Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 26s (2025)
https://www.jneonatalsurg.com

OPEN ACCESS

Prediction Of Alzheimer's Disease Using Health Records and Machine Learning Based Framework

Shimpy Harbhajanka Goyal[1], Anusha Jain[2], Priyanka Dhasal[3], Sonal Modh Bhardwaj[4]

[1]Assistant Professor, Department of Computer Science and Engineering, Medi-Caps University, Indore, Madhya Pradesh, India

Email ID: shimpy.h@gmail.com

[2]Assistant Professor, Department of Computer Science and Engineering, Medi-Caps University, Indore, Madhya Pradesh, India

Email ID: anusha.jain9@gmail.com

[3]Assistant Professor, Department of Computer Science and Engineering, Medi-Caps University, Indore, Madhya Pradesh, India

Email ID: priyanka.dhasal07@gmail.com

[4]Assistant Professor, Department of Computer Science and Engineering, Medi-Caps University, Indore, Madhya Pradesh, India

Email ID: sonalmodh@gmail.com

Cite this paper as: Shimpy Harbhajanka Goyal, Anusha Jain, Priyanka Dhasal, Sonal Modh Bhardwaj, (2025) Prediction Of Alzheimer's Disease Using Health Records and Machine Learning Based Framework, *Journal of Neonatal Surgery*, 14 (26s), 66-77

ABSTRACT

Alzheimer's Disease and Related Dementias (ADRD) are progressive neurodegenerative disorders posing significant social, clinical, and economic burdens. This study aims to enhance early diagnosis of ADRD using advanced machine learning (ML) models applied to longitudinal electronic health record (EHR) data from the University of Missouri Healthcare system. We evaluated six ML algorithms—Gradient Boosted Trees (GBT), Light Gradient Boosting Machine (LightGBM), Random Forest (RF), eXtreme Gradient Boosting (XGBoost), Logistic Regression (LR), and AdaBoost—over varying prediction windows from one to five years. The dataset included 123,735 patients aged 50 and above, with 4012 diagnosed cases of ADRD. A comprehensive set of demographics, clinical, and behavioural features were extracted and pre-processed to train the models. Among all models, GBT demonstrated the highest predictive accuracy, achieving an AUC of 0.833 in the five-year prediction window. SHAP (SHapley Additive exPlanations) analysis revealed key risk factors such as depression, age groups (70–80 and 80–90), sleep apnoea, and cardiovascular diseases. These results suggest that ML models, especially GBT, can provide high-performing, interpretable tools for early ADRD detection, informing clinical decision-making and enabling timely intervention strategies.

Keywords: *Alzheimer's Disease, ADRD, Machine Learning, EHR, Gradient Boosted Trees, Predictive Modeling, SHAP Analysis, Dementia Risk, Early Diagnosis*

1. INTRODUCTION

A collection of irreversible neurodegenerative diseases known as Alzheimer's Disease and Related Dementias (ADRD) gradually deteriorate memory, cognitive abilities, and the capacity to carry out everyday tasks [1]. Alois Alzheimer, a German psychiatrist, was the first to define Alzheimer's disease (AD). It was noted in Auguste, a patient who passed away in 1906 as a result of cognitive decline [2]. The abnormal buildup of tau proteins, amyloid-beta (A) proteins, and neurodegeneration in the brain are biologically defined as AD [3]. Frequently, these pathological alterations manifest 20 years prior to the onset of clinical symptoms. There are few alternatives for therapy as the disease worsens, progressing from mild cognitive impairment (MCI) to severe dementia [4]. AD was the eighth greatest cause of death in 2022, affecting more than 6 million Americans 65 and over. Patients, family, and society as a whole bear a heavy cost from ADRD. It is anticipated that there will be more than 75 million AD patients by 2030, and that number will quadruple by 2050 [5]. Treatment for AD and dementia will cost $321 billion in 2022, and unpaid caring will cost an additional $271 billion. By 2050, the yearly expenditures are expected to surpass $1 trillion. Because it enables treatments before significant cognitive impairment occurs, early identification of ADRD is crucial. Between 2002 and 2012, over 99 percent of clinical studies failed to provide effective

