

Development and internal validation of a prediction model for moderate-to-severe insomnia among nurses using machine learning-assisted predictor selection: a multicenter cross-sectional study

Yue Zhao

Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua Medicine, Tsinghua University

Zhe Liu

Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua Medicine, Tsinghua University

Lu Chen

Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua Medicine, Tsinghua University

Huaxin Si

Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua Medicine, Tsinghua University

Yan Ma

Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua Medicine, Tsinghua University

Mi Zhao

Peking University First Hospital

Li Song

Peking University First Hospital

Xiaoyu Zhang

Beijing Tsinghua Chang Gung Hospital

Dong Li

1da04595@btch.edu.cn

Beijing Tsinghua Chang Gung Hospital

Xu Zhang

Sun Yat-sen University

Keywords: Nurse, Moderate-to-severe insomnia, Work stress, Career adaptability, Prediction model, Machine learning

Posted Date: October 22nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11262809/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Development and internal validation of a prediction model for moderate-to-severe insomnia among nurses using machine learning-assisted predictor selection: a multicenter cross-sectional study

Yue Zhao, Zhe Liu, Lu Chen, Huaxin Si, Yan Ma, Mi Zhao, Li Song, Xiaoyu Zhang, Dong L*,
Xu Zhang*

First author :

Yue Zhao, PhD, Department of Nursing, Beijing Tsinghua Changgung Hospital, School of
Clinical Medicine, Tsinghua Medicine, Tsinghua University, No. 168 Litang Road, Changping
District, Beijing, China, 102218. Email: zhaoyue729@163.com ORCID : 0000-0002-9449-
5483

Corresponding author:

Dong Li, PhD, Department of Medical Data Science Center, Beijing Tsinghua Changgung
Hospital, School of Clinical Medicine, Tsinghua Medicine, Tsinghua University, No. 168 Litang
Road, Changping District, Beijing, China, 102218. Email: lda04595@btch.edu.cn

Xu Zhang, RN, PhD, School of Nursing, Sun Yat-sen University, No. 74, 2nd Sun Yat-sen
Road/Street, Guangzhou, China, 510080. Email: lnzhangxu@foxmail.com

Abstract

Background: Clinically significant insomnia is common among nurses. We aimed to develop and internally validate a parsimonious prediction model for moderate-to-severe insomnia among nurses, to quantify its optimism, and to test whether machine learning adds value beyond logistic regression.

Methods: In a multicenter cross-sectional survey (four hospitals in China, 2023; N = 2,179), moderate-to-severe insomnia was defined as an Insomnia Severity Index (ISI) score ≥ 15 . Candidate predictors were screened with Boruta and LASSO in a training set (n = 1,917), and a logistic regression model was evaluated in a stratified random test set (n = 262). Hospital identifiers were unavailable, so only internal validation was possible. Optimism was assessed with bootstrap correction, 500 repeated random splits and nested cross-validation including the selection step; random forest and gradient boosting served as benchmarks; calibration, net benefit and sensitivity to straight-line responding (n = 1,144 retained) were examined (TRIPOD+AI).

Results: Moderate-to-severe insomnia affected 29.8% (649/2,179). Night shifts per month, challenge and hindrance stress (CHSS) and career adaptability (CAAS) were selected in every nested fold: >8 night shifts/month versus none, odds ratio (OR) 2.37 (95% CI 1.79–3.15); hindrance stress per SD, 1.61 (1.35–1.93); challenge stress per SD, 1.33 (1.11–1.61); CAAS per 10 points, 0.84 (0.78–0.91). Test-set AUC was 0.766 (95% CI 0.708–0.825), calibration slope 1.36 (0.96–1.94) and Brier score 0.174; however, this lay in the upper range of 500 random splits (mean AUC 0.72; 2.5th–97.5th percentile 0.65–0.78), and the optimism-corrected AUC was 0.71 (nested cross-validation 0.71). Random forest (AUC 0.77) and gradient boosting (0.78) did not differ significantly from logistic regression, and the stress scales alone gave 0.75. At a risk

1 threshold of 0.30, sensitivity was 72%, specificity 63% and positive predictive value 45%.

2 Results were consistent in the screened subsample (AUC 0.72).

3 **Conclusions:** Night-shift burden, work stress and lower career adaptability were associated with
4 moderate-to-severe insomnia, but discrimination was modest once optimism was accounted for
5 (AUC about 0.71–0.72). Because predictors and outcome were self-reported at one time point
6 and only internal validation was possible, the model and its nomogram are exploratory and
7 require external validation before any use in practice.

8 **Key words:** Nurse, Moderate-to-severe insomnia, Work stress, Career adaptability, Prediction
9 model, Machine learning

10

Background

Insomnia, the most common sleep disorder globally, is a nightly complaint of sleep disturbance due to insufficient amount of sleep, difficulty in initiating or maintaining sleep, or non-restorative sleep [1]. A recent meta-analysis reported a pooled prevalence of insomnia of 22.0% among the general population [2]. Evidence has shown that chronically poor sleep is associated with elevated risks of cardiovascular, metabolic diseases, mental health disorders, and suicidal behaviors [3-4].

Prevalence of insomnia is particularly high among nurses, with studies suggesting rates between 33.0% and 69.7% [5-6]. This elevated prevalence and severity of insomnia among nurses may lead to changes in lifestyle and health risks. Evidence suggests that nurses who experience insomnia might alter their eating habits by increasing their consumption of sugary and sweetened drinks, which could elevate their body mass index and further increase their risk of obesity, diabetes, gastrointestinal disorders, and cardiovascular and oncological diseases [7]. A national survey with 19,386 full-time bed nurses in Taiwan showed that long work hours and chronic insomnia may increase the risk of needlestick and sharps injuries [8]. Furthermore, insomnia may adversely affect nurses' cognitive function, impacting decision-making, attention, and memory, which may increase the likelihood of clinical errors, endangering patient safety and care quality [9].

Notably, severe insomnia has an even greater impact on health. There is evidence that severe insomnia may be linked to more severe depression, greater functional deficits [10], reduced quality of life [11], and an increased risk of suicidal ideation [12]. Studies have shown that 11.3% to 38.3% of nurses experience severe insomnia [13-15]. However, few studies have investigated severe insomnia and its predictors among nurses, and some studies have focused on

the influencing factors of insomnia in nurses. A longitudinal study in Taiwan on newly graduated nurses found a consistent pattern of insomnia severity, which was significantly influenced by educational level and exacerbated over time by occupational stress [16]. During the COVID-19 pandemic, a study conducted in four public tertiary hospitals in China reported a high prevalence of insomnia of 52.8% among nursing staff. Key predictors of insomnia included gender, years of work experience, presence of chronic diseases, duration of midday naps, direct involvement in COVID-19 patient care, frequency of night shifts, availability of professional psychological support during the pandemic, personal experiences of severe illness or death among close contacts due to COVID-19, intense fear of the virus, fatigue, and overall perceived stress [17]. Additionally, a cross-sectional study involving 720 nurses found that factors such as work setting, weekly working hours, shift work, decision latitude, psychological job demands, and social support were differentially associated with insomnia experiences, with psychological capital significantly mediating the relationship between these occupational stressors and insomnia [18]. In addition to psychiatric symptoms such as depression and anxiety, which are well-established correlates of insomnia among nurses[19-20], a wide range of work-related and psychosocial factors have been shown to play a central role in the development and maintenance of sleep problems in this population[18, 21-23]. From an occupational health perspective, work schedule, job stressors, and career-related resources are particularly important because they are, at least in part, modifiable through organizational or managerial interventions. In the present study, we therefore adopted a workplace-centered approach and focused our predictor set on work characteristics, work-related stress, and career adaptability, rather than on diagnostic measures of depression and anxiety.

1 However, all these previously mentioned studies primarily describe correlations and cannot
2 predict the severity of insomnia, limiting their practical use in clinical settings, where
3 understanding the potential severity of insomnia could lead to more personalized and effective
4 interventions. Therefore, there is a crucial need for predictive models that can estimate insomnia
5 severity based on individual risk profiles, thereby enhancing its clinical management. Recent
6 applications of machine learning have begun to predict anxiety, depression and insomnia among
7 university students [24]. Additionally, research in China has integrated machine learning with
8 traditional Chinese medicine classifications to predict insomnia severity [25]. Despite these
9 advancements, the specific application of machine learning for predicting insomnia among
10 nurses remains limited. Machine learning algorithms are adept at analyzing complex data from
11 electronic health records and psychological evaluations, identifying subtle patterns that
12 traditional methods may overlook [26]. These capabilities are essential for developing predictive
13 models that accurately assess the likelihood of insomnia based on detailed risk profiles.

14 An earlier analysis by our group developed a predictive model in a smaller sample of 1,969
15 nurses from the same four hospitals, using hospital-based external validation. The present study
16 re-analyzes an expanded extract of the same ongoing survey ($N = 2,179$). This sample overlaps
17 with, and is therefore not independent of, the earlier one; we re-implemented the feature-
18 selection and modeling pipeline and added a data-quality screening step. Our aims were to (1)
19 develop and internally validate a parsimonious model for moderate-to-severe insomnia ($ISI \geq 15$)
20 using machine learning-assisted predictor selection; (2) quantify its optimism, calibration and
21 selection stability; and (3) test whether more flexible machine learning algorithms improve on
22 logistic regression. At most, such a model could support workplace-level risk profiling; whether
23 it does so in practice would require external validation.

Methods

Study design and objective

This study used a multicenter cross-sectional design. The primary objective was to identify predictors of moderate-to-severe insomnia in clinical nurses using machine learning methods and to develop and internally validate a predictive model. The study followed the TRIPOD+AI guidelines [27]; the completed checklist is provided in Additional file: TRIPOD+AI checklist. Predictors and outcome were measured at the same time, so the model estimates the probability of currently having moderate-to-severe insomnia given work-related and psychosocial characteristics (a cross-sectional risk-stratification model), not the probability of future incidence.

