RESEARCH ARTICLE

Predictors of Subjective Cognitive Decline in Adults Aged 45 Years and Older: Findings from the 2022 BRFSS

Juwon Lee

Groton School, 282 Farmers Row, Groton, Massachusetts, USA; japple3.141592@gmail.com

ABSTRACT: Subjective cognitive decline (SCD) is a frequent early symptom of Alzheimer's disease and other dementias. This study used national survey data and machine learning to identify factors linked to SCD. Adults aged ≥45 years who completed the 2022 Behavioral Risk Factor Surveillance System cognitive decline module were included (n=62,743). The outcome was a yes/no report of worsening confusion or memory loss. Thirty-seven variables were selected from multiple domains and analyzed using four models: least absolute shrinkage and selection operator (LASSO), random forest (RF), extreme gradient boosting (XGB), and artificial neural network (ANN). A total of 11.0% of respondents reported SCD. Model accuracy was 0.818–0.834, and the area under the receiver operating characteristic curve (AUROC) ranged from 0.751 to 0.755. Aggregated feature importance ranked depression highest, followed by chronic obstructive pulmonary disease, lack of emotional support, unemployment, arthritis, stroke, low income, no physical activity, chronic heart disease, and short sleep duration. Logistic regression confirmed these associations, with depression showing an odds ratio of 2.46 (95% CI 2.23–2.72). SCD was associated with several modifiable factors. Interventions addressing mental health, chronic disease management, social support, and adequate sleep may reduce the risk and the burden of cognitive decline.

KEYWORDS: Medical and Health Sciences, Public Health, Subjective Cognitive Decline, Depression, Chronic Diseases and Social Support.

Introduction

Subjective cognitive decline (SCD), a frequent early symptom of Alzheimer's disease (AD) and other dementias, is reported by 9.6% of adults aged ≥45 years in the United States.[1] SCD is defined as a self-reported observation of worsening confusion or memory loss and is usually identified by the Subjective Cognitive Disorder Questionnaire (SCD-Q). While the incidence of SCD is on a steady rise among older adults, the number of individuals who experience age-associated cognitive decline is expected to grow. The estimated lifetime risk of AD for a 45-year-old is currently 19.5% for women and 10.3% for men.[2] In 2023, the national cost for Alzheimer's and other dementias was an estimated $345 billion, and this figure is projected to rise to nearly $1 trillion by 2050, as stated in the Alzheimer's Association's 2023 Facts and Figures report.[3] In addition to the financial burden, the disease places a significant amount of emotional and physical distress on caregivers.[4] Given such profound impact of cognitive diseases on individuals and their families, there is an urgent need to identify modifiable risk factors that could lead to better prevention and management strategies.

A series of studies have analyzed national survey data, such as the National Health and Nutrition Examination Survey (NHANES) and the Behavioral Risk Factor Surveillance System (BRFSS), to investigate the predictors of cognitive decline. A past NHANES study suggested that certain dietary factors, such as an anti-inflammatory diet like the Mediterranean diet, could have a positive impact on cognitive performance.[5] Research using BRFSS data has identified multiple modifiable factors associated with SCD; state-level analyses revealed that poverty, diabetes, and hypertension were strongly correlated with SCD and functional impairments.[6] Smoking also showed a graded association with SCD, with the highest prevalence in current smokers compared to never smokers.[7] Sleep problems, particularly comorbid insomnia and sleep apnea, were related to almost double the odds of SCD compared to healthy sleep patterns.[8] Additionally, insufficient social and emotional support was associated with a higher prevalence of SCD.[9] These findings demonstrate the value of BRFSS data for identifying modifiable risk factors at a population level; however, the previous studies were limited by outdated datasets and a narrow focus on specific factors rather than a comprehensive evaluation of a wide range of potential predictors.

In the past decade, machine learning techniques have provided new analytic ways for large-scale survey datasets, such as NHANES and BRFSS, which enable the identification of complex relationships and the discovery of unexpected predictors in the field of epidemiology. By leveraging methods such as clustering, predictive modeling, and feature selection, machine learning can process high-dimensional data and reveal nuanced patterns often missed by traditional statistical approaches.[10] This capability is particularly valuable in exploring the interactions between socioeconomic, behavioral, and medical factors and their collective impact on cognitive health, which can lead to more tailored and effective interventions.[11]

Our study sought to apply various machine learning algorithms to identify risk factors associated with SCD, using data from the 2022 BRFSS. The findings of this research were intended to guide interventions and help reduce the burden of cognitive decline on society as well as individuals and fam-

DOI: 10.36838/IJHSR812.9

ilies. To achieve this, we selected a broad range of variables from multiple domains and applied different machine learning algorithms to determine the factors linked to SCD. Feature importance was calculated for each algorithm, and the scores were combined to identify variables that were consistently significant across all models.

