



## OPEN ACCESS

## EDITED BY

Yi Li,  
Cornell University, United States

## REVIEWED BY

Dhiraj Kumar,  
National Eye Institute (NIH), United States  
José Enrique Arriola Infante,  
Hospital de la Santa Creu i Sant Pau, Spain

## \*CORRESPONDENCE

Arrate Muñoz-Barrutia  
✉ mamunozb@ing.uc3m.es

†These authors share first authorship

RECEIVED 12 January 2024

ACCEPTED 04 June 2024

PUBLISHED 26 June 2024

## CITATION

Gaeta AM, Quijada-López M, Barbé F, Vaca R, Pujol M, Minguez O, Sánchez-de-la-Torre M, Muñoz-Barrutia A and Piñol-Ripoll G (2024) Predicting Alzheimer's disease CSF core biomarkers: a multimodal Machine Learning approach. *Front. Aging Neurosci.* 16:1369545. doi: 10.3389/fnagi.2024.1369545

## COPYRIGHT

© 2024 Gaeta, Quijada-López, Barbé, Vaca, Pujol, Minguez, Sánchez-de-la-Torre, Muñoz-Barrutia and Piñol-Ripoll. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Predicting Alzheimer's disease CSF core biomarkers: a multimodal Machine Learning approach

Anna Michela Gaeta<sup>1†</sup>, María Quijada-López<sup>2†</sup>, Ferran Barbé<sup>3,4</sup>, Rafaela Vaca<sup>3</sup>, Montse Pujol<sup>5</sup>, Olga Minguez<sup>3</sup>, Manuel Sánchez-de-la-Torre<sup>4,6</sup>, Arrate Muñoz-Barrutia<sup>2,7\*</sup> and Gerard Piñol-Ripoll<sup>5</sup>

<sup>1</sup>Servicio de Neumología, Hospital Universitario Severo Ochoa, Leganés, Spain, <sup>2</sup>Departamento de Bioingeniería, Universidad Carlos III de Madrid, Leganés, Spain, <sup>3</sup>Group of Translational Research in Respiratory Medicine, Hospital Universitari Arnau de Vilanova and Santa Maria, Institut de Recerca Biomedica de Lleida (IRBLleida), Lleida, Spain, <sup>4</sup>Centro de Investigación Biomédica en Red de Enfermedades Respiratorias (CIBERES), Madrid, Spain, <sup>5</sup>Unitat Trastorns Cognitius, Clinical Neuroscience Research, Institut de Recerca Biomedica de Lleida (IRBLleida), Hospital Universitari Santa Maria, Lleida, Spain, <sup>6</sup>Group of Precision Medicine in Chronic Diseases, Hospital Nacional de Paraplégicos, IDISCAM, Department of Nursing, Physiotherapy and Occupational Therapy, Faculty of Physiotherapy and Nursing, University of Castilla-La Mancha, Toledo, Spain, <sup>7</sup>Departamento de Bioingeniería, Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, Spain

**Introduction:** Alzheimer's disease (AD) is a progressive neurodegenerative disorder. Current core cerebrospinal fluid (CSF) AD biomarkers, widely employed for diagnosis, require a lumbar puncture to be performed, making them impractical as screening tools. Considering the role of sleep disturbances in AD, recent research suggests quantitative sleep electroencephalography features as potential non-invasive biomarkers of AD pathology. However, quantitative analysis of comprehensive polysomnography (PSG) signals remains relatively understudied. PSG is a non-invasive test enabling qualitative and quantitative analysis of a wide range of parameters, offering additional insights alongside other biomarkers. Machine Learning (ML) gained interest for its ability to discern intricate patterns within complex datasets, offering promise in AD neuropathology detection. Therefore, this study aims to evaluate the effectiveness of a multimodal ML approach in predicting core AD CSF biomarkers.

**Methods:** Mild-moderate AD patients were prospectively recruited for PSG, followed by testing of CSF and blood samples for biomarkers. PSG signals underwent preprocessing to extract non-linear, time domain and frequency domain statistics quantitative features. Multiple ML algorithms were trained using four subsets of input features: clinical variables (CLINVAR), conventional PSG parameters (SLEEPVAR), quantitative PSG signal features (PSGVAR) and a combination of all subsets (ALL). Cross-validation techniques were employed to evaluate model performance and ensure generalizability. Regression models were developed to determine the most effective variable combinations for explaining variance in the biomarkers.

**Results:** On 49 subjects, Gradient Boosting Regressors achieved the best results in estimating biomarkers levels, using different loss functions for each biomarker: least absolute deviation (LAD) for the  $A\beta_{42}$ , least squares (LS) for  $p$ -tau and Huber for  $t$ -tau. The ALL subset demonstrated the lowest training errors for all three biomarkers, albeit with varying test performance. Specifically, the SLEEPVAR subset yielded the best test performance in predicting  $A\beta_{42}$ , while the ALL subset most accurately predicted  $p$ -tau and  $t$ -tau due to the lowest test errors.

**Conclusions:** Multimodal ML can help predict the outcome of CSF biomarkers in early AD by utilizing non-invasive and economically feasible variables. The integration of computational models into medical practice offers a promising tool for the screening of patients at risk of AD, potentially guiding clinical decisions.

#### KEYWORDS

Alzheimer's disease, neurodegeneration, biomechanism, diagnosis, therapeutic target, quantitative polysomnographic signal analysis, CSF biomarkers, Machine Learning

## 1 Introduction

Alzheimer's disease (AD) is a neurodegenerative disorder that severely affects cognitive function (Masters et al., 2015). Current statistics on AD paint a concerning picture, with epidemiological forecasts suggesting that by 2060, one in three individuals over the age of 85 in the United States will be affected by the disease (Rajan et al., 2021). AD typically begins to develop in the elderly population a decade or more before clinical diagnosis (Villemagne et al., 2013). This asymptomatic stage is followed by a prodromal phase—or Mild Cognitive impairment (MCI) due to AD—characterized by preserved baseline functionality despite cognitive decline, which precedes the onset of established dementia (Masters et al., 2015). Traditionally, a definitive diagnosis of AD is established post-mortem via histological examination, revealing the hallmark neuropathological features of the disease, such as extracellular amyloid plaques and neurofibrillary tangles (Masters et al., 2015; Jack et al., 2018). Remarkably, emerging evidence strongly correlates CSF amyloid $\beta$  (A $\beta$ 42), phospho-tau (p-tau), total-tau (t-tau) levels and AD neuropathological lesions, establishing CSF biomarkers as critical for early AD detection with high diagnostic accuracy (Masters et al., 2015; Jack et al., 2018). It is known that AD progresses through multiple stages characterized by neuropathological advancements throughout the brain, with amyloid pathology preceding tau pathology, followed by neurodegeneration (Jack et al., 2018). In this context, the National Institute on Aging and Alzheimer's Association (NIA-AA) guidelines introduced in the year 2018 the AT(N) criteria based on core fluid and neuroimaging biomarkers (Jack et al., 2018), thus redefining the concept of the “Alzheimer's continuum” to represent the biological as well as the clinical progression of the disease. This paradigm shift arises from evidence indicating that the AD continuum encompasses a prolonged asymptomatic phase characterized by sequential alterations in biomarkers (Aisen et al., 2017; Jack et al., 2018). Biomarkers are grouped into three categories based on the pathologic processes each one measures, encompassing amyloid deposition (A), tau aggregates (T), and neurodegeneration (N), with CSF A $\beta$ 42 and amyloid PET for amyloid pathology (A); CSF p-tau and tau-PET of tau pathology (T); atrophy in MRI, hypometabolism in FDG-PET, and CSF t-tau for neurodegeneration (N; Jack et al., 2018). The AT(N) framework, relying solely on CSF biomarkers, has shown 85–90% sensitivity and specificity for AD, with combined assessments yielding greater diagnostic accuracy than individual CSF markers (Grøntvedt et al.,

2020). Moreover, this approach is more cost-effective compared to imaging biomarkers (Contador et al., 2023).

While neurodegeneration in AD is irreversible and lacks definitive treatment (Conti Filho et al., 2023), early diagnosis can enhance prognosis and improve disease management (Crous-Bou et al., 2017). Consequently, primary intervention strategies now focus on modifiable risk factors such as cardiovascular health and lifestyle choices (Crous-Bou et al., 2017). Notably, recent research has highlighted a history of disrupted sleep, reported years before the onset of cognitive impairment, as a potentially modifiable risk factor for AD (Macedo et al., 2017). Sleep disturbances are connected to cognitive dysfunction and AD pathogenesis, including the accumulation of amyloid proteins and tau phosphorylation (Ahmadian et al., 2018). Chronic partial sleep restriction in rodents has demonstrated an increase in A $\beta$  deposition. The mechanism behind this phenomenon may involve reduced interstitial fluid volume during sleep deprivation, potentially hindering A $\beta$  clearance (Insel et al., 2021). Additionally, acute sleep deprivation in humans has been shown to elevate overnight CSF A $\beta$  levels by 25–30% compared to sleeping controls (Lucy et al., 2018). Loss of slow-wave sleep (SWS) due to partial sleep deprivation has been associated with an acute increase in next-morning CSF A $\beta$  levels (Ju et al., 2017). It is believed that SWS plays a critical role in A $\beta$  turnover, potentially due to increased glymphatic system activity in the brain during this stage (Mander et al., 2015). On the other hand, concerning tau pathology, cerebrospinal fluid tau levels increased by over 50% following sleep deprivation (Holth et al., 2019). Furthermore, chronic sleep deficiency over 2 months has been linked to a more than 50% increase in insoluble tau in the brains of AD patients (Nunomura et al., 2001). These findings suggest that the sleep-wake cycle plays a crucial role in regulating tau levels in the brain, with sleep deprivation contributing to elevated cerebral tau and its pathological spread (Holth et al., 2019). Moreover, a meta-analysis has determined that patients with Obstructive Sleep Apnea (OSA) have double the risk of cognitive decline and/or AD compared to individuals without OSA (Bubu et al., 2017). Additionally, multiple studies support the hypothesis that the relationship between these conditions could be bidirectional and may manifest before the clinical signs of AD become apparent (Yaffe et al., 2011). Thus, identifying markers for sleep disturbances in early AD stages could pave the way for developing preventive strategies targeting neurodegeneration through sleep intervention.

Therefore, the challenge of diagnosing AD in its initial phases is significant, necessitating the discovery of early biomarkers. Recent studies affirm that core CSF biomarkers can identify preclinical AD in asymptomatic individuals (Sabbagh and DeCourt, 2023), predict progression to MCI, and distinguish AD from other dementias (Hampel et al., 2004; De Leon et al., 2006). This underscores their potential utility in clinical settings for diagnosing preclinical AD, although currently core CSF AD biomarkers are not widely utilized in clinical practice for these purposes (Sabbagh and DeCourt, 2023). Nevertheless, CSF collection requires a lumbar puncture, an invasive procedure with contraindications for some individuals, such as those on anticoagulant therapy (Kim, 2022). Additionally, determining CSF biomarkers necessitates high-cost enzyme-linked immunosorbent assay (ELISA) methods using specific antibodies (Dakterzada et al., 2021). Thus, despite their diagnostic value, these techniques may limit the feasibility of widespread screening, highlighting the need for less invasive diagnostic methods.

In this regard, past studies have proposed that electroencephalography (EEG) metrics may serve as predictive, non-invasive biomarkers of AD (Babiloni et al., 2010). These studies typically focus on analyzing awake resting-state EEG signals for spectral content, complexity, and synchronization (Gallego-Jutglá et al., 2015). Notably, certain EEG spectral power profiles have been associated with traditional CSF biomarkers in the early stages of AD (Chino-Vilca et al., 2022). Researchers have increasingly shown interest in analyzing sleep EEG. EEG changes associated with AD during sleep, such as REM sleep slowing, reduced spindle and K-complex amplitude and duration and disrupted slow-wave activity (SWA) during NREM sleep are well-documented (Petit et al., 2004). Some of these alterations were intriguingly linked to tau and amyloid protein accumulation (Winer et al., 2019).

In recent years, advances in computational neuroscience have been applied to the sleep EEG signal analysis in AD, based on quantitative EEG (qEEG) feature extraction using diverse algorithms (Geng et al., 2022; Azami et al., 2023). Notably, a seminal study demonstrated that specific qEEG measures during sleep could distinguish AD dementia patients from those with MCI and healthy controls (HC), identifying these measures and potential electrophysiological biomarkers of AD (Azami et al., 2023). However, despite the acknowledged impact of sleep on AD neuropathology (Yaffe et al., 2011; Ahmadian et al., 2018), few studies extend these methods to polysomnography (PSG; Khosroozad et al., 2023). PSG is a non-invasive examination that allows for both qualitative and quantitative assessment of various parameters (Berry et al., 2015), thus potentially offering additional insights alongside other biomarkers. Given the complex nature of AD, it is becoming increasingly clear that relying on a single type of biomarker may not suffice for accurate diagnosis.

Consequently, a multimodal approach that incorporates various categories of data is needed and Machine Learning (ML) technologies could address these challenges (Borhani et al., 2022). ML is a subset of artificial intelligence, employing algorithms to parse data and uncover underlying insights. Unlike classical statistical methods, where the focus is on predetermined data patterns, ML enables computational models to identify and learn from data patterns that may otherwise go unnoticed (Bzdok et al.,

2018). Additionally, recent advancements in artificial intelligence have made the analysis of large, multimodal datasets not only more efficient but also more clinically relevant (Obermeyer and Emanuel, 2016). Therefore, ML models hold promise in sleep medicine, where vast amounts of electrophysiological data are generated during PSG recordings (Obermeyer and Emanuel, 2016; Khosroozad et al., 2023). This innovative approach has been applied in AD research for diagnosis, progression prediction or neurodegeneration detection, integrating various potential biomarkers such as clinical and neuroimaging data, neuropsychological test results and rest EEG spectral features (Bandyopadhyay and Goldstein, 2023). Nevertheless, models that incorporate sleep-related variables to detect AD neuropathology at an early stage of the disease have not yet been fully explored.

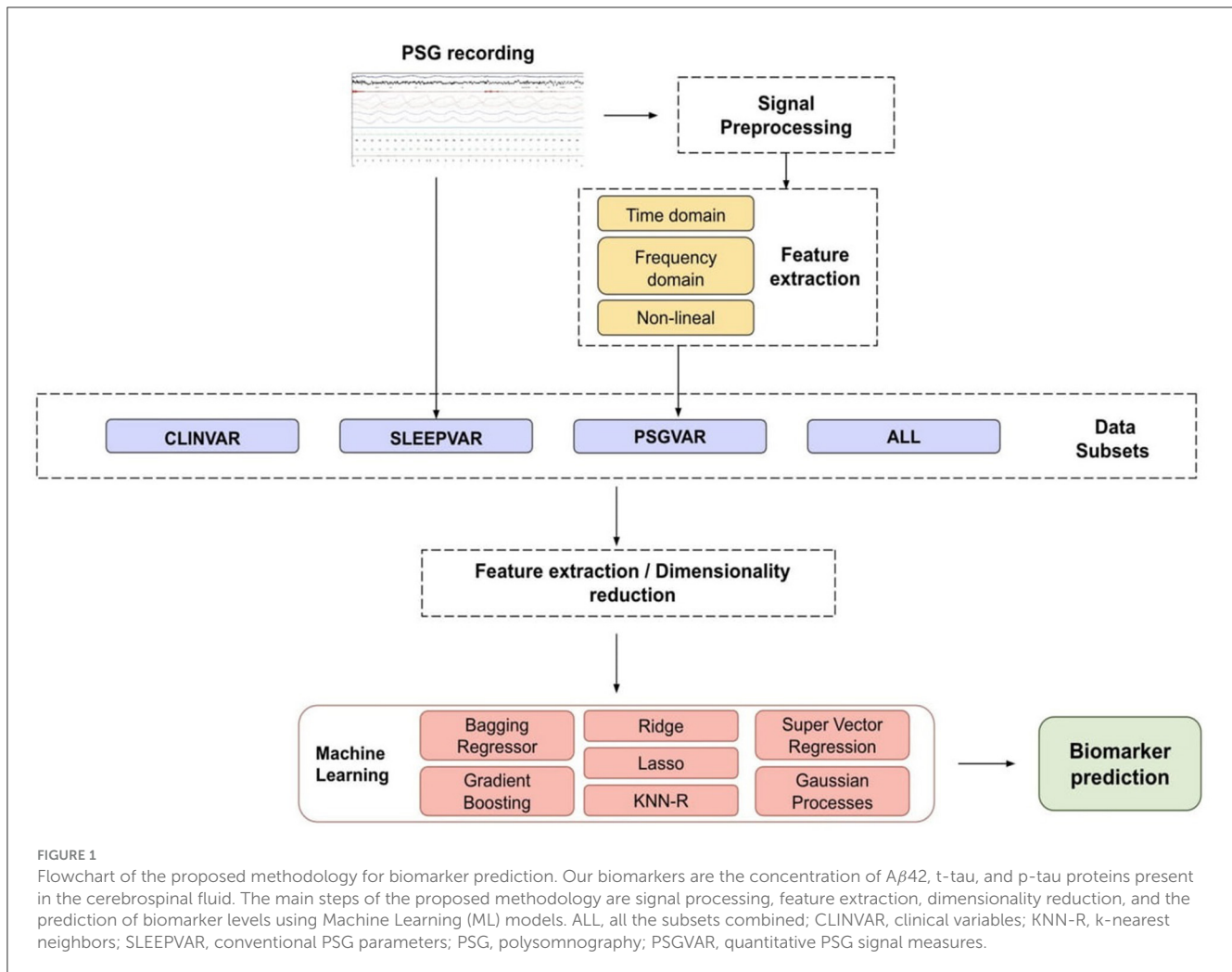
Based on these considerations, we hypothesize that ML techniques could enable us to identify, within a mild-moderate AD population, a specific array of diverse non-invasive variables—including those related to sleep—as potential indicators of AD neuropathology. Our research is thus directed at developing a multimodal ML model to serve as an early diagnostic tool that can accurately and non-invasively predict levels of CSF AD core biomarkers. This is critical for identifying AD-specific neuropathological changes and providing novel insights for a non-invasive clinical protocol to monitor AD-related neuropathology. To this end, we will train various ML models incorporating different subsets of non-invasive variables. These include conventional PSG parameters, quantitative PSG signal features and a range of clinical variables related to AD pathogenesis, such as comorbidities, sociodemographic, and sleep-related information.

