Journal of Anxiety Disorders 114 (2025) 103052

Contents lists available at ScienceDirect

Journal of Anxiety Disorders

journal homepage: www.elsevier.com/locate/janxdis

# Sleep disturbances predict nine-year panic disorder chronicity: The sleep-panic nexus theory with machine learning insights

Nur Hani Zainal [a,*], Natalia Van Doren [b]

[a] National University of Singapore, Department of Psychology, Kent Ridge Campus, Singapore
[b] University of California at San Francisco, Department of Psychiatry and Behavioral Sciences, San Francisco, CA, USA

ARTICLE INFO

*Keywords:*
Longitudinal
Machine learning
Panic disorder
Subjective sleep disturbances
Sleep actigraphy

ABSTRACT

*Background:* Panic disorder (PD) is a chronic and impairing anxiety disorder. Individuals with more sleep disturbances might be predisposed to nine-year PD chronicity. However, linearity assumptions, small predictor sets, and analytic and design limitations have hindered optimal identification of which sleep disturbance variables are distal risk factors for PD chronicity. We thus used machine learning (ML) to predict nine-year PD chronicity using high-dimensional data.

*Method:* Community-dwelling adults ($N = 1054$) completed clinical interviews, self-reports, and seven-day sleep actigraphy at Wave 1 (W1) and the same clinical interview at Wave 2 (W2) nine years later. The baseline data comprised 43 actigraphy, self-reported sleep disturbances, clinical, and demographic variables. Seven ML models were examined. Gradient boosting machine (GBM) was the best-performing algorithm. PD chronicity was defined as the presence of a PD diagnosis at both W1 and W2.

*Results:* The GBM accurately predicted PD chronicity (area under the receiver operating characteristic curve [AUC] =.764). Shapley additive explanation analysis showed that the top W1 predictors of PD chronicity were comorbid major depressive disorder, low healthcare utilization, sleep medication use, lengthier wake after sleep onset, and sleep-wake circadian disruptions based on actigraphy and self-reports. Lower household income and younger age were also top predictors. Additionally, the final multivariate model was well-calibrated.

*Conclusions:* As proposed in our *sleep-panic nexus theory*, actigraphy and subjective sleep disturbances have essential prognostic value in predicting long-term PD chronicity. Harnessing ML facilitates accurate prediction by identifying complex, nonlinear relations across high-dimensional datasets, possibly improving prevention and treatment tailoring.

## 1. Introduction

Panic disorder (PD) is a common anxiety disorder typified by repeated panic attacks, which are extreme fear surges that occur without notice and are paired with arousal sensations, such as accelerated heart rate and hot flashes (de Jonge et al., 2016). Worries about the prospect of panic attacks occurring unexpectedly tend to aggravate PD symptoms and worsen avoidance behaviors long-term (Hovenkamp-Hermelink et al., 2021). Further, greater PD severity has been related to greater comorbid major depressive disorder (MDD) and generalized anxiety disorder (GAD) symptoms (Olfson et al., 2017). Persistently high PD severity has been reliably linked to poorer quality of life and socio-occupational functioning, especially with agoraphobia present (Carta et al., 2015). Frequent panic attacks have also been correlated with more lifetime alcohol use disorder (AUD) and substance use disorder (SUD) symptoms through self-medication processes that further impair role functioning (Blakey et al., 2022; Zvolensky et al., 2006). Identifying distal risk factors (i.e., predictor variables of outcomes measured in the long term) of PD chronicity is thus critical for informing optimal prevention and treatment strategies.

Sleep disturbances could be an essential distal risk factor for the chronicity of long-term PD. Cognitively, worries about uncontrollable factors might lengthen sleep onset latency (SOL), decreasing the threshold and tolerance for panic attacks over time (Alfano, 2018). Persistent low sleep quality would fuel catastrophic thinking due to the known impact of sleep deficits on cognitive functioning, which, in turn,

* Correspondence to: National University of Singapore (NUS) Department of Psychology Faculty of Arts and Social Sciences Block AS4, Office #03-25, 9 Arts Link Singapore, 117570, Singapore.
*E-mail address:* hanizainal@nus.edu.sg (N.H. Zainal).

https://doi.org/10.1016/j.janxdis.2025.103052
Received 28 April 2025; Received in revised form 16 July 2025; Accepted 16 July 2025
Available online 17 July 2025

increases anxiety sensitivity (fear of arousal sensations; Jemcov et al., 2023). Moreover, daytime fatigue, which could lead to more napping (e.g., rest or sleep bouts during active wake phases), may disrupt circadian sleep-wake cycles over extended periods (McCallum et al., 2019). Longer wake after sleep onset (WASO) might predict future PD chronicity by fragmenting regular sleep architecture, raising somatic arousal, and predisposing one to increased anxiety sensitivity and panic attacks (Sheikh et al., 2003). Plausibly, these processes dovetail with cumulative wear-and-tear of the hypothalamic-pituitary-adrenal (HPA) axis via the buildup of allostatic load, which triggers and prolongs PD chronicity (Santamaria-Garcia et al., 2025).

Some support for these theoretical pathways exists, albeit across relatively short timescales. Sleep disturbances, as indexed by more napping or sleep bouts during active wake periods or lower total sleep time (TST), predicted PD and related anxiety symptoms several months or years later in community women (Jackson et al., 2014), primary care patients (Marcks et al., 2010), trauma-exposed youths (Geng et al., 2018), and youths with autism (Uren et al., 2019). Moreover, chronic sleep disturbances in childhood led to more anxiety disorder symptoms in adulthood (Gregory et al., 2005), plausibly indicating heterotypic continuity (i.e., discontinuity in symptom expression across time but with stable underlying susceptibility) (Speranza et al., 2023). Similarly, across two community-dwelling samples, lengthier WASO, SOL, and related problems predicted incident and recurrent anxiety disorders, including PD, one to three years later (Barbotin et al., 2022; Hysing et al., 2025). From a developmental psychopathology perspective, sleep disturbances might heighten predisposition to long-term chronicity of PD. However, no studies have examined these research questions over longer periods (i.e., beyond three years), and they have primarily focused on the developmental periods of children, adolescents, and young adulthood. This limitation leaves open the question of how earlier onset sleep problems might lead to PD chronicity over decades in midlife and older adult populations.

Based on prior theory and research, we aimed to identify the specific sleep disturbance indices that serve as distal risk factors for the chronicity of PD. We extended earlier work on this topic in the following ways. First, most research on the relationship between PD chronicity or symptoms and sleep disturbances has been cross-sectional (Belleville & Potočnik, 2019), preventing causal conclusions (Pearl, 2014). Our nine-year study addresses this shortcoming, enabling us to identify distal risk factors. Second, most research assumed linear relations between sleep disturbances and PD symptoms. However, their associations might be complex, nonlinear, and conditional on one another (Malyszczak & Janocha, 2022; Schoenberg, 2020) as PD symptoms and sleep disturbances progress over distinct developmental periods. For example, a young adult with initially minor sleep disturbances may show little to no increase in PD symptoms. However, as sleep problems worsen and persist into middle adulthood, there could be a sudden, marked escalation in PD chronicity, reflecting a nonlinear developmental trajectory. Third, using actigraphy indices regularly collected during sleep studies to predict nine-year PD chronicity could aid in the development of clinical prediction tools (Liu et al., 2015). Such tools could help inform preventive interventions to help stall or delay the progression of worsening PD symptoms.

Thus, we harnessed supervised machine learning (ML) approaches that could capture interactions and nonlinearities rather than employing linear modeling techniques, such as the standard ordinary least squares (OLS) regression. Briefly, ML represents data-driven modeling techniques that capture complex interactions and nonlinearities to generalize such patterns to unseen data (Van Lissa, 2022). To prevent the lack of generalizability of learned patterns to new data, also known as overfitting, relevant theories informed the sleep disturbance variables in our predictor set (Yarkoni & Westfall, 2017). Furthermore, by accounting for plausible, complex nonlinearities and interactions, ML methods provide a more informed basis for developing a 'prognostic calculator' to predict long-term PD chronicity, if externally validated (Collins et al., 2024). This prognostic calculator might be instrumental for prevention and treatment (Tian et al., 2024). Third, we used a high-dimensional predictor set comprising more variables than prior studies. ML techniques are well-suited to handle high-dimensional datasets compared to their OLS regression counterparts in optimizing the bias-variance trade-off (Orru et al., 2019). Fourth, sleep disturbances are typically measured using trait-level self-reports or objective polysomnography measures assessed in the laboratory (Belleville & Potočnik, 2019). We expanded on previous reports by using objective seven-day sleep actigraphy, which involves passive wearable markers embedded into daily life that capture activity in both resting and sleep phases, as well as rest and sleep bouts in active wake phases. This endeavor is essential as sleep actigraphy might offer more information than self-reported sleep disturbances and those captured by polysomnography and subjective approaches (Pastre & Lopez-Castroman, 2022).

