# Uncovering Sleep Determinants of Clinical Pain Severity Through Machine Learning


Eric Duan[1], Nandini Raghuraman[1,2,3], Rachel Massalee[1], Yang Wang[1,2,3*]

[1] Department of Pain and Translational Symptom Science, School of Nursing, University of Maryland, Baltimore, US

[2] Center to Advance Chronic Pain Research, University of Maryland, Baltimore, US

[3] Placebo Beyond Opinions Center, University of Maryland School of Nursing, Baltimore, USA

*Corresponding author e-mail: yang.wang@umaryland.edu


## Background

Sleep and circadian rhythms play a critical role in shaping individual pain experiences[1]. Poor sleep can worsen pain, creating a self-perpetuating cycle, yet it's still unclear which specific sleep parameters best predict clinical pain. Studies reveal discrepancies between subjective sleep reports (e.g., diaries, questionnaires) and objective measures (e.g., actigraphy, polysomnography), especially among individuals with chronic pain and insomnia[2].

To address this gap, we applied machine learning (ML) techniques to data from 556 adults aged 18-88 in the Cleveland Family Study (CFS)[3,4]. This dataset includes both subjective and objective sleep data along with clinical pain measures. Using Gradient Boosted Regression (GBR), Random Forest (RF), and Elastic Net Regression (ENR) models, we identified key sleep predictors of pain severity in a general population.

## Methods

Data were drawn from the CFS, a longitudinal, family-based cohort designed to investigate genetic factors related to sleep apnea. Clinical pain severity was assessed using the bodily pain subscale of the SF-36 health survey. Our analysis focused on the relationship between sleep profiles and pain severity, incorporating both subjective (e.g., sleep quality ratings) and objective (e.g., polysomnography-derived metrics) sleep parameters.

Sociodemographic and clinical data included age, sex, race, income, BMI, health behaviors (e.g., alcohol use, smoking, exercise), and comorbidities such as cardiovascular disease, depression, and anxiety. Objective sleep data were collected via overnight polysomnography and included measures like oxygen saturation and sleep architecture. Subjective sleep data were obtained from the Functional Outcomes of Sleep Questionnaire (FOSQ).