Methods

Study population:

The BRFSS is a nationwide, population-based survey that includes a large sample size and collects a wide range of demographic, socioeconomic, and health-related data. Among the most recent publicly available BRFSS datasets, the 2022 BRFSS provides the widest coverage of non-clinical risk factors, including socioeconomic indicators and health-related behavioral factors. Therefore, it was selected for this study to examine the associations between memory loss and modifiable, non-clinical risk factors among U.S. adults. In the 2022 BRFSS dataset (n = 445,133), only adults aged 45 years and older (n = 311,176) were eligible for the Cognitive Decline Module. Among these respondents, adults who answered either "yes" or "no" to the question, "*During the past 12 months, have you experienced confusion or memory loss that is happening more often or is getting worse?*" were included in this analysis. The final study population consisted of 62,743 adults.

Variable selection:

In the study, the outcome variable, SCD, was a binary dependent variable determined through a yes/no response to the CMEMLOS questionnaire item. During the variable selection process, we excluded variables that were directly related to sampling methods, inherently sex-specific (e.g., mammography, prostate-specific antigen testing), consequential to SCD (e.g., decision-making difficulties and self-reported health status), or derived or overlapping variables of other variables under consideration. As a result, a total of 37 variables were selected. These variables were classified into six categories (1) Demographic variables: sex (male or female), age groups (45–65 or 65+), race (non-Hispanic white or others), living arrangements (living alone or together), marital status (married, separated/divorced/deceased, and unmarried), veteran status, and living with children; (2) socioeconomic variables: income levels (<35,000, 35,000-70,000, or 70,000+), educational level (graduated college, some college, high school or less than high school), housing type (own or rent), and current employment status; (3) medical condition variables: the presence of chronic diseases such as asthma, chronic obstructive pulmonary disease (COPD), chronic kidney disease, cancer, stroke, coronary artery disease/myocardial infarction, diabetes, arthritis, depression, tooth loss, and obesity (body mass index > 25); (4) accessibility to medical care variables: having a personal medical care provider, health insurance, and doctor visit in the preceding year; (5) health-related behavioral variables: heavy drinking (more than 14 drinks per week for males and more than 7 drinks per week for females), smoking status (never smoker, former smoker or current smoker), flu vaccination, frequency of intended physical activity in the past 30 days, and sleep duration; and (6) social and emotional well-being variables: presence of adverse childhood experiences, frequency of feeling socially isolated (always/usually or sometimes/never), and the availability of social and emotional support (always/usually or sometimes/never).

Data preprocessing:

To address multicollinearity, predictors were evaluated using the variance inflation factor (VIF). Variables with VIF values greater than 10 were removed iteratively. Missing data was handled using multivariate imputation by chained equations (MICE), which imputes missing values while accounting for relationships between variables.[12] Variables with more than 20% missing responses were excluded from imputation to reduce bias.[13] In this process, three variables were excluded: living with children (high multicollinearity), frequency of feeling isolated, and COVID-19 vaccination status (high proportion of missing responses).

Machine learning models and feature importance aggregation:

Four advanced machine learning models were used to evaluate predictors and determine their importance. These models included least absolute shrinkage and selection operator (LASSO) regression, a random forest classifier (RF), an extreme gradient boosting algorithm (XGB), and an artificial neural network (ANN). LASSO is a linear model with L1 regularization that selects features by shrinking irrelevant coefficients to zero.[14] Both XGB and RF are tree-based models capable of exploring non-linear relationships. XGB uses boosting, a sequential process where trees are constructed to correct errors from previous ones,[15] whereas RF uses bagging. In this parallel process, multiple decision trees are independently built on bootstrapped subsets of the data.[16] ANN is a nonlinear model composed of interconnected layers of neurons that can learn complex patterns.[17] Each model was trained and validated using a 5-fold stratified cross-validation procedure to ensure robustness and reduce overfitting. Model performance was evaluated using accuracy, area under the receiver operating characteristic curve (AUROC), and F-1 score.