The present study's hypotheses were twofold. First, we hypothesized that our multivariate model with 43 sleep actigraphy and self-reported variables of sleep disturbances would accurately predict nine-year PD chronicity. This multivariate predictive accuracy was denoted by an area under the receiver operating characteristic curve (AUC) of $\geq .70$ (Gonzalez, 2021). This step was essential, as accurate multivariate prediction was a prerequisite for identifying the key sleep disturbance predictors of long-term PD chronicity. Second, we expected the final model, which included the top 20 predictors of PD chronicity across nine years, to indicate actigraphy and self-reported indices of sleep disturbances consistently. Specifically, we anticipated that sleep disturbance factors (e.g., longer WASO, SOL) and lower sleep efficiency would precede more nine-year PD symptoms, even after adjusting for various clinical and demographic covariates.

## 2. Method

### 2.1. Eligibility criteria

Community-dwelling adult participants ($N = 1054$) from the Midlife Development in the United States (MIDUS) project who offered informed consent to undergo structured psychiatric interviews, self-report survey, and sleep actigraphy protocols at Wave 1 (W1; 2004–2006; Ryff et al., 2021; Ryff et al., 2019b). They completed the same structured psychiatric interview at Wave 2 (W2; 2013–2014; Ryff et al., 2019a). Participants who did not complete the sleep actigraphy and self-report protocols, and thus did not provide sufficient data for the present research aims, were excluded from the current study.

### 2.2. Participants

Regarding demographics, the mean age at W1 was 55.35 ($SD$ = 11.80, range = 34.00–84.00). Although 577 (54.7 %) were women, the remaining 477 (45.3 %) were men. Education levels included 215 (20.4 %) with formal college, university, or post-graduate education, 300 (28.5 %) with some college, 238 (22.6 %) with high school diploma, and 301 (28.6 %) with no high school education or declined to disclose. Concerning racial identity, 990 (93.9 %) were White, and the remaining 64 (6.1 %) identified as Asian, African American, Pacific Islander/Native American, or others.

### 2.3. Procedures

#### 2.3.1. Overview

The MIDUS staff administered the Composite International Diagnostic Interview-Short Form (CIDI-SF; Kessler et al., 1998) at W1 and W2 to assess diagnoses of PD, MDD, and GAD, as detailed in the next section. Participants also completed self-reported surveys on demographics, AUD (Selzer, 1971), SUD (Turiano et al., 2012) symptoms, and Pittsburgh Sleep Quality Index (PSQI; Buysse et al., 1989) at W1. The

seven-day sleep actigraphy measures were also administered at W1 (Love et al., 2010). Only participants who completed all key procedures were retained to ensure alignment with the study's aims and hypotheses.

**Sleep actigraphy protocol.** Participants wore sleep actigraphy devices on their wrists while completing a daily sleep diary for seven consecutive nights and days (Aqua et al., 2024). Sleeping and wake-time activities were recorded in 30-second intervals before the collected data were analyzed using the Actiware software (Koninklijke Philips, NV, USA). Resting phases were identified through daily diary responses (Bhat et al., 2024). If such data were missing, successive data or event-based indices provided signals to mark resting phases. When summary statistics were computed, wake-time activity had a threshold of 40 as preset by Actiware.

Various sleep disturbance indices were captured. Sleep efficiency (the proportion of time in bed spent asleep or TST ÷ time in bed) and WASO (the duration of wakefulness following sleep onset) represent distinct measures of sleep continuity versus fragmentation (Shrivastava et al., 2014). TST was operationalized as the total minutes of sleep reported in the daily diaries (Bei et al., 2017). Relatedly, wake time was self-recorded. Activity counts, rest bouts, and sleep bouts were also captured during resting, sleep, and wake-time phases using actigraphy wearables (Teas & Friedman, 2021).

### 2.4. Measures

#### 2.4.1. W1 and W2 PD diagnosis

PD diagnoses in the past 12 months during W1 and W2 were assessed using the CIDI-SF (Kessler et al., 1998), which is congruent with the Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised (DSM-III-R; American Psychiatric Association, 1987). Symptoms included in the diagnostic criteria were (i) heart pounds; (ii) chest or stomach discomfort, pain, or tightness; (iii) sweating; (iv) shaking or trembling; (v) chills or hot flashes; and (vi) sense of unreality ('*present*' [1] or '*absent*' [0]). Prior research alluded to good construct validity of the scale scores (Zainal et al., 2024). PD chronicity outcome was defined as those with a PD diagnosis at both W1 and W2.

#### 2.4.2. W1 MDD diagnosis

Given prior literature (Wang et al., 2023), W1 MDD diagnosis was controlled for in the analysis. MDD symptoms in the past 12 months were also measured with the CIDI-SF (Kessler & Üstün, 2004). Participants responded to whether they experienced any of these symptoms most of the day or almost daily for at least two weeks: (i) feeling blue, depressed, or sad; (ii) loss of interest; (iii) atypically fatigued; (iv) appetite decreases or increases; (v) trouble falling or staying asleep; (vi) concentration difficulties; (vii) feelings of self-worthlessness; and (viii) suicide ideation ('*present*' [1] or '*absent*' [0]). Previous studies also suggested that this scale had strong construct validity (Zainal & Newman, 2022; Zainal et al., 2024).

#### 2.4.3. W1 GAD diagnosis

W1 GAD diagnosis was also adjusted based on past research (Locke et al., 2015). GAD symptoms in the past 12 months were similarly evaluated using the CIDI-SF (Kessler & Üstün, 2004). Participants reported the degree (1 = *never* to 4 = *most days*) to which they encountered these symptoms due to their worries: (i) restlessness; (ii) feeling keyed up, on edge, or nervous; (iii) irritability; (iv) difficulty falling asleep; (v) difficulty staying asleep; (vi) concentration difficulties; (vii) memory difficulties; (viii) fatigue; (ix) quickly tired; and (x) muscle aches and sores. Internal consistency was high (α = .98). Recent psychometric analyses demonstrated that the scale has good convergent validity and strong discriminant validity (Ng et al., 2024).

#### 2.4.4. W1 AUD and SUD symptoms

As PD symptoms have sometimes been found to co-occur with AUD and SUD symptoms (Barrett et al., 2021; Blakey et al., 2022), our analysis adjusted for these confounds. The Alcoholism Screening Test was used to measure AUD symptom presence in the past 12 months (Selzer, 1971). Participants endorsed if they experienced any of these symptoms: (i) high risk of injury (e.g., motor vehicle accident) under the effects of alcohol; (ii) psychological issues due to alcohol use (e.g., depression, suspicion, reality disturbances); (iii) a strong desire or urge to drink, which felt irresistible; (iv) at least a month of high alcohol use with side effects; and (v) need to drink more than before to attain the same effects. Additionally, participants endorsed if they encountered these SUD symptoms in the last 12 months from using any drugs (marijuana/hashish; nerve pills/prescription painkillers/sedatives; cocaine/crack/stimulants; LSD/other hallucinogens; heroin; Turiano et al., 2012): (i) consuming greater amounts than planned; (ii) psychosocial impairments; (iii) heightened risk of injury; (iv) emotional difficulties; (v) strong desire or urge that feels irresistible; (vi) time inefficiencies; and (vii) higher tolerance prompting more use. Internal consistency values were acceptable for the sum score of the AUD scale (α = .70; range: 0–5) and SUD scale (α = .81; range: 0–18). Prior research implied excellent construct validity of these measures (Zainal et al., 2024).

#### 2.4.5. W1 Subjective sleep quality

The PSQI comprised seven items (Buysse et al., 1989): (i) poor subjective sleep quality, (ii) sleep latency, (iii) sleep duration, (iv) habitual sleep inefficiency, (v) sleep disturbances range, (vi) sleep medication use, and (vii) daytime dysfunction. Each item ranged from 0–3, except for item (v), which ranged from 0–4. Higher values denoted a greater level of the construct. PSQI scores have shown good internal consistency, retest reliability, and construct validity (Mollayeva et al., 2016).

#### 2.4.6. W1 Mental health (MH) and medical visits

Given their importance as plausible confounders, counts of visits to MH and medical professionals in the past 12 months were measured with two continuous summary variables. Participants self-reported the frequency of visits for physical health appointments, urgent care, and medical procedures. They also stated the number of contacts with counselors, psychiatrists, primary care physicians, and spiritual advisors for psychological issues during this timeframe.