## 2 Materials and methods

Figure 1 presents a flowchart outlining the study methodology and the ML analysis process for predicting CSF biomarkers.

### 2.1 Study population

We performed an additional analysis based on the longitudinal study NCT02814045, which monitored cognitive progression in a cohort with mild-moderate AD. The study adhered to the principles of the Declaration of Helsinki and received approval from the ethics committee of Hospital Arnau de Vilanova de Lleida (CE-1218). From 2014 to 2018, the Cognitive Disorders Unit at Hospital Universitari Santa Maria (Lleida, Spain) systematically enrolled patients who were over 60 years old, drug naïve and diagnosed with mild-moderate AD. The diagnosis of AD was formulated according to clinical criteria established by the National Institute of Aging Alzheimer's Association (NIA-AA) guidelines (McKhann et al., 2011). We employed the Mini-Mental State Examination (MMSE) tests to assess the severity of cognitive impairment. Exclusively patients exhibiting  $MMSE \geq 20$  were included (Folstein et al., 1975; Perneczky et al., 2006; Lanctôt et al., 2024). Consent for participation was obtained from the patients themselves,



their responsible caregivers, or legal guardians. Exclusion criteria were set to omit individuals with any condition that could interfere with adherence to the study protocols. These exclusion criteria included the presence of sleep disorders, unexplained daytime sleepiness; severe organic or psychiatric conditions, alcohol consumption exceeding 280 g/week; significant central nervous system disease other than AD; untreated deficiencies in vitamin B12 or folate, and untreated thyroid disease. Furthermore, patients who had taken medications such as neuroleptics, hypnotics, antidepressants, or beta-blockers within 15 days before actigraphy were also excluded from the study. Eligible patients underwent thorough overnight PSG assessments. Subsequently, CSF and blood samples were collected the following morning to evaluate biomarker levels. We gathered anthropometric and clinical data including age, gender, body mass index, cardiovascular risk factors, years of education, and toxic habits. Daytime sleepiness was measured using the Epworth sleepiness Scale (ESS), with a score greater than 10 indicating excessive drowsiness (Bzdok et al., 2018). [Supplementary Table 1](#) provides a detailed listing of the sociodemographic, anthropometric, and clinical characteristics collected.

## 2.2 Cerebrospinal fluid biomarkers and ApoE genotyping

CSF samples obtained via lumbar puncture and amounting to 8–10 mL, were centrifuged at low speed (2,000  $\times$  g for 10 min at 4°C) to pellet any cellular elements. They were then aliquoted in polypropylene tubes before being frozen and stored at  $-80^{\circ}\text{C}$ . The samples were then processed for biomarker analysis. Concentrations of CSF t-tau, threonine 181 p-tau, and A $\beta$ 42 were determined using enzyme-linked immunosorbent assays (Innotest<sup>®</sup>, Fujirebio, Belgium), as directed by the manufacturer. All samples were tested in duplicate, and the results were given in parts per million (pg/ml). IRBLleida Biobank (B.0000682) and Plataforma Biobancos PT17/0015/0027 facilitated the sample collection process. The A $\beta$ 42 cut-off values were 600 pg/ml, with values below this suggesting amyloid deposition. The cut-off values for t-tau and p-tau were set at 425 and 65 pg/ml, respectively. Values exceeding these cut-offs indicated the presence of neurofibrillary tangles. Concurrent with the CSF collection, blood samples were drawn, processed immediately for analysis, and examined.

The ApoE genotype was determined using a Maxwell<sup>®</sup> RCS blood DNA kit (Promega, USA) and 20  $\mu\text{L}$  of DNA from a 2 mL

peripheral blood sample. Participants were categorized as either homozygous or heterozygous carriers of the ApoE4 allele (noted as ApoE4+), with phenotypes 2/4, 3/4, or non-carriers (noted as ApoE4-), with phenotypes 2/2, 2/3, and 3/3.

## 2.3 Polysomnography

Overnight supervised PSG recordings were performed using Philips Respironics Alice 6 equipment to determine sleep parameters. The methodology for sleep recording, technical specifications, and manual grading of raw data was established in accordance with published standards (Berry et al., 2015). The PSG study included the simultaneous recordings from the following electrophysiologic channels: six electroencephalograms (EEG) leads (F3-A2/F4-A1; C3-A2/C4-A1; O2-A1/O1-A2), two bilateral electrooculographic leads (EOG), a single chin electromyogram (EMG) channel, sensors for chest and abdominal respiratory effort, airflow measurements (obtained through both an oral and nasal thermocouple and nasal pressure records), pulse oximetry, an electrocardiogram (ECG), body position sensor, a snoring detection microphone and bilateral piezoelectric sensors detect leg movement detection. Experts manually scored the raw data, extracting sleep and respiratory parameters in line with the established literature (Berry et al., 2015). The physicians' annotations included details on any artifacts observed during the recordings. Measures of sleep quality were determined and included total sleep time (TST), sleep efficiency (SE)—calculated as the percentage of TST relative to the time spent in bed (TIB)—and arousal index (AI), defined as the total number of arousals per hour of sleep. Apneas were identified using an oronasal thermal sensor and defined as a reduction in the airflow sensor signal by 90% when compared to the pre-event baseline value for more than 10 s. Hypopneas were recognized as a decrease of at least 30% in airflow, as measured by nasal pressure, persisting for over 10 s, and associated with either arousal or a 3% decrease in oxygen saturation from the pre-event baseline. The apnea-hypopnea index (AHI) was defined as the average number of apnea and hypopnea episodes per hour of TST. The oxygen desaturation index (ODI) is the average number of instances per hour of sleep where oxygen saturation decreases by 3% or more. The time with oxygen saturation below 90%, referred to as CT90%, is expressed as a percentage of the TST. An AHI exceeding 15 events per hour is indicative of obstructive sleep apnea (OSA).

## 2.4 Signal preprocessing

The MATLAB<sup>®</sup> signal processing toolbox was used to preprocess and analyze signal data offline. To standardize the format of the signals across different PSG channels and to mitigate the effects of the noise during the analysis phase, an initial preprocessing was conducted. The preprocessing phase was divided into three main steps:

- Resampling: due to varying sample rates across the channels and subjects, signals were resampled for uniformity based

TABLE 1 Signal frequency ranges and filter settings: the original sample rate (fs0), the target sample rate (fst), the output sample rate (fsf), and the filter frequencies for the Electrooculogram (EOG), Electroencephalogram (EEG), Electromyogram (EMG), Respiratory effort (Effort), Airflow, Pulse, Oxygen saturation (SpO<sub>2</sub>), Electrocardiogram (ECG), and Snore channels.

Signal	fs0 (Hz)	fst (Hz)	fsf (Hz)	Filter (Hz)
EOG	200–500	500	500	0.3–50
EEG	200–500	500	500	0.3–50
EMG	200–500	500	200	1–90
Effort	100	100	100	-
Airflow	100	100	100	0.1–15
Pulse	1	1	1	-
SpO <sub>2</sub>	1	1	1	-
ECG	200–500	500	200	0.3–50
Snore	500	500	500	1–100

on international guidelines (Bandyopadhyay and Goldstein, 2023), as is detailed in Table 1. Subsequently, each signal was decomposed into five new signals corresponding to the sleep phases: W, N1, N2, N3, and REM phases. Segments identified as “Wake” were discarded. The remaining segments, identified as the same sleep stage in their respective hypnogram, were concatenated to form the new signals. These were then resampled at a standard rate of 128 Hz to diminish their size. The resampled signals were further divided into uniform 10-s segments.

- Filtering: the channels were filtered with a fifth-order bandpass digital Butterworth filter (Butterworth, 1930). The cutoff frequencies for each channel were set in accordance with the AASM guidelines (Berry et al., 2015), as specified in Table 1, to ensure comparably meaningful spectral content across channels.
- Artifact removal: based on the annotations provided by expert physicians, segments containing artifacts were identified and excluded from the signals.

## 2.5 Feature extraction

Utilizing various computational methods (Álvarez et al., 2012; Alvarez et al., 2013; Umut, 2016), the preprocessed PSG signals were parameterized with features from linear (time-domain and frequency domain) and non-linear features. Each feature set provides unique insights into the signals' characteristics. Time-domain statistics offer information on the signal amplitude, while frequency-domain statistics leverage the spectral information within the signals (Dressler et al., 2004). In contrast, non-linear features shed light on the temporal complexity and regularity of the signals (Abásolo et al., 2006; Furui et al., 2020). Features were calculated as the mean across each sleep stage (N1, N2, N3, and REM) for the entire duration of the overnight recording. For segments annotated as distinct sleep stages, parameters were

TABLE 2 Frequency domain statistics and spectral parameters computed for Electrooculogram (EOG), Electroencephalogram (EEG), Electromyogram (EMG), and the Airflow and Oxygen saturation (SpO<sub>2</sub>) channels of the polysomnography.

Signal	Components
EOG	MF 0.5–30 Hz, DF 0.5–30 Hz, Pw 0.5–30 Hz (TotalPw)
EEG	MF 0.5–30 Hz, DF 0.5–30 Hz, Pw 0.5–30 Hz (TotalPw),
	Pw 0.5–4 Hz (delta), Pw 4–7 Hz (theta), Pw 8–12 Hz (alpha),
	Pw 14–22 Hz (lowBeta), Pw 23–30 Hz (highBeta),
	Pw 14–30 Hz (beta), Pw 31–40 Hz (gamma), Pw 12–14 Hz (spindles)
EMG	MF 0.5–30 Hz, DF 0.5–30 Hz, Pw 0.5–30 Hz (TotalPw)
Airflow	MF 0.025–0.05 Hz, DF 0.025–0.05 Hz, Pw 0.025–0.05 Hz (TotalPw)
SpO <sub>2</sub>	MF 0.014–0.033 Hz, DF 0.014–0.033 Hz, Pw 0.014–0.033 Hz (TotalPw)

MF, mean frequency; DF, Dominant frequency; Pw, Power.

derived from the mean values within 2-min windows, employing a 50% overlap strategy.

For the time-domain analysis of all PSG signals, we calculated four time domain and four non-linear parameters. Time domain features included Root Mean Square, variance, skewness, kurtosis, and maximum value (Álvarez et al., 2012; Alvarez et al., 2013; Gerla et al., 2019). The non-linear parameters included Shannon Entropy (ShanEnt), Sample Entropy (SampEnt), Lempel Ziv (LempZiv), and Higuchi Fractal Dimension (HFD; Alvarez et al., 2006, 2013). Frequency-domain statistics and spectral parameters (Dressler et al., 2004; Alvarez et al., 2006; Pedregosa et al., 2011) were explicitly calculated for the EOG, EEG, EMG, Airflow, and SpO<sub>2</sub> channels (see Table 2). The spectral parameters were derived from the Power Spectral Density (PSD) of each normalized 2-min segment (Gerla et al., 2019). The estimates were obtained using Welch's averaged, modified periodogram method with a Hamming window of 65,536 points for EOG, 32,768 for EMG signals, 16,384 for the airflow signals, and 128 for the SpO<sub>2</sub> signal.

### 3 Statistical analysis

Statistical analyzes were conducted using R statistical software, version 3.3.1. Descriptive statistics were calculated for both normally and non-normally distributed quantitative data, with the former presented as mean and standard deviation (SD), and the latter as median with interquartile range (IQR). For categorical variables, we reported absolute and relative frequencies. In the case of quantitative variables, we computed Pearson's correlation coefficient to assess the relationship with the target biomarker. For categorical features, a one-way ANOVA test was employed to determine if there were significant differences between the means of two or more groups. To eliminate irrelevant, correlated, and noisy data, a feature selection process was undertaken. This step aimed to reduce the dataset dimensionality with minimal loss of significant information. We performed feature selection on the training partition of the dataset, organizing the variables into four subsets as follows:

- CLINVAR: sociodemographic, anthropometric, and clinical variables.
- SLEEPVAR: conventional PSG parameters.
- PSGVAR: quantitative measure derived from PSG signals.
- ALL: a combination of all the above subsets.

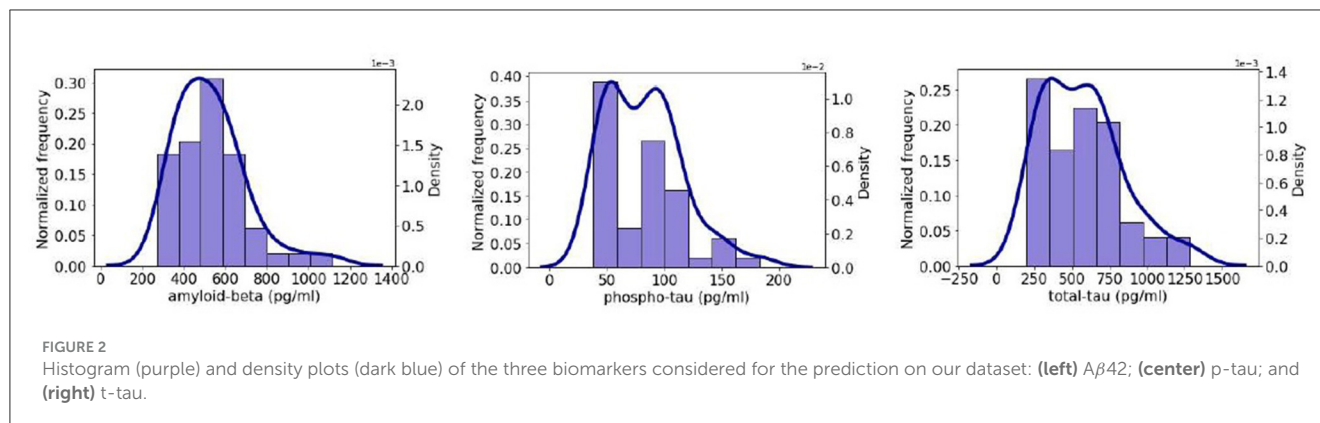
For the CLINVAR, SLEEPVAR, and PSGVAR subsets, any feature that was missing in more than 50% of the samples was excluded. An initial feature selection was conducted by retaining only those features that exhibited a correlation coefficient with the target biomarker about 0.1 for CLINVAR and SLEEPVAR and above 0.3 for PSGVAR. The ALL subset was then refined to include only features meeting these criteria. Subsequently, we calculated the correlation coefficient for each pair of features, opting to retain only one variable from pairs where the correlation exceeded 0.9, to avoid redundancy. Further dimensionality reduction was accomplished through principal component analysis (PCA). In this process, we chose to retain enough components to explain 90% of the variance, creating new subsets for CLINVAR, SLEEPVAR, PSGVAR, and ALL based on the selected principal components. Additionally, for each biomarker and subset, we identified the features that most effectively explained the variance (PCA-selected features, PCA-sel). Following this, we assessed whether the variance chosen through the initial feature selection exhibited a statistically significant linear relationship with the target biomarkers using Pearson's correlation and ANOVA tests as previously described. For feature selection and ML analyzes, the data were divided into training and testing subsets, with 25% of the data allocated to the test partition. Moreover, all subsets were standardized using the mean and standard deviation derived from the training data.