Feature importance was primarily calculated using model-specific methods if available intrinsically. For LASSO, the absolute values of the coefficients were used, as it performs variable selection by shrinking irrelevant coefficients to zero. In XGB and RF, split-based metrics were applied, with gain used for XGB and mean decrease in impurity for RF, to evaluate how features improved model splits. In contrast to the other three algorithms, ANN lacks internal measures of feature importance; therefore, permutation-based importance scores were calculated to measure the effect of each feature on model performance.[18] To ensure comparability, the importance scores were normalized by dividing each score by the total importance within each model. Then the normalized scores were summed across models to produce aggregated feature importances to identify variables consistently important across models. Additionally, to provide an interpretable measure of the relationships between predictors and the outcome (i.e.,

DOI: 10.36838/IJHSR812.9

SCD), the ten most significant predictors identified through machine learning were further evaluated using logistic regression and presented as odds ratios (ORs) with 95% confidence intervals (CIs).

Statistical analysis:

All analyses were conducted using Python (version 3.10.11). Data preprocessing was performed using pandas (version 2.2.3) and NumPy (version 1.25.2). LASSO and RF were implemented using scikit-learn (version 1.7.2). XGB was implemented using XGBoost (version 3.1.3), and ANN was implemented using PyTorch (version 2.4.1). GLM analyses were performed using statsmodels (version 0.14.3). A P-value of less than 0.05 was considered statistically significant.

Results

Study population characteristics:

The study included 62,743 adults aged 45 years and older, representing 14.1% of survey respondents. Among them, 6,912 (11.0%) reported SCD. Compared with adults without SCD, those with SCD were more likely to be aged ≥65 years, non-White or Hispanic, unmarried, veterans, and living alone (all $p<0.001$; Table 1). They had lower income, lower educational attainment, lower home ownership, and higher unemployment ($p<0.001$). Depression (43.7% vs. 16.5%), asthma, chronic obstructive pulmonary disease, chronic kidney disease, cancer, coronary heart disease, stroke, diabetes, and arthritis were more prevalent in the SCD group. In contrast, obesity showed no significant difference ($p=0.379$). Adults with SCD were less likely to report having a personal healthcare provider or health insurance and more likely to have shorter sleep duration, lower physical activity, and current smoking (all $p<0.001$). Adverse childhood experiences, lack of emotional support, and frequent social isolation were also more common in the SCD group (all $p<0.01$).

Table 1: Characteristics of adults with and without subjective cognitive decline (SCD). Adults with SCD had a higher prevalence of depression and multiple chronic diseases, along with socioeconomic disadvantage and lower levels of protective health behaviors.

	Adults without SCD (n = 55,831, 89.0%)	Adults with SCD (n=6,912, 11.0%)	P-value
Demographic variables			
Sex, male	24,894 (44.6%)	3,113 (45.0%)	0.486
Age, 65+	30,427 (54.5%)	4,003 (57.9%)	<0.001
Race, Non-white or Hispanic	7,103 (13.1%)	1,022 (15.3%)	<0.001
Veteran, yes	8,453 (15.2%)	1,324 (19.2%)	<0.001
Marital status, unmarried	21,990 (40.4%)	3,511 (52.4%)	<0.001
Living alone, yes	17,499 (31.5%)	2,587 (37.7%)	<0.001
Socioeconomic variables			
Income level			<0.001
<$35,000	18,843 (41.7%)	1,386 (24.7%)	
$35,000–70,000	14,519 (32.1%)	1,699 (30.3%)	
>$70,000	11,846 (26.2%)	2,527 (45.0%)	
Level of education			<0.001
Graduated college	24,947 (44.8%)	2,335 (33.9%)	
Some college	15,312 (27.5%)	2,085 (30.3%)	
High school	12,956 (23.3%)	1,915 (27.8%)	
Less than high school	2,464 (4.4%)	551 (8.0%)	
Home ownership, no	8,242 (14.8%)	1,715 (25.0%)	<0.001
Employment status, unemployed	32,801 (59.1%)	5,239 (76.2%)	<0.001
Medical conditions			
Depression	9,191 (16.5%)	2,991 (43.7%)	<0.001
Asthma	7,380 (13.3%)	1,489 (21.9%)	<0.001
Chronic obstructive pulmonary disease	5,333 (11.1%)	1,455 (30.7%)	<0.001
Chronic renal disease	3,199 (5.8%)	732 (10.7%)	<0.001
Cancer	8,940 (16.1%)	1,456 (21.3%)	<0.001
Chronic heart disease	6,400 (11.6%)	1,511 (22.3%)	<0.001
Stroke	2,874 (5.2%)	884 (12.9%)	<0.001
Diabetes	9,715 (17.4%)	1,873 (27.2%)	<0.001
Arthritis	24,728 (44.5%)	4,360 (63.6%)	<0.001
Teeth removal	29298 (53.9%)	4535 (67.8%)	<0.001
Obesity	37,077 (70.7%)	4,693 (71.2%)	0.379
Access to medical care			
Personal health care provider, yes	3,612 (6.5%)	399 (5.8%)	0.029
Health insurance, no	1,501 (2.8%)	217 (3.3%)	0.024
Health-related behavioral variables			
Health check-up within 1 year, no	7,066 (12.8%)	731 (10.7%)	<0.001
Heavy drinking, yes	3,531 (6.4%)	444 (6.6%)	0.731
Smoking status			<0.001
Never	31,639 (57.1%)	3,094 (45.2%)	
Former	17,954 (32.4%)	2,658 (38.8%)	
Current	5,802 (10.5%)	1,099 (16.0%)	
COVID-19 vaccination, no	2,412 (13.9%)	277 (12.9%)	0.228
Flu vaccination, no	21,661 (39.0%)	2,632 (38.4%)	0.340
Physical activity, no	13,279 (23.9%)	2,701 (39.2%)	
Sleep duration			<0.001
6 hours or less	15,739 (28.5%)	2,716 (40.1%)	
7 hours	17,402 (31.5%)	1,377 (20.3%)	
8 hours or more	22,090 (40.0%)	2,678 (39.6%)	
Social and emotional well-being variables			
Adverse childhood experience, yes	16,245 (29.1%)	2,132 (30.8%)	0.003