### 2.5. Data analyses

All data analyses were conducted using *R* (version 4.4.2; R Core Team, 2025). Feature engineering steps were conducted separately in the training and testing folds within a five-fold nested cross-validation framework to prevent data leakage and overfitting (high variance or problems generalizing patterns to unseen data; Lewis et al., 2023). These steps consisted of random forest (RF) imputation using the *missRanger* package (Mayer, 2024), normalization of continuous variables to have a mean (*M*) of 0 and a standard deviation (*SD*) of 1, and one-hot encoding of categorical predictors with the *caret* package (Kuhn, 2008). Relatedly, RF imputation was chosen as a missing data handling strategy, given its statistical superiority over standard multiple imputation in generating unbiased and precise estimates by accommodating potential nonlinearities and interactions (Shah et al., 2014). Overall, the initial predictor set in both the training and testing data folds had 43 variables. Table 1 presents the descriptive statistics for all predictor variables, including the variables of W1 and W2 PD diagnosis and severity, as well as the outcome variable of PD chronicity. Note that in our predictor set, we excluded the W1 PD diagnosis to avoid tautological arguments.

The present study examined seven ML models shown in Table 2 and focused on the model that performed well in the testing folds: the gradient boosting machine (GBM) algorithm. The GBM classification model was trained using the *caret* and *gbm* packages, which enabled optimal implementation. A stringent 5-fold nested cross-validation approach was conducted to test a GBM model for predicting PD chronicity. In each of the five outer-loop testing folds, the data were divided to ensure that no data leakage occurred while the inner-loop training

**Table 1**
Descriptive statistics of variables in the training set (n = 1054).

| | *M/n* | (*SD*)/(%) | Minimum | Maximum | Skewness | Kurtosis |
|---|---|---|---|---|---|---|
| W1 Age (years) | 55.32 | (11.78) | 34 | 84 | 0.30 | −0.67 |
| W1 Men vs. Women | 477 | (45.26) | – | – | – | – |
| White | 214 | (20.31) | – | – | – | – |
| W1 College-educated | 988 | (93.73) | – | – | – | – |
| W1 MDD severity | 0.85 | (2.34) | 0 | 14 | 2.82 | 7.36 |
| W1 GAD severity | 11.69 | (6.85) | 8 | 32 | 1.48 | 0.53 |
| W1 AUD severity | 0.07 | (0.37) | 0 | 4 | 7.03 | 56.43 |
| W1 SUD severity | 0.61 | (1.89) | 0 | 18 | 4.11 | 21.71 |
| W1 PSQI Subjective poor sleep quality | 0.97 | (0.68) | 0.00 | 3.00 | 0.46 | 0.52 |
| W1 PSQI SOL | 0.88 | (0.92) | 0.00 | 3.00 | 0.86 | −0.12 |
| W1 PSQI Sleep duration | 0.78 | (0.75) | 0.00 | 3.00 | 0.82 | 0.51 |
| W1 PSQI Habitual sleep inefficiency | 0.72 | (1.16) | 0.00 | 4.00 | 1.47 | 0.90 |
| W1 PSQI Sleep disturbances | 1.28 | (0.56) | 0.00 | 3.00 | 0.61 | 0.47 |
| W1 PSQI Sleep medication use | 0.57 | (1.07) | 0.00 | 3.00 | 1.57 | 0.79 |
| W1 PSQI Daytime dysfunction | 0.81 | (0.67) | 0.00 | 3.00 | 0.49 | 0.21 |
| W1 Total activity counts (Rest phase) | 14683.67 | (8892.75) | 3594.00 | 53063.57 | 1.70 | 3.23 |
| W1 Mean activity counts (Rest phase) | 31.44 | (18.88) | 6.90 | 120.42 | 1.82 | 4.02 |
| W1 Maximum activity counts (Rest phase) | 683.04 | (207.62) | 201.43 | 1420.33 | 0.83 | 0.82 |
| W1 Wake time (Rest phase) | 69.01 | (34.19) | 22.33 | 234.07 | 1.38 | 2.38 |
| W1 % of wake time (Rest phase) | 14.89 | (7.38) | 5.50 | 41.96 | 1.51 | 2.22 |
| W1 Mean wake bouts (Rest phase) | 38.15 | (13.52) | 10.71 | 95.71 | 1.13 | 2.59 |
| W1 Mean sleep bouts (Rest phase) | 12.50 | (6.06) | 3.29 | 81.16 | 4.32 | 35.83 |
| W1 Total activity counts (Sleep phase) | 7713.94 | (4604.40) | 1002.33 | 28696.57 | 1.76 | 4.20 |
| W1 Mean activity counts (Sleep phase) | 18.66 | (12.03) | 2.59 | 74.10 | 1.92 | 4.39 |
| W1 Maximum activity counts (Sleep phase) | 510.46 | (176.68) | 185.43 | 1297.33 | 0.91 | 1.30 |
| W1 Mean SOL (Sleep phase) | 25.77 | (23.08) | 0.21 | 128.57 | 1.87 | 3.58 |
| W1 Total sleep time (TST; Sleep phase) | 13.46 | (14.94) | 0.50 | 84.92 | 2.28 | 5.61 |
| W1 Sleep efficiency (Sleep phase) | 81.61 | (9.57) | 44.27 | 93.61 | −1.52 | 2.50 |
| W1 WASO (Sleep phase) | 45.57 | (21.87) | 8.83 | 139.86 | 1.33 | 2.51 |
| W1 Wake time (Sleep phase) | 45.65 | (22.13) | 8.83 | 139.86 | 1.37 | 2.74 |
| W1 % of wake time (Sleep phase) | 10.79 | (5.69) | 2.28 | 31.08 | 1.49 | 2.06 |
| W1 Mean wake bouts (Sleep phase) | 32.06 | (10.59) | 10.14 | 71.33 | 0.59 | 0.13 |
| W1 Mean sleep bouts (Sleep phase) | 14.25 | (7.43) | 4.26 | 81.44 | 4.37 | 32.35 |
| W1 Total activity counts (Active phase) | 328289.45 | (105549.46) | 62573.67 | 620944.17 | 0.36 | −0.01 |
| W1 Mean activity counts (Active phase) | 335.93 | (108.24) | 65.49 | 660.60 | 0.42 | 0.29 |
| W1 Maximum activity counts (Active phase) | 1376.26 | (364.51) | 429.83 | 2406.83 | 0.38 | −0.10 |
| W1 Wake time (Active phase) | 818.20 | (105.14) | 504.92 | 1085.75 | −0.49 | 0.07 |
| W1 % of wake time (Active phase) | 83.23 | (9.21) | 48.62 | 97.64 | −1.02 | 0.94 |
| W1 Mean wake bouts (Active phase) | 62.58 | (30.14) | 7.33 | 158.17 | 0.58 | −0.14 |
| W1 Mean sleep bouts (Active phase) | 2.73 | (1.46) | 1.22 | 19.29 | 7.31 | 77.14 |
| W1 Medical professional visits (12-month) | 3.64 | (4.05) | 0 | 57 | 4.59 | 39.50 |
| W1 MH professional visits (12-month) | 2.18 | (8.87) | 0 | 145 | 9.20 | 110.52 |
| W1 PD severity | 0.43 | (1.18) | 0 | 6 | 2.90 | 7.84 |
| W2 PD severity | 0.36 | (1.08) | 0 | 6 | 3.29 | 10.64 |
| W1 PD diagnosis | 69 | (6.55) | – | – | – | – |
| W2 PD diagnosis | 62 | (5.88) | – | – | – | – |
| PD chronicity (presence of W1 and W2 PD) | 109 | (10.30) | – | – | – | – |

*Note.* W1, wave 1 (2004–2006); MDD, major depressive disorder; GAD, generalized anxiety disorder; PD, panic disorder; AUD, alcohol use disorder; SUD, substance use disorder; PSQI, Pittsburgh Sleep Quality Index; SOL, sleep onset latency; WASO, wake after sleep onset; MH, mental health; W2, wave 2 (2013–2014).