Study population

Nurses from four general hospitals located in Beijing, Shanxi, and Ningxia, representing northern, central, and western regions of China, were recruited between February and October 2023. These hospitals were purposively selected according to two considerations: (1) regional diversity, to enhance the representativeness of nurses working in different healthcare contexts; and (2) established academic collaboration with our research team, which ensured feasibility and consistency of data collection. The inclusion criteria comprised participants who (1) were aged 18 years or older, (2) were licensed nurses, and (3) volunteered to participate in this study. Nurses from these hospitals were recruited through convenience sampling and completed standardized questionnaires assessing moderate-to-severe insomnia.

Of 2,323 questionnaires initially returned, 63 were excluded because the respondent did not consent to participate, a further 49 were excluded because they originated from hospitals below tertiary level, and a further 32 were excluded for participant age under 18 years, leaving a final

analytic sample of 2,179 (an expansion of the 1,969-participant sample used in the original analysis of this survey).

Sample size adequacy was examined with the criteria of Riley et al. [28], which aim for an expected shrinkage of at least 0.9. Using the Cox–Snell R^2 of the final model in the training set (0.121) and the outcome prevalence of 29.8%, 277 participants (about 83 events) would be required for four predictor parameters, but about 2,423 participants if all 35 candidate predictor parameters (The sources, definitions, and encoding of the 16 original candidate predictors, represented by 35 feature columns for feature selection, are detailed in Supplementary Table S1.) were carried into a single unpenalized model. The training set ($n = 1,917$; 571 events; 16 events per candidate parameter) satisfies the first criterion and the conventional rule of 10 events per candidate parameter, but not the second; we therefore used penalized, cross-validated selection and explicitly quantified optimism (see Statistical analysis).

Validation design

The original analysis of this survey ($N = 1,969$) used a hospital-based external validation design, in which three hospitals served as the training set and the fourth, held out entirely, served as an external test set. In the expanded dataset used for the present re-analysis, the hospital-of-origin identifier was not retained during data consolidation, which precluded reproduction of that hospital-based external validation design. We have therefore instead used a stratified random split of the full analytic sample (88% training, $n = 1,917$; 12% test, $n = 262$; stratified on the outcome, fixed random seed) for internal validation, together with 5-fold cross-validation within the training set. A stratified random split with cross-validation is an accepted minimum standard for internal validation under the TRIPOD+AI framework, but it is a methodologically weaker

test of generalizability than true cross-institution external validation, because the training and test partitions in a random split are drawn from the same underlying hospital-mix distribution. This represents a deviation from the original study protocol, and we recommend that the hospital linkage be verified or restored in future work by the research team if genuine external validation is again required. Results in the present manuscript should accordingly be interpreted as internally, not externally, validated. The split seed was fixed at 2026; to assess sensitivity to this single split, we repeated the split 500 times (see Statistical analysis). Hospital-level clustering could not be modelled, and no geographic or temporal validation was possible because neither the hospital nor the response date was available in the analysis dataset.

Data collection

Data for this study were collected using the Wenjuanxing online platform (<https://www.wjx.cn/>). The research team first established collaboration with the Department of Nursing at each study center. Detailed explanations of the survey's objectives, and the inclusion and exclusion criteria were provided to administrators of the nursing departments at each hospital. These administrators were then tasked with sending participation invitations through WeChat software to nurses based on records of nurses' induction and attendance. Nurses who volunteered to participate in the study signed an informed consent form. To ensure the completeness of the data, all survey questions were set as mandatory; the electronic questionnaire could only be submitted once all responses were filled. Additionally, the system was configured to allow only one submission per IP address to prevent multiple entries from the same user. The design of the survey and the mandatory nature of the questions ensured that there were no missing data in this study, allowing for a comprehensive analysis without the need for imputation methods.

Assessment of moderate-to-severe insomnia

Insomnia was measured using the Insomnia Severity Index (ISI), a self-reported instrument recommended by the Chinese guidelines for the diagnosis and treatment of insomnia [29]. The ISI was developed by Morin in 1993 to assess the severity of insomnia in patients over the previous two weeks, and to reflect information about the impact of insomnia on health and daytime functioning [30]. The ISI was translated into Chinese (ISI-C) by Hong Kong scholar Doris S.F. Yu in 2010 [31]. The scale consists of 7 items and is rated on a 5-point Likert scale, with 0–4 points given for each item and a total score of 0–28 points. Insomnia severity was classified as no significant insomnia for ISI-C scores of 0–7, subthreshold insomnia state for scores of 8–14, clinical insomnia for scores of 15–21, and severe insomnia for scores of 22–28. As previous studies frequently apply an ISI score of 15 as the key cut-off distinguishing non-clinical or subthreshold symptoms from clinically meaningful, moderate to severe insomnia [32–34], the present study also adopted an ISI score ≥ 15 as the threshold in the predictive model. Because scores of 15–21 correspond to clinical insomnia of moderate severity and only scores of 22–28 to severe insomnia in the ISI scoring bands, the outcome is labelled ‘moderate-to-severe insomnia’ throughout (ISI ≥ 22 : n = 160, 7.3%; ISI 15–21: n = 489, 22.4%). Cronbach’s α for this scale in the present sample was 0.917.

Candidate predictors

Predictors associated with moderate-to-severe insomnia were identified based on expertise and previous research. Key factors included nurses’ demographic and work characteristics, career adaptability, and work stress. The same survey also contained a nursing practice environment scale (PES-NWI), the Personalized Psychological Flexibility Index (PPFI) and a transition-shock scale for newly graduated nurses (TSNGNS). These were not part of the primary candidate set,

which was restricted to work characteristics, work stress and career adaptability, and were examined in a sensitivity analysis (see Statistical analysis).

Nurses' demographic and work characteristics

The research team designed a general information questionnaire for nurses by reviewing and discussing previous literature. The questionnaire included continuous variables such as age and years of experience. Other variables were categorized as follows: gender (female or male), education (college and below, undergraduate, master's and above), marital status (unmarried, married, divorced, or widowed), and parental status (yes or no). Working sections were classified into internal medicine ward; surgical ward; obstetrics and gynecology ward; pediatric ward; ear, nose, and throat (ENT) ward; intensive care unit (ICU); outpatient clinic; operating room; emergency department; or others. Employment types were described as a permanent position, contract position, or labor dispatch position. Shift situation was defined as permanent day shift, permanent night shift, rotating day and night shifts, or non-fixed rotation. The average number of night shifts per month was categorized as 0, 1–3, 4–8, or >8. Average overtime per week was classified as ≤5 h, 6–10 h, 11–15 h, or >15 h. Average monthly income was categorized as <5000, 5000–10000, or >10000 yuan. Commuting time (round trip) was categorized as <30 min, 30–60 min, 61–120 min, or >120 min.

Challenge and Hindrance Related Self-Reported Stress (CHSS)

The CHSS was developed by Cavanaugh et al. in 2000, translated and revised by Chang, and is primarily used to evaluate the level of stress experienced by healthcare workers at work [35-36]. The scale comprises two dimensions, with six items measuring challenging stressors and five items measuring hindrance stressors. The scale is rated on a 5-point Likert scale, with a score of 1 indicating “no stress” and a score of 5 indicating “extreme stress.” Thus, higher scores indicate

greater work stress. Means $\pm$ standard deviations (SDs) were used to report results independently for challenge stress and hindrance stress. In this study, Cronbach's α for challenge stressors and hindrance stressors was 0.959 and 0.936, respectively.

Career Adapt-Abilities Scale (CAAS)

CAAS was revised by Savickas et al. in 2012 and validated in 13 countries [37]. In this study, we used the Chinese version of CAAS, which was translated by Chinese scholars Hou et al. to measure nurses' level of career adaptation [38]. The scale comprises 4 dimensions—concern (6 items), control (6 items), curiosity (6 items), and confidence (6 items)—totaling 24 items. All items were scored using a 5-point Likert scale, with each item scoring 1 to 5 points for “not strong” to “very strong,” respectively. The total score ranged from 24 to 120, with higher scores representing higher career adaptability. Cronbach's α for this scale in the present sample was 0.979.

Data quality screening

Because the survey items on the CAAS and CHSS scales are unidirectionally worded (no reverse-scored items), straight-line responding, a possible marker of inattentive or careless responding, can be detected as a “long-string” response pattern: a respondent who selects the identical response option for every item on a scale [39]. We therefore flagged, for each participant, whether all items of the CAAS scale, or all items of the CHSS scale, showed zero variance (i.e., the identical rating on every item). Because CAAS and CHSS are the two psychometric scales retained as candidate predictors in the present model, this screen was restricted to those two scales rather than applied to the wider survey battery. Participants flagged on either scale were retained in the primary (full-sample) analysis, and a sensitivity analysis was subsequently performed on the subsample of participants flagged on neither scale (the screened

subsample) to assess whether the model's associations were robust to this potential source of measurement error.

Statistical analysis

Descriptive statistics and dataset comparison

For the initial statistical description of the training and test sets, continuous variables were presented as mean $\pm$ SD and median (interquartile range, IQR). Categorical variables were expressed as frequency and percentage (%). Differences between the two datasets were analyzed using the Wilcoxon rank-sum test for continuous variables and Pearson's chi-squared test for categorical variables.

Feature selection

Feature selection was conducted using two machine learning algorithms to ensure the reliable identification of predictors of moderate-to-severe insomnia in nurses. First, the Boruta algorithm was applied to the training set for initial screening. The Boruta algorithm identifies all relevant features by iteratively comparing the importance of each predictor to randomized versions of the dataset, thus providing a comprehensive assessment of variable importance [40].

