD.00000000028185

45. Uchida H, Kuroiwa C, Ohki S, Takahashi K, Tsuchiya K, Kikuchi S, et al. Assessing the smallest detectable change of the kessler psychological distress scale score in an adult population in Japan. *Psychol Res Behav Manage*. (2023) 16:2647–54. doi: 10.2147/PRBM.S417446

46. Ongeri L, Ametaj A, Kim H, Stroud RE, Newton CR, Kariuki SM, et al. Measuring psychological distress using the K10 in Kenya. J Affect Disord. (2022) 303:155-60. doi: 10.1016/j.jad.2022.02.012

47. Wu K, Li Y, Zou Y, Ren Y, Wang Y, Hu X, et al. Tai Chi increases functional connectivity and decreases chronic fatigue syndrome: A pilot intervention study with machine learning and fMRI analysis. *PloS One.* (2022) 17:e0278415. doi: 10.1371/journal.pone.0278415

48. Manzar MD, Alghadir AH, Khan M, Salahuddin M, Albougami A, Maniago JD, et al. Anxiety symptoms are associated with higher psychological stress, poor sleep, and inadequate sleep hygiene in collegiate young adults-A cross-sectional study. *Front Psychiatry.* (2021) 12:677136. doi: 10.3389/fpsyt.2021.677136

49. Azzi V, Hallit S, Malaeb D, Obeid S, Brytek-Matera A. Drunkorexia and emotion regulation and emotion regulation difficulties: the mediating effect of disordered eating attitudes. *Int J Environ Res Public Health*. (2021) 18:2690. doi: 10.3390/ijerph18052690

50. Palatty PL, Fernandes E, Suresh S, Baliga MS. Comparison of sleep pattern between medical and law students. *Sleep Hypn*. (2011) 13:1–2.