### 4 Machine Learning analyzes

ML analyzes were performed using the scikit-learn Machine Learning library in Python (Pedregosa et al., 2011). We trained several classical ML models using the four subsets of variables that, according to the PCA, best explained the variance of the dataset.

Dimensionality reduction helped to remove irrelevant, redundant, and noisy data, which in turn improved model performance and the intelligibility of the results. Model hyperparameters were fitted via five-fold cross-validation to ensure robustness. The models' effectiveness was assessed by calculating the Mean Absolute Error (MAE in pg/ml) across both the training and the test datasets. The MAE metric was preferred over the commonly used root mean square error as it provides a clearer and more direct measure of the average error, according to some studies (Willmott and Matsuura, 2005).

A Regression model based on the k-nearest neighbors (KNN) algorithm was trained, where the target prediction is based on local interpolation from the nearest neighbors in the training set. Additionally, we trained two Gaussian Process (GPs) models employing a radial basis function (RBF) kernel and a Matern 3/2 kernel. These GP models were optimized 20 times using the GPy framework in Python, with the optimal solution selected from these iterations. For ensemble learners utilizing bagging techniques, we optimized a Bagging Regressor (BR), a Random Forest Regressor (RFR), and an Extra Trees Regressor (ETR). With regards to



boosting methods, we fitted three Gradient Boosting Regressors (GBRs) with distinct loss functions: least squares (LS), least absolute deviation (LAD), and a hybrid approach (Huber).

In terms of models with strong regularization policies, we fitted a linear regression with L2-norm regularization, known as Ridge Regression, and another with L1-norm regularization (Lasso). Furthermore, we trained four Support Vector Regressions (SVRs) using different kernels: linear (LIN), polynomial (POLY), radial basis function (RBF), and sigmoid (SIG).

## 5 Results

### 5.1 Demographic and clinical characteristics

A total of 61 subjects were recruited for this study. However, three subjects that did not possess PSG recordings and 9 whose biomarker measures were missing were excluded from the analysis, reducing the dataset to 49 subjects.

The population was equally distributed by gender (females were 50.8%), and the median age of the study population was 75.0 [72.0; 78.0] years. The Body Mass Index (BMI) median was 28.0 [24.4; 31.1]. The most frequently associated comorbidities were hypertension (63.9%), dyslipidemia (49.2%), heart disease (19.7%), and diabetes mellitus (16.4%). Concerning global cognitive dysfunction, the median MMSE score was 23.0 [21.0; 25.0]. Only a few patients reported symptoms indicative of poor sleep quality, such as insomnia, non-restorative sleep, nocturnal awakenings, and daytime drowsiness. The cohort was free from significant subjective daytime sleepiness, as indicated by a median score of 5.0 [3.0–8.0] on the Epworth Sleepiness Scale.

The PSG data indicated a high prevalence of OSA, with 26 (42,62%) patients. The median (IQR) AHI was 27.0 events/hour [15.4; 52.5]. The median ODI was 18.3 [12.6; 44.0] events/hour, and a median T90 was 11.9% [0.7; 7.2]. On average, participants slept (TST) 262.1 [203.5; 326.1] min. Sleep was highly fragmented, with a median arousals index (AI) value of 36.1 [23.2; 49.5] arousals per hour. In [Figure 2](#), the density and histogram plots of the three studied biomarkers are shown. We observe that t-tau and p-tau shows a right skewed behavior while A $\beta$ 42 is closer to symmetric. [Table 3](#) presents a summary of the clinical and demographic characteristics of the study population.

### 5.2 Correlations between CSF biomarkers and selected variables of the four subsets

The statistical analysis conducted on various subsets of data revealed specific correlations between sleep-related variables and the levels of AD CSF biomarkers. The results of the statistical analyzes for the quantitative features identified as relevant by PCA are presented in [Tables 4–6](#) for the CSF A $\beta$ 42, p-tau, and t-tau biomarkers, respectively. Similarly, the results for the categorical variables are shown in [Tables 7–9](#).

**A $\beta$ 42:** For the A $\beta$ 42 biomarker, poorer sleep quality ( $\rho = -0.3853$ ;  $p = 0.0062$ ) and shorter sleep duration on holidays ( $\rho = -0.3981$ ;  $p = 0.0046$ ) within the CLINVAR subset were linked to lower A $\beta$ 42 levels. Additionally, increased A $\beta$ 42 levels were found in individuals with sleep disturbances due to heartburn ( $p = 0.0296$ ) and those carrying the ApoE genotype ( $p = 0.0123$ ). In terms of polysomnography variables (PSGVAR), certain EEG characteristics during different sleep stages and while awake were positively correlated with A $\beta$ 42 levels such as the skewness in the EEG O1-A2 channel during the N1 (EEGO1\_A2\_sk\_N1) and N2 (EEGO1\_A2\_sk\_N2) sleep stages ( $\rho = 0.3534$ ,  $p = 0.0127$  and  $\rho = 0.3778$ ,  $p = 0.0074$ , respectively). Notably, the only sleep variable (SLEEPVAR) that correlated positively with A $\beta$ 42 was the percentage of time spent with oxygen saturation below 90% during sleep ( $\rho = 0.3129$ ,  $p = 0.0322$ ).

**P-tau:** Regarding the p-tau biomarker, increases were associated with specific EEG patterns during certain sleep stages and a negative correlation was found with the maximum chin electromyography (EMG) value during the N1 sleep stage ( $\rho = -0.3147$ ,  $p = 0.0276$ ). In the clinical variables subset (CLINVAR), arterial hypertension ( $p = 0.0127$ ), pulmonary diseases ( $p = 0.0274$ ), and level of education ( $p = 0.0136$ ) were significantly associated with p-tau levels.

**T-tau:** For the t-tau biomarker, significant correlations were identified with features derived from the thoracic effort channel such as the Lempel Ziv during N2 sleep stages ( $\rho = 0.4216$ ,  $p = 0.0025$ ) and the Sample Entropy during N2 ( $\rho = 0.3927$ ,  $p = 0.0052$ ), indicating a rise in t-tau levels. However, no significant relationships were observed with the general sleep variables, while the frequency of snoring was positively correlated with t-tau in the CLINVAR subset ( $p = 0.0149$ ).

TABLE 3 Descriptive characteristics of the study population.

Feature	Mean ± std
<b>Sociodemographic data</b>	
Gender (women)	31 (50.8%)
Age (years)	75.0 [72.0; 78.0]
BMI, kg/m <sup>2</sup>	28.0 [24.4; 31.1]
<b>Comorbidities</b>	
Hypertension	39 (63.9%)
Diabetes mellitus	10 (16.4%)
Dyslipidemia	30 (49.2%)
Heart diseases	12 (19.7%)
Hypertension	3 (4.9%)
OSA	26 (42.6%)
<b>Alzheimer's disease parameters</b>	
Mini-mental state examination	23.0 [21.0; 25.0]
Aβ42 (pg/ml)	515.0 [404.25; 614.25]
t-tau (pg/ml)	555.0 [327.5; 712.0]
p-tau (pg/ml)	82.0 [52.0; 98.75]
<b>Polysomnographic parameters</b>	
Epworth sleepiness scale	5.0 [3.0; 8.0]
Time in bed (minutes)	420.1 [391.0; 446.1]
Total sleep time (minutes)	262.1 [203.5; 326.1]
Sleep efficiency (%)	66.4 [49.7; 79.4]
N1 stage (% TST)	11.96 [7.7; 18.4]
N2 stage (% TST)	23.1 [16.0; 35.1]
N3 stage (% TST)	13.6 [7.1; 22.4]
REM stage (% TST)	6.6 [3.3; 11.2]
REM sleep (% TST)	6.6 [3.3; 11.2]
Latency to N1 stage (minutes)	23.1 [11.9; 57.7]
Latency to REM sleep (minutes)	126.7 [85.1; 179.9]
AHI (events/hour)	27.0 [15.4; 52.5]
AI (events/hour)	36.1 [23.2; 49.5]
Mean SaO <sub>2</sub> (%)	93.0 [92.0; 94.0]
CT90 (%)	1.9 [0.7; 7.2]
ODI (events/h)	18.3 [12.6; 44.0]

AHI, apnoea-hypopnea index; AI, arousal index; ApoE, apolipoprotein E; Aβ42, amyloid-beta protein; BMI, body mass index; ESS, Epworth sleepiness scale; IQR, interquartile range; ODI, oxygen-hemoglobin desaturation index; SaO<sub>2</sub>, oxygen-hemoglobin saturation; TST, total sleep time; CT90, percent of TST spent below 90% oxygen-hemoglobin desaturation; p-tau, phospho-tau; t-tau, total-tau.

### 5.3 Regression models

Table 10 presents the mean absolute error (MAE) of each model family utilized in estimating the levels of CSF biomarkers. Figures 3–5 provide more comprehensive information.

TABLE 4 Pearson's correlation and the p-value between CSF Aβ42 biomarker and the quantitative variables that were considered relevant by the PCA algorithm.

Variables	Correlation coefficient	p-value
<b>PSGVAR</b>		
EEG3_A2_sk_N2	0.2630	0.0678
EEGO1_A2_sk_N1	0.3534	<b>0.0127*</b>
EEGO1_A2_sp_N1	0.3285	<b>0.0211*</b>
EEGO1_A2_sk_N2	0.3778	<b>0.0074*</b>
EEGO1_A2_alpha_N2	0.4105	<b>0.0033*</b>
EEGO1_A2_sp_N2	0.3563	<b>0.0119*</b>
<b>SLEEPVAR</b>		
N° Mixed Apnea	0.0826	0.5723
Mixed Apnea Index	0.1022	0.4843
N° Arousals	0.1331	0.3619
Total Arousal Index	0.0802	0.5837
CT90	0.3129	<b>0.0322*</b>
SpO <sub>2</sub> min	-0.1871	0.1978
<b>CLINVAR</b>		
Sleep (0–10)	-0.3853	<b>0.0062*</b>
Sleep duration (Holidays)	-0.3981	<b>0.0046*</b>

Values in bold denote statistically significant correlations: \*p < 0.05. alpha, alpha band; CT90, the percentage of TST during which oxygen saturation was <90%; EEG, Electroencephalogram; EMG, Electromyogram; N1, N1 sleep stage; N2, N2 sleep stage; O2\_A1, O2-A1 EEG channel; sk, skewness; sp, spindle; SpO<sub>2</sub>min, minimum oxygen saturation.

For the Aβ42 biomarker, the test partition yielded higher MAE values across all models, regardless of the subject of data utilized. The performance in regression was comparable across different model families. Notably, using solely quantitative PSG signal parameters (PSGVAR), the training MAE was reduced for ensemble models (50.6 ± 23.21) compared to regularized regressions (103.54 ± 0.59), Gaussian Processes (GPs; 105.29 ± 2.91), or Support Vector Regression (SVR; 118.42 ± 20.91). The test MAE for ensemble models (164.88 ± 8.12), regularized regressions (161.63 ± 0.54), GPs (157 ± 4.37), and SVR (155.30 ± 10.32) showed minimal variation, with the ensemble models even registering higher values. When incorporating the ALL subset, which combines both PSGVAR and CLINVAR features, a similar test performance was observed with a lower training MAE compared to using only PSGVAR. However, employing solely clinical features (CLINVAR) resulted in diminished performance for the ensemble models, while the MAE for the other models slightly decreased, leading to generally comparable results across all models. Conversely, utilizing conventional sleep parameter reports (SLEEPVAR) achieved the lowest test MAE, with all models yielding values under 144. An individual assessment of each model performance (Figure 3) reveals that, for Aβ42 biomarker prediction, Gradient Boosting Regressors (GBRs) consistently outperformed other models. The most efficient model, the GBR employing the least absolute deviation loss (GBR\_lad) trained with



TABLE 5 Pearson's correlation and the  $p$ -value between CSF p-tau biomarker and the quantitative variables that were considered relevant by the PCA algorithm.

Variables	Correlation coefficient	$p$ -value
<b>PSGVAR</b>		
EEGF3_A2_theta_N2	0.2533	0.0790
EEO2_A1_sk_N1	0.3077	<b>0.0314*</b>
EEO2_A1_sk_N2	0.3736	<b>0.0081*</b>
EMGChin_ShanEnt_N3	0.2159	0.1541
EMG_Chin_max_N1	-0.3147	<b>0.0276*</b>
EMGChin_max_N2	-0.2305	0.1109
EffortTHO_LempZiv_N2	0.3148	<b>0.0275*</b>
Leg2_HDF_N2	0.1616	0.2671
Snore_HFD_N2	-0.1647	0.2579
<b>SLEEPVAR</b>		
Mixed Apnea Index	0.1256	0.3898
N° Mixed Apnea	-0.1876	0.1966
Recording time	0.1439	0.3236
N3 Latency	0.2483	0.0852
SpO <sub>2</sub> min	0.1557	0.2850
<b>CLINVAR</b>		
Weight (Kg)	-0.2364	0.1057
Waist	-0.2506	0.0892
BMI	-0.2520	0.0839
Glucose	-0.1892	0.2242
Insulin	-0.2364	0.1316

Values in bold denote statistically significant correlations: \* $p < 0.05$ . BMI, Body Mass Index; EEG, Electroencephalogram; EMG, Electromyogram; EffortTHO, Thoracic Effort; HFD, Higuchi Fractal Dimension; LempelZiv, Lempel Ziv; max, maximum value; N1, N1 sleep stage; N2, N2 sleep stage; N3, N3 sleep stage; O2\_A1, O2-A1 EEG channel; ShanEnt, Shannon Entropy; sk, skewness; sp, spindle; SpO<sub>2</sub>min, minimum oxygen saturation.

the SLEEPVAR subset, had the lowest training and test MAE (Train MEA: 47.54, Test MAE: 141.8).

For predicting p-tau levels (Figure 4), the lowest MAE values were achieved with the ALL subset and GBRs, particularly the GBR<sub>ls</sub>, which demonstrated an MAE of 0.02 for training and 14.12 for testing. However, when applying this same subset (ALL) to optimize the rest of the ML models, there was a decline in performance for both training and testing partitions. This pattern was also observed in the PSGVAR subset, where the ensemble models outperformed other model families during training ( $7.81 \pm 5.80$ ), but their test performance did not significantly differ. Conversely, models using the CLINVAR and SLEEPVAR subsets generally underperformed compared to those using the ALL or PSGVAR subsets, as summarized in Table 10.

In the case of t-tau prediction, similar to p-tau, the lowest MAE was recorded with the ALL subset and GBRs. However, the GBRs using Huber loss function showed superior training

TABLE 6 Pearson's correlation and the  $p$ -value between CSF t-tau biomarker and the quantitative variables that were considered relevant by the PCA algorithm.