**Table 2**
Performance metrics of each multivariate ML classification algorithm predicting W2 PD diagnosis with 43 W1 variables.

| Model | AUC | Sensitivity | Specificity | PPV | NPV |
|---|---|---|---|---|---|
| LASSO | .739 (.688–.790) | .667 (.580–.752) | .758 (.731–.784) | .244 (.197–.293) | .951 (.936–.966) |
| Ridge | .751 (.704–.798) | .721 (.635–.806) | .703 (.674–.732) | .222 (.179–.265) | .956 (.940–.970) |
| ENR | .730 (.677–.782) | .667 (.578–.752) | .751 (.723–.776) | .239 (.191–.287) | .951 (.935–.966) |
| DCT | .678 (.620–.736) | .667 (.579–.758) | .644 (.613–.674) | .180 (.144–.219) | .943 (.925–.961) |
| RF | .729 (.681–.777) | .784 (.704–.856) | .592 (.562–.624) | .184 (.149–.219) | .959 (.943–.974) |
| GBM | .764 (.717–.811) | .613 (.523–.704) | .817 (.792–.842) | .282 (.226–.339) | .947 (.932–.962) |
| SVM | .691 (.637–.742) | .703 (.616–.791) | .627 (.596–.658) | .181 (.144–.218) | .947 (.930–.965) |

*Note*. ML, machine learning; W2, wave 2 (2013–2014); PD, panic disorder; W1, wave 1 (2004–2006); AUC, area under the receiver operating characteristic curve; PPV, positive predictive value; NPV, negative predictive value; LASSO, least absolute shrinkage and selection operator; ENR, elastic net regression; DCT, decision trees (also known as classification and regression trees [CART]); RF, random forest; GBM, gradient boosting machine; SVM, support vector machine.

folds were used for parameter tuning and model selection. The GBM tuning parameters (interaction depth [1], learning rate [shrinkage = 0.1], and number of trees [n.trees = 100]) were set to standard defaults, and these parameters were kept constant across the grid, i.e., neither grid search nor optimization was conducted within the inner loops. This method facilitated an unbiased model testing process by strictly keeping the model selection and parameter tuning processes (inner loop) and the model performance estimation procedure (outer loop) independent of each other (Fife & D'Onofrio, 2023). Prediction scores and model performance estimates were combined across all outer-loop folds.

Model performance was assessed using various recommended performance metrics for the binary outcome, namely the presence or

absence of chronic PD (Pargent et al., 2023): the area under the receiver operating characteristic curve (AUC), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). Each of their 95 % confidence intervals (CIs), which indicated the degree of performance uncertainty, was estimated across 2000 bootstrapping replications (Rosenbusch et al., 2021). This method enables a holistic understanding of the model's actions and the comparative effects of each predictor on the PD chronicity outcome, providing critical insight into the complex (and possibly nonlinear) relationships between each predictor and the outcome.

The Shapley additive explanations (SHAP) bee swarm plot was used as an interpretable ML method to interpret the multivariate GBM regression model with the *treeshap* package (Komisarczyk et al., 2024). This approach provided a participant-level method that visualized predictor importance and each predictor's effect on model predictions (Lundberg & Lee, 2017). The SHAP method harnessed game theory principles and enabled understanding of local and global predictor effects on the model's predicted outcomes (Ponce-Bobadilla et al., 2024). The bee swarm plot displayed participant-level SHAP values for each predictor (Lundberg et al., 2020). Whereas the x-axis showed the effect of each predictor on the model output, the y-axis listed the relative importance of predictors, with the most essential predictor appearing first. Each SHAP plot data point indicated a participant, with the color denoting the predictor value (blue = low, red = high), and horizontal position representing the direction and strength of its effect on the model's prediction. This granular visualization enables the detection of potential nonlinearities and interactions among predictors, providing insights that standard OLS regression does not (Molnar, 2022).

Model calibration analyses were also performed to evaluate the degree of alignment between predicted probabilities and observed scores in a classification model, using three indices: (i) Brier score; (ii) expected calibration error (ECE); (iii) integrated calibration index (ICI; Austin & Steyerberg, 2019). The Brier score indicated the mean squared difference between the predicted probabilities and the actual observed scores, with lower scores suggesting better prediction accuracy (range = 0–1) (Steyerberg et al., 2010). Predicted probabilities were quantified as deciles, and the average differences between the predicted probabilities and the actual observed scores were calculated for each decile (Steyerberg, 2019). ECE quantified the mean absolute differences between these scores weighted by bin proportions. ICI calculated identical difference scores but weighted by bin counts. Lower ECE and ICI values signified better model calibration. The 95 % CIs of these metrics were also calculated by aggregating estimates across 2000 bootstrap resampling iterations.

### 2.6. Availability of data and materials

The authors are unable to share the database directly. However, the datasets can be retrieved from the publicly available repository (https://tinyurl.com/icpsr-midus). Data analytics scripts using the *R* programming language can be provided upon reasonable request.

## 3. Results

### 3.1. Model performance (Hypothesis 1)

Table 2 presents the performance metrics of the multivariate models. The GBM algorithm performed better than other algorithms in the testing fold, with the highest AUC value (AUC: .764; 95 % CI: [0.717–0.811]; Sensitivity: .613 [0.523–0.704]; Specificity: .817 [0.792–0.842]). Hypothesis 1 thus received full support.

### 3.2. Relative importance analyses (Hypothesis 2)

Fig. 1 displays the SHAP bee swarm plot of the top 20 multivariate predictors of PD chronicity, given their mean absolute SHAP values, which indicate each predictor's marginal contribution to the multivariate model's output at the participant level. Higher SHAP values reflected stronger relative importance (marked by the #). The color scheme in the figure depicted the directionality of associations, such that red denoted positive associations and blue denoted negative correlations.

Top psychopathology predictors of higher PD chronicity probability included MDD diagnosis (#1), fewer visits to MH professionals (#4), and medical professionals in the past 12 months (#7). Regarding PSQI-based variables, sleep medication use (#2), more sleep disturbances (#8), and daytime dysfunction (#19) were predictive of PD chronicity. Active phase actigraphy predictors of PD chronicity included higher percentage

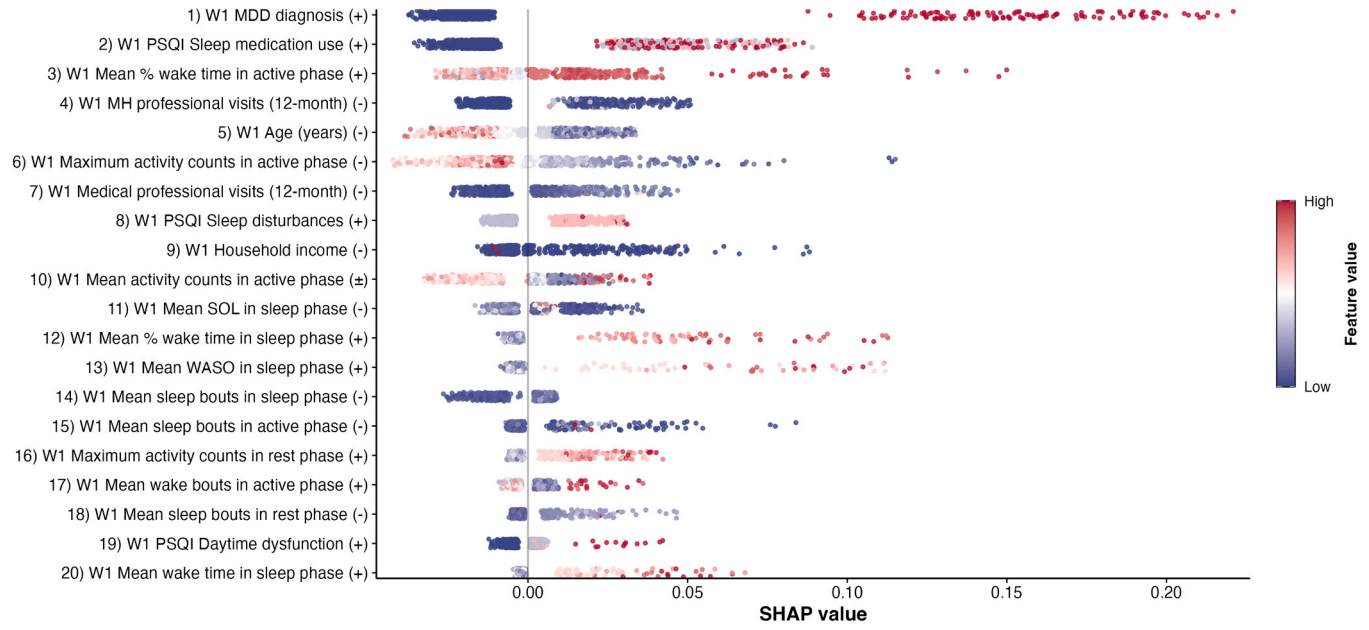


**Fig. 1.** Shapley additive explanation (SHAP) bee swarm plots depicting the top 20 multivariate predictors of PD chronicityNote. PD, panic disorder; W1, wave 1 (2004–2006); PSQI, Pittsburgh sleep quality index; MH, mental health; TST, total sleep time; WASO, wake after sleep onset; AUD, alcohol use disorder. PD chronicity is defined as the presence of a PD diagnosis at both W1 and wave 2 (W2; 2013–2014).

of wake time (#3), both high and low activity counts (#10), fewer sleep bouts (#15), and higher wake bouts (#17). Sleep phase actigraphy predictors of PD chronicity included shorter SOL (#11), higher percentage of wake time (#12), longer WASO (#13), fewer sleep bouts (#14), and longer wake time (#20). Rest phase actigraphy predictors of PD chronicity included more maximum activity counts (#16) and fewer sleep bouts (#18). Regarding sociodemographic variables, younger age (#5) and lower household income (#9) were predictive of PD chronicity.