Performance of the prediction model:

The performance of the four machine learning models was evaluated for predicting SCD using accuracy, AUROC, and F1-score (Table 2). Overall, model performance was comparable across algorithms. Accuracy ranged from 0.818 to 0.834, and AUROC values ranged from 0.751 to 0.757. F1-scores were similar across models, ranging from 0.359 to 0.363.

Identification of significant predictors across all models:

Figure 1 shows both consistent and model-specific patterns in variable importance measures. Depression was the most important predictor, ranked first across all models. Lack of emotional support and unemployment were also consistently ranked among the top five across models. Arthritis, stroke, low income, and no physical activity were generally ranked within the top 10, indicating their relevance across multiple approaches. Some variables displayed model-specific differences. For example, COPD was ranked second in XGB and RF but had lower importance in LASSO (rank 17) and ANN (rank 14). Similarly, short sleep duration was ranked 9–10 in XGB and RF but lower in LASSO and ANN (ranks 19 and 10). Variables such as no flu vaccination, no health insurance, and heavy alcohol consumption were consistently among the least important predictors across all models.

The aggregated feature importance analysis identified depression as the most important variable, with a higher importance score (0.7625) compared to other variables (Figure 2). COPD, lack of emotional support, unemployment, and arthritis were among the top five variables, with scores ranging from 0.2200 to 0.3127. Other predictors in the top 10 included stroke, low income, no physical activity, heart disease, and short sleep duration, with scores ranging from 0.1065 to 0.1854.

DOI: 10.36838/IJHSR812.9

Table 2: Comparison of performance to predict subjective cognitive decline among the four machine-learning models. All four models demonstrated comparable predictive performance, indicating stable and reliable identification of factors associated with subjective cognitive decline.

	Accuracy	AUROC	F1-score
LASSO	0.827 (0.811–0.842)	0.752 (0.748–0.757)	0.360 (0.354–0.367)
XGB	0.834 (0.832–0.836)	0.757 (0.752–0.761)	0.363 (0.355–0.370)
RF	0.832 (0.829–0.834)	0.751 (0.748–0.754)	0.359 (0.353–0.366)
ANN	0.818 (0.798–0.837)	0.755 (0.751–0.760)	0.361 (0.354–0.367)

Data are presented as means with 95% confidence intervals in parentheses.

Abbreviations: AUROC, area under the receiver operating curve; LASSO, least absolute shrinkage and selection operator; XGB, extreme gradient boosting algorithm; RF, random forest; ANN, artificial neural network.