Variables	Correlation coefficient	$p$ -value
<b>PSGVAR</b>		
EffortTHO_SampEnt_N2	0.3927	<b>0.0052*</b>
EffortTHO_LempZiv_N2	0.4216	<b>0.0025*</b>
<b>SLEEPVAR</b>		
Central apnea index	0.1724	0.2359
Recording time	0.1443	0.3225
Sleep efficiency	-0.1029	0.4815
N3 latency	0.2097	0.1479
N3 time	-0.1030	0.4811
% N3	-0.1407	0.3348
Time awake	0.1490	0.3067
<b>CLINVAR</b>		
Waist	-0.2420	0.1012
Insulin levels	-0.2409	0.1242

Values in bold denote statistically significant correlations: \* $p < 0.05$ . EffortTHO, Thoracic Effort; LempelZiv, Lempel Ziv; max, maximum value; N3, N3 sleep stage; SampEnt, Sample Entropy.

performance (Train MAE: 7.33, Test MAE: 166.83), while the best test performance was achieved using the LAD loss functions (Train MAE: 19.55, Test MAE: 140.02), as shown in Figure 5. On the other hand, the highest test errors occurred with models trained solely on conventional sleep parameters (SLEEPVAR) evaluated by trained physicians. For instance, the ensemble models' test MAE using the SLEEPVAR subset was  $212.37 \pm 14.28$ , compared to  $159.90 \pm 12$  (63) with the CLINVAR subset, as reported in Table 10. Conversely, for the training partition, a different pattern emerges, with the CLINVAR subset presenting a higher MAE ( $127.62 \pm 42.83$ ) compared to the SLEEPVAR subset ii ( $99.80 \pm 42.75$ ). Notably, the Gradient Boosting Regressors (GBRs) trained on the CLINVAR subset demonstrated lower MAE values, as depicted in Figure 5. Furthermore, models calibrated solely with the quantitative parameters derived from PSG recordings (PSGVAR) displayed performance closely paralleling those utilizing the CLINVAR subset, as indicated in Table 10.

The models that yielded the most accurate estimations for the levels of the three biomarkers were the GBRs, specifically employing a least absolute deviation (LAD) loss for A $\beta$ 42, a least squares (LS) loss for p-tau, and a Huber loss for t-tau. The alternative ML models, in addressing the scarcity of training data, employed regularization techniques to construct simpler models, which resulted in higher training errors. For all three biomarkers, minimal training errors were observed when utilizing the comprehensive ALL subset, though the test evaluations varied. In terms of test performance, for the A $\beta$ 42 biomarker, the lowest MAE was recorded with the SLEEPVAR subset, that is, the subset comprising only the standard PSG parameters.

TABLE 7 ANOVA *p*-value categorical features that were considered relevant by the PCA algorithm for predicting A $\beta$ 42 biomarker.

CLINVAR variable	<i>P</i> -value
Physical activity	0.3624
Depression	0.2251
Epilepsy	0.2675
Pulmonary disease	0.1456
Non-restorative sleep	0.8501
Asphyxia crises	0.7328
Nocturia	0.6854
Headache	0.1863
Insomnia	0.1073
Cataplexy	0.1821
Real or vivid dreams	0.2154
Sleep interrupted by heartburn	<b>0.0296*</b>
APOE4+	<b>0.0125*</b>
Nap	0.6240

Values in bold indicate statistically significant correlations: \**p* < 0.05.

TABLE 8 ANOVA *p*-value categorical features that were considered relevant by the PCA algorithm for predicting p-tau biomarkers.

CLINVAR variable	<i>P</i> -value
Scholarship	<b>0.0136*</b>
Depression	0.1557
Arterial hypertension	<b>0.0127*</b>
Pulmonary disease	<b>0.0274*</b>
Snore	0.1058
Asphyxia crises	0.8299
Insomnia	0.1476
Memory disorder	0.2441
Concentration disorder	0.1147
Cataplexy	0.1151
Grind teeth	0.2325
Nap	0.5223
APOE4+	0.1613

Values in bold indicate statistically significant correlations: \**p* < 0.05.

## 6 Discussion

In this study, we trained various ML models to evaluate which combination of selected subsets of non-invasive variables could accurately predict CSF A $\beta$ 42, p-tau, and t-tau biomarker levels in a cohort of patients with mild-moderate AD(AD). The subsets included clinical variables (CLINVAR) previously established as relevant to AD pathogenesis, conventional PSG parameters (SLEEPVAR), quantitative PSG features (PSGVAR) derived from advanced signal analysis, and a combination of all these variables (ALL). Variable selection was performed using a

TABLE 9 ANOVA *p*-value categorical features that were considered relevant by the PCA algorithm for predicting t-tau biomarkers.

CLINVAR variable	<i>p</i> -value
Sex	0.3163
Arterial hypertension	0.0523
Heart disease	0.0905
Epilepsy	0.2752
Snore	<b>0.0149*</b>
Memory disorder	0.5125
Cataplexy	0.1596
Grind teeth	0.2824
Nap	0.2119

Values in bold indicate statistically significant correlations: \**p* < 0.05.

range of statistical approaches, with an emphasis on those with the highest discriminating power. These newly curated subsets, containing only relevant variables, were then used as input for the models.

Moreover, we investigated the correlations between the selected relevant variables and the CSF biomarkers to determine the nature of their relationships, linear or otherwise. The Gradient Boosting Regressors (GBRs) emerged as the most effective models in estimating the levels of the three biomarkers. In summary, the lowest training errors for all three biomarkers were observed when employing the ALL subset, although the test evaluations showed differing results. Regarding test performance, for the A $\beta$ 42 biomarker, the lowest mean average error (MAE) was achieved using the subset consisting solely of conventional PSG parameters, which was thereby identified as the best predictor of the CSF A $\beta$ 42 levels. Furthermore, it was found that the combination of clinical variables, conventional PSG parameters, and quantitative PSG features most effectively predicted CSF p-tau and t-tau levels. Interestingly, not all the selected variables demonstrated linear relationships with the biomarkers in question.

### 6.1 Best data subset for CSF A $\beta$ 42 prediction

According to our ML models, the most significant variables for predicting CSF A $\beta$ 42 levels were standard PSG parameters (SLEEPVAR), which included CT90, minimum SaO<sub>2</sub>, the number of Mixed Apneas, Mixed Apnea Index, the number of arousals, and Total Arousal Index.

The influence of hypoxemia on CSF A $\beta$ 42 is not unexpected, as numerous studies have indicated that cerebral hypoxia can escalate amyloid deposition. This is achieved through both genetic and epigenetic modifications by influencing the expression levels of enzymes critical for protein synthesis and degradation (Lall et al., 2019). Furthermore, hypoxia has been linked to calcium dysregulation in neurons and glial cells. This occurs through the formation of calcium-permeable pores, disruption of glutamate

TABLE 10 Mean and standard deviation of the mean average error (MAE) of Machine Learning (ML) models for the prediction of A $\beta$ 42, p-tau, and t-tau, biomarkers grouped by the family of models (EM, Ensemble models; RLR, Regularized Linear Regressions; GP, Gaussian Processes; KNR, regression based on k-nearest neighbors; SVR, Support Vector Regression).

	A $\beta$ 42							
	CLINVAR		SLEEPVAR		PSGVAR		ALL	
	Train	Test	Train	Test	Train	Test	Train	Test
EM	62.22 $\pm$ 7.33	170.43 $\pm$ 6.41	88.28 $\pm$ 32.71	143.5 $\pm$ 11.42	50.6 $\pm$ 23.21	164.88 $\pm$ 8.12	40.88 $\pm$ 18.72	160.67 $\pm$ 10.63
RLR	81.21 $\pm$ 0.16	154.33 $\pm$ 0.41	116.78 $\pm$ 1.12	142.35 $\pm$ 2.14	103.54 $\pm$ 0.59	161.63 $\pm$ 0.54	84.65 $\pm$ 10.74	154.97 $\pm$ 11.92
GP	91.08 $\pm$ 0.58	152.98 $\pm$ 0.41	115.44 $\pm$ 0.04	140.37 $\pm$ 0.05	105.29 $\pm$ 2.9	157 $\pm$ 4.37	77.02 $\pm$ 4.16	150.96 $\pm$ 1.91
KNR	93.86	168.58	119.59	147.13	94.67	142.83	109.03	146.15
SVR	78.52 $\pm$ 10.66	151.65 $\pm$ 5.68	111.73 $\pm$ 2.64	138.64 $\pm$ 2.4	120.73 $\pm$ 20.91	155.3 $\pm$ 10.32	91.12 $\pm$ 20.03	158.32 $\pm$ 14.96
	p-tau							
	CLINVAR		SLEEPVAR		PSGVAR		ALL	
	Train	Test	Train	Test	Train	Test	Train	Test
EM	16.22 $\pm$ 3.97	21.63 $\pm$ 4.06	12.65 $\pm$ 6.51	24.34 $\pm$ 2.47	7.81 $\pm$ 5.8	24.69 $\pm$ 3.42	6.95 $\pm$ 5.63	18.06 $\pm$ 2.9
RLR	19.01 $\pm$ 0.18	18.87 $\pm$ 0.02	25.49 $\pm$ 0.49	24.77 $\pm$ 0.05	16.75 $\pm$ 0.77	24.15 $\pm$ 3.74	19.64 $\pm$ 6.95	25.24 $\pm$ 3.15
GP	20.68 $\pm$ 0.02	15.61 $\pm$ 0.06	27.96 $\pm$ 0.02	22.93 $\pm$ 0.09	11.35 $\pm$ 5.4	29 $\pm$ 2.56	9.62 $\pm$ 6.03	20.16 $\pm$ 1.07
KNR	28.03	18.01	28.55	21.27	15.98	26.66	16.93	17.65
SVR	20.06 $\pm$ 5.97	18.21 $\pm$ 0.75	23.05 $\pm$ 15.47	24.8 $\pm$ 4.07	15.07 $\pm$ 5.09	23.5 $\pm$ 1.5	11.31 $\pm$ 6.72	26.39 $\pm$ 1.76
	t-tau							
	CLINVAR		SLEEPVAR		PSGVAR		ALL	
	Train	Test	Train	Test	Train	Test	Train	Test
EM	127.62 $\pm$ 42.83	159.9 $\pm$ 12.63	99.8 $\pm$ 42.75	212.37 $\pm$ 14.28	135.38 $\pm$ 43	182.84 $\pm$ 36.33	71.86 $\pm$ 57.53	176.69 $\pm$ 22.85
RLR	188.72 $\pm$ 12.03	140.95 $\pm$ 16.46	200.41 $\pm$ 3.55	202.34 $\pm$ 4.99	176.73 $\pm$ 0.05	154.43 $\pm$ 0.02	125.48 $\pm$ 4.63	177.74 $\pm$ 2.86
GP	191.49 $\pm$ 25.43	156.53 $\pm$ 34.57	225.73 $\pm$ 0.56	170.56 $\pm$ 12.81	190.18 $\pm$ 0.02	154.97 $\pm$ 0.19	168.32 $\pm$ 11.96	146.72 $\pm$ 21.1
KNR	226.55	158.10	152.89	196.31	203.38	189.49	159.87	161.65
SVR	193.99 $\pm$ 30.78	139.78 $\pm$ 8.45	188.53 $\pm$ 53.44	227.01 $\pm$ 34.48	172.85 $\pm$ 14.93	165.94 $\pm$ 15.26	116.96 $\pm$ 38.37	188.48 $\pm$ 0.6

We tried with four different datasets: CLINVAR, clinical variables; SLEEPVAR, conventional sleep parameters assessed by expert physicians; PSGVAR, Parameters that were computed over the PSG signals; ALL, the three previous ones combined.

transmission, malfunction of intracellular calcium stores, and neuroinflammation (Lall et al., 2019). More closely, while the relationship between SaO<sub>2</sub> min and A $\beta$ 42 is not linear, our results revealed a positive correlation between CSF A $\beta$ 42 and CT90, surprisingly indicating that more severe hypoxemia is associated with higher CSF A $\beta$ 42 levels. This finding could be underpinned by the protective effect of hypoxic preconditioning (Shah et al., 2013), a phenomenon known to activate endogenous neuroprotective mechanisms in models of cerebral hypoxic and ischemic conditions (Wang et al., 2017). In line with our observations, animal studies have suggested that hypoxic preconditioning can mitigate memory impairment and A $\beta$  pathology (Zhang et al., 2020). However, these findings need to be corroborated with alternative analytical approaches and through longitudinal follow-up of our cohort.

Additionally, our research indicates that the Mixed Apnea Index is another standard PSG parameter that is predictive of CSF A $\beta$ 42 levels, once again with low correlation coefficients. This highlights the importance of specifically characterizing respiratory events that compose the AHI to predict various

outcomes and develop customized therapeutic strategies accurately. Incorporating both apneas and hypopneas into a single measure, such as the AHI, for diagnosing and evaluating the severity of OSA, implies that these events are equivalent in their clinical implications. However, the existing evidence is not univocal (Kulkas et al., 2017; Spector et al., 2019). Interestingly, recent research appears to support our hypothesis, showing that in a population with mild-to-moderate AD, hypopnea and apnea events are distinctively associated with sleep architecture, levels of pathological AD markers, and cognitive decline (Targa et al., 2023).

Similarly, the predictive role of the AI for CSF A $\beta$ 42 that we identified—despite the association not being linear—underscores the importance of detailed and personalized sleep analysis in early AD. Corroborating our findings, studies have demonstrated that arousals differ in their oscillatory composition and have various associations with early AD-related amyloid neuropathology and cognitive function (Chylinski et al., 2021). Future studies are essential to explore these observations further and to determine if there are causal relationships.

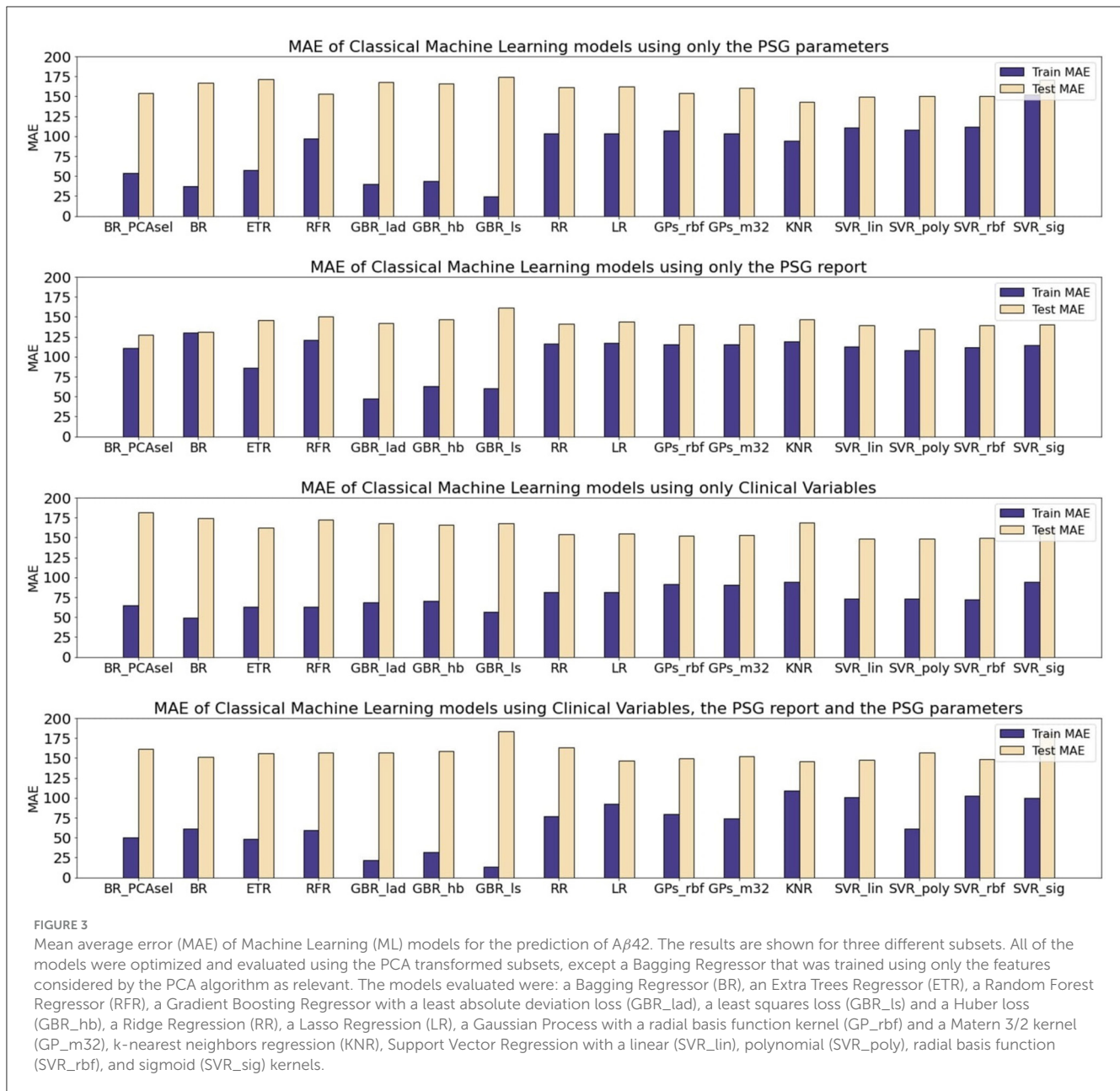


FIGURE 3

Mean average error (MAE) of Machine Learning (ML) models for the prediction of  $A\beta_{42}$ . The results are shown for three different subsets. All of the models were optimized and evaluated using the PCA transformed subsets, except a Bagging Regressor that was trained using only the features considered by the PCA algorithm as relevant. The models evaluated were: a Bagging Regressor (BR), an Extra Trees Regressor (ETR), a Random Forest Regressor (RFR), a Gradient Boosting Regressor with a least absolute deviation loss (GBR\_lad), a least squares loss (GBR\_ls) and a Huber loss (GBR\_hb), a Ridge Regression (RR), a Lasso Regression (LR), a Gaussian Process with a radial basis function kernel (GP\_rbf) and a Matern 3/2 kernel (GP\_m32), k-nearest neighbors regression (KNR), Support Vector Regression with a linear (SVR\_lin), polynomial (SVR\_poly), radial basis function (SVR\_rbf), and sigmoid (SVR\_sig) kernels.