therapies for ADRD, despite intensive research efforts. However, the U.S. healthcare system might save up to $7.9 trillion if dementia was diagnosed at the moderate cognitive impairment (MCI) stage rather than the late stage [6-9]. Only six medications have been licensed by the U.S. Food and Drug Administration (FDA) since 1998 to treat the symptoms of ADRD: aducanumab, galantamine, donepezil, memantine, rivastigmine, and a combination of donepezil and memantine [10]. The early to middle stages of the illness are when these FDA-approved drugs work best, providing symptomatic relief without stopping the disease's development. These medicines' poor efficacy emphasises the need for more potent therapeutic approaches.

The development of machine learning (ML) presents fresh hope for enhancing ADRD early detection. Utilising extensive datasets from administrative claims, electronic health record (EHR) systems, and neuroimaging, machine learning (ML) models can reveal trends and insights that conventional approaches might miss. Recent research has shown that machine learning (ML) can be used to manage various neurological disorders and forecast the development of AD [11-14]. Prior research has shown that machine learning models are very flexible for various datasets, demonstrating their resilience and suitability for other serious conditions including heart disease, breast cancer, and prostate cancer [15]. Public health actions can also be informed by a better understanding of ADRD demographics, medication use, and comorbidities. Depression and cardiovascular disease, for instance, have been repeatedly associated with a higher risk of ADRD. Amyloid buildup and compromised cognitive performance have been linked to sleep disturbances. Designing age and gender-sensitive preventive efforts can be aided by knowledge of demographic diversity in ADRD risk, such as variances by sex and higher relationships in older persons. Better patient outcomes can result from modifying prescription use, such as avoiding anticholinergic medications, which are linked to an increased risk of dementia [16]. Patients with ADRD may benefit from individualised therapies made possible by early identification of high-risk individuals.

Gradient-Boosted Trees (GBT) [17], Light Gradient-Boosting Machine (LightGBM) [18], Random Forest (RF) [19], eXtreme Gradient-Boosting (XGBoost) [20], Logistic Regression (LR) [21], and Adaptive Boosting (AdaBoost) [22] are six different machine learning models that we used in our study to classify ADRD using de-identified EHR data from the University of Missouri (MU) Healthcare, which is a member of the National Patient-Centered Clinical Research Network (PCORnet). We sought to diagnose ADRD accurately and early by finding critical predictive indicators and optimising the machine learning models for accuracy, sensitivity, Area Under the Receiver Operating Characteristic Curve (AUC-ROC), and F1-scores. A wide variety of EHR characteristics were included in our study as predictors in the machine learning models. These variables included a wide range of comorbidities, medical diagnoses, behavioural risk factors (like smoking history), vital signs (like systolic and diastolic blood pressure), and demographic characteristics (like age, race, marital status, and sex). These included cardiovascular diseases (e.g., hypertension, heart disease), neurological disorders (e.g., stroke), psychiatric disorders (e.g., depression, anxiety), metabolic disorders (e.g., diabetes, obesity), and other medical conditions (e.g., sleep disorders, kidney disease, vitamin D deficiency). The Methods section has a thorough list of these variables. We experimented with several one- to five-year forecast frames. There is still room to enhance predictive machine learning models for the early detection of ADRD, even in the face of notable advancements in data analytics and healthcare. Issues with current methods include poor interpretability for clinical application and restricted integration of complete EHR data. By using advanced machine learning models and SHAP (SHapley Additive exPlanations) analysis to pinpoint important risk variables and increase the accuracy of ADRD diagnosis, this work seeks to close these gaps. We will examine the methodology, the findings, and the importance of the findings in the remaining portion of the study

2. METHODOLOGY

2.1. Study participants

The goal of this retrospective case-control research was to predict the diagnosis of ADRD in persons 50 years of age and older. We started with 380,269 patients who satisfied the following requirements: (1) they had to be at least 40 years old on January 1, 2010 (the date records in the MU EHR system began to be kept), (2) they had to be admitted between January 1, 2010, and December 31, 2023, and (3) they had to have had at least two documented contacts at MU Healthcare.

