Nature and Science of Sleep

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Exploring Sleep Duration and Insomnia Among Prospective University Students: A Study with Geographical Data and Machine Learning Techniques

Firoj Al-Mamun[1–3], Mohammed A Mamun[1–3], Md Emran Hasan[1,4], Moneerah Mohammad ALmerab[5], David Gozal[6]

[1]CHINTA Research Bangladesh, Savar, Dhaka, Bangladesh; [2]Department of Public Health and Informatics, Jahangirnagar University, Savar, Dhaka, Bangladesh; [3]Department of Public Health, University of South Asia, Dhaka, Bangladesh; [4]Software College, Northeastern University, Shenyang, 110169, People's Republic of China; [5]Department of Psychology, College of Education and Human Development, Princess Nourah Bint Abdulrahman University, Riyadh, 11671, Saudi Arabia; [6]Department of Pediatrics and Office of the Dean, Joan C. Edwards School of Medicine, Marshall University, Huntington, WV, 25701, USA

Correspondence: Mohammed A Mamun, CHINTA Research Bangladesh, Savar, Dhaka, 1342, Bangladesh, Email mamunphi46@gmail.com; David Gozal, Joan C. Edwards School of Medicine, Marshall University, Huntington, WV, 25701, USA, Email gozal@marshall.edu

Background: Sleep disruptions among prospective university students are increasingly recognized for their potential ramifications on academic achievement and psychological well-being. But, information regarding sleep issues among students preparing for university entrance exams is unknown. Thus, this study aimed to investigate the prevalence and factors associated with sleep duration and insomnia among university entrance test-takers in Bangladesh, utilizing both traditional statistical analyses and advanced geographic information system and machine learning techniques for enhanced predictive capability.

Methods: A cross-sectional study was conducted in June 2023 among 1496 entrance test-takers at Jahangirnagar University, Dhaka. Structured questionnaires collected data on demographics, academic information, and mental health assessments. Statistical analyses, including chi-square tests and logistic regression, were performed using SPSS, while machine learning models were applied using Python and Google Colab.

Results: Approximately 62.9% of participants reported abnormal sleep duration (<7 hours/night or >9 hours/night), with 25.5% experiencing insomnia. Females and those dissatisfied with mock tests were more likely to report abnormal sleep duration, while repeat test-takers, those with unsatisfactory mock test results, or anxiety symptoms had a higher risk of insomnia. Machine learning identified satisfaction with previous mock tests as the most significant predictor of sleep disturbances, while higher secondary school certificate GPA had the least influence. The CatBoost model achieved maximum accuracy rates of 61.27% and 73.46%, respectively, for predicting sleep duration and insomnia, with low log loss values indicating robust predictive performance. Geographic analysis revealed regional variations in sleep disturbances, with higher insomnia prevalence in some southern districts and abnormal sleep duration in northern and eastern districts.

Conclusion: Sleep disturbances are prevalent among prospective university students and are associated with various factors including gender, test-taking status, mock test satisfaction, and anxiety. Targeted interventions, including sleep education and psychological support, hold promise in ameliorating sleep health and overall well-being among students, potentially enhancing entrance test performance.

Keywords: sleep disturbances, abnormal sleep duration, insomnia, university entrance exams, mental health

Introduction

Sleep is a fundamental component of health and well-being, and is critically important for cognitive functioning, emotional regulation, and overall physical health.[1] Abnormal sleep duration, defined as either short (less than 7 hours/night) or long

Received: 25 June 2024
Accepted: 14 August 2024
Published: 20 August 2024

(more than 9 hours/night) sleep, has been repeatedly associated with significant adverse effects on health-related quality of life.[2–9] Altered sleep patterns in adolescents and young adults are frequent, multifactorial, and entail disturbed appetite and physical activity that promote the risk of obesity along with emotional challenges and enhanced risk of depression, mood and psychiatric disorders.[10–18]

University entrance exams are extremely competitive in most countries, are often marked by intense study schedules and heightened stress levels, and consequently may lead to alterations in sleep patterns, and frequently induce the emergence of insomnia.[19] In a highly competitive educational environment, students are subjected to immense pressure in their aspirations to perform well on these exams. This pressure is further ignited by societal expectations and the high lifelong stakes associated with securing admission to prestigious academic institutions. Consequently, students will often opt for extending study duration at the expense curtailing sleep. Notably, poor sleep quality, irregular sleep schedules or insufficient sleep are associated with an increased risk of mental health issues, including anxiety and depression, which are frequently encountered among students facing academic pressures.[19]

In Bangladesh, evidence related to sleep-related problems among students preparing for university entrance exams remains under-explored. However, mental health problems, including suicide, burnout, digital addiction have all been evaluated.[20–23] Studies focused on sleep characteristics in university students in Bangladesh are scarce. For example, a study conducted among 332 university students reported a prevalence of 66.6% poor sleep quality, with students sleeping less than 7 hours nightly having poorer sleep quality compared to those sleeping more 7 hours per night.[24] In another study in which 450 undergraduate students were surveyed, 58.4% were poor sleepers and 26.2% had sleep disturbances, with cigarette smokers having significantly poorer sleep quality.[25] However, none of these studies specifically assess insomnia and sleep duration in the period of preparation for prospective university students who are attending entrance tests.

Moreover, the use of advanced analytical techniques such as Geographic Information Systems (GIS) and machine learning has been limited in this area of research. GIS and machine learning offer powerful tools for identifying and analyzing both spatial and non-spatial factors influencing sleep disturbances, providing a more comprehensive understanding of the issue. GIS can be utilized to map the spatial distribution of insomnia, revealing regional patterns and potential environmental influences. Machine learning, on the other hand, can handle large datasets to identify complex, non-linear relationships between various predictors and sleep outcomes, enhancing the accuracy and robustness of predictive models.

This study aims to address these gaps by (i) investigating the prevalence of insomnia and sleep duration issues among prospective university students in Bangladesh; (ii) identifying factors associated with these sleep disturbances using advanced analytical techniques; (iii) employing GIS to map the spatial distribution of insomnia and sleep duration, revealing regional patterns; and (iv) utilizing machine learning to identify significant predictors of sleep disturbances and enhance predictive accuracy. Thus, this comprehensive approach will provide valuable insights into the determinants of sleep health in this population and inform the development of effective interventions and policies to support students during this critical period.

Methods

Study Design, Participants and Procedures

A cross-sectional study was carried out among entrance test-takers at Jahangirnagar University, Dhaka, Bangladesh, in June 2023. Data were collected using a convenience sampling technique, targeting participants who resided in the university dormitories during the entrance test period. Data collection was conducted by a trained research team with prior experience in similar studies, ensuring reliability and consistency in the data-gathering process.

Measures

Sociodemographic and Admission-Related Factors

The study collected information related to sex, permanent residence (rural vs urban), and family type (nuclear vs joint). Besides, information related to the admission test included whether they were taking the test for the first time or were repeat test-takers. Data were gathered on their Grade Point Average (GPA) in the Secondary School Certificate (SSC) and Higher Secondary School Certificate (HSC), their satisfaction with mock tests during admission test preparation, and whether they were coached for the test by a professional or attended a coaching center. Additionally, their high school background (Science, Arts, and Commerce) was documented. Information on mental health problems was also collected (see below).