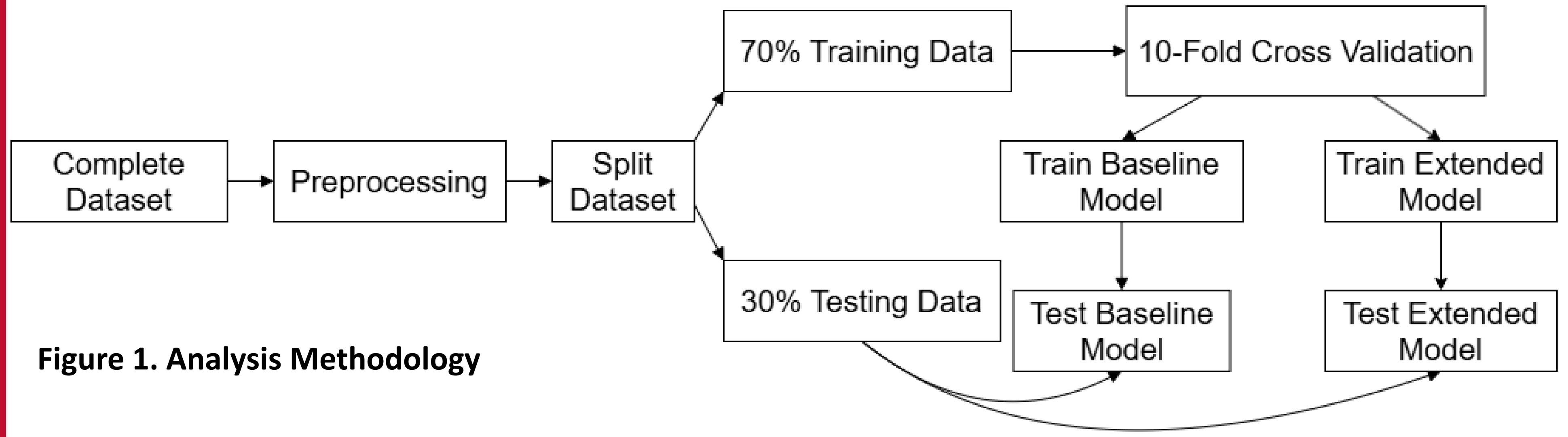


**Figure 1. Analysis Methodology**

Missing data were evaluated using Little's MCAR test and addressed with median imputation. Categorical variables were dummy-coded for compatibility with machine learning algorithms. Three machine learning models—GBR, RF, and ENR—were applied to predict pain severity. Each model was trained using 70% of the data and validated on the remaining 30%, with nested 10-fold cross-validation used to tune hyperparameters and minimize overfitting.

Two model versions were compared: a baseline model including only sociodemographic and clinical predictors, and an extended model that incorporated sleep parameters. Model performance was evaluated using $R^2$ to quantify explained variance in pain scores. Sensitivity analyses were conducted to assess consistency of results across the three machine learning approaches.

## Results

The final sample included 556 adults (mean age = 46.77 ± 16.30 years), with 55.5% identifying as female and 55.2% as African American. Participants reported an average SF-36 pain score of 65.87 (SD = 25.90), indicating low to moderate pain severity with notable individual variation.