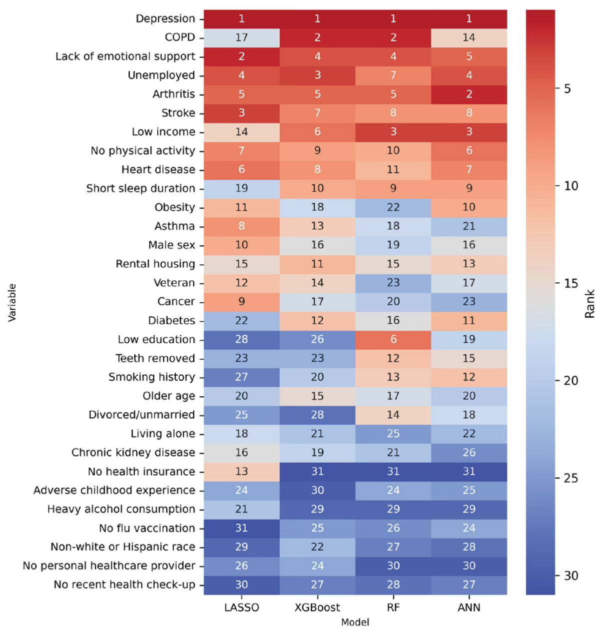

Figure 1: Variable importance ranks across models. Several factors, such as depression and lack of emotional support, maintained high importance across all models, whereas others varied substantially among models. This pattern indicates both shared and model-specific determinants of subjective cognitive decline. Darker red indicates higher importance (lower rank), and darker blue indicates lower importance (higher rank). COPD = chronic obstructive pulmonary disease; LASSO = least absolute shrinkage and selection operator; XGB = extreme gradient boosting algorithm; RF = random forest; ANN = artificial neural network.

Confirmation of significant predictors using logistic regression:

Logistic regression was conducted to evaluate the association between the top 10 variables identified through aggregated feature importance and SCD (Table 3). Depression was the strongest predictor (OR: 2.461; 95% CI: 2.229–2.717). Lack of emotional support (1.936; 1.773–2.114) and stroke (1.698; 1.556–1.854) also showed strong associations. COPD, arthritis, unemployment, and chronic heart disease showed moderate associations with ORs ranging from 1.289 to 1.516, all statistically significant ($p<0.001$). Low income, no physical activity, and short sleep duration were also significant, with smaller effect sizes (OR, 1.119–1.361).

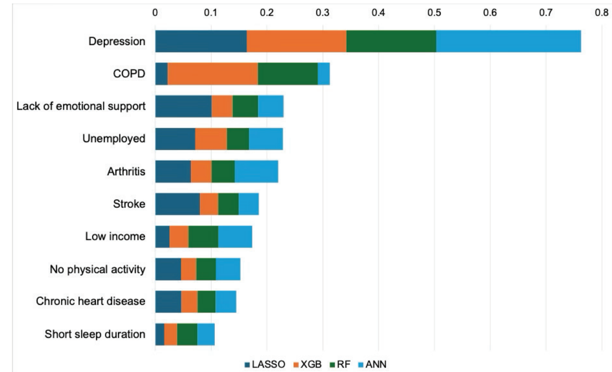

Figure 2: Top 10 variables by aggregated feature importance across the four advanced machine-learning models. Depression showed the highest overall importance, followed by COPD, lack of emotional support, unemployment, and arthritis. Stroke, low income, no physical activity, chronic heart disease, and short sleep duration also ranked among the top ten. COPD = chronic obstructive pulmonary disease; LASSO = least absolute shrinkage and selection operator; XGB = extreme gradient boosting algorithm; RF = random forest; ANN = artificial neural network.

Table 3: Odds ratios estimated by logistic regression for the top 10 predictors identified by aggregated feature importance. These results indicate that the highest-importance factors identified by the machine-learning models were statistically significant predictors of subjective cognitive decline.

Variable	Odds ratio (95% Confidence intervals)	P-value
Depression	2.46 (2.23–2.72)	<0.001
Chronic obstructive pulmonary disease	1.29 (1.17–1.42)	<0.001
Lack of emotional support	1.94 (1.77–2.11)	<0.001
Arthritis	1.46 (1.38–1.54)	<0.001
Unemployment	1.52 (1.42–1.62)	<0.001
Stroke	1.70 (1.56–1.85)	<0.001
Low-income level	1.23 (1.18–1.27)	<0.001
No physical activity	1.36 (1.29–1.44)	<0.001
Chronic heart disease	1.45 (1.35–1.55)	<0.001
Short sleep duration	1.12 (1.08–1.15)	<0.001

Discussion

Main Findings:

This study analyzed predictors of subjective cognitive decline (SCD) among adults aged 45 years and older using the 2022 BRFSS. The prevalence of SCD was 11.0%, similar to prior BRFSS reports.[19] Depression was the most significant predictor. Other significant predictors included lack of emotional support, COPD, arthritis, unemployment, stroke, low income, no physical activity, chronic heart disease, and short sleep duration.