Features with importance scores significantly higher than the maximum importance score of shadow features were retained. Ordinal predictors (education, night-shift frequency, overtime, income, and commuting time) were integer-encoded to preserve their natural ordering during this step, while unordered categorical predictors were dummy-encoded. Boruta was run on the training set with a pre-specified random seed (2026; random forest with 200 trees of maximum depth 5 and balanced class weights, up to 100 iterations) and repeated with 30 further seeds to assess the stability of the classification. Although Boruta's algorithm has significant advantages

in efficient feature selection, as the algorithm is based on random forests, random-forest importance measures are known to favour correlated predictors and do not account for collinearity between features [41]. Therefore, in this study, the important features identified by Boruta were further analyzed using 10-fold cross-validated least absolute shrinkage and selection operator (Lasso) regression to screen for predictor variables of moderate-to-severe insomnia. Lasso regression, a linear regression model with L1 regularization, imposes penalties on the model's coefficients, driving some of them to zero, shrinking and selecting among correlated features [42]. The “lambda+1se” criterion was used to select the model with the minimal set of predictors that still provided good performance.

Development and validation of the predictive model

The European Commission and the World Health Organization (WHO) recognize transparency, explainability and robustness as fundamental values, particularly in the context of artificial intelligence (AI) [43-44]. To align with these principles and to facilitate clinical interpretability, we selected logistic regression as the final modeling approach because of its transparent and easily interpretable structure.

Using the predictors retained after Lasso regression, a multivariable logistic regression model was developed to predict moderate-to-severe insomnia in nurses. Prior to model building, multicollinearity among the predictors was assessed using the variance inflation factor (VIF). Predictors with a VIF greater than 10 were to be excluded; no predictor exceeded this threshold. Model performance was evaluated in two steps, as detailed under Validation design: 5-fold cross-validation within the training set, followed by evaluation in the held-out random-split test set. Discrimination was quantified by the area under the receiver operating characteristic curve

(AUC). Calibration was assessed with the calibration slope and calibration-in-the-large (intercept with the slope fixed at 1) [45], the observed-to-expected ratio, the Brier score [46] and a calibration plot with a fitted quadratic curve. Ninety-five percent confidence intervals for test-set metrics were obtained by bootstrap (2,000 resamples). Clinical utility was examined with decision curve analysis (DCA) [47, 48] and with sensitivity, specificity and predictive values at prespecified risk thresholds (0.20, 0.25, 0.30 and 0.40) and at the Youden-optimal threshold derived in the training set. A nomogram was constructed from the logistic coefficients to make the model transparent. Because odds ratios for continuous predictors depend on the measurement scale, we also report odds ratios per 10 points of CAAS and per standard deviation of each continuous predictor.

Optimism, stability and benchmarking

A single random split can give an optimistic or pessimistic impression of performance, and estimates are inflated when the selection step is left outside validation. We therefore (i) repeated the stratified 88:12 split 500 times, refitting the final four-predictor model each time, and summarized the distribution of test-set AUC, Brier score and calibration slope; (ii) estimated the optimism of the final model with Harrell's bootstrap procedure (1,000 resamples) [49]; (iii) reran the complete selection pipeline (Boruta followed by LASSO with the 1-SE rule) inside each outer fold of a repeated 5-fold cross-validation (3 repeats, 15 outer folds) within the training set, to obtain a performance estimate that includes the selection step and the selection frequency of each candidate; and (iv) compared the final model in the same test set with a random forest (minimum leaf size chosen by 5-fold cross-validation), gradient boosting and L1-penalized logistic regression, each using all candidate variables [50], and with single-domain models (night

shifts only; CHSS scales only). Paired differences in AUC were summarized with bootstrap percentile intervals.

Sensitivity analyses

To test the robustness of the findings, we (i) repeated the model with alternative ISI cut-offs (≥ 8 , ≥ 10 , ≥ 15 and ≥ 22); (ii) added the PES-NWI, PPFi and TSNGNS to the candidate set and repeated selection; (iii) tested linearity of the continuous predictors with likelihood-ratio tests against restricted cubic splines (4 degrees of freedom) and tested the interaction between the two CHSS scales and between any night-shift exposure and hindrance stress; and (iv) repeated the analysis in the screened subsample, as described above.

All statistical analyses were conducted using Python (version 3.11; pandas, statsmodels, scikit-learn and the Boruta package). Two-sided tests were employed with a significance level of $\alpha = 0.05$.

Ethics approval and consent to participate

The study was conducted according to the guidelines of the Declaration of Helsinki, and was reviewed and approved by the Medical Ethics Committee of the Peking University First Hospital (2022415002). Participants signed an informed consent form prior to the start of this study.

Results

Participant characteristics

Of 2,323 questionnaires returned, 2,179 met the inclusion criteria after excluding non-consenting respondents, respondents from non-tertiary hospitals, and respondents under 18 years of age (see Study population). A stratified random split allocated 1,917 participants to the training set and 262 to the test set. In the training set, 571 (29.8%) nurses met criteria for moderate-to-severe insomnia, closely matching the 29.8% prevalence in the test set. Statistical comparison showed no significant difference between the training and test sets for any characteristic examined (all $P > .05$; Table 1).

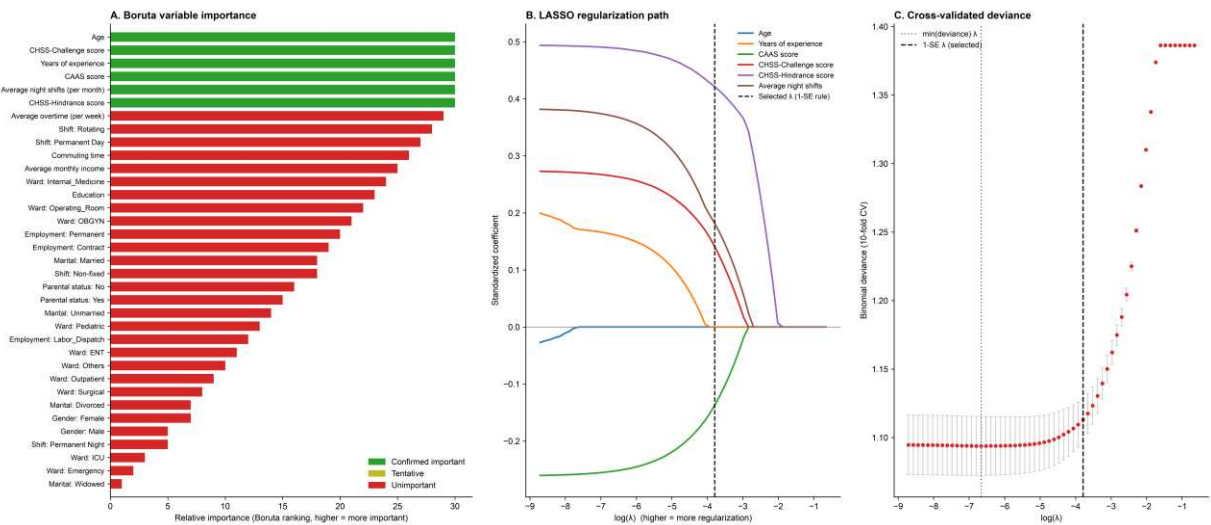
Identification of predictive factors

The Boruta algorithm and Lasso regression were employed for feature selection to address data dimensionality within the sample. With a pre-specified random seed (2026), Boruta confirmed six of the 35 candidate parameters as important (Figure 1A): age, years of experience, CAAS score, CHSS-challenge score, CHSS-hindrance score and average night shifts per month; none was classified as tentative, and all remaining candidates, including average overtime per week, commuting time, monthly income, education, working section, employment type, gender, marital status, parental status and shift situation, were classified as unimportant. Because Boruta is stochastic, it was repeated with 30 further random seeds: years of experience, CAAS, CHSS-challenge, CHSS-hindrance and night shifts were confirmed in all 30 runs, whereas age was confirmed in 17 (57%) and tentative in 13 (43%), and average overtime was confirmed in 8 (27%) (Supplementary Table S2). Marital status and parental status, which had been confirmed in the original, smaller cohort's Boruta analysis, and commuting time and shift situation, which had been tentative there, were never confirmed or tentative here, which may reflect differences in

sample composition or the corrected ordinal-versus-nominal predictor-encoding scheme used in the present analysis (see Feature selection). The variables that were consistently relevant across runs and analyses are therefore night shifts, CAAS and the two CHSS scales, whereas the classification of age and overtime is borderline.

Following Boruta's initial screening, Lasso regression was applied for further refinement (Figure 1B and 1C), confirming average night shifts per month, CHSS-challenge score, CHSS-hindrance score, and CAAS score as the four predictors retained at the 1-SE regularization strength; age and years of experience, although confirmed by Boruta, were shrunk to zero at this strength (Figure 1B). Notably, this four-predictor structure is identical to that identified in the original, smaller cohort ($N = 1,969$) that used a different, independently implemented feature-selection pipeline, consistent with the earlier finding, although the two samples overlap. In the nested analysis, in which the selection pipeline was re-run within each of 15 outer folds, these four predictors were selected in every fold, whereas years of experience (3/15 folds), overtime (2/15) and age (1/15) were selected only occasionally (Supplementary Table S2).

Figure 1. Feature selection for the moderate-to-severe insomnia predictive model (training set, N = 1,917).



(A) Boruta variable importance ranking (random seed 2026; 35 candidate parameters, six confirmed); ordinal predictors (education, night-shift frequency, overtime, income, commuting time) were integer-encoded and unordered categorical predictors were dummy-encoded prior to this step. (B) LASSO regularization path (standardized coefficients) for the six Boruta-confirmed predictors; the dashed line marks the selected regularization strength (1-SE rule). (C) Ten-fold cross-validated binomial deviance as a function of the regularization strength, with the minimum-deviance and selected (1-SE) values marked.