2.2. Selection criteria for cases and controls

Two primary prediction factors were used in this investigation to characterise ADRD patients. First, patients with an ADRD diagnosis based on codes 331.0, 290.0, 290.1, 290.2, 290.3, 290.4, 290.43, 331.82, 294.1, G30.0, G30.1, G30.8, G31.83, F00, F00.2, F01, F02, and F00.9 from the International Classification of Diseases, 9th or 10th Revisions (ICD-9/ICD-10). Early onset, late onset, and confirmed ADRD cases (those that do not fall into the early or late onset categories) are all covered by these ICD codes. Patients who were administered dementia-related drugs, such as rivastigmine, galantamine, donepezil, memantine, aducanumab, and brexpiprazole, which are often used for AD, were also included in our study. Additionally, patients need to have at least two interactions documented in the MU EHR system. Because brexpiprazole was approved by the FDA in 2023 as the first pharmaceutical therapy for agitation linked to dementia in Alzheimer's disease, it was included in the research [17]. We included it because of its usefulness in addressing a prevalent neuropsychiatric

symptom in ADRD, even though it is used as an additional therapy for major depressive disorder and schizophrenia.

We chose members of the control group (non-ADRD) if they (1) had at least two documented contacts in the MU EHR system, and (2) had no ADRD-related diagnoses based on ICD-9/10 codes or prescriptions for dementia-related drugs. Of the 123,735 distinct individuals in our final dataset, 4012 had an ADRD diagnosis (cases), and 119,723 had neither an ADRD diagnosis nor a diagnosis associated to ADRD (controls).

2.3. Study design

We separated the period into two parts for our study: (1) an observation window and (2) a forecast window. For every instance, we designate an index date that corresponds to the earliest date of either an ADRD diagnosis or the initial prescription of a medicine associated to dementia. To illustrate the time periods in which a case had its index date, we established several prediction windows: one year, two years, three years, four years, and five years. For both case and control data, the observation period ran from January 1, 2010 (the day when records in the MU EHR system began) to December 31, 2018. We were able to assess the possibility of forecasting ADRD at various time intervals since the models were trained solely using data from the observation window.

As seen in Fig. 1, we employed distinct case datasets for every prediction window in addition to a fixed control dataset [2010–2019]. Those with an index date in [2019–2020] were included in the case data for the one-year forecast window. The case data for the two-year prediction window comprised those whose index date fell inside [2019–2021), and so forth. Unlike a shorter pre-onset case window, our method uses a substantial quantity of control data to generate a solid "healthy" reference.

Behavioural risk factors (e.g., smoking history), demographic variables (age, race, marital status, and sex), and vital variables (diastolic blood pressure, or DBP) were among the variables that were recorded by the EHR system and included in our dataset.