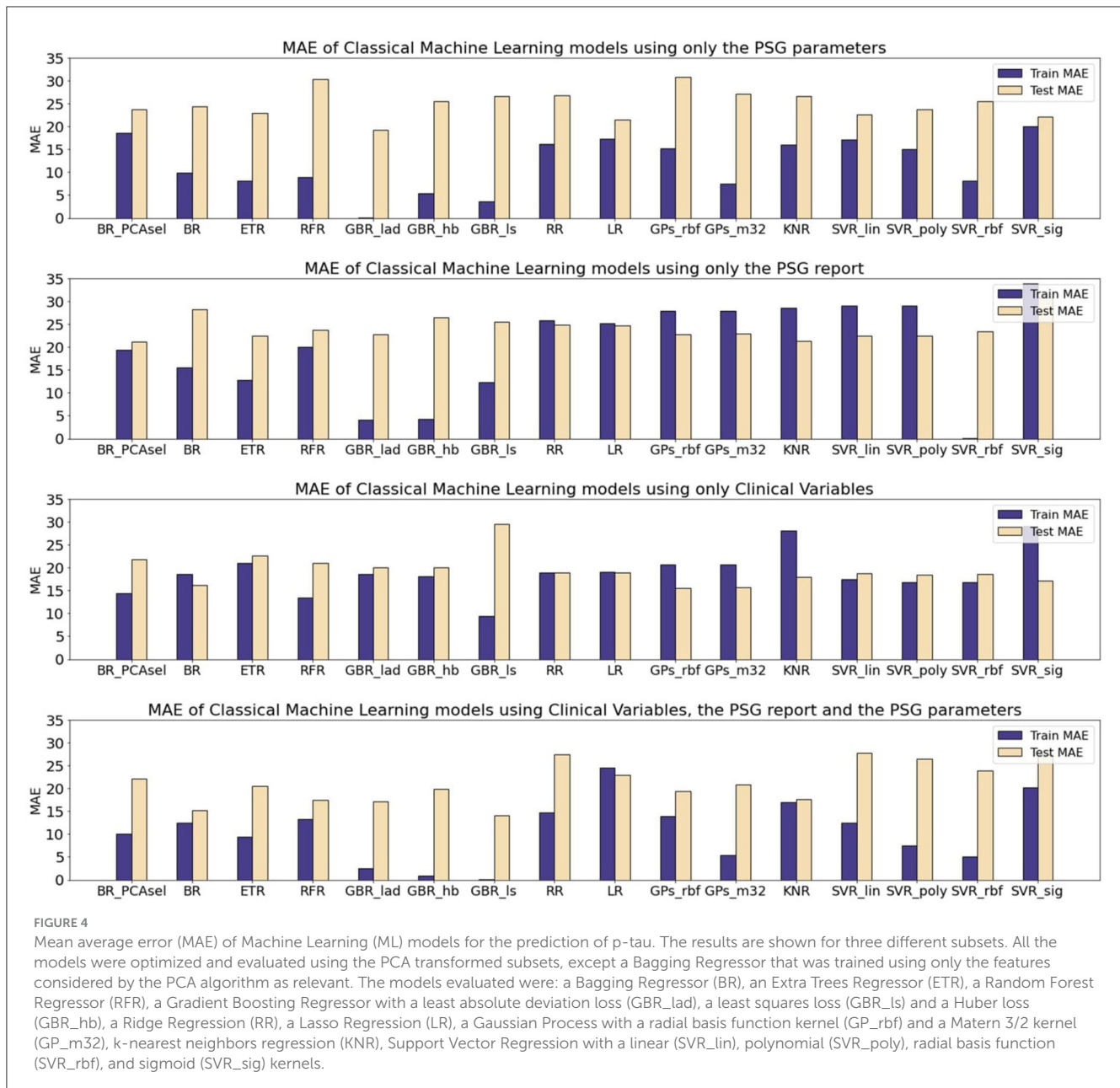
## 6.2 Best data subset for CSF p-tau

The most effective combination of feature groups for predicting CSF p-tau levels overall included quantitative PSG features, conventional PSG parameters, and clinical variables (ALL subset).

Regarding the role of quantitative PSG signal features (PSGVAR), our findings underscored the significance of the skewness of the EEG signal from the O1-A2 channel during the N1 and N2 sleep stages. We noted a significant positive linear correlation with p-tau levels.

Levels of p-tau have been shown in previous studies to correlate well with the spread of tau pathology from temporal regions to

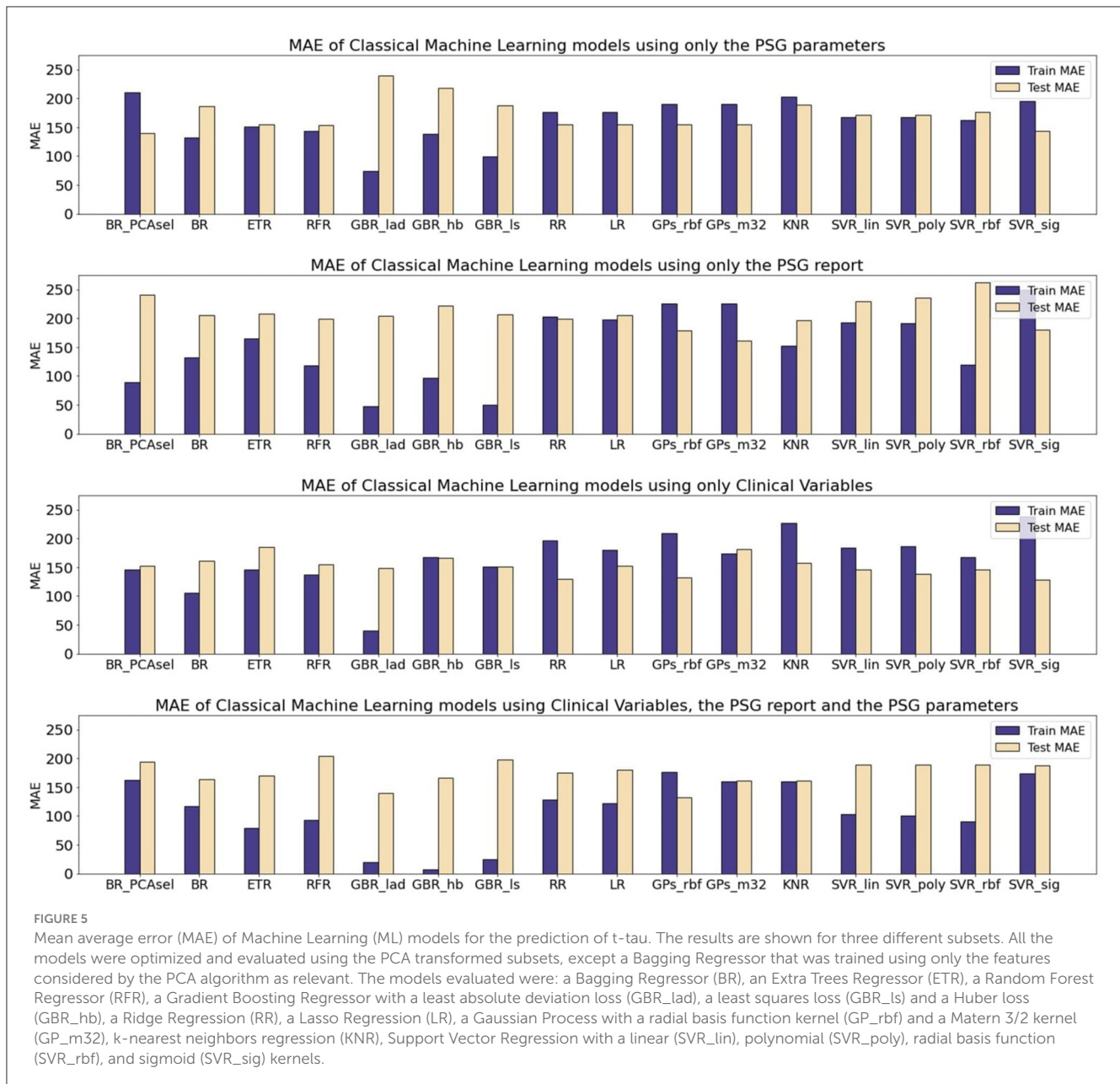
other areas of the brain. Tau pathology, more than amyloid, is closely associated with observed atrophy in MRI, hypometabolism in FDG-PET, and cognitive symptoms in AD (Jack et al., 2018). Additionally, tau pathology typically spreads from the temporal to more posterior structures, particularly the parietal regions (Masters et al., 2015). Consequently, we hypothesize that higher levels of p-tau may correlate with increased tau pathology in these posterior structures. Moreover, skewness, as a measure of asymmetry in the probability distribution of a signal (Xiang et al., 2020), may reflect this increased pathology, potentially resulting in more pronounced alterations in EEG signals from these regions. This could render the signal more unpredictable and with distinct spatial patterns, given the positive correlation between the two variables.



This discovery may represent one of the most significant and unforeseen findings of our study, suggesting specific quantitative PSG features, such as sleep EEG, could serve as potential early biomarkers of AD. Prior research has supported the potential of EEG recordings for the early detection of AD. Furthermore, in line with our findings, other studies (Jeong, 2004; Ghorbanian et al., 2015; Liu et al., 2016) have indicated that a characteristic EEG abnormality in AD patients is a generalized slowing of the rhythms and a reduction in complexity across various brain regions. Interestingly, these EEG abnormalities in AD have been correlated with disease severity, as they directly reflect the anatomical and functional changes in the cerebral cortex that are affected by the disease (Kowalski et al., 2001).

Among the conventional PSG variables (SLEEPVAR), we find it noteworthy to highlight that our models selected minimum SaO<sub>2</sub> as one of the parameters predictive of CSF p-tau levels, suggesting that more severe hypoxemia may lead to an increase in CSF p-tau levels, given the positive relationship. This finding aligns with previous results indicating that hypoxia triggers tau hyperphosphorylation and memory deficits in rats (Zhang et al., 2014).

Besides, our results regarding the clinical variables subset (CLINVAR) appear to corroborate existing evidence that arterial hypertension is a statistical predictor of p-tau accumulation, in line with previous findings, suggesting that hypertension may elevate the risk for AD (Lennon et al., 2021). Meanwhile, our data also indicate that metabolic clinical markers, including BMI, waist circumference, and serum insulin and glucose levels, are implicated



in the prediction of p-tau protein levels in the CSF in AD, as explored in previous studies (Gonçalves et al., 2019; Lee et al., 2023). However, in our study, these relationships did not follow a linear pattern.

### 6.3 Best data subset for CSF t-tau

As observed for the prediction of CSF p-tau, the ALL subset also emerged as the best subset of variables for CSF t-tau levels.

Interestingly, regarding the PSGVAR subset, we found that the Lempel-Ziv and Sample Entropy—both of which are measures of signal complexity (Grassberger and Procaccia, 1983)—are pertinent to predict t-tau levels. Our results indicated that the complexity of

the thoracic effort signals, increases with the levels of t-tau during NREM sleep.

Increased complexity of the thoracic effort signal may be associated with OSA syndrome, which is highly prevalent in our study population (Gaeta et al., 2020). It is known that increased respiratory effort against collapsed airways leads to elevated intrathoracic disturbances (Farré et al., 2004; Sánchez-de-la Torre et al., 2013). Consequently, a higher complexity in the thoracic effort signal may imply a more severe manifestation of OSA, thus, the positive correlation between Lempel-Ziv complexity and Sample Entropy of thoracic effort and tau CSF biomarkers could indicate an indirect link between OSA syndrome and the advancement of AD tau pathology. However, we did not find any correlation between t-tau and the AIH, affirming the hypothesis that a more precise definition of sleep-related breathing disorders in AD is necessary, as previously discussed (Targa et al., 2023).

Several mechanisms have been proposed to explain this phenomenon (Andrade et al., 2018). Previous investigations suggest that the effort to breathe against a collapsed upper airway during obstructive sleep apnea episodes may cause repetitive high-pressure fluctuations, potentially impacting the glymphatic system, thereby altering the concentration of CSF metabolites and biomarkers (Andrade et al., 2018). The glymphatic system operates optimally during SWA, facilitating the removal of  $A\beta$  and other metabolites from the interstitial space. Studies conducted with mouse models of AD have demonstrated that sleep deprivation acutely increases soluble  $A\beta$  levels and promotes chronic amyloid deposition (Ju et al., 2017). In human subjects, sleep deprivation has been observed to elevate cerebrospinal fluid (CSF)  $A\beta$  levels. Consequently, it has been theorized that frequent arousals in obstructive sleep apnea (OSA) contribute to reduced SWA, leading to elevated  $A\beta$  levels and ultimately increasing  $A\beta$  aggregation into amyloid, thereby heightening the risk of AD (Ju et al., 2016). Furthermore, persistent sleep fragmentation resulting from OSA disrupts CSF-ISF (Interstitial Spinal Fluid) exchange, leading to an accumulation of  $A\beta$  (Ju et al., 2016).

Nonetheless, contradictory findings have been documented in the literature. For example, Ju et al. (2016) observed lower concentrations of  $A\beta_{42}$ , tau, and other derived metabolites in the CSF of patients with OSA. However, they noted that total protein levels, which are primarily derived from blood albumin, did not decrease in patients with severe OSA compared to control subjects (Ju et al., 2016), suggesting that glymphatic clearance processes may be compromised in OSA (Aspelund et al., 2015). This discrepancy underscores the need for future research that includes larger populations to compare OSA patients with non-OSA individuals.

We hypothesize that akin to the effect on CSF amyloid, the increased intrathoracic and intracranial pressures induced by OSA acutely and repeatedly may hinder the circulation of brain metabolites from ISF into CSF, thereby fostering tau accumulation. Our hypothesis is supported by prior studies revealing compromised CSF-ISF exchange polarization in a mouse model of tauopathy, indicating the potential of this clearance pathway to worsen or even initiate pathogenic tau accumulation (Harrison et al., 2020). Comprehensive analyzes are required.

Moreover, regarding the predictive value of the conventional PSG parameter (SLEEPVAR subset), our findings concur with earlier research, which has shown a negative association between tau pathology and time spent in N3 sleep stage. Remarkably, SWA during N3 sleep is widely reported to be disrupted in AD (Mander et al., 2015). Furthermore, consistent with our findings, has been demonstrated an inverse correlation between SWA and the AD tau pathology, as evidenced by PET imaging and CSF tau biomarker levels in the early stages of the disease (Lucey et al., 2019).

On the other hand, the results concerning the CLINVAR subset, while not displaying a significant linear correlation, overlap with those observed for p-tau, confirming the importance of metabolic measures, particularly insulin levels, in tau pathology. Insulin, a key regulator of glucose homeostasis and metabolism, plays a critical role in various brain functions, including synaptic plasticity, learning, and memory. Besides its traditional function as a microtubule-stabilizing protein, tau also serves as a scaffold protein interacting with components of the brain's insulin signaling

pathway (Gonçalves et al., 2019). Impairment in brain insulin signaling has been consistently linked to cognitive decline in animal models and humans. Insoluble fractions of post-mortem AD and other tauopathies have revealed hyperphosphorylated tau-containing neurons with accumulated insulin oligomers (Gonçalves et al., 2019). While the exact influence of tau pathology on insulin signaling remains unclear, it's established that insulin resistance can lead to tau hyperphosphorylation and cognitive decline in both human and animal models (Gonçalves et al., 2019). However, the negative relationship between insulin serum levels and CSF t-tau we observed, remains to be more extensively investigated.