Development of a predictive model

A multivariable logistic regression model was developed within the training set to predict the risk of moderate-to-severe insomnia in nurses. The model incorporated four independent variables: average night shifts (per month), CAAS score, CHSS-challenge score, and CHSS-hindrane score. Table 2 presents the results. All VIFs were below 2.7, confirming the absence of material multicollinearity. Nurses working an average of 4–8 night shifts per month had an odds ratio (OR) of 1.36 (95% CI 1.05–1.76) for experiencing moderate-to-severe insomnia relative to those with no night shifts, and those working more than 8 nights per month had a higher OR of 2.37 (95% CI 1.79–3.15); the 1–3 night-shifts category did not differ significantly

from the reference (OR 1.17, 95% CI 0.77–1.80, $P = .46$). Higher scores on the CHSS-challenge (OR 1.39, 95% CI 1.12–1.73) and CHSS-hindrance (OR 1.68, 95% CI 1.38–2.04) scales were also associated with increased risk of moderate-to-severe insomnia. Conversely, higher CAAS scores were a protective factor (OR 0.983 per point, 95% CI 0.976–0.990; OR 0.84 per 10 points, 95% CI 0.78–0.91). Expressed per standard deviation, the odds ratios were 1.61 (95% CI 1.35–1.93) for CHSS-hindrance, 1.33 (95% CI 1.11–1.61) for CHSS-challenge and 0.76 (95% CI 0.68–0.85) for CAAS. The two CHSS scales were strongly correlated ($r = 0.78$), so their individual estimates should be interpreted jointly. Based on the model parameters derived from the training set, a points-based nomogram (Figure S1) was constructed for clinical application.

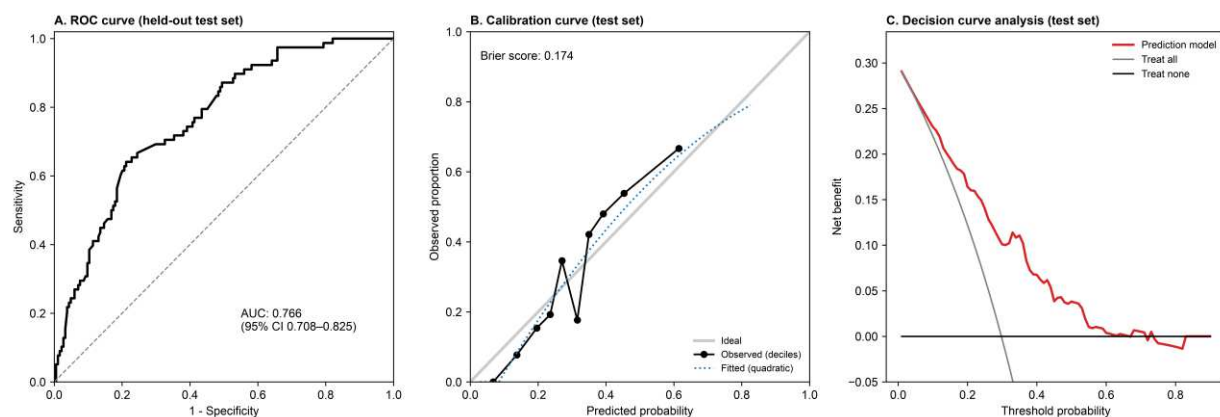
Validation of the predictive model

The performance of the predictive model was evaluated using discrimination, calibration, and DCA. The average AUC from 5-fold cross-validation within the training set was 0.709; in the held-out (random-split) test set, the model achieved an AUC of 0.766 (95% CI 0.708–0.825; Figure 2A), indicating acceptable, but not high, discrimination (see Optimism-corrected performance and stability). This test-set AUC is somewhat higher than the AUC observed in the original, hospital-based external validation of the smaller cohort (0.710); this is plausibly attributable to the larger sample size improving the precision of the estimate and, more importantly, to the fact that a random split — unlike true cross-institution external validation — draws its training and test partitions from the same underlying hospital-mix distribution, which would be expected to yield somewhat higher apparent performance than genuine external validation.

Calibration was assessed with the Brier score and calibration measures (Figure 2B). Calibration-in-the-large was close to zero (intercept -0.03 , 95% CI -0.33 to 0.24 ; observed-to-expected ratio

1 1.02, 95% CI 0.87–1.23) and the calibration slope was 1.36 (95% CI 0.96–1.94), with a Brier
2 score of 0.174 (95% CI 0.154–0.197) [46]. The wide interval around the slope reflects the small
3 test set (78 events), and the decile-level plot is noisy in the sparse high-risk range. DCA was
4 performed on the test set to evaluate net clinical benefit across risk thresholds compared with
5 treating all or none [47] (Figure 2C). At threshold probabilities of 0.10, 0.20, 0.30, 0.40 and 0.50,
6 the net benefit of the model was 0.230, 0.164, 0.101, 0.067 and 0.038, compared with 0.220,
7 0.122, –0.003, –0.170 and –0.405 for treating all; the advantage over treat-all was marginal at
8 low thresholds and clearer between 0.20 and 0.50.

9 **Figure 2. Performance of the final logistic regression model in the held-out (random-split)**
10 **test set (n = 262).**



12 (A) Receiver operating characteristic (ROC) curve, with area under the curve (AUC) and bootstrap 95% confidence interval. (B)
13 Calibration plot comparing predicted probability of moderate-to-severe insomnia (x-axis) with observed proportion (y-axis)
14 across deciles of predicted risk, with a fitted quadratic smoother; Brier score shown. (C) Decision curve analysis comparing the
15 net benefit of the prediction model with treat-all and treat-none strategies across a range of threshold probabilities.

16 Optimism-corrected performance and stability

17 The test-set AUC of 0.766 was favorable relative to other possible splits. Across 500 repeated
18 stratified random splits, the mean test AUC was 0.717 (SD 0.032; 2.5th–97.5th percentile 0.652–

0.778), the mean calibration slope 0.99 (SD 0.21) and the mean Brier score 0.182; the AUC of the reported split ranked at the 94th percentile. Harrell's bootstrap estimated an optimism of only 0.004 for the four-predictor model (apparent AUC 0.715 and optimism-corrected AUC 0.711 in the training set; 0.721 and 0.717 in the full analytic sample; optimism-corrected calibration slope 0.98). When the complete selection pipeline was re-run inside each outer fold, the mean AUC was 0.708 (SD 0.023), versus 0.710 (SD 0.023) for the fixed four predictors, indicating little additional optimism from the selection step. We therefore regard an AUC of about 0.71–0.72 as the best estimate of the model's discrimination in comparable nurses from these hospitals, and the 0.766 in the reported split as an upper-range result (Figure 3B; Supplementary Table S2).

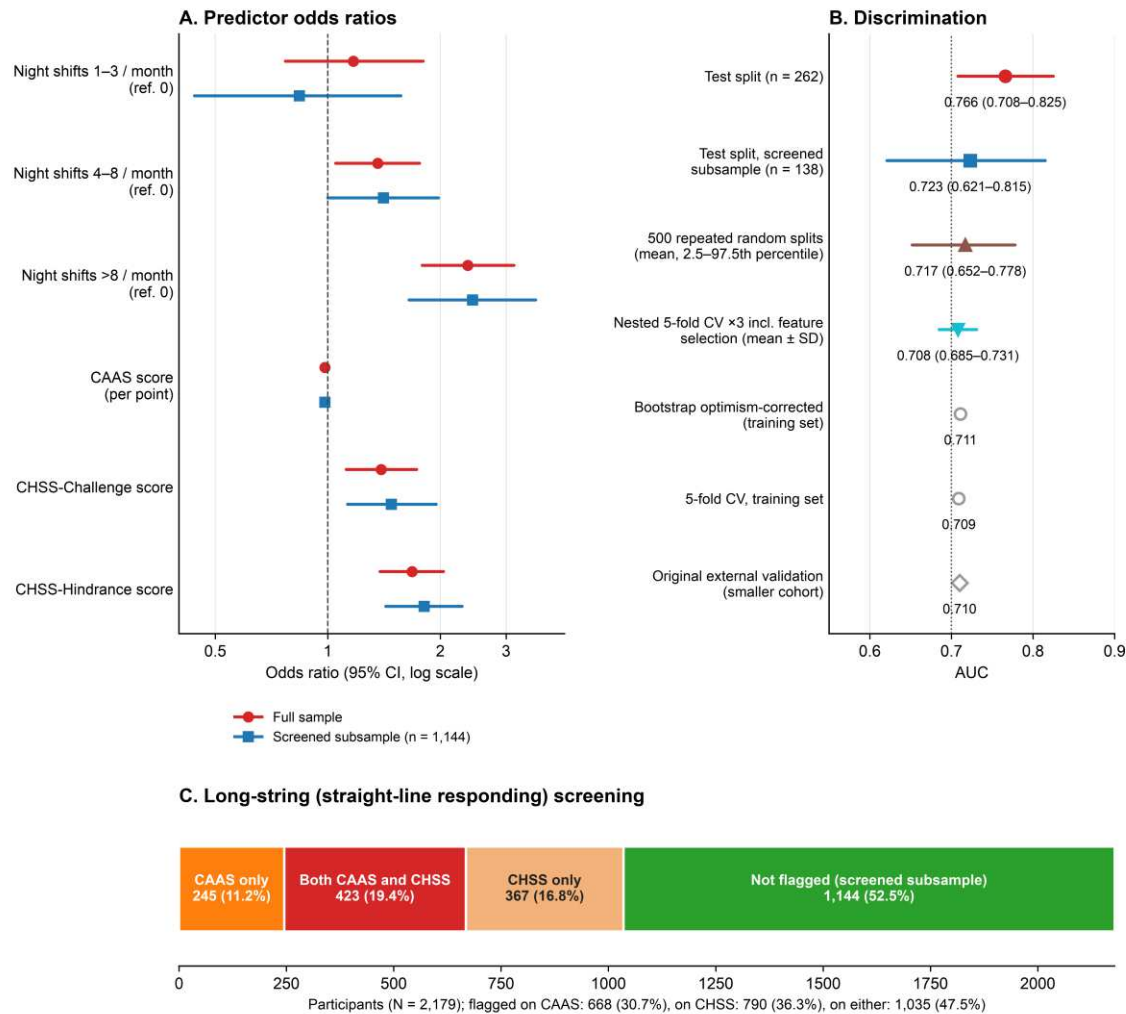
Comparison with flexible models and classification performance