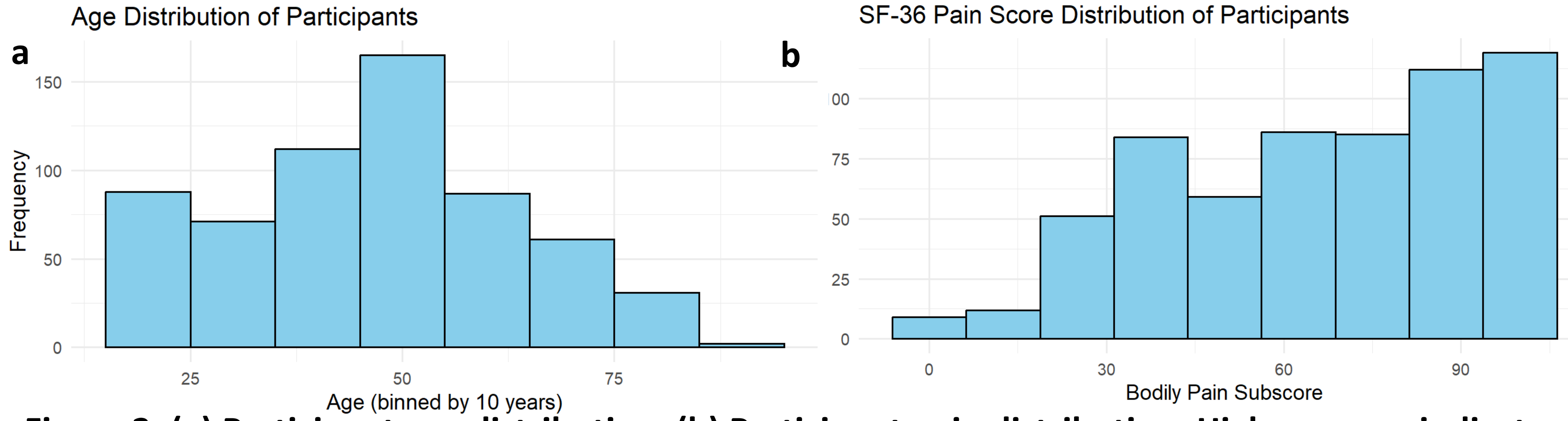


**Figure 2. (a) Participant age distribution. (b) Participant pain distribution. Higher scores indicate less pain.**

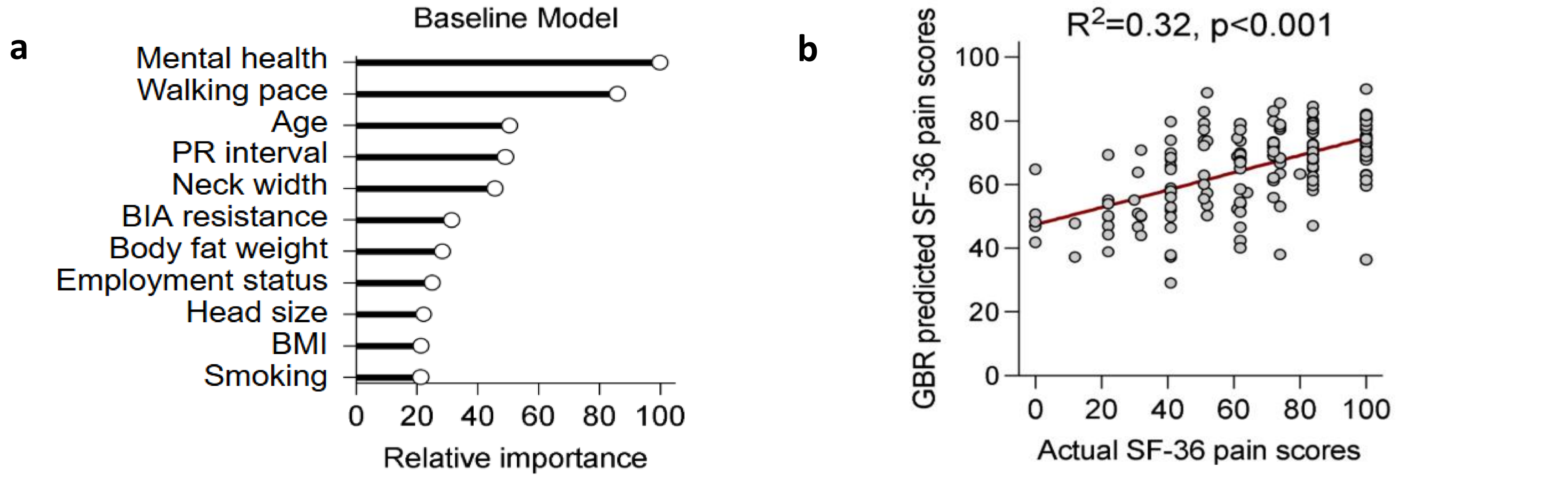


**Figure 3. (a) Key predictors in the baseline GBR included mood status, walking pace, age, PR interval, and employment status. (b) The baseline model explained 32% of the variance in pain scores ($R^2$ = 0.32). These findings were largely consistent across Random Forest (RF) and Elastic Net Regression (ENR) models.**