Anxiety (Generalized Anxiety Disorder)

Anxiety was assessed using the seven items of Generalized Anxiety Disorder (GAD-7).[26] Participants responded to the items based on the experience of the past two weeks on a four-point Likert scale (0 = not at all, 1 = several days, 2 = more than half of the days, and 3 = nearly every day). A total score ranging from 0 to 21, with a score of 10 or higher indicated anxiety among participants.[26]

Sleep Duration

Sleep duration was estimated based on the report and the National Sleep Foundation guidelines to define normal and abnormal sleep duration were used. According to the expert panel, 7 to 9 hours of sleep per night is considered normal for youth and young adults. Accordingly, sleeping less than 7 hours per night (short sleep duration) or more than 9 hours per night (long sleep duration) were categorized as abnormal sleep duration.[2]

Insomnia

Insomnia was evaluated using the two-item Insomnia Severity Index (ISI-2).[27] Participants responded to the items based on their experiences over the past two weeks on a five-point Likert scale (0=not at all to 4= very much). A total score ranging from 0 to 8 was calculated, with a score of 6 or higher indicating the presence of insomnia among participants.

Machine Learning Models

K-Nearest Neighbors (KNN)

An instance-based learning technique termed K-Nearest Neighbors (KNN) is predicated on delaying computation until after classification. Regression and classification tasks are carried out by locally approximating functions in this non-parametric method. KNN finds the K training examples in the feature space that are closest to the object for classification. A majority vote among the object's K closest neighbors determines the object's class membership; K is usually a modest positive number. The item is placed in the class of its single closest neighbor when K = 1.[28]

Random Forest (RF)

Renowned for its resilience in addressing classification and regression issues, Random Forest is an ensemble learning methodology. This approach entails building a large number of decision trees during training, from which the mean prediction (in the case of regression) or the mode of the classes (in the case of classification) is obtained. By bagging training models using a variety of subsets of the training data Random Forests reduce overfitting in decision trees and improve prediction accuracy. As a result, the diverse forest of trees yields a more reliable and comprehensive model.[29]

Support Vector Machine (SVM)

A reliable supervised learning approach that is mostly employed for classification problems is called Support Vector Machine (SVM). It functions by determining the best hyperplane to maximally divide various class memberships in a dataset. Support vectors data points that are closest to the decision boundary are used by SVM to improve classification accuracy. The kernel trick adds to its versatility by making it possible to handle nonlinear data effectively.[30]

Gradient Boosting Machines (GBM)

Gradient Boosting Machines (GBM) are a potent ensemble machine learning technology that build a strong learner by adding weak learners (usually decision trees) one after the other. This increases predicted accuracy. This strategy continuously focuses on areas where prior models underperformed by utilizing gradient descent to minimize the loss function. For a variety of predictive applications, GBMs are adaptable and efficient, although careful hyperparameter adjustment is needed to avoid overfitting and control computing demands. Their capacity to handle complicated, nonlinear data efficiently makes them widely utilized in a variety of sectors.[31]

Categorical Boosting (CatBoost)

CatBoost is a state-of-the-art machine learning technique that excels at managing data with categories. Decision tree gradient boosting is the main topic. Yandex created CatBoost, a novel technique that reduces typical problems with categorical data without requiring a large amount of pre-processing. In order to accomplish this, it combines one-hot encoding with an innovative algorithmic method that raises prediction accuracy while lowering overfitting. CatBoost is widely recognized for its effectiveness and expandability, rendering it an immensely beneficial instrument for numerous uses, including predictive modeling and recommendation systems. In the emerging field of machine learning, the method is significant because to its well-documented contribution to improving model performance, especially in datasets with strong categorical characteristics.[32]

Gradient Boosting (XGBoost)

XGBoost is a significant advancement in ensemble machine learning approaches as it outperforms classic gradient-boosting with its gradient-boosting architecture. The breakthroughs in algorithmic design and systems optimization that this methodology offers make it noteworthy. By learning from the mistakes made by earlier trees, it enables the sequential construction of decision trees, hence lowering errors.[33] Because of its methodical approach to improving speed and performance and its deft handling of large-scale data, XGBoost is a breakthrough in predictive modeling that beats its GBM predecessors in terms of accuracy.[31]

Ethics Statement

The study adhered to the guidelines provided in the Helsinki Declaration, as well as approved by the CHINTA Research Bangladesh [ref: chinta/2023/5]. Before initiating data collection, the purposes and objectives of the study were briefed to the participants. They were assured about the confidentiality of their data throughout the research process as well as their right to withdraw from participation at any point of the time. There were no participants under the age of 18, and written informed consent was required for participating in the study.

Statistical Analysis

Data were initially entered into Google Forms and subsequently analyzed using the IBM SPSS Statistics for Windows version 25.0, Armonk, NY: IBM Corp, USA. Frequencies and percentages were calculated for categorical variables, while chi-square tests were utilized to measure the association between the study variables and sleep duration and insomnia. To further explore relationships, a logistic regression model was developed, incorporating all covariates used in the study, with sleep duration and insomnia as the outcome variables. Model fit was assessed using the Hosmer-Lemeshow significant criteria, where a non-significant value indicated adequate model fit. In addition, Python was employed to apply machine learning models to the data. Google Colab served as the platform for analysis. Supervised classification algorithms, including K-Nearest Neighbor (KNN), Random Forest (RF), Support Vector Machine (SVM), Gradient Boosting Machines (GBM), Extreme Gradient Boosting (XGBoost), and Categorical Boosting (CatBoost), were utilized to classify the learning outcomes of sleep duration and insomnia. For all statistical tests, a two-tailed significance level of $p<0.05$ was set, with a 95% confidence interval.

Results

Characteristics of the Participants

A total of 1496 participants completed the survey and were included in the final analysis. Table 1 reports the distribution characteristics of the cohort. Most of the participants were females (56.7%), resided in rural areas (68.1%), and were from a nuclear family (82.2%). Regarding admission-related variables, 63% were first-time test-takers, had a high GPA, 77.2% were coached by a professional or were being trained in a coaching center, 36.5% were satisfied with their previous mock tests, and 70.2% were from a science background. Moreover, 45.2% participants reported anxiety symptoms (Table 1).

Table 1 Associations Between the Study Variables with Sleep Duration and Insomnia

Study Variables	Total	Sleep Duration			Insomnia	
	n (%)	Normal (557, 37.2%);	Abnormal (939, 62.8%);	χ^2 (*p*)	(n=381; 25.5%)	χ^2 (*p*)
		n (%)	n (%)		Yes; n (%)	
Sociodemographic Variables						
Sex						
Male	648 (43.3)	271; 41.8%	377; 58.2%	10.298 **(0.001)**	168; 25.9%	0.126 (0.722)
Female	848 (56.7)	286; 33.7%	562; 66.3%		213; 25.1%	
Permanent residence						
Rural	1009 (68.1)	371; 36.8%	638; 63.2%	0.673 (0.412)	249; 24.7%	1.010 (0.315)
Urban	472 (31.9)	184; 39%	288; 61%		128; 27.1%	
Family Type						
Nuclear	1210 (82.2)	446; 36.9%	764; 63.1%	0.158 (0.691)	300; 24.8%	1.353 (0.245)
Joint	262 (17.8)	100; 38.2%	162; 61.8%		74; 28.2%	
Admission-related Variables						
Student Status						
First-time test takers	942 (63)	341; 36.2%	601; 63.8%	1.162 (0.281)	200; 21.2%	24.052 **(<0.001)**
Repeat test takers	554 (37)	216; 39%	338; 61%		181; 32.7%	
Secondary School Certificate GPA						
Poor (<4.5)	215 (14.8)	75; 34.9%	140; 65.1%	1.214 (0.545)	69; 32.1%	14.026 **(0.001)**
Moderate	428 (29.4)	155; 36.2%	273; 63.8%		126; 29.4%	
High (5)	814 (55.9)	313; 38.5%	501; 61.5%		178; 21.9%	
Higher Secondary School Certificate GPA						
Poor (<4.5)	83 (5.7)	32; 38.6%	51; 61.4%	2.176 (0.337)	24; 28.9%	2.720 (0.257)
Moderate	371 (25.5)	127; 34.2%	244; 65.8%		105; 28.3%	
High (5)	1002 (68.8)	386; 38.5%	616; 61.5%		244; 24.4%	
Coached by professional or coaching center						
Yes	1134 (77.2)	415; 36.6%	719; 63.4%	0.115 (0.735)	273; 24.1%	6.587 **(0.010)**
No	335 (22.8)	126; 37.6%	209; 62.4%		104; 31%	
Satisfied with previous mock tests						
Yes	497 (36.5)	159; 32%	338; 68%	8.865 **(0.003)**	95; 19.1%	15.483 **(<0.001)**
No	863 (63.5)	346; 40.1%	517; 59.9%		248; 28.7%	

(*Continued*)

Table 1 (Continued).