### 3.3. Model calibration analyses

Model calibration analyses examined the Brier score, ECE, and ICI values. The Brier score was close to 0 (0.086, 95 % CI [0.073–0.098]). Similarly, both the ECE and ICI values (0.023, 95 % CI [0.016–0.043]) approached 0. Therefore, the model was well-calibrated.

## 4. Discussion

The current study aimed to determine how baseline sleep disturbances, as indexed by both actigraphy and self-report, predict the long-term chronicity of PD. If externally validated (Collins et al., 2024), a prognostic calculator could be constructed, informed by ML, to detect high-risk individuals and identify changeable risk factors. Plausible accounts were proposed via a novel framework called the *sleep-panic nexus theory,* as illustrated in Fig. 2. Beyond theoretical integration, this model holds clinical pertinence by proposing that disruptions in sleep-wake rhythms and cadences, when coinciding with other mental disorders, underutilized healthcare, and social determinants of health, increase the risk of chronic PD. With further validation and refinements, this theory might inform improvements to scalable, sleep-focused preventive interventions.

Notably, comorbid MDD (but not GAD) and higher AUD symptom severity were predictive of PD chronicity. Plausibly, comorbid MDD and higher AUD severity may amplify panic attacks, avoidance, and related PD symptoms by aggravating stress reactivity (Cook et al., 2021) and weakening treatment-seeking motivation (Bifftu et al., 2018), thereby contributing to the persistence of PD. PD diagnosis was excluded from the high-dimensional predictor set altogether, thus setting sleep variables apart as longitudinal predictors of PD chronicity instead of co-occurring symptom correlates. In the final model, several PSQI- and actigraphy-indexed markers emerged as key predictors even after adjusting for these psychiatric comorbidities. Actigraphy-indexed metrics are not stated in formal diagnostic criteria (American Psychiatric Association, 2013) and indicate unique physiological processes. Thus, we construe their prognostic utility as evidence that sleep disturbances serve as changeable, mechanistic risk factors, rather than duplicative proxies of baseline diagnoses or symptoms. Our outcomes emphasize the clinical relevance of incorporating sleep variables into psychiatric comorbidity profiles. Even after adjusting for comorbid MDD and AUD severity, the emergence of several actigraphy- and PSQI-indexed sleep disturbances underscores the value of ML in detecting non-intuitive (e.g., nonlinear) changeable risk factors of chronic PD.

Additionally, why was PD chronicity strongly predicted by limited engagement with MH and medical professionals, as well as self-reported sleep medication use, increased disturbances, and daytime dysfunction? Fewer visits to MH and medical professionals might indicate treatment barriers, underutilization, or chronic avoidance patterns (Horenstein & Heimberg, 2020), i.e., factors that likely correlate with and maintain PD symptoms across long durations. Self-reported sleep variables, including any use of sleep medication, sleep disturbances, and heightened daytime dysfunction, likely indicate the bidirectional associations between subjective sleep quality and PD (Harvey et al., 2011). Serotonin serves as a precursor to melatonin synthesis, and disruptions in serotonin pathways could compromise circadian rhythms and sleep modulation (Sundberg et al., 2020). Appraisals of poor sleep quality might increase moments of emotional dysregulation at both behavioral and physiological levels (Riemann et al., 2024), thereby prolonging PD symptoms and hindering remission. Together, the *sleep-panic nexus theory* proposes that these factors emphasize the importance of multidisciplinary care targeting physical health, MH comorbidities, and sleep disturbances, reducing barriers, and improving engagement with healthcare.

Moreover, what might explain the correlation between actigraphy-indexed sleep-wake rhythms and the chronicity of PD? During the sleep phase, a combination of fewer integrated sleep events, longer wake times, increased durations of WASO, and shorter SOL predicted the development of chronic PD. During the active phase, individuals with chronic PD exhibited elevated wakefulness, varied physical activity, fewer sleep adjustments, such as naps (sleep bouts), and frequent awakenings (wake bouts). During the rest phase, fewer sleep bouts and

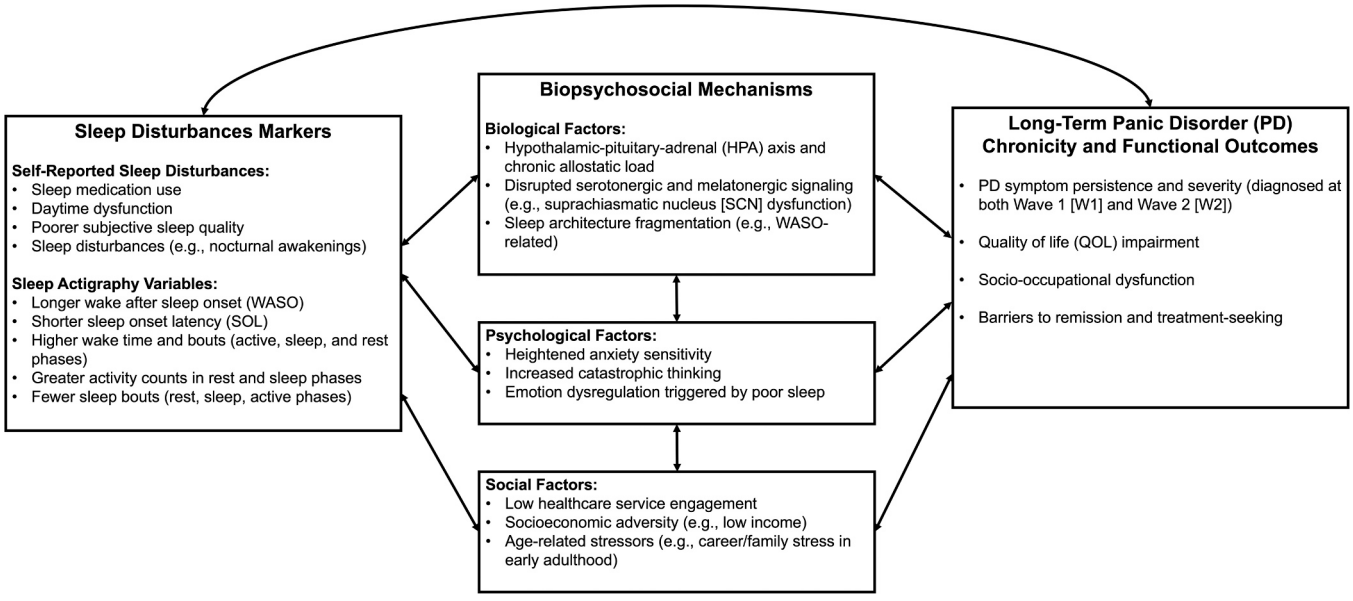


**Fig. 2.** Theoretical tenets of the sleep-panic nexus modelNote. Double-headed arrows indicated plausible bidirectional relationships across time. The *sleep-panic nexus model* proposes that actigraphy-indexed and self-reported sleep disturbances predict panic disorder (PD) chronicity in the long run via biopsychosocial mechanisms, including cognitive-affective disturbances, neurobiologically-associated stress reactivity, and socio-environmental risk factors.

higher physical activity peaks were predictive of chronic PD. This configuration indicated that chronic PD might be intimately linked to somatic arousal, sleep discontinuities, and fragmentation that compromise daytime and nighttime routine activities. Such irregularities in sleep-wake rhythms would likely prolong PD symptoms by compromising reparative sleep and sustaining vicious cycles of hypervigilance and sleep fragmentation (Carbone et al., 2023). Tenable accounts for these outcomes were that increased ruminative thinking styles could aggravate somatic arousal and adversely affect rapid eye movement (REM) sleep, fragmenting sleep architecture and continuity (Yang et al., 2024). Furthermore, biologically associated sleep indices, such as decreased slow-wave sleep marked by delta and theta waves (Belleville & Potočnik, 2019), coupled with increased maximum resting heart rate (Tsai et al., 2024), could be mechanisms linking actigraphy indices to the chronicity of PD.

Simultaneously, lower household income and younger age were correlated with the chronicity of PD. Perhaps in earlier adulthood, multiple age-related life stressors, including career pressures, family-building duties, and financial commitments (Collins et al., 2023; LaMontagne et al., 2024), might worsen PD symptoms over time. Low income may perpetuate PD symptoms by compromising emotion regulation, maintaining sleep disturbances, and raising anxiety sensitivity, in the context of economic stressors (Etindele Sosso et al., 2022). Collectively, these outcomes underscore the importance of considering demographic and socioeconomic factors in predicting long-term PD chronicity.

The results provide three central clinical advancements. First, the prognostic value of changeable, unique sleep attributes, including daytime dysfunction, SOL, and WASO, underscores the significance of sleep-based treatments in reducing the long-term risk of PD chronicity. Second, comorbid MDD and low healthcare engagement were critical predictors, emphasizing the importance of coordinated care frameworks that remedy barriers to access and MH comorbidities. Third, leveraging ML facilitated the identification of non-intuitive, nonlinear risk factors that might not have been easily detected through standard OLS regression. This statistical method is promising for constructing prognostic instruments to identify individuals at high risk of PD chronicity, thereby improving precision prevention in clinical practice and community settings.