## 6.4 Related works

In light of the existing body of research, our study is among the first to evaluate a variety of ML models trained to non-invasively predict AD neuropathology by integrating sleep data, clinical variables crucial for AD development, conventional PSG parameters, and a broad spectrum of quantitative signal features from sleep data. Our approach stands out by employing a multi-domain quantitative signal feature analysis, encompassing both linear and non-linear aspects, which extends beyond the conventional spectral features typically utilized in prior studies. Additionally, our analysis encompasses the entirety of PSG signals and is not limited solely to EEG data. The landscape of ML applications in AD diagnosis and progression encompasses a diverse range of techniques and input variables. To illustrate, Table 11 presents a summary of pertinent studies, highlighting the ML techniques and proposed biomarkers at various stages of AD diagnosis. Key variables such as educational level, professional experience, social engagement, dietary habits, APOE genotype, and age have been identified as significant predictors for the conversion to MCI in a substantial elderly cohort over a 5-year timeframe, utilizing Random Forest and permutation-based Machine Learning methods (Gómez-Ramírez et al., 2020).

Recent studies have increasingly focused on sleep-related variables, parameters, and quantitative signal features, utilizing techniques. However, many of these studies did not simultaneously incorporate various types of sleep measures, where not conducted explicitly on AD populations, or limited their quantitative signal analysis to resting state EEG data. For example, one study employed advanced ML models to assess the impact of sleep disturbances on the development of AD and the transition from normal cognitive (NC) functioning to MCI. Employing Random Tree classifiers, the most effective model achieved a 70% accuracy rate in predicting the transition from MCI to AD, with factors such as caregiver distress, frequency of sleep disturbances, and excessive daytime sleepiness playing significant roles. Logistic regression and KNN models predicted with 60% accuracy re-conversion from MCI to NC, with caregiver distress, sleep disturbance frequency, and OSA most pronounced among males (Bonfa et al., 2023).

Moreover, research has confirmed the link between dementia and sleep disturbance, with gradient boosting emerging as the most accurate method, boasting a 92.9% accuracy rate, and F1-score of 0.926, and ROC AUC of 0.974, and a Brier score of 0.056. The

TABLE 11 Relevant studies on biomarkers and Machine Learning (ML) models of Alzheimer's disease (AD).

References	Proposed AD biomarkers/ input variables	Machine Learning models	Results	Normal cognition/ Alzheimer's disease stage
Gómez-Ramírez et al. (2020)	Clinical variables	Random Forest (RF) Permutation-based Machine Learning (PBML)	Educational level, working experience, social life, subjective cognitive decline, diet rich in sweets, APOE and age have been selected as some of the most important variables for MCI conversion in a large elderly population, using RF that showed a high predictive performance with an accuracy of 0.851	Conversion from NC to MCI
Bonfa et al. (2023)	Sleep disturbances	Random Tree classifiers Logistic regression K-nearest neighbors (KNN)	Based on Random Tree classifiers, the best model predicted a 70% accuracy change from MCI to AD, with caregiver distress, sleep disturbance frequency, and excessive daytime sleepiness playing significant roles. Logistic regression and KNN models predicted with 60% accuracy re-conversion from MCI to NC, with caregiver distress, sleep disturbance frequency, and OSA most pronounced among males.	Conversion from NC to MCI and from MCI to other stages of AD
Nyholm et al. (2024)	Sleep and clinical variables	Gradient Boosting (GB) Logistic regression Gaussian naive Bayes (GNB) Random Forest (RF) Support Vector Machine (SVM)	GB resulted the most accurate method, with 92.9% accuracy, 0.926 f1-score, 0.974 ROC AUC, and a Brier score of 0.056. Sleeping more than 2 h per day, gender, education level, age, waking up throughout the night, and snoring were the variables with the greatest feature importance across all ML algorithms.	Age > 60 years from the Swedish National Study on Aging and Care (SNAC) including NC, MCI and moderate dementia
Abate et al. (2020)	Clinical variables and plasma p53	Regression Tree (RT)	Plasma p53 was found more relevant for identifying accurately classify (AUC = 0.92) A $\beta$ <sup>+</sup> /amnesic Mild Cognitive Impairment (aMCI) patients who will develop AD.	MCI
Chang et al. (2021)	Plasma D-glutamate	Support Vector Machine (SVM) Logistic regression Random Forest (RF) Naïve Bayes	The naïve Bayes model and RF model were the best models for determining MCI and AD susceptibility based on D-glutamate, respectively (area under the receiver operating characteristic curve: 0.8207 and 0.7900; sensitivity: 0.8438 and 0.6997; and specificity = 0.8158 and 0.9188, respectively).	MCI Others stages of AD
Jo et al. (2020)	Tau- Positron Emission Tomography (PET)	Convolutional Neural Network (CNN) Layer-wise relevance propagation (LRP) algorithm	Deep learning-based classification model framework combining 3D CNN and LRP algorithms of AD from NC yielded an average accuracy of 90.8%	Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort including HC, MCI and others AD stages
Popuri et al. (2020)	Magnetic Resonance Images (MRI) t-tau A $\beta$ 1-42 CSF	Ensemble-learning (Multi-Kernel classifier, Variational Bayes Probabilistic Multi-Kernel Learning)	MDART (MRI-based on Dementia Alzheimer's type DAT) score achieved a classification performance on stable vs. progressive MCI groups with an AUC of 0.81 for TTC of 6 months and 0.73 for TTC of up to 7 years	NC MCI and others AD stages
Pirrone et al. (2022)	Resting-state EEG power spectral density features	Decision Trees (DT) Support Vector Machines (SVM) K-nearest neighbor (KNN)	K-NN was the best classification algorithm as concerning the accuracy reaching 97, 95, and 83% accuracy when considering binary classifications (HC vs. AD, HC vs. MCI, and MCI vs. AD) and an accuracy of 75% when dealing with the three classes (HC vs. AD vs. MCI).	HC, MCI, others stage AD
Kim et al. (2023)	Resting-state EEG power spectral density features	Support Vector Machine (SVM) Logistic regression K-nearest neighbors (KNN) Gaussian Naïve Bayes (GNB) Random Forest (RF) AdaBoost (Ada) GBMBoost (GBM) XGBoost (XGB)	The best model demonstrated 90.9% sensitivity, 76.7% specificity, and 82.9% accuracy in MCI+SCD (33 A $\beta$ <sup>+</sup> , 43 A $\beta$ <sup>-</sup> ). Limited to SCD, 92.3% sensitivity, 75.0% specificity, and 81.1% accuracy (13 A $\beta$ <sup>+</sup> , 24 A $\beta$ <sup>-</sup> ). 90% sensitivity, 78.9% specificity, and 84.6% accuracy for MCI (20 A $\beta$ <sup>+</sup> , 19 A $\beta$ <sup>-</sup> ).	Subjective cognitive decline (SCD), MCI
Kim et al. (2021)	Resting-state EEG power spectral density features	Multi-Model Ensembles Genetic Algorithms Support Vector Machine (SVM)	Genetic Algorithm Heuristic achieved 85.7% sensitivity, 89.3% specificity, and 88.6% accuracy in (SCD) amyloid positive/negative classification, and 83.3% sensitivity, 85.7% specificity, and 84.6% accuracy in MCI amyloid positive/negative classification.	Subjective cognitive decline (SCD), MCI
Gaubert et al. (2021)	Clinical variables Resting-state EEG power spectral density features (qEEG) Magnetic Resonance Imaging (MRI)	Random Forest (RF) Logistic regression Support Vector Machine (SVM)	qEEG was the strongest predictor of neurodegeneration, with 82% negative predictive value (NPV), 38% positive predictive value (PPV), 77% specificity and 45% sensitivity. The combination of demographic, neuropsychological data, ApoE $\epsilon$ 4, and MRI hippocampal volumetry most strongly predicted amyloid (80% NPV, 41% PPV, 70% specificity, 58% sensitivity) and most strongly predicted decline to prodromal AD at 5 years (97% NPV, 14% PPV, 83% specificity, and 50% sensitivity).	French INSIGHT-preAD cohort, which includes baseline data cognitively normal individuals, between 70 and 85 years old, with subjective memory complaints but unimpaired cognition, prodromal AD

(Continued)



TABLE 11 (Continued)

References	Proposed AD biomarkers/ input variables	Machine Learning Models	Results	Normal cognition/ Alzheimer's disease stage
Geng et al. (2022)	Sleep slow-waves and spindles data, spectral and complexity features from sleep EEG.	Support Vector Machine (SVM) Gate Recurrent Unit network techniques (GRU)	The MCI classification accuracy of the GRU network based on features extracted from sleep EEG was the highest at 93.46%.	MCI, HC
Khosroozad et al. (2023)	Power Spectral Density features from PSG	Neural Networks (NN) Kernel algorithms	Time-Lag (TL), extracted from high-frequency movements and respiratory alterations during sleep early detect MCI in AD, resulting in high sensitivity (86.75% for NN and 65% for Kernel), specificity (89.25 and 100%), and accuracy (88 and 82.5%).	NC, MCI

Ada, AdaBoost; AD, Alzheimer's Disease; AUC, Area under curve; CNN, Convolutional Neural Network; DT, Decision Trees; EEG, Electroencephalography; GNB, Gaussian Naive Bayes; GRU, Gate Recurrent Unit network techniques; GBM, GBMBoost; GB, Gradient Boosting; KNN, K-nearest neighbors; LRP algorithm, Layer-wise relevance propagation; MRI, Magnetic Resonance Images; MCI, Mild Cognitive Impairment; NPV, Negative predictive value; NN, Neural Networks; NC, normal cognition; PBML, Permutation-based Machine Learning; PET, Positron Emission Tomography; PSG, Polysomnography; PPV, Positive predictive value; RF, Random Forest; RT, Regression Tree; SCD, Subjective cognitive decline; SVM, Support Vector Machine; SNAC, Swedish National Study on Aging and Care; TL, Time-Lag; XGB, XGBoost.

relevant factors differed amongst ML algorithms, while sleeping more than 2 h per day, gender, education level, age, waking up throughout the night, and snoring were the variables with the greatest feature importance across all ML algorithms (Nyholm et al., 2024).

Interestingly, Kim et al. (2021) developed a spectral qEEG-ML algorithm to predict AD defined by  $A\beta$ -PET positivity among patients with subjective cognitive decline (SCD) and MCI. Genetic Algorithm Heuristic achieved 85.7% sensitivity, 89.3% specificity, and 88.6% accuracy in SCD amyloid positive/negative classification, and 83.3% sensitivity, 85.7% specificity, and 84.6% accuracy in MCI amyloid positive/negative classification, confirming previous results. The authors used other types of EEG-ML algorithms to detect brain  $A\beta$  pathology validated with  $A\beta$  PET. The best model demonstrated 90.9% sensitivity, 76.7% specificity, and 82.9% accuracy in MCI + SCD (33  $A\beta^+$ , 43  $A\beta^-$ ), suggesting that qEEG could serve as a promising biomarker for  $A\beta$  amyloid pathology (Kim et al., 2023).

Furthermore, Gaubert et al. (2021) evaluated the efficacy of an ML approach in achieving early AD diagnosis based on  $A\beta$ -PET and MRI measurement, introducing a multimodal non-invasive biomarker strategy. Importantly, qEEG was the strongest predictor of neurodegeneration, with 82% negative predictive value, 38% positive predictive value, 77% specificity, and 45% sensitivity. Besides the combination of demographic, neuropsychological data, ApoE4 and MRI hippocampal volumetry most strongly predicted amyloid (80% NPV, 41% PPV, 70% specificity, and 58% sensitivity).

Remarkably, innovative research utilizing Support Vector Machine classifiers and Gate Recurrent Unit network techniques has demonstrated the potential of sleep EEG data, including slow waves and spindles, to distinguish MCI from HC with high accuracy (Geng et al., 2022). The MCI classification accuracy of the GRU network based on features extracted from sleep EEG was the highest at 93.46%. Their experimental findings demonstrate that, as compared to awake EEG, sleep EEG can give more helpful information for identifying MCI and HC.

Additionally, pioneering work has proposed a novel diagnostic approach for MCI in AD, combining sleep-related movements with advanced signal processing and ML techniques (Khosroozad et al.,

2023). This approach, which explores the relationship between high-frequency movements and respiratory changes during sleep, highlights the potential of advanced quantitative sleep signal feature analyzes in a multimodal ML setting for early detection of AD neuropathology. The approach employs Neural Networks and Kernel algorithms, resulting in high sensitivity (86.75% for NN and 65% for Kernel), specificity (89.25 and 100%), and accuracy (88 and 82.5%) for early detection of MCI in AD.

Despite these advancements, the extensive exploration of advanced quantitative sleep signal features to non-invasively detect the AD neuropathology within a multimodal ML framework remains a promising area for future research.

## 6.5 Strengths and limitations of the study

To our knowledge, our study for the first time integrated clinical variables, conventional PSG parameters, and quantitative PSG signal features for the non-invasive prediction of core CSF biomarkers of AD using various ML models.

This study is novel in that it applies computational engineering to extract a wide range of quantitative signal features from thoroughly preprocessed PSG signals, identifying potential biomarkers for the early detection of AD. To date, no research has delineated the relationship between these specific quantitative PSG features and AD-related CSF biomarkers of neurodegeneration, suggesting that sleep qEEG could demonstrate a topographic specificity in their associations. Our approach involved working with spectral, time-domain, and non-linear parameters directly extracted from electrophysiological signals, indicating that PSG recordings may contain additional crucial information for the diagnosis of AD.

A significant aspect of this research is the rigorous cross-validation ML methodology employed. We trained multiple ML models to determine which could deliver the best performance for our outcomes. The models most effective at addressing the overfitting issue were Gradient Boosting Regressors, utilizing a least absolute deviation loss function for  $A\beta_{42}$ , a least squares

loss for p-tau and a Huber loss for t-tau. The other optimized ML models generally handled the scarcity of training data by using regularization to construct simpler models that accept higher training errors.

Moreover, in assessing correlations between the selected variables and the CSF biomarkers, we found that not all variables showed significant linear associations with the targets. This demonstrates how ML facilitates the exploration of non-linear relationships that may not be detectable using conventional statistical methods, in addition to providing several other advantages like as fault tolerance and real-time operation, making them suitable for complex applications (Chang et al., 2021). In the realm of modeling, linearity denotes a system or model's characteristic where the output varies directly with the input, while non-linearity suggests a more intricate relationship between input and output, eluding simple linear expressions attainable through conventional statistical means. ML prioritizes precise predictions, while traditional statistical models concentrate on uncovering interrelations among variables. ML offers advantages in flexibility and scalability over conventional statistical techniques, rendering it suitable for diverse tasks such as diagnosis, classification, and survival prognosis. A fundamental disparity between ML and traditional statistical methods lies in their primary objectives: the former emphasizes predictive accuracy, while the latter focuses on deducing relationships among variables. Additionally, ML adeptly addresses interactions, which are difficult to investigate with traditional statistical methods that primarily handle interactions between the principal determinant and single potential confounders (Rajula et al., 2020).

We acknowledge the limitations of our study, which are important to consider when interpreting the results. One of the primary limitations is the reliance on a relatively small sample size of 49 participants, which may limit the generalizability of our findings. Furthermore, ethical considerations prevented us from conducting a comparative analysis with a control group. The absence of neuroimaging assessments, such as MRI, which are known to correlate with disease progression and serve as critical predictive markers for AD in patients with MCI (Jack et al., 2018), is another limitation that merits attention. Moreover, we recognize a notable limitation in our study, as patients were included solely based on the 2011 NIA-AA clinical criteria (McKhann et al., 2011), without considering the updated 2018 biological criteria (Jack et al., 2018), due to the trial registration in 2016. This approach may not fully capture the evolving understanding of the disease, where biomarkers exhibit changes years before symptom onset. The presence of normal biomarker levels, especially amyloid pathology, in individuals with dementia suggests the potential for an alternative diagnosis (Jack et al., 2018). Given these limitations, we recommend a cautious interpretation of our results, although we believe they contribute valuable insights.