Study Variables	**Total**	**Sleep Duration**			**Insomnia**	
	n (%)	**Normal (557, 37.2%);**	**Abnormal (939, 62.8%);**	**χ^2 (*p*)**	**(n=381; 25.5%)**	**χ^2 (*p*)**
		n (%)	**n (%)**		**Yes; n (%)**	
Sociodemographic Variables						
Educational background						
Science	1046 (70.2)	385; 36.8%	661; 63.2%	0.228 (0.892)	247; 23.6%	10.542 **(0.005)**
Commerce	59 (2.8)	16; 39%	25; 61%		7; 17.1%	
Arts	559 (27)	153; 38%	250; 62%		126; 31.3%	
Mental Health Problems						
Anxiety						
No	766 (54.8)	296; 38.6%	470; 61.4%	0.856 (0.355)	174; 22.7%	7.800 **(0.005)**
Yes	632 (45.2)	229; 36.2%	403; 63.8%		185; 29.3%	

Note: Significant p-values are bolded.

Of all responders, 37.2% reported a normal sleep duration of 7–9 hours per night, while 62.8% reported an abnormal sleep duration, with 58.8% indicating short sleep duration and 4% long sleep duration. Insomnia was reported among 25.5% of the participants (Table 1).

Relationship of the Variables with Sleep Duration and Insomnia

Table 2 presents significant associations between some of the study variables and sleep duration and insomnia. A significant association was found between sex and sleep duration patterns, with females reporting more abnormal sleep duration than males (χ^2=10.298, p=0.001). Additionally, participants who expressed satisfaction with their previous

Table 2 Logistic Regression Between Study Variables and Abnormal Sleep Duration and Insomnia

Study Variables	**Abnormal Sleep Duration (HL, *p* = 0.247)**		**Insomnia (HL, *p* = 0.478)**	
	AOR (95% CI)	***p*-value**	**AOR (95% CI)**	***p*-value**
Socio-demographic variables				
Sex				
Male	Reference	**0.031**	Reference	0.552
Female	1.31 (1.02–1.68)		0.91 (0.69–1.21)	
Permanent residence				
Rural	Reference	0.399	Reference	0.314
Urban	0.89 (0.68–1.16)		1.16 (0.86–1.55)	

(*Continued*)

Table 2 (Continued).

Study Variables	Abnormal Sleep Duration (HL, p = 0.247)		Insomnia (HL, p = 0.478)	
	AOR (95% CI)	***p*-value**	**AOR (95% CI)**	***p*-value**
Family type				
Nuclear	Reference	0.766	Reference	0.416
Joint	1.04 (0.76–1.43)		1.15 (0.81–1.62)	
Admission-related variables				
Student status				
First-time test takers	Reference	0.117	Reference	**0.001**
Repeat test takers	0.80 (0.61–1.05)		1.61 (1.20–2.16)	
Secondary School Certificate GPA				
Poor (<4.5)	Reference	0.159	Reference	0.328
Moderate (4.5–4.99)	0.99 (0.65–1.50)		0.90 (0.58–1.40)	
High (5)	0.75 (0.48–1.16)		0.73 (0.46–1.17)	
Higher Secondary School Certificate GPA				
Poor (<4.5)	Reference	0.348	Reference	0.969
Moderate (4.5–4.99)	1.35 (0.76–2.40)		0.95 (0.51–1.78)	
High (5)	1.10 (0.63–1.92)		0.93 (0.50–1.70)	
Coached by professional or coaching center				
Yes	Reference	0.758	Reference	0.640
No	1.05 (0.75–1.47)		1.09 (0.76–1.56)	
Satisfied with previous mock tests				
Yes	Reference	**0.016**	Reference	**<0.001**
No	0.73 (0.56–0.94)		1.74 (1.29–2.33)	
Educational background				
Science	Reference	0.247	Reference	0.146
Commerce	1.05 (0.48–2.33)		0.51 (0.19–1.40)	
Arts	0.77 (0.56–1.05)		1.24 (0.88–1.74)	
Mental health problems				
Anxiety				
No	Reference	0.573	Reference	**0.022**
Yes	1.07 (0.84–1.36)		1.36 (1.04–1.78)	

Note: Significant p-values are bolded.
Abbreviations: CI, confidence interval; HL, Hosmer-Lemeshow test.

mock tests exhibited a higher prevalence of abnormal sleep duration compared to those who were dissatisfied (χ^2=8.865, p=0.003).

Regarding insomnia, a significant relationship was observed with student status, as repeat test-takers exhibited a higher rate of insomnia symptoms compared to first-time test-takers (χ^2=24.052, p<0.001). Furthermore, students with poor SSC GPA experienced significantly more insomnia problems than those with moderate and high SSC GPAs (χ^2=14.026, p=0.001). However, insomnia was not significantly associated with HSC GPA (p>0.05). Moreover, participants who were coached by a professional or in a coaching center had significantly lower levels of insomnia than those who were not (χ^2=6.587, p=0.010). Those who did not have satisfactory mock test results also experienced a significantly higher rate of insomnia problems compared to those counterparts with satisfactory mock test results (χ^2=15.483, p<0.001). Additionally, suffering from mental health problems showed a significant relationship with insomnia (χ^2=7.800, p=0.005) (Table 1).

Factors Associated with Sleep Duration and Insomnia

Table 2 presents factors associated with sleep duration and insomnia. Females had a significantly higher risk of abnormal sleep duration compared to males [OR=1.31 (1.02–1.68); p=0.031]. Participants with dissatisfactory mock test results were 0.73 times less likely to experience abnormal sleep duration [OR=0.73 (0.56–0.94); p=0.016]. In terms of insomnia, repeat test-takers were at 1.61 times higher risk of suffering from insomnia compared to first-time test-takers [OR=1.61 (1.20–2.16); p=0.001], and insomnia risk was higher among participants with unsatisfactory mock test results [OR=1.74 (1.29–2.33); p<0.001]. Moreover, the risk of insomnia increased 1.36-fold among participants with anxiety symptoms [OR=1.36 (1.04–1.78); p=0.022] (Table 2).

Feature Selection

XGBoost SHAP values were used to assess each feature impact while feature correlations were examined using Cramer's V criterion to decide which features to include in the machine learning model. Since each attribute had a considerable impact on the model's output and little correlation with the others, they were all eventually added to the model. Satisfied with previous mock tests was the most significant predictor and higher secondary school certificate GPA had the least influence (Figure 1). Besides, some characteristics had a relatively higher influence on the model's output than others, such as the student status, anxiety, secondary school certificate GPA, educational background and gender.