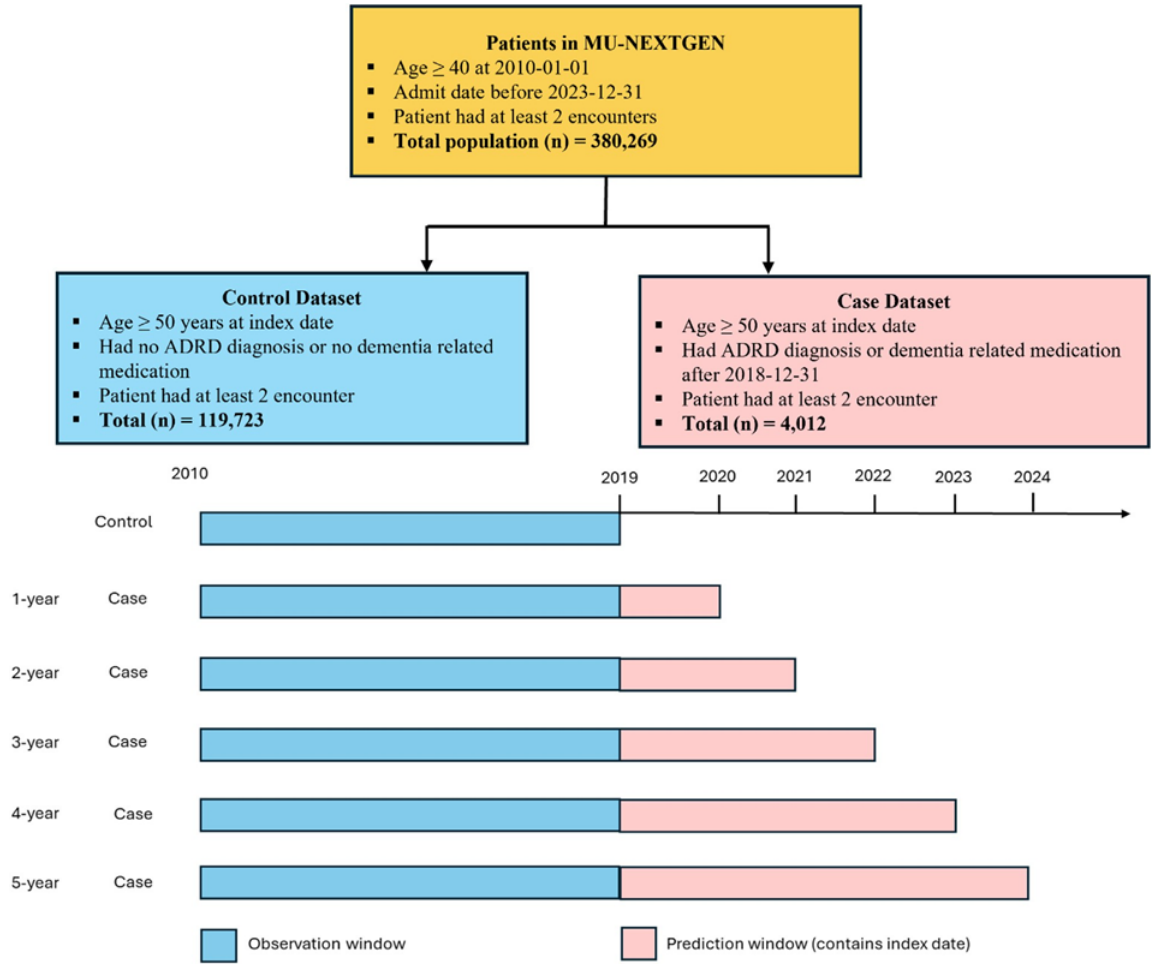

Fig. 1. Flowchart for preparing the case-control study

Furthermore, we identified comorbidities as risk factors by doing a comprehensive assessment of the literature [23, 44].

Diabetes, epilepsy, depression, obesity, stroke, anxiety, hypertension, hyperlipidaemia, cardiovascular disease, headache, sleep disorder, concussion, heart disease, sleep apnoea, insomnia, kidney disease, cholesterol, vitamin D deficiency, enlarged prostate, bone disease, and depressive disorder are among the co-morbidities and medical diagnoses.

2.4. Preparing the data

We eliminated feature variables with a missing rate of 30% or more in order to remedy missing data. Patients with less than 20% missing data were imputed, while those with more than 20% missing data were eliminated from the study. The ultimate number of cases and controls is displayed in Fig. 1. Initially, our dataset contained 4012 unique patients in the positive class (cases) and 232,795 unique patients in the negative class (controls). All categorical variables were subjected to one-hot encoding during the data preparation stage. The handling of continuous variables was based on their unique properties. Age, for instance, was divided into five different groups: [50,60], [60,70], [70,80], [80,90], and above 90 years old. Clinical standards were used to classify DBP and SBP into three levels: "normal," "high," and "critically high." Normal (<80 mmHg), high (80–90 mmHg), and critically high (>90 mmHg) were the classifications given to DBP, whilst normal (<120 mmHg), high (120–140 mmHg), and critically high (>140 mmHg) were given to SBP [45].

Binary values made up the resultant feature vector, with 0s and 1s denoting the presence or absence of each category. The feature vector for medical diagnostic or comorbidity conditions was constructed using one-hot encoding. Binary numbers were used to encode the smoking history: 0 represented never smoking, and 1 represented all other smoking categories, including current and previous smokers. This method made it easier to distinguish between those who have never smoked and those who had eighty percent of the dataset was used for training, while the remaining twenty percent was used for testing. Models were created using the training set, and their performance was evaluated using the testing set.