Using all candidate variables, random forest, gradient boosting and L1-penalized logistic regression achieved test-set AUCs of 0.773, 0.778 and 0.766, respectively; differences from the four-predictor logistic model were small and their bootstrap intervals included zero (−0.007, 95% CI −0.038 to 0.020; −0.012, −0.042 to 0.020; 0.000, −0.020 to 0.018). A model using night shifts alone reached an AUC of 0.593, whereas the two CHSS scales alone reached 0.753 (difference from the full model 0.013, 95% CI −0.031 to 0.060), indicating that most of the discrimination derives from work stress. At a risk threshold of 0.30, the model flagged 47.7% of test-set nurses, with sensitivity 71.8% (95% CI 61.0–80.6), specificity 62.5% (55.3–69.2), positive predictive value 44.8% (36.4–53.5) and negative predictive value 83.9% (76.9–89.1); thresholds of 0.20 and 0.40 shifted this balance toward sensitivity (93.6%) or specificity (85.9%), respectively (Supplementary Table S3).

Sensitivity analysis in the screened subsample

Of the 2,179 participants, 1,035 (47.5%) were flagged for straight-line (long-string) responding on the CAAS or CHSS item batteries, leaving a screened subsample of 1,144 participants (Supplementary Table S4). Refitting the model in this screened subsample yielded odds ratios that were consistent with, or modestly stronger than, those observed in the full sample for all four predictors, and discrimination was similar (AUC 0.723, 95% CI 0.621–0.815, in a held-out subset of the screened subsample). This indicates that the associations reported above are not an artifact of inattentive responding, and if anything are attenuated rather than inflated by its inclusion. Participants flagged and not flagged for straight-line responding had a similar prevalence of moderate-to-severe insomnia (28.8% vs. 30.7%, $P = .36$). Figure 3 summarizes the predictor odds ratios in the full and screened samples (panel A), the discrimination estimates obtained with different validation approaches (panel B) and the composition of the long-string screen (panel C).

Figure 3. Model effects, discrimination and data-quality screening.



(A) Odds ratios (95% CI, log scale) for the four retained predictors in the full sample (Table 2) and in the subsample excluding participants flagged for straight-line responding ($n = 1,144$; Supplementary Table S4). Night-shift categories are relative to no night shifts; CAAS is per one-point increase. (B) Area under the ROC curve (AUC) by validation approach: the reported test split ($n = 262$) and the screened-subsample test split ($n = 138$) with 95% confidence intervals; the mean and 2.5th–97.5th percentile over 500 repeated random splits; the mean $\pm$ SD over 15 outer folds of nested cross-validation including feature selection; and point estimates for the bootstrap optimism-corrected AUC and the 5-fold cross-validated AUC in the training set and for the original hospital-based external validation of the smaller cohort. (C) Composition of the long-string screen across the CAAS and CHSS item batteries ($N = 2,179$); participants flagged on neither battery formed the screened subsample retained for the sensitivity analysis. CAAS = Career Adapt-Abilities Scale; CHSS = Challenge- and Hindrance-Related Self-Reported Stress.

Other sensitivity analyses

1 With alternative ISI cut-offs, the test-set AUC was 0.780 ($ISI \geq 8$), 0.787 (≥ 10) and 0.787 (≥ 22),
2 compared with 0.766 at ≥ 15 , and the associations of >8 night shifts and hindrance stress
3 remained positive (Supplementary Table S5). When the PES-NWI, PPFI and TSNGNS were
4 added to the candidate set, LASSO additionally retained PPFI, the practice-environment score
5 and years of experience, but adding PPFI to the four-predictor model did not improve the test-set
6 AUC (0.754 vs. 0.755 for the equivalent model with linear night-shift coding). Likelihood-ratio
7 tests suggested mild departures from linearity for CHSS-challenge ($P = .037$) and CHSS-
8 hindrance ($P = .033$) but not for CAAS ($P = .12$), and the interaction terms were not significant
9 ($P = .19$); linear terms were retained for interpretability.

10

1 **Table 1. Participant characteristics: training vs. internal validation (test) dataset**

Characteristic		Training dataset, N = 1,917	Internal test dataset, N = 262	P-value
Age				0.12b
	Mean $\pm$ SD	32 $\pm$ 7	31 $\pm$ 7	
	Median (IQR)	30 (27, 36)	30 (27, 35)	
Years of experience				0.15b
	Mean $\pm$ SD	10 $\pm$ 8	9 $\pm$ 7	
	Median (IQR)	8 (3, 14)	8 (3, 13)	
CAAS scores				0.40b
	Mean $\pm$ SD	88 $\pm$ 16	89 $\pm$ 16	
	Median (IQR)	89 (75, 96)	91 (74, 96)	
CHSS-Challenge scores				0.96b
	Mean $\pm$ SD	3.17 $\pm$ 0.86	3.18 $\pm$ 0.85	
	Median (IQR)	3.00 (3.00, 3.83)	3.00 (3.00, 3.83)	
CHSS-Hindrane scores				0.67b
	Mean $\pm$ SD	2.89 $\pm$ 0.92	2.92 $\pm$ 0.89	
	Median (IQR)	3.00 (2.40, 3.40)	3.00 (2.45, 3.40)	
Gender				0.72c
	Female	1,838 (95.9%)	253 (96.6%)	
	Male	79 (4.1%)	9 (3.4%)	
Education				0.10c

	College and below	334 (17.4%)	35 (13.4%)	
	Undergraduate	1,550 (80.9%)	219 (83.6%)	
	Master and above	33 (1.7%)	8 (3.1%)	
Marital status				0.29c
	Unmarried	623 (32.5%)	97 (37.0%)	
	Married	1,250 (65.2%)	162 (61.8%)	
	Divorced	38 (2.0%)	2 (0.8%)	
	Widowed	6 (0.3%)	1 (0.4%)	
Parental status				0.76c
	Yes	1,070 (55.8%)	143 (54.6%)	
	No	847 (44.2%)	119 (45.4%)	
Working section				
	Internal Medicine	569 (29.7%)	88 (33.6%)	
	Ward			
	Surgical Ward	286 (14.9%)	41 (15.6%)	
	Obstetrics and	206 (10.7%)	19 (7.3%)	
	Gynecology Ward			
	Pediatric Ward	227 (11.8%)	34 (13.0%)	
	ENT (Ear, Nose, and	36 (1.9%)	4 (1.5%)	
	Throat) Ward			
	Intensive Care Unit	130 (6.8%)	18 (6.9%)	
	(ICU)			
	Outpatient Clinic	148 (7.7%)	19 (7.3%)	
	Operating Room	85 (4.4%)	10 (3.8%)	
	Emergency	48 (2.5%)	7 (2.7%)	

	Department			
	Others	182 (9.5%)	22 (8.4%)	
Employment type				0.12c
	Permanent Position	602 (31.4%)	78 (29.8%)	
	Contract Position	1,115 (58.2%)	166 (63.4%)	
	Labor Dispatch	200 (10.4%)	18 (6.9%)	
	Position			
Shift situation				0.19c
	Permanent Day Shift	644 (33.6%)	81 (30.9%)	
	Permanent Night Shift	80 (4.2%)	10 (3.8%)	
	Rotating Day and Night Shifts	905 (47.2%)	118 (45.0%)	
	Non-fixed Rotation	288 (15.0%)	53 (20.2%)	
Average night shifts (per month)				0.68c
	0	683 (35.6%)	85 (32.4%)	
	1-3	146 (7.6%)	18 (6.9%)	
	4-8	686 (35.8%)	99 (37.8%)	
	>8	402 (21.0%)	60 (22.9%)	
Average overtime (per week)				0.85c
	≤5h	1,092 (57.0%)	156 (59.5%)	
	6h-10h	488 (25.5%)	65 (24.8%)	
	11h-15h	184 (9.6%)	22 (8.4%)	

	>15h	153 (8.0%)	19 (7.3%)	
Average monthly income (yuan)				0.96c
	<5000	429 (22.4%)	59 (22.5%)	
	5000-10000	1,092 (57.0%)	147 (56.1%)	
	>10000	396 (20.7%)	56 (21.4%)	
Commuting time				0.54c
	<30min	587 (30.6%)	74 (28.2%)	
	30-60min	788 (41.1%)	116 (44.3%)	
	61-120min	381 (19.9%)	55 (21.0%)	
	>120min	161 (8.4%)	17 (6.5%)	
Moderate-to-Severe Insomnia				1.00c
	Yes	571 (29.8%)	78 (29.8%)	
	No	1,346 (70.2%)	184 (70.2%)	

1 Data are n (%) unless otherwise indicated. b: Wilcoxon rank-sum test. c: Pearson's chi-squared test. CAAS = Career Adapt-
2 Abilities Scale. CHSS = Challenge- and Hindrance-Related Self-Reported Stress. Moderate-to-severe insomnia = Insomnia
3 Severity Index (ISI) ≥ 15 . Working section comparison omitted (10-level nominal variable with several small expected cell counts
4 across the two split groups; not tested).