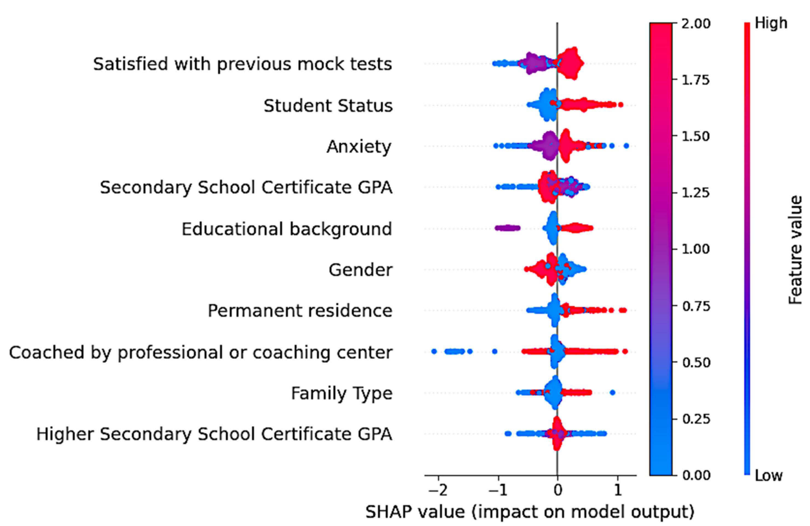

Figure 1 Features impact on the model by XGBoost SHAP value.

Table 3 Evaluation of Machine Learning Model Performances for Sleep Duration and Insomnia

Model	Sleep Duration				Insomnia			
	ACC	Precision	F1 score	Log loss	ACC	Precision	F1 score	Log loss
KNN	55.76	53.00	53.64	2.27	70.12	61.88	63.89	2.31
RF	54.76	51.95	52.67	1.05	70.95	64.13	65.36	0.69
SVM	57.76	52.77	52.75	0.73	71.12	64.92	65.99	0.66
GBM	61.10	56.18	52.41	0.68	72.45	61.24	62.99	0.58
XGBoost	60.10	54.92	53.02	0.70	71.95	61.98	63.48	0.59
CatBoost	61.27	53.95	48.89	0.67	73.46	54.17	62.36	0.56

Evaluation of Machine Learning Model Performances

Table 3 presents the predictive performance metrics of the machine learning models for sleep duration and insomnia. After assessments that included accuracy, precision, F1 score, and log-loss metrics, each model was found to have somewhat differing levels of ability to predict sleep duration and insomnia. Interestingly, all algorithms achieved robust degrees of accuracy. For example, CatBoost exhibited the highest accuracy score of 61.27% for sleep duration, while GBM demonstrated an accuracy score of 61.10%. Similarly, CatBoost yielded an accuracy of 73.46% for insomnia. Regarding F1 scores, KNN scored 53.64% for sleep duration, while SVM achieved K1 score of 65.99% for insomnia. Additionally, all algorithms demonstrated logarithmic loss values of less than 3% for both sleep duration and insomnia, indicating highly accurate and confident model predictions. Notably, CatBoost reached the maximum forecast accuracy with a minimum log loss of 0.67 and 0.56 for sleep duration and insomnia, respectively. Taken together, the CatBoost model exhibited overall better performance than the other models. Figure 2 displays the ROC-AUC of the various machine learning-derived algorithms for sleep duration and insomnia. In the case of sleep duration, the AUC score for the SVM was 0.52 and in the case of insomnia, an AUC score of 0.62 was recorded for the CatBoost model. Despite these findings, none of the models, including SVM and CatBoost, showed sufficient discriminatory power to effectively and accurately predictive individuals with abnormal sleep duration or insomnia.

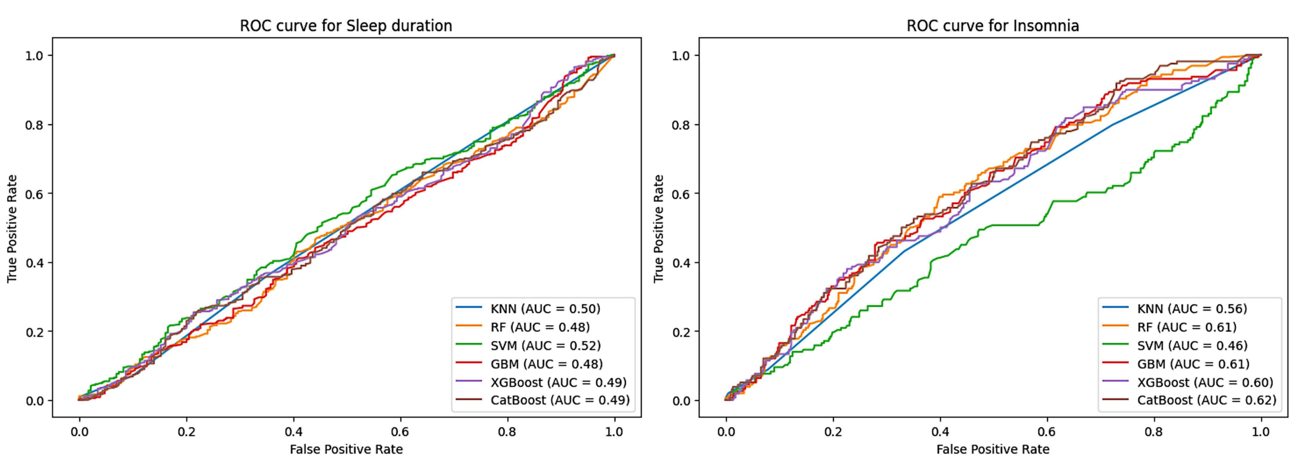

Figure 2 ROC-AUC curve of the sleep duration and insomnia.

https://doi.org/10.2147/NSS.S481786

Sleep Related Problems Across Districts

District residence was significantly associated with abnormal sleep duration (χ^2=179.997, p=0.001), but not with Insomnia (χ^2=70.418, p=0.243). However, in some southern districts, including Magura, Narail, Jessore, Bagerhat, Satkhira, Jhalokati, a higher prevalence of insomnia was detected when compared to other parts of the country (Figure 3). Regarding abnormal sleep duration, some northern districts, including Dinajpur, Joypurhat, Jamalpur, and some eastern districts, including Netrakona, Sunamganj, Narsighdi exhibited significantly more prevalent short sleep duration (Figure 4). Long sleep duration was more prevalent among some southern districts, including Panchagarh, Rangpur, Gaibandha, and some eastern districts, including Munshiganj, Bagerhat, Borguna, and Patuakhali (Figure 5).

Discussion

The present study examined abnormal sleep duration and insomnia among prospective university students. The findings revealed that 62.9% had abnormal sleep duration, and 25.5% had insomnia, with female, repeat-test-takers, those who were dissatisfied with their mock test results, and those reporting anxiety symptoms being at significantly higher risk of suffering from sleep-related problems. In addition, some regional divergence in the prevalence of sleep issues was noted and associated with the aforementioned risk factors. Machine learning modeling to enable prediction of individual test takers at risk of sleep problems incorporated several of the relevant risk factors, but their predictive ability was relatively modest. The current findings are anticipated to inform potential educational and interventional initiatives aimed at reducing sleep related abnormalities among participants and possibly improving their entrance exam scores.

The majority of the participants had abnormal sleep duration, with short sleep duration being clearly predominant, and insomnia was also quite frequent. The prevalence rates of sleep problems are concerning and suggest the presence of a widespread problem that could adversely affect both short-term and long-term academic performance and mental health. Previous studies conducted among university students reported similar findings with a preponderance of the students having poor sleep quality.[24,25] For instance, 61.7% university students reported less than 7 hours nightly sleep duration.[24] In another study involving a large cohort sample, 8.9% slept less than 7 hours nightly and 3.7% slept more than 9 hours nightly, clearly a markedly reduced prevalence of abnormal sleep duration when compared to our current findings.[34] We surmise that these differences may be due to the high-pressure academic environment and societal and familial expectations of being admitted into a university.

Notwithstanding, university students have consistently been found to suffer from a higher weighted mean prevalence of insomnia than the general population.[35] Such difficulties in falling asleep, maintaining sleep or both can reflect substantial problems with sleep hygiene (ie, markedly irregular bedtimes and other healthy sleep habits) but could also indicate higher stress. Here, approximately a quarter of students preparing the entrance exams had symptoms compatible with insomnia, which is comparatively higher than a previous nationwide study conducted in the general population whereby 13.1% had moderate and 2.8% had severe insomnia symptoms.[36] Furthermore, another study conducted among university students reported that 10.6% had clinically significant insomnia.[37] Considering the fact that university entrance test-takers suffer from comparatively high frequency of mental health problems,[20,21] the enhanced vulnerability to manifest insomnia symptoms is not surprising.