2.5. Model validation and analysis

In order to predict ADRD early on, we trained and evaluated six distinct machine learning classification models: GBT, LightGBM, RF, XGBoost, LR, and AdaBoost. These models were selected because previous research on ADRD prediction and associated medical classification tasks showed them to be useful. Particularly effective for ADRD risk assessment, GBT, LightGBM, and XGBoost are well known for their robust prediction performance and capacity to manage intricate, non-linear interactions in structured healthcare data [46]. An Alzheimer's disease study using RF.

2.6. Descriptive Statistics in Control and Case

Research for image classification because of its capacity to manage high-dimensional datasets and its resilience to overfitting [47]. Because of its ease of use and interpretability, LR is used as a baseline model and has been widely used in ADRD research to simulate the course of the illness using clinical data [48]. AdaBoost has been effectively used in ADRD prediction using deep learning techniques [49] and improves classification performance by strengthening weak classifiers, especially in unbalanced datasets.

These models were optimised and assessed using a layered cross-validation technique. Every model was integrated into a pipeline that first comprised the corresponding classifier and then a StandardScaler for feature normalisation. A 5-fold StratifiedKFold inner cross-validation loop with grid search (GridSearchCV) was used to do hyperparameter tweaking. The model's performance was then evaluated using an outer 5-fold StratifiedK-fold cross-validation with the optimal hyperparameters applied. A 0.5 threshold was used to classify predicted probabilities, and confusion matrices were created for each fold. Metrics including accuracy, precision, sensitivity, F1-score, AUC-ROC, and specificity were used to assess the model's performance. To estimate point values, 1000 rounds of bootstrapping were used. The model that performed the best on all criteria was chosen, and its generalisation was further assessed using a hold-out test set.

We created SHAP values for each of the five prediction windows (1, 2, 3, 4, and 5 years) and used SHAP (Shapley Additive exPlanations) techniques [50] to evaluate the ML models' predictions. Summary charts were created using these values, showing risk variables across time and offering insights into the interpretability of the model. In particular, summary plots and SHAP bar plots were produced for the top 12 risk factors in the model with the highest performance. While features with negative SHAP values were connected to a decreased risk, those with positive values were linked to a higher chance of ADRD. Each SHAP value's magnitude indicated how important that feature was to the model overall.

3. RESULTS

3.1. Sample characteristics

The case and control groups' descriptive data are displayed in Table 1. Our data set spans the years January 1, 2010, through December 31, 2023, and comprises 4012 cases and 119,723 control subjects who are all 50 years of age or older. The main demo-graphic traits of both groups were the focus of the investigation. Indicating an older population, the case group's mean age (77.50 ± 9.25 years) was greater than that of the control group (75.51 ± 10.18 years). In both categories, the proportion of female patients was higher than that of male patients. Additionally, considering the demographics of patients that visited

MU Healthcare, the White race predominated in both categories.

Table 1 Control and case dataset

Variables	Control		Case	
	Mean	Standard Deviation	Mean	Standard Deviation
Age	75.51	10.18	77.5	9.25
Female	64072	53.5 %	2415	60.2 %
Male	55651	46.5 %	1597	39.8 %
White	109568	92.3 %	3778	94.3 %
Asian	1058	0.9 %	19	0.5 %
Unknown	448	0.5 %	12	0.3 %

3.2. Performance evaluation of model prediction

LR, GBT, LightGBM, XGBoost, RF, and AdaBoost are the six ML classification models that we trained using hyperparameter tweaking on the training set. As previously mentioned, the whole list of predictors was used to train all models, with no variables being left out. These predictors included comorbidities, behavioural risk factors, vital signs, and demographics. The models were trained to forecast the occurrence of ADRD over prediction windows of one, two, three, four, and five years. We divided the unseen test set into two classes: ADRD (case, positive class) and non-ADRD (control, negative class) patients, using the top-performing model for each classifier.