The present study has several limitations. First, unassessed potential confounds, such as genetic vulnerabilities (Moraes et al., 2024), should be investigated further. Second, replication efforts should utilize measures consistent with Diagnostic and Statistical Manual-Fifth Edition (DSM-5) criteria, although few differences exist between these and the DSM-III-R measure used (Asmundson et al., 2014). Third, extending this study to a more culturally diverse sample could provide more evidence about the generalizability of these patterns. Fourth, although the present sample size was sufficient for the study's aims, replication studies should increase the sample sizes to reduce variance, enhance generalizability, and strengthen model calibration (Goldenholz et al., 2023). Despite these limitations, the study's strengths include robust nested CV ML approaches that prevent data leakage, a nine-year longitudinal timeframe, and a relatively large predictor set comprising key demographics and sleep actigraphy variables.

Pending external validation (Collins et al., 2024), the *sleep-panic nexus theory* provides several actionable implications for clinical psychological science and healthcare policy across levels of care. At the patient level, therapists might preemptively screen for self-reported and actigraphy-based sleep disturbances, such as brief SOL, daytime dysfunction, and lengthier WASO, among clients with or at risk for PD to guide preventive sleep-focused treatments. Policy-wise, healthcare providers should give precedence to coordinated stepped-care that seamlessly links behavioral sleep medicine, primary care, and psychiatry, particularly for those with comorbid MDD and high AUD severity who underuse services. Technologically, this model facilitates the construction of ML-based 'prognostic calculators' utilizing passive sensor wearables to identify sleep-wake discontinuity patterns as plausible mechanisms and risk factors. These instruments could facilitate scalable, tailored prevention treatments that target underserved populations, such as emerging adults and low-income individuals, across various contexts. Future trials of universal, selective, or indicated prevention programs should, thus, evaluate sleep-based digital interventions (e.g., cognitive-behavioral therapy for insomnia [CBT-I] apps or actigraphy-driven coaching) to improve sleep-wake rhythms and routines. The upshot is that these strategies across various levels of care might evaluate and enhance the *sleep-panic nexus theory* as a mechanisms-informed framework to alleviate PD chronicity via changeable sleep-focused targets.

## CRediT authorship contribution statement

**Nur Hani Zainal:** Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Natalia Van Doren:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Methodology, Data curation, Conceptualization.

## Declaration of Generative AI and AI-assisted technologies in the writing process

None. The authors did not use generative AI tools in the writing process. The authors are solely responsible for the scientific content and integrity of this work.

## Declaration of Competing Interest

The authors hereby declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

The authors do not have permission to share data directly. Data are publicly available at https://www.icpsr.umich.edu/web/ICPSR/series/203.

## References

Alfano, C. A. (2018). Re)conceptualizing sleep among children with anxiety disorders: Where to next? *Clinical Child and Family Psychology Review, 21*(4), 482–499. https://doi.org/10.1007/s10567-018-0267-4

American Psychiatric Association. (1987). *Diagnostic and Statistical Manual of Mental Disorders, 3rd Edition (DSM-III-R)*. American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.).). American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

Aqua, J. K., Barnum, O., & Johnson, D. A. (2024). Sleep as a contributor to socioeconomic disparities in hypertension: The Midlife in the United States (MIDUS II) Study. *Sleep, 47*(9). https://doi.org/10.1093/sleep/zsae142

Asmundson, G. J., Taylor, S., & Smits, J. A. (2014). Panic disorder and agoraphobia: An overview and commentary on DSM-5 changes. *Depression and Anxiety, 31*(6), 480–486. https://doi.org/10.1002/da.22277

Austin, P. C., & Steyerberg, E. W. (2019). The Integrated Calibration Index (ICI) and related metrics for quantifying the calibration of logistic regression models. *Statistics in Medicine, 38*(21), 4051–4065. https://doi.org/10.1002/sim.8281

Barbotin, B., Hoertel, N., Olfson, M., Blanco, C., Sanchez-Rico, M., Lejoyeux, M., & Geoffroy, P. A. (2022). Sleep complaints among adults with major depressive episode are associated with increased risk of incident psychiatric disorders: Results from a population-based 3-year prospective study. *Journal of Clinical Psychiatry, 84*(1). https://doi.org/10.4088/JCP.21m14236

Barrett, A. J., Taylor, S. L., Kopak, A. M., & Hoffmann, N. G. (2021). PTSD, panic disorder and alcohol use disorder as a triple threat for violence among male jail detainees. *Journal of Criminal Psychology, 11*(1), 21–29. https://doi.org/10.1108/JCP-07-2020-0029

Bei, B., Seeman, T. E., Carroll, J. E., & Wiley, J. F. (2017). Sleep and physiological dysregulation: A closer look at sleep intraindividual variability. *Sleep, 40*(9). https://doi.org/10.1093/sleep/zsx109

Belleville, G., & Potočnik, A. (2019). A meta-analysis of sleep disturbances in panic disorder. In In. R. Woolfolk, L. Allen, F. Durbano, & F. Irtelli (Eds.), *Psychopathology -*