Our findings align with those of studies conducted in various regions worldwide, indicating a heightened risk of sleep-related issues among female students. It is now well-established that the complex interplay between biological, psychological, and social factors predisposes women to be more susceptible to the emergence of sleep disorders.[38–42] Similar findings have been also depicted in a previous study, thereby reinforcing the concept of enhanced susceptibility among women to develop and manifest poor sleep quality.[24] Given these gender disparities, targeted interventions focused to the specific needs of female students may be essential for early detection and intervention in cases of sleep disturbances. By implementing gender-specific approaches, we can effectively address sleep-related challenges and promote better overall well-being among female students.

Repeat-test takers had a higher risk of suffering from sleep-related problems. These findings likely reflect the increased pressure to improve their performance in subsequent test-taking attempts. Previous studies conducted among university entrance test-takers reported that repeat test-takers had a significantly higher rate of mental health problems,

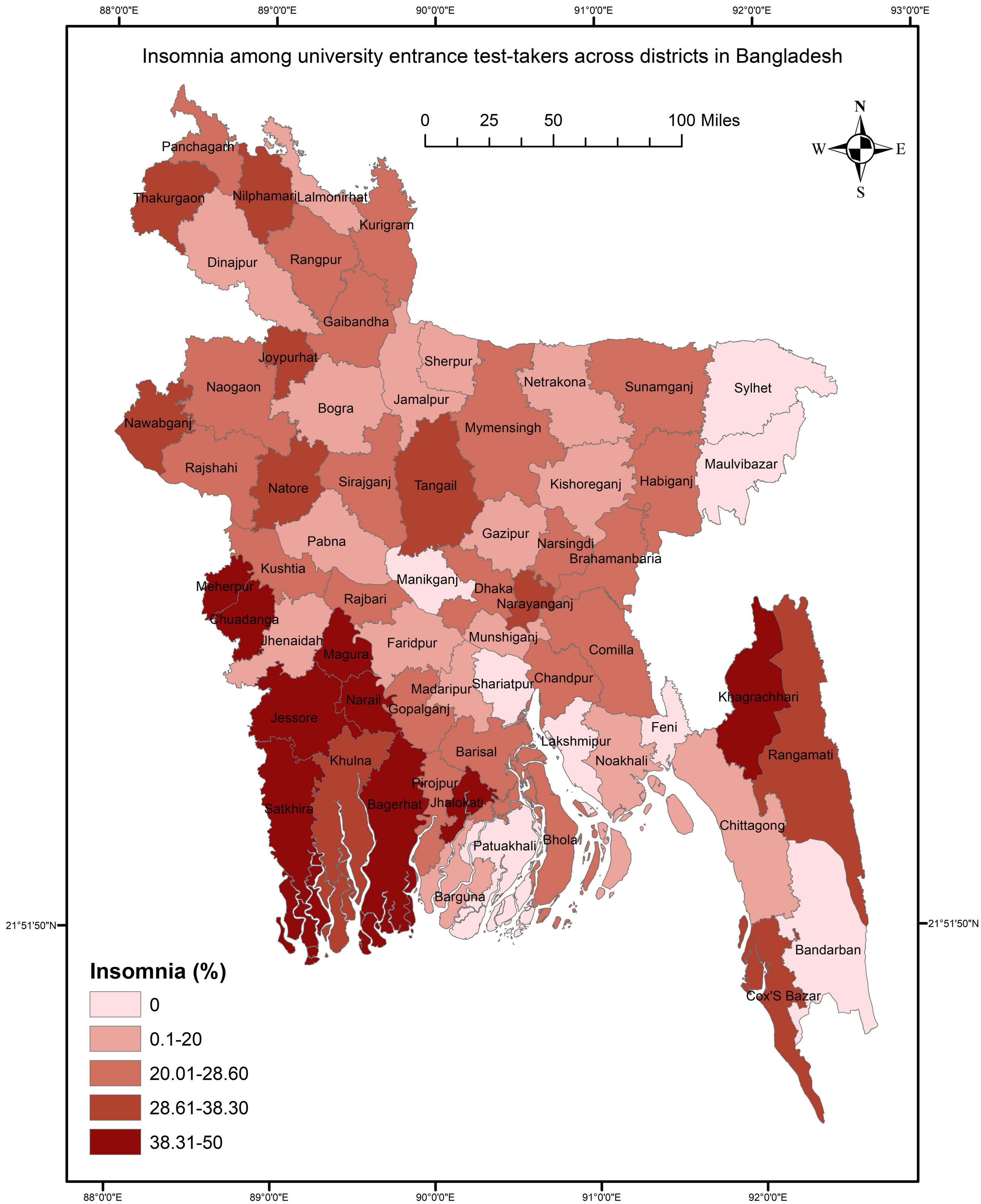

Figure 3 Distribution of insomnia among prospective university students across districts in Bangladesh.