Table 2 Performance of ADRD Predictive Models (a best model according to AUC score)

Prediction Window	Model	Accuracy	AUC	Precision	Sensitivity	Specificity	F-1
1 Year	LR	0.696	0.775	0.97	0.696	0.695	0.801
1 Year	GBT	0.978	0.809a	0.97	0.978	0.999	0.968
1 Year	LightGBM	0.97	0.808	0.964	0.97	0.999	0.957
1 Year	XGBoost	0.979	0.799	0.975	0.979	1	0.97
1 Year	RF	0.978	0.792	0.978	0.978	1	0.967
1 Year	AdaBoost	0.978	0.782	0.969	0.978	1	0.968
2 Years	LR	0.684	0.787	0.965	0.684	0.682	0.789
2 Years	GBT	0.974	0.821a	0.969	0.974	1	0.962
2 Years	LightGBM	0.973	0.821	0.963	0.973	0.999	0.962
2 Years	XGBoost	0.976	0.818	0.972	0.976	0.999	0.967
2 Years	RF	0.974	0.816	0.974	0.974	1	0.961
2 Years	AdaBoost	0.974	0.796	0.964	0.974	0.999	0.962
3 Years	LR	0.688	0.784	0.961	0.688	0.687	0.789
3 Years	GBT	0.974	0.822a	0.971	0.971	0.999	0.965
3 Years	LightGBM	0.971	0.817	0.964	0.971	0.999	0.958
3 Years	XGBoost	0.974	0.798	0.97	0.974	0.998	0.966

3 Years	RF	0.971	0.807	0.972	0.971	1	0.958
3 Years	AdaBoost	0.97	0.8	0.941	0.97	1	0.955
4 Years	LR	0.698	0.763	0.956	0.698	0.698	0.795
4 Years	GBT	0.97	0.808a	0.964	0.97	0.999	0.957
4 Years	LightGBM	0.97	0.808a	0.964	0.97	0.999	0.957
4 Years	XGBoost	0.971	0.807	0.966	0.971	0.999	0.96
4 Years	RF	0.969	0.797	0.968	0.969	0.999	0.954
4 Years	AdaBoost	0.969	0.795	0.96	0.969	0.999	0.956
5 Years	LR	0.688	0.782	0.955	0.688	0.687	0.787
5 Years	GBT	0.97	0.833a	0.968	0.97	0.999	0.96
5 Years	LightGBM	0.969	0.831	0.966	0.969	0.999	0.958
5 Years	XGBoost	0.968	0.829	0.961	0.969	0.999	0.953
5 Years	RF	0.968	0.823	0.967	0.968	0.999	0.954
5 Years	AdaBoost	0.967	0.818	0.959	0.967	0.999	0.953

The performance of the ML model is then reported. The GBT model's Area under the Curve (AUC) performed best in forecasting ADRD during a 5-year period, as seen in Table 2. We used SHAP analysis to guarantee inter-pretability and pinpoint the key determinants impacting ADRD risk. The part "SHAP Analysis and Model Interpretability" provides a thorough explanation of feature contributions and how they relate to the model's predictions. In all five prediction windows, the GBT model continuously beat the other model in terms of accuracy and AUC. For predicting ADRD at 5-year, 4-year, 3-year, 2-year, and 1-year prediction windows, the GBT model obtained the best AUC scores of 0.833, 808, 0.822, 0.821, and 0.809, respectively. In the 5-year prediction window, LightGBM and XGBoost both demonstrated excellent performance, with AUC values of 0.831 and 0.829, respectively. On the other hand, within the 5-year prediction window, the LR model's low-est AUC was 0.782. As the forecast window grew from one to five years, the AUC score steadily rose. Other performance criteria, such as accuracy, sensitivity, specificity, and F1 scores, showed comparable outcomes.

Fig. 2 (right) shows the AUC-ROC curves for the 5-year forecast window employing six different machine learning models. The AUC score of the GBT model, for instance, was 0.833. The LightGBM, XG-Boost, RF, AdaBoost, and LR models came next, with AUC scores of 0.831, 0829, 0.822, 0.818, and 0.782, respectively. Furthermore, the AUC-ROC curves for predictions employing the best model, GBT, for 1-, 2-, 3-, 4-, and 5-year prediction windows were shown in Fig. 2 (left). This increasing trend from one to five years indicates that the GBT model's prediction accuracy aids in the incorporation of longitudinal data. Other parameters, such as specificity and sensitivity, showed consistent results.