The models that excelled in predicting the levels of the three biomarkers were the Gradient Boosting Regressors, which utilized a LAD loss for A $\beta$ 42, an LS loss for p-tau, and a Huber loss for t-tau. These are among the most common loss functions for regression challenges. The distinctions between them are as follows:

1. LAD loss is less sensitive to outliers but backs a closed-form solution due to its non-continuous derivatives;
2. LS loss, on the other hand, penalizes large deviations more severely, making it less robust to outliers, yet it tends to yield more stable solutions (Natekin and Knoll, 2013);
3. The Huber loss integrates aspects of both LAD and LS losses (Natekin and Knoll, 2013).

For instance, in predicting A $\beta$ 42 levels, outliers could be disregarded with relative safety. In contrast, outliers played a more critical role in optimizing the models for the tau-related biomarkers. Examining Figure 1 reveals that p-tau and t-tau have similar density distributions. Additionally, boosting techniques are generally recognized for their low bias, which is the discrepancy between the average prediction of the model and the actual value. Put simply, a high-bias model tends to be overly simplistic and does not fit the training data well, leading to significant errors in both the training and the test datasets. However, it is essential to consider that a low-bias model often has high variance, increasing the risk of overfitting the training data.

Thus, despite the Gradient Boosting Regressors achieving some of the lowest training errors, their MAEs on the test set were still considerable. This suggests that, even with cross-validation to fine-tune the models for an optimal balance in the bias-variance trade-off, the models may still not generalize well beyond the training data. In terms of ensemble models employing bagging techniques, they typically displayed test errors comparable to those of boosting methods but with increased training errors. Models such as the Bagging Regressor (BR), Random Forest Regressor (RFR), and Extra Trees Regressor (ETR) function by aggregating the decisions for T individual decision trees, each trained on a bootstrapped subset of the data. The final prediction is the majority vote across these T trees. This approach reduces the variance associated with single decision trees, resulting in a model that is less prone to overfitting and demonstrates improved generalization. Furthermore, decision trees, like those in a random forest, provide insight into the importance of features during training. However, the application of PCA for dimensionality reduction, while preventing noise overfitting, obscures the interpretability of model decisions. Other models generally address overfitting by deliberately allowing higher training errors, thereby enhancing their generalization capabilities to new data at the expense of complexity.

Gaussian Process (GP) models, in contrast, exhibited higher training errors than test errors, signifying potential underfitting. This could indicate that the selected kernel was insufficient in capturing the complexity of the data, potentially necessitating a more sophisticated kernel for optimization. Although GPs can assess feature importance through kernel length scales, this interpretability is lost when PCA is applied. Support Vector Regressors (SVRs) struck a better balance between training and test errors, often with training errors marginally exceeding those of other models. The K-Nearest Neighbors Regressors (KNN) also managed the bias-variance trade-off effectively, sometimes with a training error even surpassing the test error.

Regarding the analysis of the three biomarkers, the lowest training errors were associated with the "ALL" subset (incorporating clinical, PSG, and quantitative PSG-derived variables), though test evaluations varied. This could be attributed to the retention of a greater number of principal components

for the “ALL” subset compared to others, potentially facilitating overfitting. Additionally, the principal components of the “ALL” subset represent linear combinations of a broader range of features, thus encapsulating more detailed information about each sample. For the A $\beta$ 42 biomarker, the “SLEEPVAR” subset—comprising solely sleep parameters evaluated by experts—yielded the lowest MAE during testing. However, it is important to note that none of the models achieved optimal performance. Therefore, while the findings are significant, they should be interpreted with caution. To ensure the robustness of our findings across various populations, we advocate for subsequent validation studies to be conducted with independent cohorts of cognitively normal subjects. Additionally, enhancing the predictive accuracy of our models through incorporating data from multiple preclinical cohorts and applying longitudinal, multimodal measures within a nested cross-validation framework represents a promising avenue for future research.

An intriguing prospect for further investigation is the integration of neuropsychological assessments and MRI data as well as blood biomarkers into our analytical framework. This approach would allow for a thorough longitudinal evaluation of the clinical progression toward prodromal AD, leading to a deeper comprehension of neurodegenerative mechanisms. Additionally, it would improve the predictive accuracy of our models and provide less invasive methods for early AD detection. Upon validation of our initial results, a critical next step would be to identify the most predictive set of features for neurodegeneration, using performance metrics as a guide. Advanced artificial intelligence techniques could then be employed to determine the ability of these features to differentiate between AD, MCI and HC groups. Moreover, the exploration of automated analysis of PSG recordings through neural network integration into our model offers a promising direction. Such automation could potentially eliminate the need for manual annotation by enabling the automatic detection of artifacts and sleep stages.

## 7 Conclusions

This study highlights the potential of ML to assess asymptomatic individuals at risk for AD through the analysis of non-invasive and cost-effective biomarkers, underscoring the ability of ML to uncover complex non-linear relationships within intricate datasets, that may elude traditional statistical methods. This could offer supplementary insights alongside other biomarkers, hinting at AD pathology in asymptomatic individuals, or functioning as an additional diagnostic tool for those ineligible for CSF biomarkers determination.

Ours results also suggest that relying solely on a single type of biomarker may not suffice for a reliable AD early detection. Notably, we emphasize the importance of specific quantitative PSG signal features such as EEG skewness, Lempel Ziv and Sample Entropy of thoracic effort signals, as reliable markers for predicting neurodegeneration, along with conventional PSG parameters. The utilization of portable PSG devices may establish a groundwork for their utilization in clinical environments.

The demonstration of the viability of these innovative approaches underlines their potential contribution to the early

diagnosis of AD, particularly through the prediction of core CSF biomarkers and the exploration of their relationships with sleep patterns.

## Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author.

## Ethics statement

The studies involving humans were approved by the study adhered to the principles of the Declaration of Helsinki and received approval from the Ethics Committee of Hospital Arnau de Vilanova de Lleida (CE-1218). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

## Author contributions

AG: Conceptualization, Data curation, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. MQ-L: Conceptualization, Formal analysis, Validation, Visualization, Writing – original draft, Writing – review & editing. FB: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing. RV: Data curation, Writing – review & editing. MP: Data curation, Writing – review & editing. OM: Data curation, Writing – review & editing. MS: Conceptualization, Investigation, Methodology, Writing – review & editing. AM-B: Conceptualization, Funding acquisition, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing. GP-R: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Validation, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was partially funded by Ministerio de Ciencia, Innovación y Universidades, Agencia Estatal de Investigación, under grant PID2019-109820RB-I00, MCIN/AEI/10.13039/501100011033/, co-financed by European Regional Development Fund (ERDF), “A way of making Europe” and Spanish Ministry of Economy and Competitiveness, Institute of Health Carlos III (grant number P114/00328); by Generalitat of Catalonia, Department of Health (PERIS 2019 SLT008/18/00050) and Fundació La Marató TV3 (464/C/2014) to GP-R. Co-financed by FEDER funds from the European Union (“A way to build Europe”). IRBLleida is a CERCA Programme/Generalitat of Catalonia.

## Acknowledgments

We would like to thank the patients and personnel of the Hospital Universitari Santa Maria's Sleep and Dementia Unit, as well as the IRBLleida Biobank and Plataforma Biobancos, for their help.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

- Abásolo, D., Hornero, R., Gómez, C., García, M., and López, M. (2006). Analysis of EEG background activity in Alzheimer's disease patients with Lempel–Ziv complexity and central tendency measure. *Med. Eng. Phys.* 28, 315–322. doi: 10.1016/j.medengphy.2005.07.004
- Abate, G., Vezzoli, M., Polito, L., Guaita, A., Albani, D., Marizzoni, M., et al. (2020). A conformation variant of p53 combined with machine learning identifies Alzheimer disease in preclinical and prodromal stages. *J. Personal. Med.* 11:14. doi: 10.3390/jpm11010014
- Ahmadian, N., Hejazi, S., Mahmoudi, J., and Talebi, M. (2018). Tau pathology of Alzheimer disease: possible role of sleep deprivation. *Basic Clin. Neurosci.* 9:307. doi: 10.32598/bcn.9.5.307
- Aisen, P. S., Cummings, J., Jack, C. R., Morris, J. C., Sperling, R., Frölich, L., et al. (2017). On the path to 2025: understanding the Alzheimer's disease continuum. *Alzheimer's Res. Ther.* 9, 1–10. doi: 10.1186/s13195-017-0283-5
- Alvarez, D., Hornero, R., Abásolo, D., Del Campo, F., and Zamarrón, C. (2006). Nonlinear characteristics of blood oxygen saturation from nocturnal oximetry for obstructive sleep apnoea detection. *Physiol. Measur.* 27:399. doi: 10.1088/0967-3334/27/4/006
- Álvarez, D., Hornero, R., Marcos, J. V., and Del Campo, F. (2012). Feature selection from nocturnal oximetry using genetic algorithms to assist in obstructive sleep apnoea diagnosis. *Med. Eng. Phys.* 34, 1049–1057. doi: 10.1016/j.medengphy.2011.11.009
- Alvarez, D., Hornero, R., Marcos, J. V., Wessel, N., Penzel, T., Glos, M., et al. (2013). Assessment of feature selection and classification approaches to enhance information from overnight oximetry in the context of apnea diagnosis. *Int. J. Neural Syst.* 23:1350020. doi: 10.1142/S0129065713500202
- Andrade, A. G., Bubu, O. M., Varga, A. W., and Osorio, R. S. (2018). The relationship between obstructive sleep apnea and Alzheimer's disease. *J. Alzheimer's Dis.* 64, S255–S270. doi: 10.3233/JAD-179936
- Aspelund, A., Antila, S., Proulx, S. T., Karlsen, T. V., Karaman, S., Detmar, M., et al. (2015). A dural lymphatic vascular system that drains brain interstitial fluid and macromolecules. *J. Exp. Med.* 212, 991–999. doi: 10.1084/jem.20142290
- Azami, H., Moguilner, S., Penagos, H., Sarkis, R. A., Arnold, S. E., Gomperts, S. N., et al. (2023). EEG entropy in REM sleep as a physiologic biomarker in early clinical stages of Alzheimer's disease. *J. Alzheimer's Dis.* 2023, 1–16. doi: 10.3233/JAD-221152
- Babiloni, C., Visser, P. J., Frisoni, G., De Deyn, P. P., Bresciani, L., Jelic, V., et al. (2010). Cortical sources of resting EEG rhythms in mild cognitive impairment and subjective memory complaint. *Neurobiol. Aging* 31, 1787–1798. doi: 10.1016/j.neurobiolaging.2008.09.020
- Bandyopadhyay, A., and Goldstein, C. (2023). Clinical applications of artificial intelligence in sleep medicine: a sleep clinician's perspective. *Sleep Breath.* 27, 39–55. doi: 10.1007/s11325-022-02592-4
- Berry, R. B., Gamaldo, C. E., Harding, S. M., Brooks, R., Lloyd, R. M., Vaughn, B. V., et al. (2015). AASM scoring manual version 2.2 updates: new chapters for scoring infant sleep staging and home sleep apnea testing. *J. Clin. Sleep Med.* 11, 1253–1254. doi: 10.5664/jcs.m.5176
- Bonfa, A., Zhang, W., Keller, Y., Barbisch, F., Schürmann, P., Moguel, L. F., et al. (2023). Exploring the relationship between sleep disturbances and Alzheimer's disease using machine learning. *Sleep* 46(Suppl.1):A370. doi: 10.1093/sleep/zsad077.0840
- Borhani, N., Ghaisari, J., Abedi, M., Kamali, M., and Gheisari, Y. (2022). A deep learning approach to predict inter-omics interactions in multi-layer networks. *BMC Bioinform.* 23, 1–17. doi: 10.1186/s12859-022-04569-2
- Bubu, O. M., Brannick, M., Mortimer, J., Umasabor-Bubu, O., Sebastião, Y. V., Wen, Y., et al. (2017). Sleep, cognitive impairment, and Alzheimer's disease: a systematic review and meta-analysis. *Sleep* 40:zsw032. doi: 10.1093/sleep/zsw032
- Butterworth, S. (1930). On the theory of filter amplifiers. *Wireless Eng.* 7, 536–541.
- Bzdok, D., Altman, N., and Krzywinski, M. (2018). Statistics versus machine learning. *Nat. Methods* 15, 233–234. doi: 10.1038/nmeth.4642
- Chang, C. H., Lin, C. H., Liu, C. Y., Huang, C. S., Chen, S. J., et al. (2021). Plasma d-glutamate levels for detecting mild cognitive impairment and Alzheimer's disease: machine learning approaches. *J. Psychopharmacol.* 35, 265–272. doi: 10.1177/0269881120972331
- Chino-Vilca, B., Rodríguez-Rojo, I. C., Torres-Simón, L., Cuesta, P., Vendrell, A. C., Piñol-Ripoll, G., Huerto, R., et al. (2022). Sex specific EEG signatures associated with cerebrospinal fluid biomarkers in mild cognitive impairment. *Clin. Neurophysiol.* 142, 190–198. doi: 10.1016/j.clinph.2022.08.007
- Chylinski, D., Van Egroo, M., Narbutas, J., Grignard, M., Koshmanova, E., Berthomier, C., et al. (2021). Heterogeneity in the links between sleep arousals, amyloid- $\beta$ , and cognition. *JCI Insight* 6:e152858. doi: 10.1172/jci.insight.152858
- Contador, J., Vargas-Martínez, A., Sánchez-Valle, R., Trapero-Bertran, M., and Lladó, A. (2023). Cost-effectiveness of Alzheimer's disease CSF biomarkers and amyloid-PET in early-onset cognitive impairment diagnosis. *Eur. Arch. Psychiat. Clin. Neurosci.* 273, 243–252. doi: 10.1007/s00406-022-01439-z
- Conti Filho, C. E., Loss, L. B., Marcolongo-Pereira, C., Rossoni Junior, J. V., Barcelos, R. M., Chiarelli-Neto, O., et al. (2023). Advances in Alzheimer's disease's pharmacological treatment. *Front. Pharmacol.* 14:1101452. doi: 10.3389/fphar.2023.1101452
- Crous-Bou, M., Minguillón, C., Gramunt, N., and Molinuevo, J. L. (2017). Alzheimer's disease prevention: from risk factors to early intervention. *Alzheimer's Res. Ther.* 9, 1–9. doi: 10.1186/s13195-017-0297-z
- Dakterzada, F., López-Ortega, R., Arias, A., Riba-Llena, I., Ruiz-Julián, M., Huerto, R., (2021). Assessment of the concordance and diagnostic accuracy between Elecsys and Lumipulse fully automated platforms and Innostest. *Front. Aging Neurosci.* 13:604119. doi: 10.3389/fnagi.2021.604119
- De Leon, M. J., DeSanti, S., Zinkowski, R., Mehta, P. D., Pratico, D., Segal, S., et al. (2006). Longitudinal CSF and MRI biomarkers improve the diagnosis of mild cognitive impairment. *Neurobiol. Aging* 27, 394–401. doi: 10.1016/j.neurobiolaging.2005.07.003
- Dressler, O., Schneider, G., Stockmanns, G., and Kochs, E. F. (2004). Awareness and the EEG power spectrum: analysis of frequencies. *Br. J. Anaesth.* 93, 806–809. doi: 10.1093/bja/ae270
- Farré, R., Montserrat, J. M., and Navajas, D. (2004). Noninvasive monitoring of respiratory mechanics during sleep. *Eur. Respirat. J.* 24, 1052–1060. doi: 10.1183/09031936.04.00072304
- Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnagi.2024.1369545/full#supplementary-material>