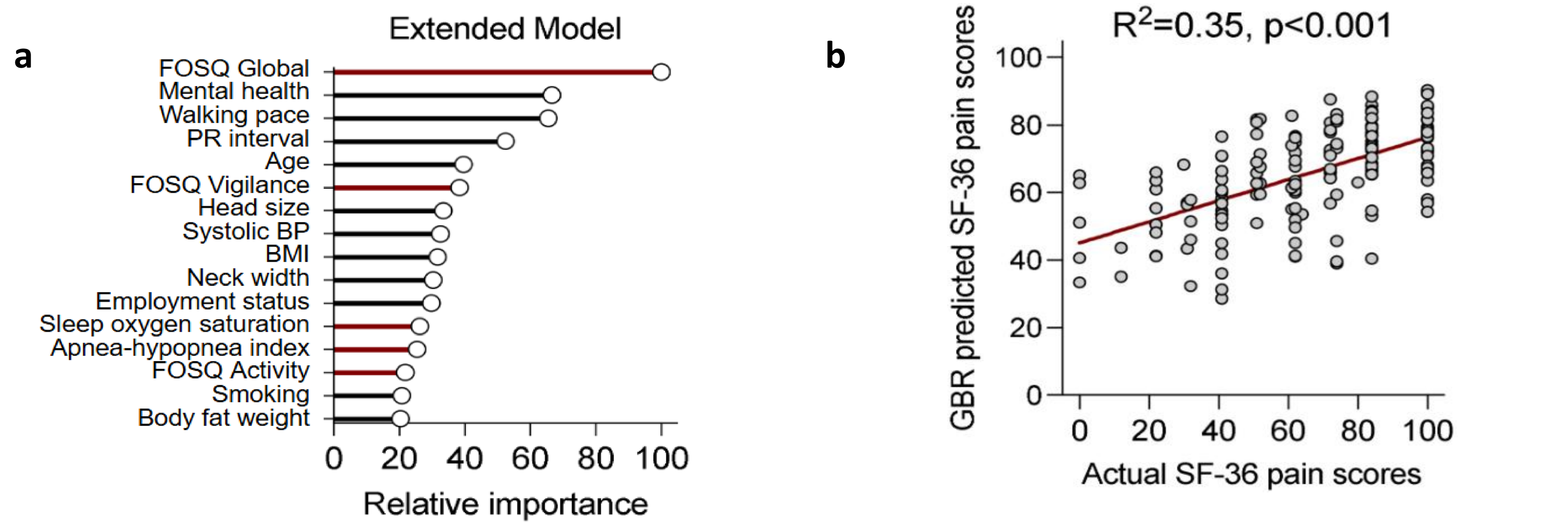


**Figure 4. (a) In the extended model, self-reported sleep factors (FOSQ activity and vigilance scores) were among the most predictive of pain severity, surpassing objective metrics such as apnea-hypopnea index (AHI). (b) The addition of sleep variables improved model performance by about 9.4% ($R^2$ = 0.35).**

## Results II

Among the predictors identified in the baseline GBR model, all features except neck circumference, BIA resistance, head circumference, and total years of smoking were consistently identified as necessary in both the RF and ENR models.

In the extended GBR model, FOSQ activity and vigilance scores were replicated as robust sleep predictors of pain in both the ENR and RF models, emphasizing the importance of self-reported sleep quality over objective sleep disturbances such as AHI.

## Conclusion

- Subjective sleep impairments—particularly difficulties with daily activities and vigilance—were the strongest sleep-related predictors of clinical pain severity.
- Objective sleep measures, such as the apnea-hypopnea index (AHI), showed limited predictive value compared to self-reported sleep quality.
- These findings support growing evidence that perceived sleep quality plays a more critical role in pain outcomes than objective metrics.
- Sleep variables, together with age and clinical factors, explained ~35% of the variance in pain severity.
- Due to the observational nature of this secondary analysis, causal relationships between sleep and pain cannot be inferred.
- Further research using multidimensional pain assessments and more comprehensive sleep metrics (e.g., sleep spindles) is needed to validate these findings.

## Bibliography


[1] Finan, P. H., Goodin, B. R. & Smith, M. T. The association of sleep and pain: an update and a path forward. The journal of pain 14, 1539-1552 (2013).

[2] Bilterys, T. et al. Relationship, differences, and agreement between objective and subjective sleep measures in chronic spinal pain patients with comorbid insomnia: A cross-sectional study. Pain 164, 2016-2028 (2023).

[3] Redline, S. et al. The familial aggregation of obstructive sleep apnea. American journal of respiratory and critical care medicine 151, 682-687 (1995).

[4] Zhang, G.-Q. et al. The National Sleep Research Resource: towards a sleep data commons. Journal of the American Medical Informatics Association 25, 1351-1358 (2018).


## Acknowledgements


The Cleveland Family Study (CFS) was supported by grants from the National Institutes of Health (HL46380, M01 RR00080-39, T32-HL07567, RO1-46380). The National Sleep Research Resource was supported by the National Heart, Lung, and Blood Institute (R24 HL114473, 75N92019R002). The PI's start-up funds (YW) from the University of Maryland School of Nursing supported this secondary data analysis.