Depression and Dementia Risk:

A series of large population-based studies has demonstrated that depression significantly increases the risk of cognitive decline and dementia.[20-22] Additionally, each episode of depression is associated with an increased risk of all types of dementia, with a 14% increase per episode.[23] Multiple biological mechanisms are known to link depression to dementia, including steroid and neurotransmitter dysregulation, hippocampal atrophy, inflammation, increased deposition of beta-amyloid, and cerebrovascular changes.[24,25] Animal and human studies have suggested that antidepressant medication may reduce dementia risk by attenuating hippocampal structural damage, enhancing neuroplasticity, and reducing amyloid and tau pathology.[26] Depression should therefore be considered an important modifiable risk factor in strategies to prevent or delay dementia.

DOI: 10.36838/IJHSR812.9

Chronic Medical Conditions:

Managing chronic medical conditions such as arthritis, COPD, and stroke may reduce the risk of SCD.[27-29] These conditions are associated with a higher prevalence of SCD and may contribute to cognitive impairment through mechanisms including inflammation and vascular changes. Addressing chronic diseases may also alleviate depression, an established risk factor for SCD. Stroke, identified as an important factor in our analysis, is a chronic condition that can be both prevented and managed. Physical activity has also been reported to influence the relationship between depression and cognition.[30] Recent studies have examined this association further. A randomized clinical trial showed that cognitive remediation combined with transcranial direct current stimulation slowed cognitive decline in older adults with remitted major depressive disorder or mild cognitive impairment.[31] These findings suggest that targeted interventions may mitigate the cognitive effects of depression, although the link between depression treatment and dementia prevention remains complex and requires further study.

Emotional and Social Support:

Emotional support is a crucial factor for mental and physical well-being. Studies have demonstrated that adequate emotional support can improve cognitive function and reduce memory decline in different populations.[32,33] Social connections may also protect against depression, which is closely related to SCD.[34] In addition, individuals with stronger social engagement often maintain positive relationships, which may lower the risk of dementia.[35] In this context, strengthening emotional and social support may help reduce the risk of SCD and dementia.

Sleep and Cognitive Function:

Inadequate sleep duration was identified as a predictor of SCD in our analysis. Sleep is a modifiable behavior and may represent a potential target for intervention, unlike predictors such as economic hardship and education level, which are less amenable to change. An inverted U-shaped association has been reported between sleep duration and cognitive function, with both short and long sleep linked to poorer cognitive outcomes and subsequent decline.[31,36] Extreme sleep durations (≤4 or ≥10 hours per night) have been associated with lower baseline cognitive function and with faster decline during follow-up. These findings underscore the need to monitor and adjust sleep duration. At the same time, further studies are required to clarify underlying mechanisms and to assess the role of sleep quality as well as duration.

Data Processing Considerations:

Regarding the preprocessing of our data, the cohort was highly imbalanced, with only 11% of respondents classified as having SCD. In addition, the overall F1-scores were relatively low across models, which probably reflected the difficulties in accurately identifying the minority SCD group, i.e., participants with SCD, despite high overall accuracy and AUROC. Oversampling techniques such as the Synthetic Minority Over-sampling Technique combined with Tomek Links (SMOTE-Tomek) are often used in previous studies to address imbalance by generating synthetic minority samples.[37] While these methods can enhance predictive performance, they may also distort the original associations between variables and outcomes, which was a central concern in this study.[38] Our primary aim was not to optimize model performance but to identify and evaluate predictors of SCD. To this end, we aggregated feature importance scores across models, which helped reduce bias inherent to any single method and provided more robust rankings of predictors. This approach also has limitations, as normalization and aggregation may dilute model-specific signals and obscure predictors that are important in one model but not in others. In some cases, consensus across models may overemphasize shared predictors while underrepresenting unique insights from individual models.

Limitations:

This study has several limitations. The cross-sectional design of the BRFSS limits causal inference. Reliance on self-reported measures of both SCD and predictor variables may also introduce recall bias. Excluding variables with a high proportion of missing responses, such as adverse childhood experiences, may have reduced the scope of the analysis. Future research should use longitudinal designs to clarify causal pathways and validate these findings.

■ Conclusion

In this study, we identified key predictors of SCD using machine learning models applied to 2022 BRFSS data. Depression was the most consistent risk factor, and chronic medical conditions, inadequate emotional support, and insufficient sleep were also significant contributors, indicating the multifactorial nature of SCD. These results clarify the associations between multiple predictors and SCD. Combining statistical approaches with machine learning may improve the interpretability of large survey data and inform public health strategies.