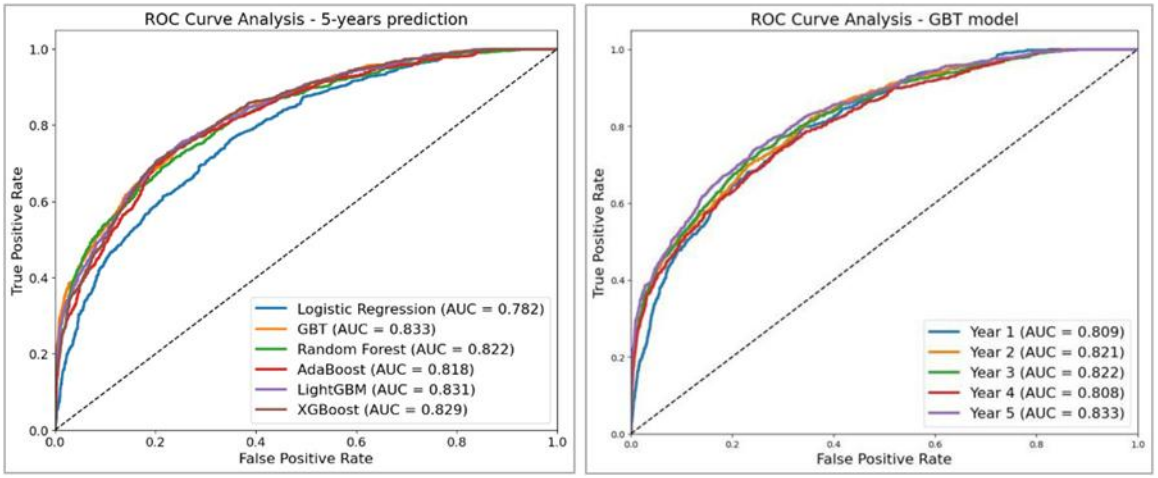

Fig. 2. Performance assessment of ML models in ADRD prediction. (Left) ROC curve analysis for the 5-year prediction window using six different ML models. (Right) ROC curve analysis for predictions across 1-, 2-, 3-, 4-, and 5-year windows. The GBT model, being the best performer, was used for the ROC plot

3.3. SHAP analysis and model interpretability

To determine the main risk variables affecting ADRD prediction and their correlation with results, we used the SHAP. In light of the GBT model's outstanding performance, SHAP values shed light on how interpretable the model was. The top 12 characteristics that had the biggest impact on the model's predictions are shown in Fig. 3, which displays the SHAP analysis for the 1- to 5-year prediction window. The most significant predictors of ADRD risk were consistently characteristics such a history of depressive disorder, higher age groups (70–80 and 80–90 years), anxiety, sleep apnoea, heart disease, headache, and high DBP. The model's interpretability and the intricate correlations between many risk variables are highlighted by the SHAP plot, which offers a thorough description of how each characteristic influences the model's prediction. The probability of an ADRD forecast rose with positive contributions (shown by the red segments) and fell with negative contributions (represented by the blue segments).

4. DISCUSSION

To determine the main risk variables affecting ADRD prediction and their correlation with results, we used the SHAP. In light of the GBT model's outstanding performance, SHAP values shed light on how interpretable the model was. The top 12 characteristics that had the biggest impact on the model's predictions are shown in Fig. 3, which displays the SHAP analysis for the 1- to 5-year prediction window. The most significant predictors of ADRD risk were consistently characteristics such a history of depressive disorder, higher age groups (70–80 and 80–90 years), anxiety, sleep apnoea, heart disease, headache, and high DBP. The model's interpretability and the intricate correlations between many risk variables are highlighted by the SHAP plot, which offers a thorough description of how each characteristic influences the model's prediction. The probability of an ADRD forecast rose with positive contributions (shown by the red segments) and fell with negative contributions (represented by the blue segments).