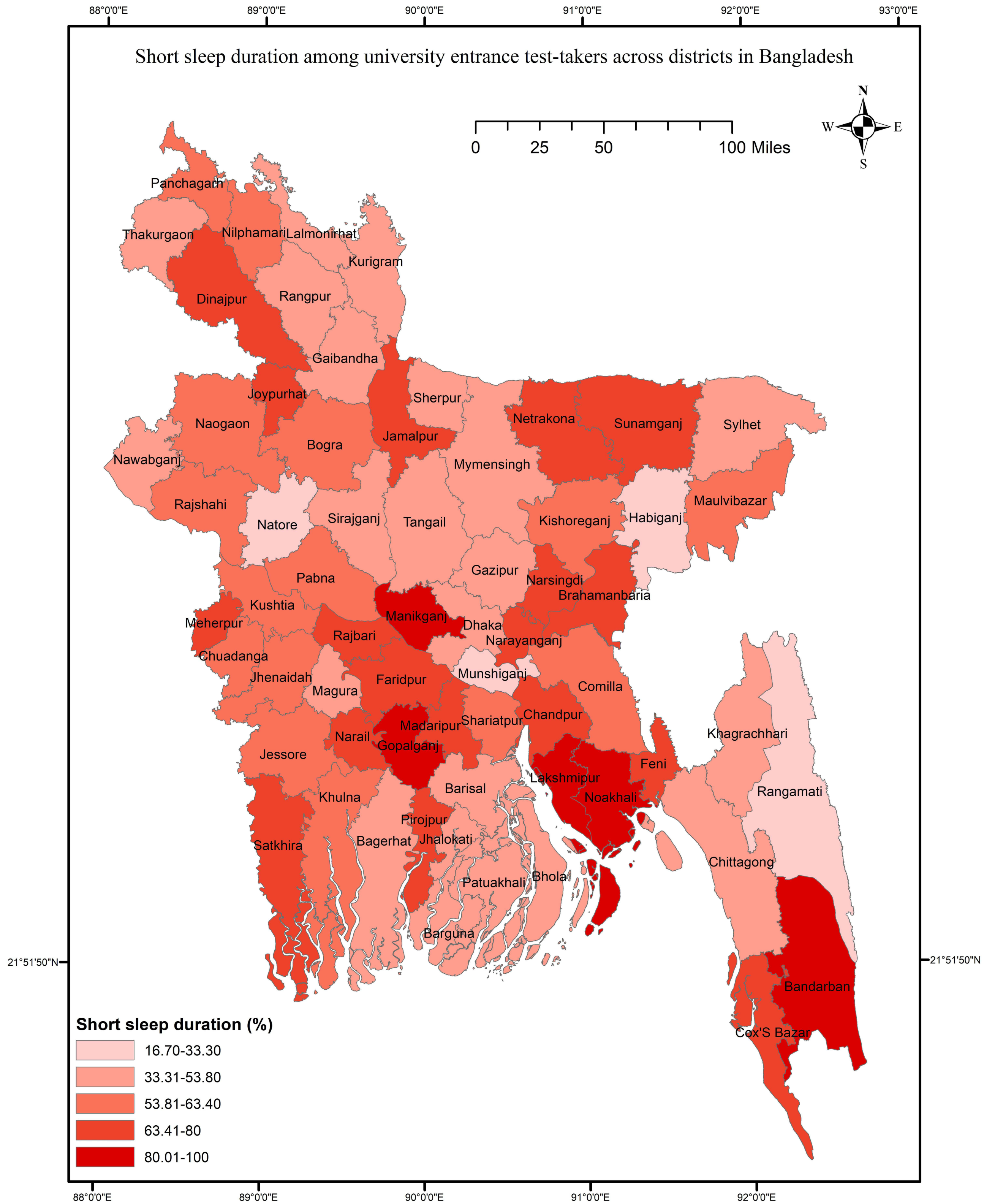

Figure 4 Distribution of short sleep duration [<7h] among prospective university students across districts in Bangladesh.

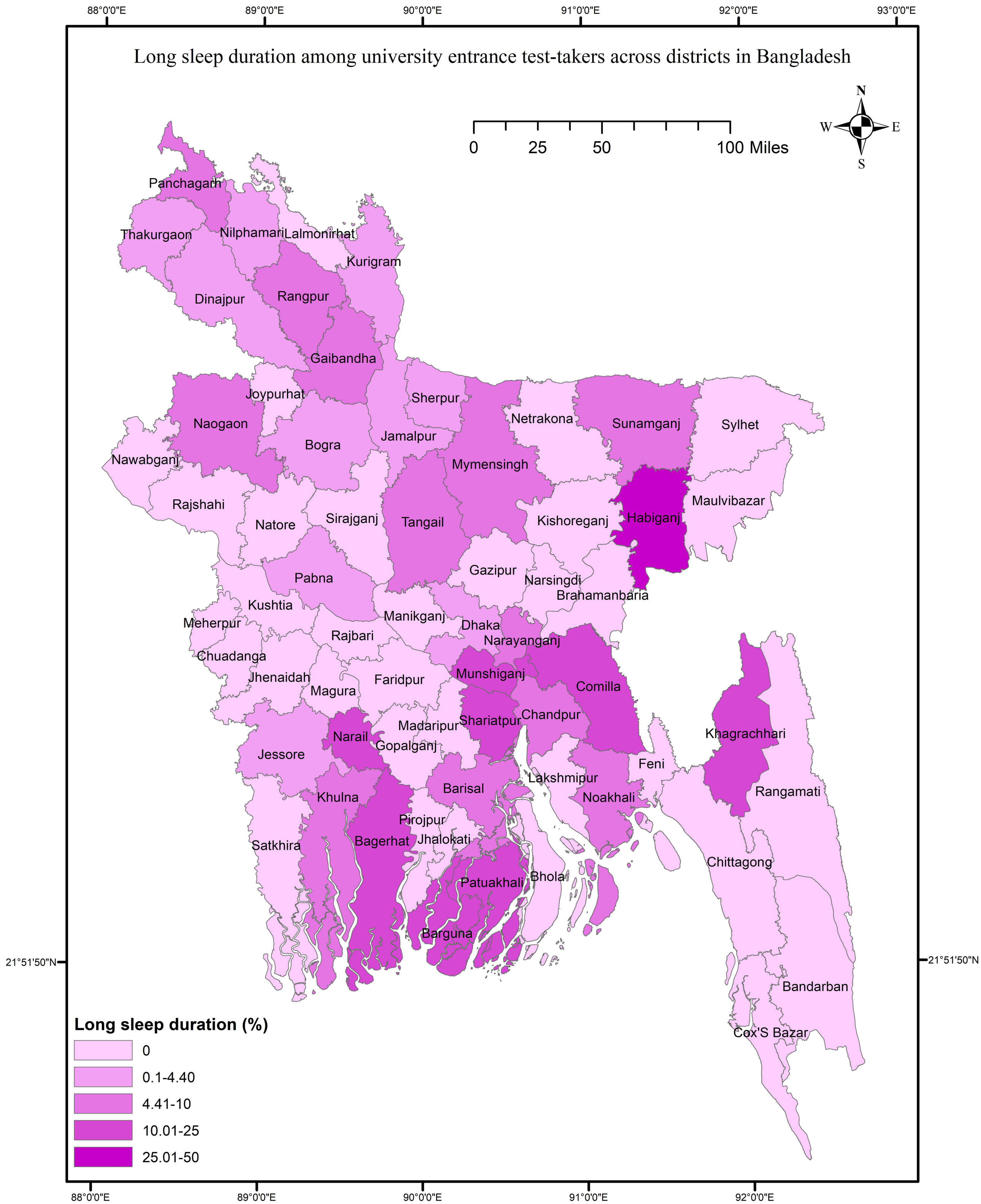

Figure 5 Distribution of long sleep duration [>9h] among prospective university students across districts in Bangladesh.

such as depression, anxiety, and burnout, compared to first-time test-takers.[21] Mental health problems are significantly associated with sleep disorders[43] and may even foster the risk of suicide.[22] Consequently, repeat test-takers should be prioritized in any screening for inordinate stress or sleep symptoms. The compounded stress these students face due to the pressure to achieve better results can lead to chronic stress and anxiety, further deteriorating sleep quality. This vicious cycle not only impairs academic performance but also poses serious threats to their mental health. Addressing these issues requires targeted interventions focusing on stress management, mental health support, and sleep hygiene education focused specifically for repeat test-takers. By providing these students with adequate resources and support, educational institutions can help mitigate the adverse effects of repeated test attempts on both sleep and overall mental well-being.

Dissatisfaction with previous mock test performance emerged as the most significant predictor of sleep problems according to the machine learning analyses. This dissatisfaction likely reflects underlying academic stress and anxiety, which are well-documented contributors to insomnia. Notably, a previous study among university entrance test takers found a significant relationship between dissatisfaction with mock test performance and burnout.[20] These findings suggest that the stress and anxiety associated with perceived academic inadequacies can significantly impact sleep quality. To mitigate these effects, it is crucial to provide students with constructive feedback and strategies for improvement, rather than focusing solely on their shortcomings. Such supportive interventions could help reduce academic stress, enhance coping mechanisms, and ultimately improve sleep quality.

Insomnia is intricately intertwined with mental health, particularly anxiety disorders.[44] This association highlights the bidirectional nature of the relationship, wherein insomnia can exacerbate anxiety symptoms, and heightened anxiety can perpetuate sleep difficulties. Consistent with findings from previous studies among university students in Lebanon,[37] this study reaffirms that anxiety is significantly more prevalent among students experiencing insomnia. Thus, the co-occurrence of insomnia and anxiety emphasizes the need for holistic approaches to mental health support in educational settings. Students experiencing sleep problems are not only at risk of academic difficulties but also face challenges related to emotional well-being. Addressing anxiety symptoms alongside sleep disturbances is crucial for promoting overall mental health resilience among university students. By recognizing and addressing the interconnected nature of these issues, educational institutions can implement comprehensive interventions that target both sleep hygiene and anxiety management, ultimately fostering a supportive environment conducive to student success and well-being.