*An International and Interdisciplinary Perspective*. IntechOpen. https://doi.org/10.5772/intechopen.86306.
Bhat, A. C., Diaz, J. A., Lee, S. A., Almeida, D. M., & Lee, S. (2024). Associations between recession hardships and subjective and objective sleep measures in the Midlife in the United States Study: Race and gender differences. *Frontiers in Sleeping, 3*. https://doi.org/10.3389/frsle.2024.1403818
Bifftu, B. B., Takele, W. W., Guracho, Y. D., & Yehualashet, F. A. (2018). Depression and its help seeking behaviors: A systematic review and meta-analysis of community survey in Ethiopia. *Depression Research and Treatment, 2018*, Article 1592596. https://doi.org/10.1155/2018/1592596
Blakey, S. M., Campbell, S. B., & Simpson, T. L. (2022). Associations between lifetime panic attacks, posttraumatic stress disorder, and substance use disorders in a nationally representative sample. *Journal of Dual Diagnosis, 18*(1), 21–32. https://doi.org/10.1080/15504263.2021.2013096
Buysse, D. J., Reynolds, C. F., 3rd, Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Research, 28*(2), 193–213. https://doi.org/10.1016/0165-1781(89)90047-4
Carbone, E. A., Menculini, G., de Filippis, R., D'Angelo, M., Zebi, L., & Steardo, L., Jr (2023). Sleep disturbances in panic disorder with comorbid complex PTSD: A possible relationship and different psychopathology? *Life, 13*(8). https://doi.org/10.3390/life13081636
Carta, M. G., Moro, M. F., Aguglia, E., Balestrieri, M., Caraci, F., Dell'Osso, L., & Faravelli, C. (2015). The attributable burden of panic disorder in the impairment of quality of life in a national survey in Italy. *International Journal of Social Psychiatry, 61*(7), 693–699. https://doi.org/10.1177/0020764015573848
Collins, G. S., Dhiman, P., Ma, J., Schlussel, M. M., Archer, L., Van Calster, B., & Riley, R. D. (2024). Evaluation of clinical prediction models (part 1): From development to external validation. *BMJ, 384*, Article e074819. https://doi.org/10.1136/bmj-2023-074819
Collins, S., Hoare, E., Allender, S., Olive, L., Leech, R. M., Winpenny, E. M., & Lotfalian, M. (2023). A longitudinal study of lifestyle behaviours in emerging adulthood and risk for symptoms of depression, anxiety, and stress. *Journal of Affective Disorders, 327*, 244–253. https://doi.org/10.1016/j.jad.2023.02.010
Cook, M. A., Newins, A. R., & Dvorak, R. D. (2021). Coping motivated alcohol use: The role of social anxiety and dissociation. *Substance Use and Misuse, 56*(2), 275–285. https://doi.org/10.1080/10826084.2020.1861630
de Jonge, P., Roest, A. M., Lim, C. C., Florescu, S. E., Bromet, E. J., Stein, D. J., & Scott, K. M. (2016). Cross-national epidemiology of panic disorder and panic attacks in the world mental health surveys. *Depression and Anxiety, 33*(12), 1155–1177. https://doi.org/10.1002/da.22572
Etindele Sosso, F. A., Kreidlmayer, M., Pearson, D., & Bendaoud, I. (2022). Towards a socioeconomic model of sleep health among the Canadian population: A systematic review of the relationship between age, income, employment, education, social class, socioeconomic status and sleep disparities. *European Journal of Investigation in Health, Psychology and Education, 12*(8), 1143–1167. https://doi.org/10.3390/ejihpe12080080
Fife, D. A., & D'Onofrio, J. (2023). Common, uncommon, and novel applications of random forest in psychological research. *Behavior Research Methods, 55*(5), 2447–2466. https://doi.org/10.3758/s13428-022-01901-9
Geng, F., Liu, X., Liang, Y., Shi, X., Chen, S., & Fan, F. (2018). Prospective associations between sleep problems and subtypes of anxiety symptoms among disaster-exposed adolescents. *Sleeping Medicine, 50*, 7–13. https://doi.org/10.1016/j.sleep.2018.05.017
Goldenholz, D. M., Sun, H., Ganglberger, W., & Westover, M. B. (2023). Sample size analysis for machine learning clinical validation studies. *Biomedicines, 11*(3). https://doi.org/10.3390/biomedicines11030685
Gonzalez, O. (2021). Psychometric and machine learning approaches for diagnostic assessment and tests of individual classification. *Psychological Methods, 26*(2), 236–254. https://doi.org/10.1037/met0000317
Gregory, A. M., Caspi, A., Eley, T. C., Moffitt, T. E., Oconnor, T. G., & Poulton, R. (2005). Prospective longitudinal associations between persistent sleep problems in childhood and anxiety and depression disorders in adulthood. *Journal of Abnormal Child Psychology, 33*(2), 157–163. https://doi.org/10.1007/s10802-005-1824-0
Harvey, A. G., Murray, G., Chandler, R. A., & Soehner, A. (2011). Sleep disturbance as transdiagnostic: Consideration of neurobiological mechanisms. *Clinical Psychology Review, 31*(2), 225–235. https://doi.org/10.1016/j.cpr.2010.04.003
Horenstein, A., & Heimberg, R. G. (2020). Anxiety disorders and healthcare utilization: A systematic review. *Clinical Psychology Review, 81*, Article 101894. https://doi.org/10.1016/j.cpr.2020.101894
Hovenkamp-Hermelink, J. H. M., Jeronimus, B. F., Myroniuk, S., Riese, H., & Schoevers, R. A. (2021). Predictors of persistence of anxiety disorders across the lifespan: a systematic review. *Lancet Psychiatry, 8*(5), 428–443. https://doi.org/10.1016/S2215-0366(20)30433-8
Hysing, M., Harvey, A. G., Skrindo Knudsen, A. K., Skogen, J. C., Reneflot, A., & Sivertsen, B. (2025). Mind at rest, mind at risk: A prospective population-based study of sleep and subsequent mental disorders. *Sleeping Medicine: X, 9*, Article 100138. https://doi.org/10.1016/j.sleepx.2025.100138
Jackson, M. L., Sztendur, E. M., Diamond, N. T., Byles, J. E., & Bruck, D. (2014). Sleep difficulties and the development of depression and anxiety: A longitudinal study of young Australian women. *Archives of Women's Mental Health, 17*(3), 189–198. https://doi.org/10.1007/s00737-014-0417-8
Jemcov, A., Olthuis, J. V., Watt, M. C., & Stewart, S. H. (2023). Do anxiety sensitivity cognitive concerns and/or depression symptoms independently explain sleep disturbances in a high anxiety sensitive treatment-seeking sample? *Journal of Anxiety Disorders, 97*, Article 102731. https://doi.org/10.1016/j.janxdis.2023.102731
Kessler, R. C., Andrews, G., Mroczek, D., Ustun, B., & Wittchen, H.-U. (1998). The World health organization composite international diagnostic interview - short-form (CIDI-SF). *International Journal of Methods in Psychiatric Research, 7*(4), 171–185. https://doi.org/10.1002/mpr.47
Kessler, R. C., & Üstün, T. B. (2004). The World Mental Health (WMH) survey initiative version of the world health organization (WHO) composite international diagnostic interview (CIDI). *International Journal of Methods in Psychiatric Research, 13*(2), 93–121. https://doi.org/10.1002/mpr.168
Komisarczyk, K., Kozminski, P., Maksymiuk, S., & Biecek, P. (2024). Treeshap: Compute SHAP values for your tree-based models using the 'TreeSHAP' algorithm. *In R Package Version 0 3 1*. ⟨https://CRAN.R-project.org/package=treeshap⟩.
Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software, 28*(5), 1–26. https://doi.org/10.18637/jss.v028.i05
LaMontagne, A. D., Too, L. S., Witt, K., Evans-Whipp, T., Owen, P. J., & Toumbourou, J. W. (2024). Does adolescent depression modify the association between psychosocial job stressors and mental health in emergent adulthood? *American Journal of Industrial Medicine, 67*(1), 44–54. https://doi.org/10.1002/ajim.23547
Lewis, M. J., Spiliopoulou, A., Goldmann, K., Pitzalis, C., McKeigue, P., & Barnes, M. R. (2023). nestedcv: An R package for fast implementation of nested cross-validation with embedded feature selection designed for transcriptomics and high-dimensional data. *Bioinformatics Advances, 3*(1), Article vbad048. https://doi.org/10.1093/bioadv/vbad048
Liu, Y., Sareen, J., Bolton, J., & Wang, J. (2015). Development and validation of a risk-prediction algorithm for the recurrence of panic disorder. *Depression and Anxiety, 32*(5), 341–348. https://doi.org/10.1002/da.22359
Locke, A. B., Kirst, N., & Shultz, C. G. (2015). Diagnosis and management of generalized anxiety disorder and panic disorder in adults. *American Family Physician, 91*(9), 617–624.
Love, G. D., Seeman, T. E., Weinstein, M., & Ryff, C. D. (2010). Bioindicators in the MIDUS national study: Protocol, measures, sample, and comparative context. *Journal of Aging and Health, 22*(8), 1059–1080. https://doi.org/10.1177/0898264310374355
Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., & Lee, S. I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence, 2*(1), 56–67. https://doi.org/10.1038/s42256-019-0138-9
Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems, 30*, 4768–4777. https://doi.org/10.5555/3295222.3295230
Malyszczak, K., & Janocha, A. (2022). The circadian rhythm of blood pressure in patients with panic disorder and generalized anxiety disorder. *Journal of Psychosomatic Research, 161*, Article 110999. https://doi.org/10.1016/j.jpsychores.2022.110999
Marcks, B. A., Weisberg, R. B., Edelen, M. O., & Keller, M. B. (2010). The relationship between sleep disturbance and the course of anxiety disorders in primary care patients. *Psychiatry Research, 178*(3), 487–492. https://doi.org/10.1016/j.psychres.2009.07.004
Mayer, M. (2024). missRanger: Fast Imputation of Missing Values. In (Version 2.5.0) https://CRAN.R-project.org/package=missRanger.
McCallum, S. M., Batterham, P. J., Calear, A. L., Sunderland, M., Carragher, N., & Kazan, D. (2019). Associations of fatigue and sleep disturbance with nine common mental disorders. *Journal of Psychosomatic Research, 123*, Article 109727. https://doi.org/10.1016/j.jpsychores.2019.05.005
Mollayeva, T., Thurairajah, P., Burton, K., Mollayeva, S., Shapiro, C. M., & Colantonio, A. (2016). The Pittsburgh sleep quality index as a screening tool for sleep dysfunction in clinical and non-clinical samples: A systematic review and meta-analysis. *Sleeping Medicine Reviews, 25*, 52–73. https://doi.org/10.1016/j.smrv.2015.01.009
Molnar, C. (2022). Interpretable machine learning: A guide for making black-box models explainable (2nd ed.). https://christophm.github.io/interpretable-ML-book/cite.html.
Moraes, A. C. N., Wijaya, C., Freire, R., Quagliato, L. A., Nardi, A. E., & Kyriakoulis, P. (2024). Neurochemical and genetic factors in panic disorder: A systematic review. *Translational Psychiatry, 14*(1), 294. https://doi.org/10.1038/s41398-024-02966-0
Ng, M. H. S., Zainal, N. H., & Newman, M. G. (2024). Positive reappraisal mediates childhood experiences of parental abuse and affection on adulthood generalized anxiety severity. *Journal of Anxiety Disorders, 102*, Article 102826. https://doi.org/10.1016/j.janxdis.2024.102826
Olfson, M., Mojtabai, R., Merikangas, K. R., Compton, W. M., Wang, S., Grant, B. F., & Blanco, C. (2017). Reexamining associations between mania, depression, anxiety and substance use disorders: Results from a prospective national cohort. *Molecular Psychiatry, 22*(2), 235–241. https://doi.org/10.1038/mp.2016.64
Orru, G., Monaro, M., Conversano, C., Gemignani, A., & Sartori, G. (2019). Machine learning in psychometrics and psychological research. *Frontiers in Psychology, 10*, 2970. https://doi.org/10.3389/fpsyg.2019.02970
Pargent, F., Schoedel, R., & Stachl, C. (2023). Best practices in supervised machine learning: A tutorial for psychologists. *Advances in Methods and Practices in Psychological Science, 6*(3), Article 25152459231162559. https://doi.org/10.1177/25152459231162559
Pastre, M., & Lopez-Castroman, J. (2022). Actigraphy monitoring in anxiety disorders: A mini-review of the literature. *Frontiers in Psychiatry, 13*, Article 984878. https://doi.org/10.3389/fpsyt.2022.984878
Pearl, J. (2014). Interpretation and identification of causal mediation. *Psychological Methods, 19*(4), 459–481. https://doi.org/10.1037/a0036434
Ponce-Bobadilla, A. V., Schmitt, V., Maier, C. S., Mensing, S., & Stodtmann, S. (2024). Practical guide to SHAP analysis: Explaining supervised machine learning model

predictions in drug development. *Clinical and Translational Science, 17*(11), Article e70056. https://doi.org/10.1111/cts.70056

R Core TeamR Core Team. (2025). R: A language and environment for statistical computing. In (Version 4.4.1) R Foundation for Statistical Computing. https://www.R-project.org/.