5 **Table 2. Final multivariable logistic regression model for moderate-to-severe insomnia**
6 **(training dataset)**

Characteristic	OR	95% CI	P-value	VIF
Average night shifts (per month)				1.299

	0	—	—	
	1-3	1.17	0.77, 1.80	0.46
	4-8	1.36	1.05, 1.76	0.018
	>8	2.37	1.79, 3.15	<.001
CAAS scores		0.983	0.976, 0.990	<.001
CHSS-Challenge scores		1.39	1.12, 1.73	0.002
CHSS-Hindrance scores		1.68	1.38, 2.04	<.001

OR = odds ratio; CI = confidence interval; VIF = variance inflation factor. All VIFs < 3, indicating no material multicollinearity. Reference category for average night shifts (per month) is '0'. Per 10 points of CAAS, OR 0.84 (95% CI 0.78–0.91); per standard deviation, OR 1.61 (1.35–1.93) for CHSS-hindrance, 1.33 (1.11–1.61) for CHSS-challenge and 0.76 (0.68–0.85) for CAAS (SDs in the training set: CAAS 15.8, CHSS-challenge 0.86, CHSS-hindrance 0.92).

Discussion

In this study, we developed and internally validated a parsimonious model to estimate the probability of moderate-to-severe insomnia among nurses, using machine learning-assisted predictor selection in an expanded multicenter sample, and we quantified how much of its apparent performance can be expected to persist.

Night-shift frequency, challenge- and hindrance-related stress, and career adaptability were selected as predictors in the training-set analysis and in every outer fold of a nested cross-validation that included the selection step, and they match the four predictors identified in an earlier analysis of this survey. Because the two samples overlap and were collected by the same team with the same instruments, this agreement demonstrates the stability of the selection pipeline rather than independent replication. In total, 29.8% of nurses had moderate-to-severe insomnia ($ISI \geq 15$), exceeding the prevalence of insomnia in the general population, yet lower than figures from prior studies in mainland China (69.7%) [5], Taiwan (65.6%) [16], and Italy (65.4%) [51], which primarily focused on insomnia in general. Several factors may influence the estimation of insomnia prevalence, including the characteristics of the study population, the definition of insomnia used, and the instruments used. We used an ISI cutoff of 15, targeting moderate-to-severe cases; a higher cutoff reduces false positives but omits milder cases [32].

Discrimination was modest. The test-set AUC of 0.766 lay in the upper range of what random splitting of these data can produce (mean 0.717 across 500 splits), and optimism-corrected estimates converged around 0.71. Random forest and gradient boosting did not outperform logistic regression, in line with systematic evidence that machine learning rarely confers a benefit over logistic regression for clinical prediction models built on a few strong, well-measured predictors [50]; this supports retaining an interpretable model. Most of the

1 discrimination came from the two stress scales, which are strongly correlated ($r = 0.78$), whereas
2 night-shift frequency, despite the largest odds ratio, added little to the AUC because it separates
3 only a minority of nurses.

4 Frequency of night shifts is a predictor of moderate-to-severe insomnia among nurses. Nurses
5 working 4–8 night shifts per month had 36% higher odds of moderate-to-severe insomnia,
6 whereas those working more than 8 night shifts per month had more than double the odds (OR
7 2.37). This escalation aligns with research from Saudi Arabia, which found that nurses working
8 more than eight night shifts per month had the highest prevalence of insomnia among the sample
9 population [52]. The impact of rotating and night-shift work on sleep disorders has been well-
10 documented in various studies [53-54]. The detrimental effects of such work schedules on sleep
11 are primarily attributed to the disruption of circadian rhythms, the internal biological clocks
12 regulating sleep-wake cycles and physiological processes. This disruption leads to a
13 misalignment between an individual's internal clock and their external environment, which
14 compromises sleep quality, duration, and overall health [55]. Additionally, a study exploring the
15 relationship between shift intensity and insomnia among hospital nurses reported that the risk
16 increased significantly if participants worked three consecutive night shifts more than once a
17 month, as short recovery times from the night shift do not provide adequate opportunity for
18 nurses to recuperate and restore their natural circadian rhythms [23]. Therefore, it is imperative
19 for healthcare managers to limit the number of night shifts per month to mitigate the adverse
20 effects of disrupted circadian rhythms on nurses' sleep quality and overall health [56].

21 In this study, significant associations were found between higher CHSS scores and an increased
22 severity of insomnia among nurses. In the context of healthcare, nurses frequently encounter
23 environmental and psychological stressors such as shift work, prolonged working hours, and

high-stress conditions [57]. Work-related stress has been shown to activate the body's stress response systems, which can disrupt sleep architecture and decrease sleep quality [58-60]. These stressors can disrupt circadian rhythms and degrade sleep quality by activating the body's stress response systems, including elevated cortisol and adrenaline levels, ultimately impairing the ability to initiate and maintain sleep [61]. Previous research has linked hindrance stress to several negative outcomes, including numerous physical [62] and mental health conditions [63-64], thus increasing the risk of moderate-to-severe insomnia. Meanwhile, hindrances that block personal achievement and job satisfaction can exacerbate feelings of frustration and anxiety, further impairing sleep [65]. On the other hand, challenges at work, although potentially offering opportunities for growth, can lead to increased cognitive and physiological arousal, which interferes with the ability to initiate and maintain sleep [66]. These findings underscore the importance of interventions aimed at reducing work-related stress to mitigate its effects on sleep among healthcare professionals. Such interventions could include stress management training, organizational changes to reduce job hindrances, and support systems to help individuals cope with job challenges more effectively. Policymakers and healthcare administrators should also consider these factors when designing job roles and workflows to minimize adverse impacts on employees' sleep health and overall well-being. Because stress, adaptability and insomnia were reported in the same questionnaire at the same time, part of these associations may reflect shared method variance [67] or reverse causation, as insomnia itself can heighten perceived stress and erode perceived adaptability; the predictors should therefore be read as correlates rather than causes.

Our study found an inverse relationship between CAAS scores and the incidence of insomnia, with higher scores reflecting stronger career adaptability and a correspondingly lower likelihood

of moderate-to-severe insomnia. This relationship underscores the potential protective role of career adaptability against moderate-to-severe insomnia among nurses. Individuals with strong adaptability in their careers can learn clinical skills faster and adapt to the pace of clinical practice, thus demonstrating higher resilience [68]. In addition, career adaptability has a positive effect on resilience and plays a mediating role between the clinical learning environment and resilience among nursing interns [69]. Evidence has shown that by maintaining emotional equilibrium and physiological stability, resilience helps to mitigate the onset of insomnia [70]. Therefore, high levels of career adaptability may enable nurses to enhance their resilience, adjust more fluidly to workplace stressors, and, in turn, reduce the risk of sleep disturbances. Given the compelling association between resilience, career adaptability, and reduced insomnia among nurses, healthcare organizations should prioritize strategies that foster these qualities. The per-point odds ratio (0.983) looks small only because CAAS spans 96 points; per 10 points the odds were 16% lower (OR 0.84), and per standard deviation 24% lower (OR 0.76). Longitudinal work is needed to establish whether career adaptability protects against insomnia or whether insomnia erodes career adaptability. Moreover, most research on this topic has focused on novice nurses and interns, who are particularly vulnerable to healthcare pressures. Our findings, drawn from a broader cross-section of the nursing workforce, suggest a need to extend this research across an even wider range of nursing experiences and settings.

The findings of this study offer several considerations for nursing management. First, given that night-shift frequency showed the largest odds ratio, optimizing shift scheduling deserves attention; flexible shift systems that limit the accumulation of night shifts may be worth testing, although this study cannot show that reducing night shifts reduces insomnia. Second, the model was expressed as a nomogram (Figure S1) to make its structure transparent. Its practical value is

limited: it requires 35 questionnaire items (CAAS 24, CHSS 11) plus shift information, whereas the 7-item ISI measures the outcome directly; at a threshold of 0.30 it would flag nearly half of nurses with a positive predictive value of 45%, and its net-benefit advantage over treating everyone was marginal at low thresholds. It is therefore better seen as a summary of which modifiable work characteristics cluster with insomnia than as a screening tool, and it should not be used for individual decisions; any organizational use would need to be voluntary, non-punitive and strictly decoupled from performance evaluation. Third, regarding psychosocial management, training programs designed to mitigate work stress and enhance career adaptability merit evaluation. Since empirical evidence characterizes career adaptability as a malleable resource [71], professional development workshops focusing on cultivating these adaptive capabilities may prove advantageous, and the integration of stress management and resilience training into routine education merits consideration, as previous studies have demonstrated the efficacy of such interventions in reducing occupational stress among healthcare professionals [72, 73].