In terms of our efforts to develop predictive algorithms using machine learning approaches, the CatBoost model yielded overall good accuracy rates of 61.27% and 73.46%, respectively for sleep duration and insomnia, along with low log loss values of 0.67% and 0.56% for sleep duration and insomnia, respectively. However, despite reasonably robust AUC values, the overall accuracy performance of such models would possibly be useful in the context of screening but would not meet the standards required for clinical diagnostics. Nonetheless, it is possible that with expanded utilization of such artificial intelligence approaches in ever growing cohorts, improved predictive performance of such algorithms will be achievable and enable systematic detections of vulnerable students even before they become symptomatic.

The district-wise data analyses revealed that geographical location was significantly associated with abnormal sleep duration but not with insomnia. Regional heterogeneity was apparent, with specific districts showing increased risk. A prior Bangladeshi study during the COVID-19 pandemic found notable district-wise differences in insomnia prevalence, with districts such as Joypurhat, Bogra, Bandarban, and Patuakhali reporting higher rates of insomnia; and it was suggested that areas heavily impacted by COVID-19 might have experienced increased insomnia rates.[36] Similarly, the current study's findings might reflect regional differences due to local environmental factors, cultural practices, or socioeconomic conditions that could exacerbate stress and anxiety, thus affecting sleep quality.[45] For instance, areas with higher noise pollution, inadequate housing conditions, or limited access to healthcare might experience greater sleep disturbances.[46,47] Besides, regional variations in cultural norms regarding bedtime routines and societal pressures can further influence sleep patterns.[48,49] Future research is needed to explore these regional disparities in greater depth to identify specific local factors influencing sleep health. Such studies will enable targeted interventions that address the unique needs of students from different districts, promoting better sleep hygiene and overall mental well-being.

This study benefits from a multifaceted approach, utilizing GIS and machine learning techniques to analyze sleep disturbances among prospective university students. The large sample size enhances the generalizability of findings, while identification of key risk factors like gender, repeat test-taking, and anxiety symptoms provides valuable insights for further research and practice. Innovative predictive modeling with machine learning highlights the potential of advanced analytics in identifying at-risk individuals. However, several limitations of this study deserve mention. First, the cross-sectional design precludes any inferences regarding causal relationships. Second, other sleep parameters such as week and weekend sleep duration, sleep-onset latency, wake after sleep onset, and daytime sleepiness were not included. Furthermore, the assessment of sleep duration relied on self-report rather than be based on objective measures such as actigraphic recordings, Moreover, using a convenience sampling method with self-rated surveys is inherently fraught with potentially limited generalizability. Thus, future research should aim to address these issues using more diverse and representative sampling methods, longitudinal designs, and complementary data collection techniques to minimize biases.

Conclusions

In summary, we found alarmingly high rates of abnormal sleep duration and insomnia symptoms among prospective university students in Bangladesh. Factors such as being female, a repeat test-taker, dissatisfaction with mock test results, and experiencing anxiety symptoms emerged as significant risk factors for sleep disturbances. It is possible that instituting comprehensive sleep education programs aimed at this particular at-risk group, along with providing psychological support when needed, may help alleviate the burden of sleep-related issues. However, despite efforts to generate algorithms for accurate prediction and detection of at-risk individuals, the performance of such approaches remained insufficiently robust to justify their implementation. Future research should focus on refining predictive models and exploring targeted interventions to support the mental and physical well-being of students during this critical period.

Data Sharing Statement

The dataset will be made available to appropriate academic parties upon appropriate request from the corresponding author/s.

Ethics Statement

The study adhered to the guidelines provided in the Helsinki Declaration, as well as approved by the CHINTA Research Bangladesh [ref: chinta/2023/5]. Before initiating data collection, the purposes and objectives of the study were briefed to the participants. They were assured about the confidentiality of their data throughout the research process as well as their right to withdraw from participation at any point of the time. There were no participants under the age of 18, and written informed consent was required for participating in the study.

Acknowledgment

The authors thank all the participants and team members involved in the project. Dr. ALmerab has to acknowledge the funding support currently receiving from Princess Nourah bint Abdulrahman University Researchers Supporting Project (PNURSP2024R563), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. Professor Gozal is supported in part by National Institutes of Health grants HL166617 and HL169266.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Watson NF, Badr MS, Belenky G, et al. Joint Consensus Statement of the American Academy of Sleep Medicine and Sleep Research Society on the recommended amount of sleep for a healthy adult: methodology and discussion. *J Clin Sleep Med*. 2024;11:931–952. doi:10.5664/jcsm.4950
2. Hirshkowitz M, Whiton K, Albert SM, et al. National Sleep Foundation's updated sleep duration recommendations: final report. *Sleep Health*. 2015;1(4):233–243. doi:10.1016/j.sleh.2015.10.004
3. Wang Y, Li B, Zhang C, Buxton OM, Redline S, Li X. Group-based sleep trajectories in children and adolescents: a systematic review. *Sleep Med Rev*. 2024;75:101916. doi:10.1016/j.smrv.2024.101916