Riemann, D., Dressle, R. J., Benz, F., Spiegelhalder, K., Johann, A. F., Nissen, C., & Feige, B. (2024). Chronic insomnia, REM sleep instability and emotional dysregulation: A pathway to anxiety and depression? *Journal of Sleeping Research, 34*(2), Article e14252. https://doi.org/10.1111/jsr.14252

Rosenbusch, H., Soldner, F., Evans, A. M., & Zeelenberg, M. (2021). Supervised machine learning methods in psychology: A practical introduction with annotated R code. *Social and Personality Psychology Compass, 15*(2). https://doi.org/10.1111/spc3.12579

Ryff, C., Almeida, D., Ayanian, J., Binkley, N., Carr, D.S., Coe, C., Williams, D. (2019a). *Midlife in the United States (MIDUS 3), 2013-2014* Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/ICPSR36346.v7.

Ryff, C.D., Almeida, D.M., Ayanian, J.Z., Carr, D.S., Cleary, P.D., Coe, C., Williams, D.R. (2021). *Midlife in the United States (MIDUS 2), 2004-2006* Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/ICPSR04652.v8.

Ryff, C.D., Seeman, T., & Weinstein, M. (2019b). Midlife in the United States (MIDUS 2): Biomarker Project, 2004-2009 Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/ICPSR29282.v9.

Santamaria-Garcia, H., Migeot, J., Medel, V., Hazelton, J. L., Teckentrup, V., Romero-Ortuno, R., & Ibanez, A. (2025). Allostatic interoceptive overload across psychiatric and neurological conditions. *Biological Psychiatry, 97*(1), 28–40. https://doi.org/10.1016/j.biopsych.2024.06.024

Schoenberg, P. L. A. (2020). Linear and Nonlinear EEG-Based Functional Networks in Anxiety Disorders. In In. Y.-K. Kim (Ed.), *Anxiety Disorders: Rethinking and Understanding Recent Discoveries* (pp. 35–59). Springer Singapore. https://doi.org/10.1007/978-981-32-9705-0_3.

Selzer, M. L. (1971). The Michigan alcoholism screening test: the quest for a new diagnostic instrument. *American Journal of Psychiatry, 127*(12), 1653–1658. https://doi.org/10.1176/ajp.127.12.1653

Shah, A. D., Bartlett, J. W., Carpenter, J., Nicholas, O., & Hemingway, H. (2014). Comparison of random forest and parametric imputation models for imputing missing data using MICE: A CALIBER study. *American Journal of Epidemiology, 179*(6), 764–774. https://doi.org/10.1093/aje/kwt312

Sheikh, J. I., Woodward, S. H., & Leskin, G. A. (2003). Sleep in post-traumatic stress disorder and panic: convergence and divergence. *Depression and Anxiety, 18*(4), 187–197. https://doi.org/10.1002/da.10066

Shrivastava, D., Jung, S., Saadat, M., Sirohi, R., & Crewson, K. (2014). How to interpret the results of a sleep study. *Journal of Community Hospital Internal Medicine Perspectives, 4*(5), 24983. https://doi.org/10.3402/jchimp.v4.24983

Speranza, A. M., Liotti, M., Spoletini, I., & Fortunato, A. (2023). Heterotypic and homotypic continuity in psychopathology: A narrative review. *Frontiers in Psychology, 14*. https://doi.org/10.3389/fpsyg.2023.1194249

Steyerberg, E. W. (2019). Evaluation of performance. In E. W. Steyerberg (Ed.), *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating* (pp. 277–308). Springer International Publishing. https://doi.org/10.1007/978-3-030-16399-0_15.

Steyerberg, E. W., Vickers, A. J., Cook, N. R., Gerds, T., Gonen, M., Obuchowski, N., & Kattan, M. W. (2010). Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology, 21*(1), 128–138. https://doi.org/10.1097/EDE.0b013e3181c30fb2

Sundberg, I., Rasmusson, A. J., Ramklint, M., Just, D., Ekselius, L., & Cunningham, J. L. (2020). Daytime melatonin levels in saliva are associated with inflammatory markers and anxiety disorders. *Psychoneuroendocrinology, 112*, Article 104514. https://doi.org/10.1016/j.psyneuen.2019.104514

Teas, E., & Friedman, E. (2021). Sleep and functional capacity in adults: Cross-sectional associations among self-report and objective assessments. *Sleeping Health, 7*(2), 198–204. https://doi.org/10.1016/j.sleh.2020.12.001

Tian, X., Li, H., Li, F., Jimenez-Herrera, M. F., Ren, Y., & Shang, H. (2024). Development and validation of a web-based calculator for determining the risk of psychological distress based on machine learning algorithms: A cross-sectional study of 342 lung cancer patients. *Supportive Care in Cancer, 33*(1), 63. https://doi.org/10.1007/s00520-024-09127-5

Tsai, C. H., Christian, M., Kuo, Y. Y., Lu, C. C., Lai, F., & Huang, W. L. (2024). Sleep, physical activity and panic attacks: A two-year prospective cohort study using smartwatches, deep learning and an explainable artificial intelligence model. *Sleeping Medicine, 114*, 55–63. https://doi.org/10.1016/j.sleep.2023.12.013

Turiano, N. A., Whiteman, S. D., Hampson, S. E., Roberts, B. W., & Mroczek, D. K. (2012). Personality and substance use in midlife: Conscientiousness as a moderator and the effects of trait change. *Journal of Research in Personality, 46*(3), 295–305. https://doi.org/10.1016/j.jrp.2012.02.009

Uren, J., Richdale, A. L., Cotton, S. M., & Whitehouse, A. J. O. (2019). Sleep problems and anxiety from 2 to 8 years and the influence of autistic traits: A longitudinal study. *European Child Adolescent Psychiatry, 28*(8), 1117–1127. https://doi.org/10.1007/s00787-019-01275-y

Van Lissa, C. J. (2022). Developmental data science: How machine learning can advance theory formation in Developmental Psychology. *Infant and Child Development, 32*(6). https://doi.org/10.1002/icd.2370

Wang, Z., Luo, Y., Zhang, Y., Chen, L., Zou, Y., Xiao, J., & Zou, Z. (2023). Heart rate variability in generalized anxiety disorder, major depressive disorder and panic disorder: A network meta-analysis and systematic review. *Journal of Affective Disorders, 330*, 259–266. https://doi.org/10.1016/j.jad.2023.03.018

Yang, S., Liu, G. X., Zhou, Y. S., Ma, Z. J., Xia, L., Zhang, H. L., & Chen, G. H. (2024). Relationships between rumination and different types of rapid eye movement sleep in patients with chronic insomnia disorder. *Sleeping Medicine, 117*, 162–168. https://doi.org/10.1016/j.sleep.2024.03.032

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science, 12*(6), 1100–1122. https://doi.org/10.1177/1745691617693393

Zainal, N. H., & Newman, M. G. (2022). Inflammation mediates depression and generalized anxiety symptoms predicting executive function impairment after 18 years. *Journal of Affective Disorders, 296*, 465–475. https://doi.org/10.1016/j.jad.2021.08.077

Zainal, N. H., Soh, C. P., & Van Doren, N. (2024). Daily stress reactivity and risk appraisal mediates childhood parental abuse predicting adulthood psychopathology severity: An 18-year longitudinal mediation analysis. *Journal of Affective Disorders, 358*, 138–149. https://doi.org/10.1016/j.jad.2024.04.068

Zvolensky, M. J., Bernstein, A., Marshall, E. C., & Feldner, M. T. (2006). Panic attacks, panic disorder, and agoraphobia: Associations with substance use, abuse, and dependence. *Current Psychiatry Reports, 8*(4), 279–285. https://doi.org/10.1007/s11920-006-0063-6