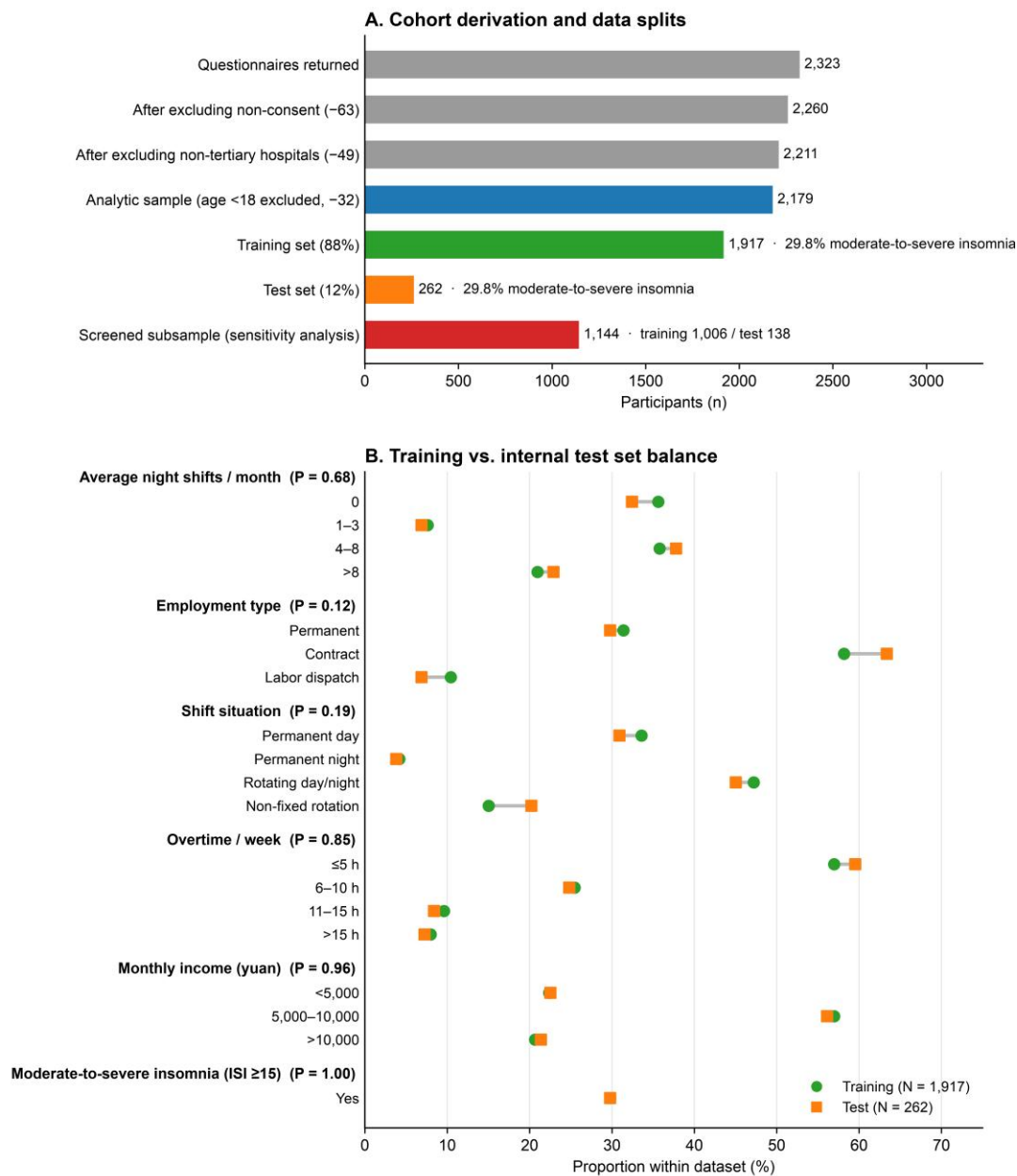
Strengths and limitations

This study has several strengths. First, it used a multicenter sample from four hospitals in three regions of China. Second, predictor selection was stable: the same four predictors were selected in all outer folds of a nested cross-validation. Third, we transparently quantified optimism (bootstrap correction, 500 repeated splits, nested cross-validation), calibration and classification performance, and compared the model with flexible machine learning algorithms. Fourth, a data-quality screen for straight-line responding was added, and the associations were consistent in the screened subsample. Finally, reporting followed TRIPOD+AI.

1 This study also has several limitations. First, the cross-sectional design, single-time-point self-
2 report and measurement of predictors and outcome in the same questionnaire preclude causal or
3 prognostic inference and expose the associations to shared method variance and reverse
4 causation. Second, only internal validation was possible: the hospital identifier was not linked in
5 the expanded dataset (Figure 4), so neither cross-hospital nor temporal validation could be
6 performed, hospital-level clustering was not modeled, and the repeated-split analysis shows that
7 a single split can overstate performance; the original hospital-based external validation of the
8 smaller cohort (AUC 0.710) remains the only cross-institution estimate. Third, the expanded
9 sample overlaps with the sample of an earlier analysis by our group, so agreement between the
10 analyses is not independent replication. Fourth, nurses were recruited by convenience sampling
11 through nursing departments, the response rate could not be determined, and the sample was
12 heavily female, which limits generalizability. Fifth, the outcome threshold of ISI ≥ 15 denotes
13 moderate-to-severe (clinically significant) insomnia, and only 7.3% of participants scored ≥ 22 ;
14 the self-reported ISI is neither a diagnostic interview nor an objective sleep measure. Sixth,
15 discrimination was modest, the test set was small (78 events), leaving wide intervals for the
16 calibration slope, and mild non-linearity of the CHSS scales was not modeled. Seventh, the
17 candidate set focused on work characteristics, work stress and career adaptability; psychological
18 flexibility and practice environment, which were collected but excluded from the primary set,
19 were additionally selected by LASSO in a sensitivity analysis, and depression, anxiety and
20 lifestyle factors were not measured. Finally, 47.5% of participants gave identical responses to all
21 items of at least one predictor scale; although the results were consistent after their exclusion,
22 such responding may still introduce measurement error.

23

Figure 4. Cohort derivation and training–test balance.



(A) Stepwise derivation of the analytic sample (2,323 questionnaires returned; 63 non-consenting, 49 non-tertiary-hospital and 32 under-18 respondents excluded; $N = 2,179$), the stratified random split into training ($n = 1,917$) and test ($n = 262$) sets with identical moderate-to-severe insomnia prevalence (29.8%), and the screened subsample used for the sensitivity analysis ($n = 1,144$; training $n = 1,006$, test $n = 138$). (B) Proportion of participants in each category of selected work-related characteristics in the training (circles) and test (squares) sets; P -values are from Table 1. Moderate-to-severe insomnia = Insomnia Severity Index ≥ 15 .

Conclusions

Night-shift frequency, challenge- and hindrance-related work stress and lower career adaptability were consistently associated with moderate-to-severe insomnia among nurses. A four-predictor logistic model showed modest discrimination (AUC about 0.71–0.72 after accounting for optimism; 0.766 in the reported test split) and adequate average calibration, and more complex machine learning algorithms offered no advantage. Given the cross-sectional, self-reported design and internal validation only, the model and its nomogram are exploratory and should not be used for individual-level decisions. Priorities for future work are external validation in independent hospitals and time periods with hospital identifiers retained, prospective designs with objective sleep measures, and tests of whether shift and stress-reduction interventions reduce insomnia.

1 **Declarations**

2 **Ethics approval and consent to participate**

3 The study was conducted according to the guidelines of the Declaration of Helsinki, and was
4 reviewed and approved by the Medical Ethics Committee of the Peking University First Hospital
5 (2022415002). Participants signed an informed consent form prior to the start of this study.

6 **Consent for publication**

7 Not applicable.

8 **Clinical trial number**

9 Not applicable.

10 **Availability of data and materials**

11 The datasets analyzed in the present study are available from the corresponding author upon
12 reasonable request.

13 **Competing interests**

14 The authors report no competing interests.

15 **Funding**

16 This research received no specific grant from any funding agency in the public, commercial, or
17 not-for-profit sectors.

18 **Authors' contributions**

19 YZ ,XZ and DL conceived and designed the study. YZ, ZL, LC, MZand LS contributed to
20 material preparation and data collection. XYZ, XZ and DL performed the data analysis and
21 interpreted the results. YZ drafted the initial manuscript. ZL, HXS and YM critically revised the

1 manuscript for important intellectual content. All authors reviewed and commented on previous
2 versions of the manuscript, read and approved the final manuscript.

3 **Acknowledgements**

4 We gratefully acknowledge all nurses that participated in this study, and the nursing
5 administrators from the hospitals for supporting this investigation.

6

References

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. American Psychiatric Association; 2013.
2. Zeng LN, Zong QQ, Yang Y, et al. Gender difference in the prevalence of insomnia: a meta-analysis of observational studies. *Front Psychiatry*. 2020;11:577429.
3. Wu TT, Zou YL, Xu KD, et al. Insomnia and multiple health outcomes: umbrella review of meta-analyses of prospective cohort studies. *Public Health*. 2023;215:66-74.
4. Zhan YX, Zhang F, Lu LH, et al. Prevalence of dyslipidemia and its association with insomnia in a community-based population in China. *BMC Public Health*. 2014;14:1050.
5. An FR, Qi YK, Zeng JY, et al. The prevalence of insomnia, its demographic correlates, and treatment in nurses working in Chinese psychiatric and general hospitals. *Perspect Psychiatr Care*. 2016;52:88-94.
6. Kowalczyk K, Krajewska-Kułak E, Sobolewski M. Relationships between sleep problems and stress coping strategies adopted by nurses including socio-occupational factors. *Front Psychiatry*. 2021;12:660776.
7. Vitale E. Correlations between insomnia and sex, work experience, shift and body mass index in Italian nurses: a scoping profile statement. *Endocr Metab Immune Disord Drug Targets*. 2022;22:1303-12.
8. Lo WY, Chiou ST, Huang N, et al. Long work hours and chronic insomnia are associated with needlestick and sharps injuries among hospital nurses in Taiwan: a national survey. *Int J Nurs Stud*. 2016;64:130-6.
9. Dong H, Zhang Q, Sun Z, et al. Sleep disturbances among Chinese clinical nurses in general hospitals and its influencing factors. *BMC Psychiatry*. 2017;17:241.
10. O'Brien EM, Chelminski I, Young D, et al. Moderate-to-severe insomnia is associated with more severe presentation and greater functional deficits in depression. *J Psychiatr Res*. 2011;45:1101-5.
11. Hajak GO, Study Group. Epidemiology of moderate-to-severe insomnia and its consequences in Germany. *Eur Arch Psychiatry Clin Neurosci*. 2001;251:49-56.
12. Soberay KA, Hanson JE, Dwyer M, et al. The relationship between suicidal responses and traumatic brain injury and moderate-to-severe insomnia in active duty, veteran, and civilian populations. *Arch Suicide Res*. 2019;23:391-410.
13. Cai Z, Cui Q, Liu Z, et al. Nurses endured high risks of psychological problems under the epidemic of COVID-19 in a longitudinal study in Wuhan China. *J Psychiatr Res*. 2020;131:132-7.
14. Vitale E, Mea R, Di Dio F, et al. Anxiety, insomnia and body mass index scores in Italian nurses engaged in the care of COVID-19 patients. *Endocr Metab Immune Disord Drug Targets*. 2021;21:1604-12.
15. Zhan N, Xu Y, Pu J, et al. The interaction between mental resilience and insomnia disorder on negative emotions in nurses in Guangdong Province, China. *Front Psychiatry*. 2024;15:1396417.
16. Huang CY, Chen CJ, Lee YF, et al. Effects of individual characteristics on insomnia severity trajectory among nurses: a prospective longitudinal study. *J Nurs Manag*. 2019;27:1640-7.
17. Zhan Y, Liu Y, Liu H, et al. Factors associated with insomnia among Chinese front-line nurses fighting against COVID-19 in Wuhan: a cross-sectional survey. *J Nurs Manag*. 2020;28:1525-35.
18. Wang MF, Shao P, Wu C, et al. The relationship between occupational stressors and insomnia in hospital nurses: the mediating role of psychological capital. *Front Psychol*. 2023;13:1070809.
19. Pachi A, et al. Stress, anxiety and depressive symptoms, burnout and insomnia among Greek nurses one year after the end of the pandemic: a moderated chain mediation model. *J Clin Med*. 2025;14:1145.
20. Bennaroch K, Shochat T. Psychobiological risk factors for insomnia and depressed mood among hospital female nurses working shifts. *Front Sleep*. 2023;2:1206101.
21. Deng X, Liu X, Fang R, et al. Evaluation of the correlation between job stress and sleep quality in community nurses. *Medicine (Baltimore)*. 2020;99:e18822.
22. Chen X, et al. Prevalence and associated factors of insomnia symptoms among doctors and nurses under the context of high prevalence of multiple infectious diseases: a cross-sectional study. *Front Public Health*. 2024;12:1423216.