4. Gale EL, James Williams A, Cecil JE. The relationship between multiple sleep dimensions and obesity in adolescents: a systematic review. *Sleep Med Rev*. 2024;73:101875. doi:10.1016/j.smrv.2023.101875
5. Gadam S, Pattinson CL, Rossa KR, et al. Interventions to increase sleep duration in young people: a systematic review. *Sleep Med Rev*. 2023;70:101807. doi:10.1016/j.smrv.2023.101807
6. van de Langenberg SCN, Kocevska D, Luik AI. The multidimensionality of sleep in population-based samples: a narrative review. *J Sleep Res*. 2022;31(4):e13608. doi:10.1111/jsr.13608
7. Musshafen LA, Tyrone RS, Abdelaziz A, et al. Associations between sleep and academic performance in US adolescents: a systematic review and meta-analysis. *Sleep Med*. 2021;83:71–82. doi:10.1016/j.sleep.2021.04.015
8. LeBourgeois MK, Hale L, Chang A-M, Akacem LD, Montgomery-Downs HE, Buxton OM. Digital media and sleep in childhood and adolescence. *Pediatrics*. 2017;140(Supplement_2):S92–S96. doi:10.1542/peds.2016-1758J
9. Owens JA, Weiss MR. Insufficient sleep in adolescents: causes and consequences. *Minerva Pediatr*. 2017;69(4):326–336. doi:10.23736/S0026-4946.17.04914-3
10. Grasaas E, Ostojic S, Jahre H. Adherence to sleep recommendations is associated with higher satisfaction with life among Norwegian adolescents. *BMC Public Health*. 2024;24(1):1288. doi:10.1186/s12889-024-18725-1
11. Dong R, Dou K, Luo J. Construction of a model for adolescent physical and mental health promotion based on the multiple mediating effects of general self-efficacy and sleep duration. *BMC Public Health*. 2023;23(1):2293. doi:10.1186/s12889-023-17197-z
12. Park G-R, Kim J. The gendered health consequences of persistent exposure to short sleep duration during adolescence. *J Adolesc*. 2023;95 (7):1311–1320. doi:10.1002/jad.12203
13. Cao Y, Yang Z, Yu Y, Huang X. Physical activity, sleep quality and life satisfaction in adolescents: a cross-sectional survey study. *Front Public Health*. 2022;10. doi:10.3389/fpubh.2022.1010194
14. Gao M, Li X, Lee C-Y, et al. Sleep duration and depression among adolescents: mediation effect of collective integration. *Front Psychol*. 2022:13. 10.3389/fpsyg.2022.1015089
15. Perez-Chada D, Gozal D. Sleep irregularity and duration in teenagers: a complex constellation of cause, consequence, both, and interdependencies. *Sleep*. 2023;46(2):zsac300. doi:10.1093/sleep/zsac300
16. Estevan I, Sardi R, Tejera AC, Silva A, Tassino B. Should I study or should I go (to sleep)? The influence of test schedule on the sleep behavior of undergraduates and its association with performance. *PLoS One*. 2021;16(3):e0247104. doi:10.1371/journal.pone.0247104
17. Okoli A, Hanlon EC, Brady MJ. The relationship between sleep, obesity, and metabolic health in adolescents: a review. *Curr Opin Endocr Metab Res*. 2021;17:15–19. doi:10.1016/j.coemr.2020.10.007
18. Li Y, Bai W, Zhu B, et al. Prevalence and correlates of poor sleep quality among college students: a cross-sectional survey. *Health Qual Life Outcom*. 2020;18(1):210. doi:10.1186/s12955-020-01465-2
19. Lund HG, Reider BD, Whiting AB, Prichard JR. Sleep patterns and predictors of disturbed sleep in a large population of college students. *J Adolesc Health*. 2010;46(2):124–132. doi:10.1016/j.jadohealth.2009.06.016
20. Mamun MA, Safiq MB, Hosen I, Al Mamun F. Burnout, does the university entrance test failing attribute? a Bangladeshi exploratory study. *PLoS One*. 2021;16(10):e0258100. doi:10.1371/journal.pone.0258100
21. Mamun MA, Misti JM, Hosen I, Al Mamun F. Suicidal behaviors and university entrance test-related factors: a Bangladeshi exploratory study. *Perspect Psychiatr Care*. 2022;58(1):278–287. doi:10.1111/ppc.12783
22. Nahrin R, Al-Mamun F, Kaggwa MM, Al Mamun M, Mamun MA. Prevalence and factors associated with suicidal ideation among students taking university entrance tests: revisited and a study based on Geographic Information System data. *BJPsych Open*. 2023;9(4):e129. doi:10.1192/bjo.2023.526
23. Al-Mamun F, Hasan ME, Mostofa NB, et al. Prevalence and factors associated with digital addiction among students taking university entrance tests: a GIS-based study. *BMC Psychiatry*. 2024;24(1):322. doi:10.1186/s12888-024-05737-9
24. MdS A, Khan S, Hsan K, Sen LC, Yunus FM, Griffiths MD. Factors affecting sleep quality among the university students in Bangladesh: a cross-sectional structured interview study. *Sleep Vigil*. 2020;4(2):177–184. doi:10.1007/s41782-020-00106-4
25. Mamun MA, Hossain MS, Kamruzzaman M, et al. Prevalence of poor sleep quality and its determinants among Bangladeshi students: a pilot study. *Sleep Vigil*. 2020;4(2):185–193. doi:10.1007/s41782-020-00109-1
26. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing Generalized Anxiety Disorder: the GAD-7. *Arch Intern Med*. 2006;166(10):1092–1097. doi:10.1001/archinte.166.10.1092
27. Kraepelien M, Blom K, Forsell E, et al. A very brief self-report scale for measuring insomnia severity using two items from the Insomnia Severity Index - development and validation in a clinical population. *Sleep Med*. 2021;81:365–374. doi:10.1016/j.sleep.2021.03.003
28. Keller JM, Gray MR, Givens JA. A fuzzy k-nearest neighbor algorithm. *IEEE Trans Syst Man Cybern*. 1985;SMC-15(4):580–585. doi:10.1109/TSMC.1985.6313426
29. Belgiu M, Drăguţ L. Random forest in remote sensing: a review of applications and future directions. *ISPRS J Photogram Remote Sensing*. 2016;114:24–31. doi:10.1016/j.isprsjprs.2016.01.011
30. Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995;20(3):273–297. doi:10.1007/BF00994018
31. Friedman JH. Greedy Function Approximation: a Gradient Boosting Machine. *Ann Stat*. 2001;29(5):1189–1232.
32. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A. CatBoost: unbiased boosting with categorical features. *Adv Neural Inf Process Syst*. 2018;31. doi:10.48550/arXiv.1706.09516
33. Chen T, Guestrin C. Xgboost: a scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. 2016. pp. 785–794. doi:10.1145/2939672.2939785.
34. Yunus FM, Khan S, Akter T, et al. How many hours do people sleep in Bangladesh? A country-representative survey. *J Sleep Res*. 2016;25 (3):365–376. doi:10.1111/jsr.12381
35. Jiang X-L, Zheng X-Y, Yang J, et al. A systematic review of studies on the prevalence of Insomnia in university students. *Public Health*. 2015;129 (12):1579–1584. doi:10.1016/j.puhe.2015.07.030
36. Al-Mamun F, Gozal D, Hosen I, Misti JM, Mamun MA. Predictive factors of insomnia during the COVID-19 pandemic in Bangladesh: a GIS-based nationwide distribution. *Sleep Med*. 2022;91:219–225. doi:10.1016/j.sleep.2021.04.025

37. Choueiry N, Salamoun T, Jabbour H, El Osta N, Hajj A, Rabbaa Khabbaz L. Insomnia and relationship with anxiety in university students: a cross-sectional designed study. *PLoS One*. 2016;11(2):e0149643. doi:10.1371/journal.pone.0149643
38. Marver JE, McGlinchey EA. Sex differences in insomnia and risk for psychopathology in adolescence. *Curr Opin Psychol*. 2020;34:63–67. doi:10.1016/j.copsyc.2019.09.004
39. Zhang J, Chan NY, Lam SP, et al. Emergence of sex differences in insomnia symptoms in adolescents: a large-scale school-based study. *Sleep*. 2016;39(8):1563–1570. doi:10.5665/sleep.6022
40. de Zambotti M, Goldstone A, Colrain IM, Baker FC. Insomnia disorder in adolescence: diagnosis, impact, and treatment. *Sleep Med Rev*. 2018;39:12–24. doi:10.1016/j.smrv.2017.06.009
41. Suh S, Cho N, Zhang J. Sex differences in insomnia: from epidemiology and etiology to intervention. *Curr Psychiatry Rep*. 2018;20(9):69. doi:10.1007/s11920-018-0940-9
42. Sivertsen B, Vedaa Ø, Harvey AG, et al. Sleep patterns and insomnia in young adults: a national survey of Norwegian university students. *J Sleep Res*. 2019;28(2):e12790. doi:10.1111/jsr.12790
43. Scott AJ, Webb TL, Martyn-St James M, Rowse G, Weich S. Improving sleep quality leads to better mental health: a meta-analysis of randomised controlled trials. *Sleep Med Rev*. 2021;60:101556. doi:10.1016/j.smrv.2021.101556
44. Gardani M, Bradford DRR, Russell K, et al. A systematic review and meta-analysis of poor sleep, insomnia symptoms and stress in undergraduate students. *Sleep Med Rev*. 2022;61:101565. doi:10.1016/j.smrv.2021.101565
45. Al-Mamun F, Hussain N, Sakib N, et al. Sleep duration during the COVID-19 pandemic in Bangladesh: a GIS-based large sample survey study. *Sci Rep*. 2023;13(1):3368. doi:10.1038/s41598-023-30023-1
46. Halperin D. Environmental noise and sleep disturbances: a threat to health? *Sleep Sci*. 2014;7(4):209–212. doi:10.1016/j.slsci.2014.11.003
47. Monazzam MR, Shamsipour M, Zaredar N, Bayat R. Evaluation of the relationship between psychological distress and sleep problems with annoyance caused by exposure to environmental noise in the adult population of Tehran Metropolitan City, Iran. *J Environ Health Sci Eng*. 2022;20(1):1–10. doi:10.1007/s40201-021-00703-z
48. Jeon M, Dimitriou D, Halstead EJ. A systematic review on cross-cultural comparative studies of sleep in young populations: the roles of cultural factors. *Int J Environ Res Public Health*. 2021;18(4):2005. doi:10.3390/ijerph18042005
49. Pal E, Blackwell JE, Ball HL, Collings PJ. Sociodemographic, temporal and bedtime routine correlates of sleep timing and duration in South Asian and white children: a Born in Bradford study. *Sleep Med*. 2023;5:100068. doi:10.1016/j.sleepx.2023.100068