■ References

1. Wooten, K. G.; McGuire, L. C.; Olivari, B. S.; Jackson, E. M. J.; Croft, J. B., Racial and Ethnic Differences in Subjective Cognitive Decline - United States, 2015-2020. *MMWR Morb Mortal Wkly Rep* **2023**, *72* (10), 249-255.
2. Murman, D. L., The Impact of Age on Cognition. *Semin Hear* **2015**, *36* (3), 111-21.
3. 2023 Alzheimer's disease facts and figures. *Alzheimer's & Dementia* **2023**, *19* (4), 1598-1695.
4. Zhang, J.; Wang, J.; Liu, H.; Wu, C., Association of dementia comorbidities with caregivers' physical, psychological, social, and financial burden. *BMC Geriatrics* **2023**, *23* (1), 60.
5. Xie, S.; Li, Z.; Yao, Q.; Zhang, Y.; Ou, Y., Adherence to Mediterranean diet and female urinary incontinence: Evidence from the NHANES database. *PLoS One* **2024**, *19* (10), e0311771.
6. de Havenon, A.; Stulberg, E. L.; Littig, L.; Wong, K. H.; Sarpong, D.; Li, V.; Sharma, R.; Falcone, G. J.; Williamson, J. D.; Pajewski, N. M.; Gottesman, R. F.; Brickman, A. M.; Sheth, K. N., Socioeconomic and medical determinants of state-level subjective cognitive

DOI: 10.36838/IJHSR812.9

decline in the United States. *Alzheimer's Dement* **2024**, *20* (11), 7567-7579.