23. Chung Y, Kim H, Koh DH, et al. Relationship between shift intensity and insomnia among hospital nurses in Korea: a cross-sectional study. *J Prev Med Public Health*. 2021;54:46-54.
24. Chowdhury AH, Rad D, Rahman MS. Predicting anxiety, depression, and insomnia among Bangladeshi university students using tree-based machine learning models. *Health Sci Rep*. 2024;7:e2037.
25. Li S, Zhu P, Cai G, et al. Application of machine learning models in predicting insomnia severity: an integrative approach with constitution of traditional Chinese medicine. *Front Med (Lausanne)*. 2023;10:1292761.
26. Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. *Science*. 2015;349:255-60.
27. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*. 2024;385:e078378.
28. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. *BMJ*. 2020;368:m441.
29. China Sleep Research Society. Guidelines for the diagnosis and treatment of insomnia in China. *Natl Med J China*. 2017;97:1844-56.
30. Morin CM. *Insomnia: psychological assessment and management*. Guilford Press; 1993.
31. Yu DS. Insomnia Severity Index: psychometric properties with Chinese community-dwelling older people. *J Adv Nurs*. 2010;66:2350-9.
32. Morin CM, Belleville G, Bédaride L, et al. The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep*. 2011;34:601-8.
33. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Med*. 2001;2:297-307.
34. Takami M, et al. Quality of life, depression, and productivity of city government employees in Japan: a comparison study using the Athens insomnia scale and insomnia severity index. *Sleep Sci Pract*. 2018;2:4.
35. Cavanaugh MA, Boswell WR, Roehling MV, et al. An empirical examination of self-reported work stress among US managers. *J Appl Psychol*. 2000;85:65-74.
36. Chang SY. *Research on the current status and impact of challenge-hindrance stress among healthcare professionals*. Central China Normal University; 2012.
37. Savickas ML, Porfeli EJ. Career Adapt-Abilities Scale: construction, reliability, and measurement equivalence across 13 countries. *J Vocat Behav*. 2012;80:661-73.
38. Hou ZJ, Leung SA, Li XX, et al. Career adapt-abilities scale-China form: construction and initial validation. *J Vocat Behav*. 2012;80:686-91.
39. Meade AW, Craig SB. Identifying careless responses in survey data. *Psychol Methods*. 2012;17:437-55.
40. Kursu MB, Rudnicki WR. Feature selection with the Boruta package. *J Stat Softw*. 2010;36:1-13.
41. Strobl C, Boulesteix AL, Kneib T, et al. Conditional variable importance for random forests. *BMC Bioinformatics*. 2008;9:307.
42. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Series B Stat Methodol*. 1996;58:267-88.
43. Panigutti C, Hamon R, Hupont I, et al. The role of explainable AI in the context of the AI Act. *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. 2023.
44. Goodman KW, Litewka SG, Malpani R, et al. Global health and big data: the WHO's artificial intelligence guidance. *S Afr J Sci*. 2023;119:1-4.
45. Van Calster B, McLernon DJ, van Smeden M, et al. Calibration: the Achilles heel of predictive analytics. *BMC Med*. 2019;17:230.
46. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21:128-38.
47. Vickers AJ, Van Calster B, Wynants L, et al. Decision curve analysis: confidence intervals and hypothesis testing for net benefit. *Diagn Progn Res*. 2023;7:11.

48. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006;26:565-74.
49. Steyerberg EW, Harrell FE Jr, Borsboom GJ, et al. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol*. 2001;54:774-81.
50. Christodoulou E, Ma J, Collins GS, et al. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol*. 2019;110:12-22.
51. Ielapi N, Andreucci M, Bracale UM, et al. Insomnia prevalence among Italian night-shift nurses. *Nurs Rep*. 2021;11:530-5.
52. Alqahtani JS, Arowosegbe A, Oyelade T, et al. The effect of cumulative night shift duties on insomnia, fatigue, and mental health in intensive care unit. *Heliyon*. 2024;10:e31066.
53. Cheng WJ, Cheng Y. Night shift and rotating shift in association with sleep problems, burnout and minor mental disorder in male and female employees. *Occup Environ Med*. 2017;74:483-8.
54. Fink AM. Shift work sleep disorder and mental health: an integrative review of neurobiological, sociological, and psychological perspectives with public policy implications. *Policy Polit Nurs Pract*. 2024;25:94-102.
55. Meyer N, Harvey AG, Lockley SW, et al. Circadian rhythms and disorders of the timing of sleep. *Lancet*. 2022;400:1061-78.
56. Knutsson A. Methodological aspects of shift-work research. *Chronobiol Int*. 2004;21:1037-47.
57. Gong Z, Li T. Relationship between feedback environment established by mentor and nurses' career adaptability: a cross-sectional study. *J Nurs Manag*. 2019;27:1568-75.
58. Vgontzas AN, Tsigos C, Bixler EO, et al. Chronic insomnia and activity of the stress system: a preliminary study. *J Psychosom Res*. 1998;45:21-31.
59. Meerlo P, Sgoifo A, Suchecki D. Restricted and disrupted sleep: effects on autonomic function, neuroendocrine stress systems and stress responsivity. *Sleep Med Rev*. 2008;12:197-210.
60. French KA, Allen TD, Henderson TG. Challenge and hindrance stressors in relation to sleep. *Soc Sci Med*. 2019;222:145-53.
61. Lo Martire V, Caruso D, Palagini L, et al. Stress and sleep: a relationship lasting a lifetime. *Neurosci Biobehav Rev*. 2020;117:65-77.
62. Vgontzas AN, Liao D, Pejovic S, et al. Insomnia with objective short sleep duration is associated with type 2 diabetes: a population-based study. *Diabetes Care*. 2009;32:1980-5.
63. Hsu CY, Chen YT, Chen MH, et al. The association between insomnia and increased future cardiovascular events: a nationwide population-based study. *Psychosom Med*. 2015;77:743-51.
64. Baglioni C, Battagliese G, Feige B, et al. Insomnia as a predictor of depression: a meta-analytic evaluation of longitudinal epidemiological studies. *J Affect Disord*. 2011;135:10-9.
65. Sonnentag S, Fritz C. Recovery from job stress: the stressor-detachment model as an integrative framework. *J Organ Behav*. 2015;36:S72-S103.
66. Cao Q, Wu H, Tang X, et al. Effect of occupational stress and resilience on insomnia among nurses during COVID-19 in China: a structural equation modelling analysis. *BMJ Open*. 2024;14:e080058.
67. Podsakoff PM, MacKenzie SB, Lee JY, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. *J Appl Psychol*. 2003;88:879-903.
68. Zhang H, Jiang JX, Zhong MH, et al. Career adaptability of newly graduated nurses at an obstetrics and gynaecology hospital in China: a qualitative study. *J Nurs Manag*. 2022;30:2046-53.
69. Xu Y, Zhang W, Wang J, et al. The effects of clinical learning environment and career adaptability on resilience: a mediating analysis based on a survey of nursing interns. *J Adv Nurs*. 2026;82:2903-12.
70. Gil-Almagro F, Carmona-Monge FJ, García-Hedraera FJ, et al. Post-pandemic insomnia in healthcare workers: a prospective study including sociodemographic, occupational and psychosocial variables. *J Clin Med*. 2024;13:3498.
71. Koen J, Klehe UC, Van Vianen AE. Training career adaptability to facilitate a successful school-to-work transition. *J Vocat Behav*. 2012;81:395-408.

- 1 72. Ruotsalainen JH, et al. Preventing occupational stress in healthcare workers. Cochrane Database Syst Rev.
- 2 2014;2014:CD002892.
- 3 73. Cleary M, Kornhaber R, Thapa DK, et al. The effectiveness of interventions to improve resilience among health
- 4 professionals: a systematic review. Nurse Educ Today. 2018;71:247-63.
- 5

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterial.docx](#)