- Furui, A., Onishi, R., Takeuchi, A., Akiyama, T., and Tsuji, T. (2020). Non-gaussianity detection of EEG signals based on a multivariate scale mixture model for diagnosis of epileptic seizures. *IEEE Trans. Biomed. Eng.* 68, 515–525. doi: 10.1109/TBME.2020.3006246
- Gaeta, A. M., Benítez, I. D., Jorge, C., Torres, G., Dakterzada, F., Minguez, O., et al. (2020). Prevalence of obstructive sleep apnea in Alzheimer's disease patients. *J. Neurol.* 267, 1012–1022. doi: 10.1007/s00415-019-09668-4
- Gallego-Jutglá, E., Solé-Casals, J., Vialatte, F. B., Elgendi, M., Cichocki, A., and Dauwels, J. (2015). A hybrid feature selection approach for the early diagnosis of Alzheimer's disease. *J. Neural Eng.* 12:e016018. doi: 10.1088/1741-2560/12/1/016018
- Gaubert, S., Houot, M., Raimondo, F., Ansart, M., Corsi, M. C., Naccache, L., et al. (2021). A machine learning approach to screen for preclinical Alzheimer's disease. *Neurobiol. Aging* 105, 205–216. doi: 10.1016/j.neurobiolaging.2021.04.024
- Geng, D., Wang, C., Fu, Z., Zhang, Y., Yang, K., and An, H. (2022). Sleep EEG-based approach to detect mild cognitive impairment. *Front. Aging Neurosci.* 14:865558. doi: 10.3389/fnagi.2022.865558
- Gerla, V., Kremen, V., Macas, M., Dudysova, D., Mladek, A., Sos, P., et al. (2019). Iterative expert-in-the-loop classification of sleep PSG recordings using a hierarchical clustering. *J. Neurosci. Methods* 317, 61–70. doi: 10.1016/j.jneumeth.2019.01.013
- Ghorbanian, P., Devilbiss, D. M., Hess, T., Bernstein, A., Simon, A. J., and Ashrafuon, H. (2015). Exploration of EEG features of Alzheimer's disease using continuous wavelet transform. *Med. Biol. Eng. Comput.* 53, 843–855. doi: 10.1007/s11517-015-1298-3
- Gómez-Ramírez, J., Ávila-Villanueva, M., and Fernández-Blázquez, M. Á. (2020). Selecting the most important self-assessed features for predicting conversion to mild cognitive impairment with random forest and permutation-based methods. *Sci. Rep.* 10:20630. doi: 10.1038/s41598-020-77296-4
- Gonçalves, R. A., Wijesekera, N., Fraser, P. E., and De Felice, F. G. (2019). The link between tau and insulin signaling: implications for Alzheimer's disease and other tauopathies. *Front. Cell. Neurosci.* 13:17. doi: 10.3389/fncel.2019.00017
- Grassberger, P., and Procaccia, I. (1983). Measuring the strangeness of strange attractors. *Phys. D Nonlin. Phenom.* 9, 189–208.
- Grøntvedt, G. R., Lauridsen, C., Berge, G., White, L. R., Salvesen, Ø., Bråthen, G., et al. (2020). The amyloid, tau, and neurodegeneration (A/T/N) classification applied to a clinical research cohort with long-term follow-up. *J. Alzheimer's Dis.* 74, 829–837. doi: 10.3233/JAD-191227
- Hampel, H., Buerger, K., Zinkowski, R., Teipel, S. J., Goernitz, A., Andreasen, N., et al. (2004). Measurement of phosphorylated tau epitopes in the differential diagnosis of Alzheimer disease: a comparative cerebrospinal fluid study. *Archiv. Gen. Psychiat.* 61, 95–102. doi: 10.1001/archpsyc.61.1.95
- Harrison, I. F., Ismail, O., Machhada, A., Colgan, N., Ohene, Y., Nahavandi, P., et al. (2020). Impaired glymphatic function and clearance of tau in an Alzheimer's disease model. *Brain* 143, 2576–2593. doi: 10.1093/brain/awaa179
- Holth, J. K., Fritsch, S. K., Wang, C., Pedersen, N. P., Cirrito, J. R., Mahan, T. E., et al. (2019). The sleep-wake cycle regulates brain interstitial fluid tau in mice and csf tau in humans. *Science* 363, 880–884. doi: 10.1126/science.aav2546
- Insel, P. S., Mohlenhoff, B. S., Neylan, T. C., Krystal, A. D., and Mackin, R. S. (2021). Association of sleep and  $\beta$ -amyloid pathology among older cognitively unimpaired adults. *J. Am. Med. Assoc. Netw. Open* 4:e2117573. doi: 10.1001/jamanetworkopen.2021.17573
- Jack Jr, C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., et al. (2018). NIA-AA research framework: toward a biological definition of Alzheimer's disease. *Alzheimer's Dement.* 14, 535–562. doi: 10.1016/j.jalz.2018.02.018
- Jeong, J. (2004). EEG dynamics in patients with Alzheimer's disease. *Clin. Neurophysiol.* 115, 1490–1505. doi: 10.1016/j.clinph.2004.01.001
- Jo, T., Nho, K., Risacher, S. L., Saykin, A. J., and Initiative, A. N. (2020). Deep learning detection of informative features in tau PET for Alzheimer's disease classification. *BMC Bioinform.* 21, 1–13. doi: 10.1186/s12859-020-03848-0
- Ju, Y. E. S., Finn, M. B., Sutphen, C. L., Herries, E. M., Jerome, G. M., Ladenson, J. H., et al. (2016). Obstructive sleep apnea decreases central nervous system-derived proteins in the cerebrospinal fluid. *Ann. Neurol.* 80, 154–159. doi: 10.1002/ana.24672
- Ju, Y. E. S., Ooms, S. J., Sutphen, C., Macauley, S. L., Zangrilli, M. A., Jerome, G., et al. (2017). Slow wave sleep disruption increases cerebrospinal fluid amyloid- $\beta$  levels. *Brain* 140, 2104–2111. doi: 10.1093/brain/awx148
- Khosroozad, S., Abedi, A., and Hayes, M. J. (2023). Sleep signal analysis for early detection of Alzheimer's disease and related dementia (ADRD). *IEEE J. Biomed. Health Informat.* 27, 2264–2275. doi: 10.1109/JBHI.2023.3235391
- Kim, K. T. (2022). Lumbar puncture: considerations, procedure, and complications. *Encephalitis* 2:93. doi: 10.47936/encephalitis.2022.00045
- Kim, N. H., Park, U., Yang, D. W., Choi, S. H., Youn, Y. C., and Kang, S. W. (2023). PET-validated EEG-machine learning algorithm predicts brain amyloid pathology in pre-dementia Alzheimer's disease. *Sci. Rep.* 13:10299. doi: 10.1038/s41598-023-36713-0
- Kim, N. H., Yang, D. W., Choi, S. H., and Kang, S. W. (2021). Machine learning to predict brain amyloid pathology in pre-dementia Alzheimer's disease using QEEG features and genetic algorithm heuristic. *Front. Comput. Neurosci.* 15:755499. doi: 10.3389/fncom.2021.755499
- Kowalski, J. W., Gawel, M., Pfeffer, A., and Barcikowska, M. (2001). The diagnostic value of EEG in Alzheimer disease: correlation with the severity of mental impairment. *J. Clin. Neurophysiol.* 18, 570–575. doi: 10.1097/00004691-200111000-00008
- Kulkas, A., Duce, B., Leppänen, T., Hukins, C., and Töyräs, J. (2017). Severity of desaturation events differs between hypoxemia and obstructive apnea events and is modulated by their duration in obstructive sleep apnea. *Sleep Breath.* 21, 829–835. doi: 10.1007/s11325-017-1513-6
- Lall, R., Mohammed, R., and Ojha, U. (2019). What are the links between hypoxia and Alzheimer's disease? *Neuropsychiat. Dis. Treat.* 2019, 1343–1354. doi: 10.2147/NDT.S203103
- Lancôt, K. L., Boada, M., Tariot, P. N., Dabbous, F., Hahn-Pedersen, J., Udayachalerm, S., et al. (2024). Association between clinical dementia rating and clinical outcomes in Alzheimer's disease. *Alzheimer's Dement.* 16:e12522. doi: 10.1002/dad2.12522
- Lee, S., Byun, M. S., Yi, D., Kim, M. J., Jung, J. H., Kong, N., et al. (2023). Body mass index and two-year change of *in vivo* Alzheimer's disease pathologies in cognitively normal older adults. *Alzheimer's Res. Ther.* 15:108. doi: 10.1186/s13195-023-01259-w
- Lennon, M. J., Koncz, R., and Sachdev, P. S. (2021). Hypertension and Alzheimer's disease: is the picture any clearer? *Curr. Opin. Psychiat.* 34, 142–148. doi: 10.1097/YCO.0000000000000684
- Liu, X., Zhang, C., Ji, Z., Ma, Y., Shang, X., Zhang, Q., et al. (2016). Multiple characteristics analysis of Alzheimer's electroencephalogram by power spectral density and Lempel–Ziv complexity. *Cogn. Neurodyn.* 10, 121–133.
- Lucey, B. P., Hicks, T. J., McLeland, J. S., Toedebusch, C. D., Boyd, J., Elbert, D. L., et al. (2018). Effect of sleep on overnight cerebrospinal fluid amyloid  $\beta$  kinetics. *Ann. Neurol.* 83, 197–204. doi: 10.1002/ana.25117
- Lucey, B. P., McCullough, A., Landsness, E. C., Toedebusch, C. D., McLeland, J. S., Zaza, A. M., et al. (2019). Reduced non-rapid eye movement sleep is associated with tau pathology in early Alzheimer's disease. *Sci. Transl. Med.* 11:eaau6550. doi: 10.1126/scitranslmed.aau6550
- Macedo, A. C., Balouch, S., and Tabet, N. (2017). Is sleep disruption a risk factor for Alzheimer's disease? *J. Alzheimer's Dis.* 58, 993–1002. doi: 10.3233/JAD-161287
- Mander, B. A., Marks, S. M., Vogel, J. W., Rao, V., Lu, B., Saletin, J. M., et al. (2015).  $\beta$ -amyloid disrupts human NREM slow waves and related hippocampus-dependent memory consolidation. *Nat. Neurosci.* 18, 1051–1057. doi: 10.1038/nn.4035
- Masters, C. L., Bateman, R., Blennow, K., Rowe, C. C., Sperling, R. A., and Cummings, J. L. (2015). Alzheimer's disease. *Nat. Rev. Dis. Prim.* 1, 1–18. doi: 10.1038/nrdp.2015.56
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack Jr, C. R., Kawas, C. H., et al. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement.* 7, 263–269. doi: 10.1016/j.jalz.2011.03.005
- Natekin, A., and Knoll, A. (2013). Gradient boosting machines, a tutorial. *Front. Neurobot.* 7:21. doi: 10.3389/fnbot.2013.00021
- Nunomura, A., Perry, G., Aliev, G., Hirai, K., Takeda, A., Balraj, E. K., et al. (2001). Oxidative damage is the earliest event in Alzheimer disease. *J. Neuropathol. Exp. Neurol.* 60, 759–767. doi: 10.1093/jnen/60.8.759
- Nyholm, J., Ghazi, A. N., Ghazi, S. N., and Berglund, J. S. (2024). Prediction of dementia based on older adults—sleep disturbances using machine learning. *Comput. Biol. Med.* 171:108126. doi: 10.1016/j.combiomed.2024.108126
- Obermeyer, Z., and Emanuel, E. J. (2016). Predicting the future—big data, machine learning, and clinical medicine. *N. Engl. J. Med.* 375:1216. doi: 10.1056/NEJMp1606181
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in Python. *J. Machine Learn. Res.* 12, 2825–2830. doi: 10.1097/01.JGP.0000192478.82189.a8
- Pernecky, R., Wagenpfeil, S., Komossa, K., Grimmer, T., Diehl, J., and Kurz, A. (2006). Mapping scores onto stages: mini-mental state examination and clinical dementia rating. *Am. J. Geriatr. Psychiat.* 14, 139–144.
- Petit, D., Gagnon, J. F., Fantini, M. L., Ferini-Strambi, L., and Montplaisir, J. (2004). Sleep and quantitative EEG in neurodegenerative disorders. *J. Psychosomat. Res.* 56, 487–496. doi: 10.1016/j.jpsychores.2004.02.001
- Pirrone, D., Weitschek, E., Di Paolo, P., De Salvo, S., and De Cola, M. C. (2022). EEG signal processing and supervised machine learning to early diagnose Alzheimer's disease. *Appl. Sci.* 12:5413. doi: 10.3390/app12115413
- Popuri, K., Ma, D., Wang, L., and Beg, M. F. (2020). Using machine learning to quantify structural MRI neurodegeneration patterns of Alzheimer's disease into dementia score: independent validation on 8,834 images from ADNI, AIBL, OASIS, and MIRIAD databases. *Hum. Brain Mapp.* 41, 4127–4147. doi: 10.1002/hbm.25115
- Rajan, K. B., Weuve, J., Barnes, L. L., McAninch, E. A., Wilson, R. S., and Evans, D. A. (2021). Population estimate of people with clinical Alzheimer's disease and mild cognitive impairment in the United States (2020–2060). *Alzheimer's Dement.* 17, 1966–1975. doi: 10.1002/alz.12362

- Rajula, H. S. R., Verlato, G., Manchia, M., Antonucci, N., and Fanos, V. (2020). Comparison of conventional statistical methods with machine learning in medicine: diagnosis, drug development, and treatment. *Medicina* 56:455. doi: 10.3390/medicina56090455
- Sabbagh, M. N., and DeCourt, B. (2023). Biomarker-based diagnosis of preclinical Alzheimer disease: time for the clinic? *Nat. Rev. Neurol.* 19, 71–72. doi: 10.1038/s41582-022-00767-x
- Sánchez-de-la Torre, M., Campos-Rodríguez, F., and Barbé, F. (2013). Obstructive sleep apnoea and cardiovascular disease. *Lancet Respirat. Med.* 1, 61–72. doi: 10.1016/S2213-2600(12)70051-6
- Shah, N., Redline, S., Yaggi, H. K., Wu, R., Zhao, C. G., Ostfeld, R., et al. (2013). Obstructive sleep apnea and acute myocardial infarction severity: ischemic preconditioning? *Sleep Breath.* 17, 819–826. doi: 10.1007/s11325-012-0770-7
- Spector, A. R., Loriaux, D., and Farjat, A. E. (2019). The clinical significance of apneas versus hypopneas: is there really a difference? *Cureus* 11:e4560. doi: 10.7759/cureus.4560
- Targa, A. D. S., Benítez, I. D., Moncusí-Moix, A., Dakterzada, F., Minguez, O., Vaca, R., et al. (2023). Breathing cessation events that compose the apnea-hypopnea index are distinctively associated with the adverse outcomes in Alzheimer's disease. *Alzheimer's Res. Ther.* 15:123. doi: 10.1186/s13195-023-01266-x
- Umut, I. (2016). PSGMiner: a modular software for polysomnographic analysis. *Comput. Biol. Med.* 73, 1–9. doi: 10.1016/j.compbiomed.2016.03.023
- Villemagne, V. L., Burnham, S., Bourgeat, P., Brown, B., Ellis, K. A., Salvado, O., et al. (2013). Amyloid  $\beta$  deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. *Lancet Neurol.* 12, 357–367. doi: 10.1016/S1474-4422(13)70044-9
- Wang, Y., Meng, R., Song, H., Liu, G., Hua, Y., Cui, D., et al. (2017). Remote ischemic conditioning may improve outcomes of patients with cerebral small-vessel disease. *Stroke* 48, 3064–3072. doi: 10.1161/STROKEAHA.117.017691
- Willmott, C. J., and Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* 30, 79–82. doi: 10.3354/cr030079
- Winer, J. R., Mander, B. A., Helfrich, R. F., Maass, A., Harrison, T. M., Baker, S. L., et al. (2019). Sleep as a potential biomarker of tau and  $\beta$ -amyloid burden in the human brain. *J. Neurosci.* 39, 6315–6324. doi: 10.1523/JNEUROSCI.0503-19.2019
- Xiang, J., Maue, E., Fan, Y., Qi, L., Mangano, F. T., Greiner, H., et al. (2020). Kurtosis and skewness of high-frequency brain signals are altered in paediatric epilepsy. *Brain Commun.* 2:fcaa036. doi: 10.1093/braincomms/fcaa036
- Yaffe, K., Laffan, A. M., Harrison, S. L., Redline, S., Spira, A. P., Ensrud, K. E., et al. (2011). Sleep-disordered breathing, hypoxia, and risk of mild cognitive impairment and dementia in older women. *J. Am. Med. Assoc.* 306, 613–619. doi: 10.1001/jama.2011.1115
- Zhang, C. E., Yang, X., Li, L., Sui, X., Tian, Q., Wei, W., et al. (2014). Hypoxia-induced tau phosphorylation and memory deficit in rats. *Neurodegener. Dis.* 14, 107–116. doi: 10.1159/000362239
- Zhang, J., Zhang, J., Li, X. J., Xiao, J., and Ye, F. (2020). Hypoxic preconditioning ameliorates amyloid- $\beta$  pathology and longterm cognitive decline in a $\beta$ pp/ps1 transgenic mice. *Curr. Alzheimer Res.* 17, 626–634. doi: 10.2174/1567205017666201007121730