7. Rajczyk, J. I.; Ferketich, A.; Wing, J. J., Relation Between Smoking Status and Subjective Cognitive Decline in Middle Age and Older Adults: A Cross-Sectional Analysis of 2019 Behavioral Risk Factor Surveillance System Data. *J Alzheimers Dis* **2023**, *91* (1), 215-223.
8. Huang, J.; Spira, A. P.; Perrin, N. A.; Ellis, A.; Hsu, E. C.; Kaufmann, C. N.; Li, J., Latent classes of sleep problems and subjective cognitive decline among middle-aged and older adults in the United States. *Arch Gerontol Geriatr* **2025**, *129*, 105657.
9. Weng, X.; George, D. R.; Jiang, B.; Wang, L., Association between subjective cognitive decline and social and emotional support in US adults. *American Journal of Alzheimer's Disease & Other Dementias®* **2020**, *35*, 1533317520922392.
10. Ngiam, K. Y.; Khor, W., Big data and machine learning algorithms for health-care delivery. *The Lancet Oncology* **2019**, *20* (5), e262-e273.
11. Topol, E. J., High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* **2019**, *25* (1), 44-56.
12. Van Buuren, S.; Groothuis-Oudshoorn, K., mice: Multivariate imputation by chained equations in R. *Journal of statistical software* **2011**, *45*, 1-67.
13. Jakobsen, J. C.; Gluud, C.; Wetterslev, J.; Winkel, P., When and how should multiple imputation be used for handling missing data in randomised clinical trials–a practical guide with flowcharts. *BMC Medical Research Methodology* **2017**, *17* (1), 162.
14. Tibshirani, R., Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **1996**, *58* (1), 267-288.
15. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, August 13-17, 2016; pp 785–794. DOI: 10.1145/2939672.2939785.
16. Breiman, L., Random forests. *Machine learning* **2001**, *45* (1), 5-32.
17. LeCun, Y.; Bengio, Y.; Hinton, G., Deep learning. *Nature* **2015**, *521* (7553), 436-444.
18. Khan, A.; Ali, A.; Khan, J.; Ullah, F.; Faheem, M., Using Permutation-Based Feature Importance for Improved Machine Learning Model Performance at Reduced Costs. *IEEE Access* **2025**.
19. Olivari, B. S.; Baumgart, M.; Taylor, C. A.; McGuire, L. C., Population measures of subjective cognitive decline: A means of advancing public health policy to address cognitive health. *Alzheimer's & Dementia: Translational Research & Clinical Interventions* **2021**, *7* (1), e12142.
20. Mirza, S. S.; Wolters, F. J.; Swanson, S. A.; Koudstaal, P. J.; Hofman, A.; Tiemeier, H.; Ikram, M. A., 10-year trajectories of depressive symptoms and risk of dementia: a population-based study. *Lancet Psychiatry* **2016**, *3* (7), 628-35.
21. Elser, H.; Horváth-Puhó, E.; Gradus, J. L.; Smith, M. L.; Lash, T. L.; Glymour, M. M.; Sørensen, H. T.; Henderson, V. W., Association of Early-, Middle-, and Late-Life Depression With Incident Dementia in a Danish Cohort. *JAMA Neurol* **2023**, *80* (9), 949-958.
22. Li, C.; Wang, W.; Wei, Y. X.; Lu, K.; Wang, J. Y.; Yao, M. H.; Du, Q. Q.; Li, X. L.; Li, S.; Tian, X. Y.; Yin, F.; Zhang, T.; Ma, Y., Association between cognitive decline and depression in middle-aged and older adults: Findings from six large cohorts in different countries. *J Affect Disorders* **2025**, *371*, 215-223.
23. Dotson, V. M.; Beydoun, M. A.; Zonderman, A. B., Recurrent depressive symptoms and the incidence of dementia and mild cognitive impairment. *Neurology* **2010**, *75* (1), 27-34.
24. Dafsari, F. S.; Jessen, F., Depression-an underrecognized target for prevention of dementia in Alzheimer's disease. *Transl Psychiat* **2020**, *10* (1).
25. Byers, A. L.; Yaffe, K., Depression and risk of developing dementia. *Nat Rev Neurol* **2011**, *7* (6), 323-331.
26. Stephan, B. C.; Brayne, C.; Savva, G. M.; Matthews, F. E., Occurrence of medical co-morbidity in mild cognitive impairment: implications for generalisation of MCI research. *Age and ageing* **2011**, *40* (4), 501-507.
27. Jinshil, K.; Eunok, P.; Minjeong, A., The cognitive impact of chronic diseases on functional capacity in community-dwelling adults. *Journal of Nursing Research* **2019**, *27* (1), e3.
28. Bakouni, H.; Guerra, S. G.; Chudzinski, V.; Berbiche, D.; Vasiliadis, H. M., One-year prospective study on the presence of chronic diseases and subsequent cognitive decline in older adults. *J Public Health-Uk* **2017**, *39* (4), E170-E178.
29. Singh, B.; Mielke, M. M.; Parsaik, A. K.; Cha, R. H.; Roberts, R. O.; Scanlon, P. D.; Geda, Y. E.; Christianson, T. J.; Pankratz, V. S.; Petersen, R. C., A Prospective Study of Chronic Obstructive Pulmonary Disease and the Risk for Mild Cognitive Impairment. *Jama Neurology* **2014**, *71* (5), 581-588.
30. Hu, L.; Smith, L.; Imm, K. R.; Jackson, S. E.; Yang, L., Physical activity modifies the association between depression and cognitive function in older adults. *J Affect Disorders* **2019**, *246*, 800-805.
31. Ma, Y. J.; Liang, L. R.; Zheng, F. F.; Shi, L.; Zhong, B. L.; Xie, W. X., Association Between Sleep Duration and Cognitive Decline. *Jama Netw Open* **2020**, *3* (9).
32. Fratiglioni, L.; Paillard-Borg, S.; Winblad, B., An active and socially integrated lifestyle in late life might protect against dementia. *Lancet Neurol* **2004**, *3* (6), 343-353.
33. Weng, X. R.; George, D. R.; Jiang, B. B.; Wang, L., Association Between Subjective Cognitive Decline and Social and Emotional Support in US Adults. *Am J Alzheimers Dis* **2020**, *35*.
34. Schwarzbach, M.; Luppa, M.; Forstmeier, S.; König, H. H.; Riedel-Heller, S. G., Social relations and depression in late life: A systematic review. *Int J Geriatr Psych* **2014**, *29* (1), 1-21.
35. Wang, H. X.; Karp, A.; Winblad, B.; Fratiglioni, L., Late-life engagement in social and leisure activities is associated with a decreased risk of dementia: A longitudinal study from the Kungsholmen project. *Am J Epidemiol* **2002**, *155* (12), 1081-1087.
36. Leng, Y.; Yaffe, K., Sleep Duration and Cognitive Aging-Beyond a U-Shaped Association. *Jama Netw Open* **2020**, *3* (9).
37. Wang, Z.; Wu, C. H.; Zheng, K. F.; Niu, X. X.; Wang, X. J., SMOTETomek-Based Resampling for Personality Recognition. *IEEE Access* **2019**, *7*, 129678-129689.
38. Buda, M.; Maki, A.; Mazurowski, M. A., A systematic study of the class imbalance problem in convolutional neural networks. *Neural Networks* **2018**, *106*, 249-259.

Author

Juwon Lee is a student at Groton School, passionate about neurological disorders and brain science. She is eager to explore behavioral and biological factors associated with these conditions. She hopes to contribute to research that advances prevention, early detection, effective intervention, and equitable access to health care.