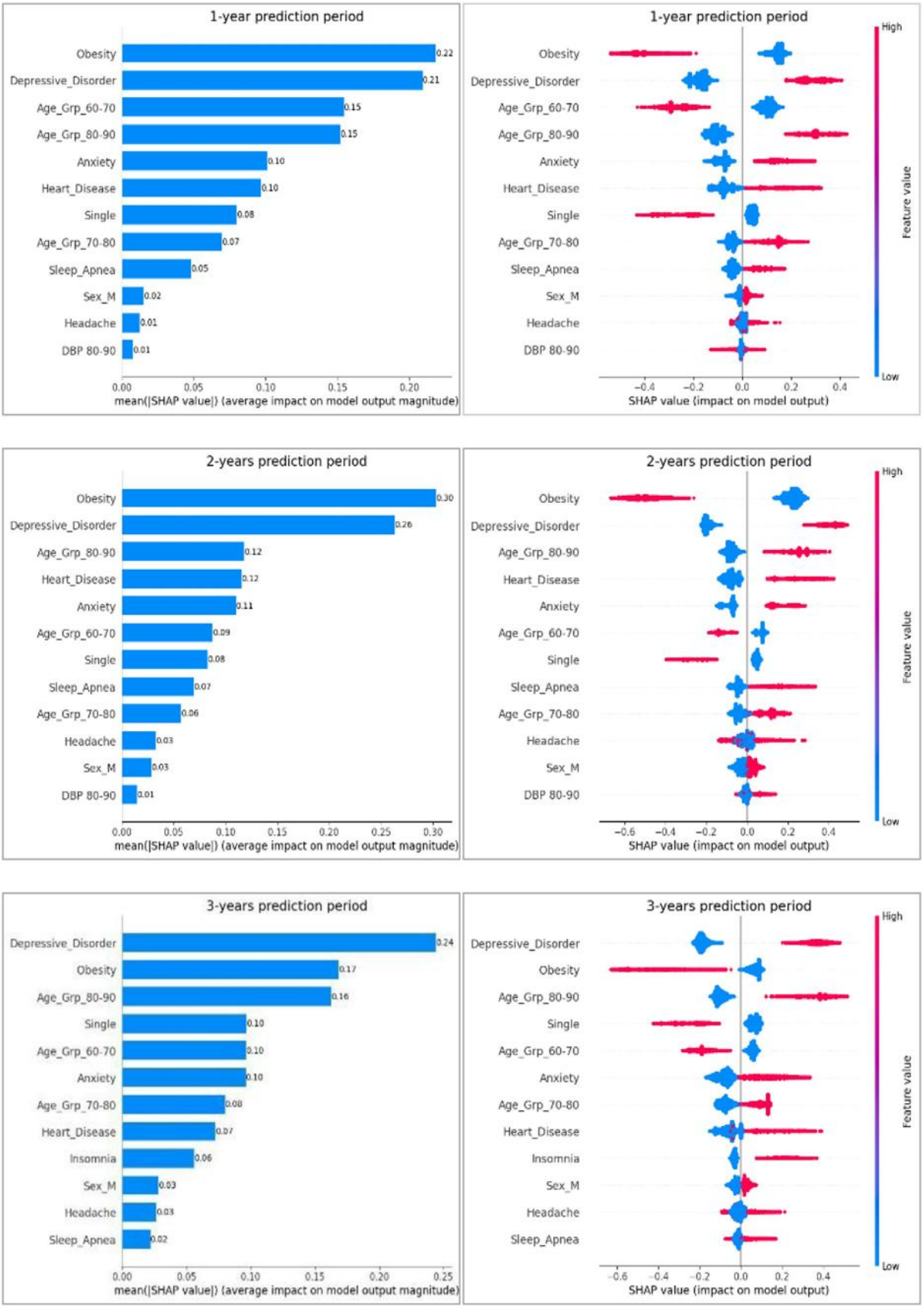

Fig. 3. SHAP plots of the top-12 features for the GBT models (1-year - 5-year prediction windows)

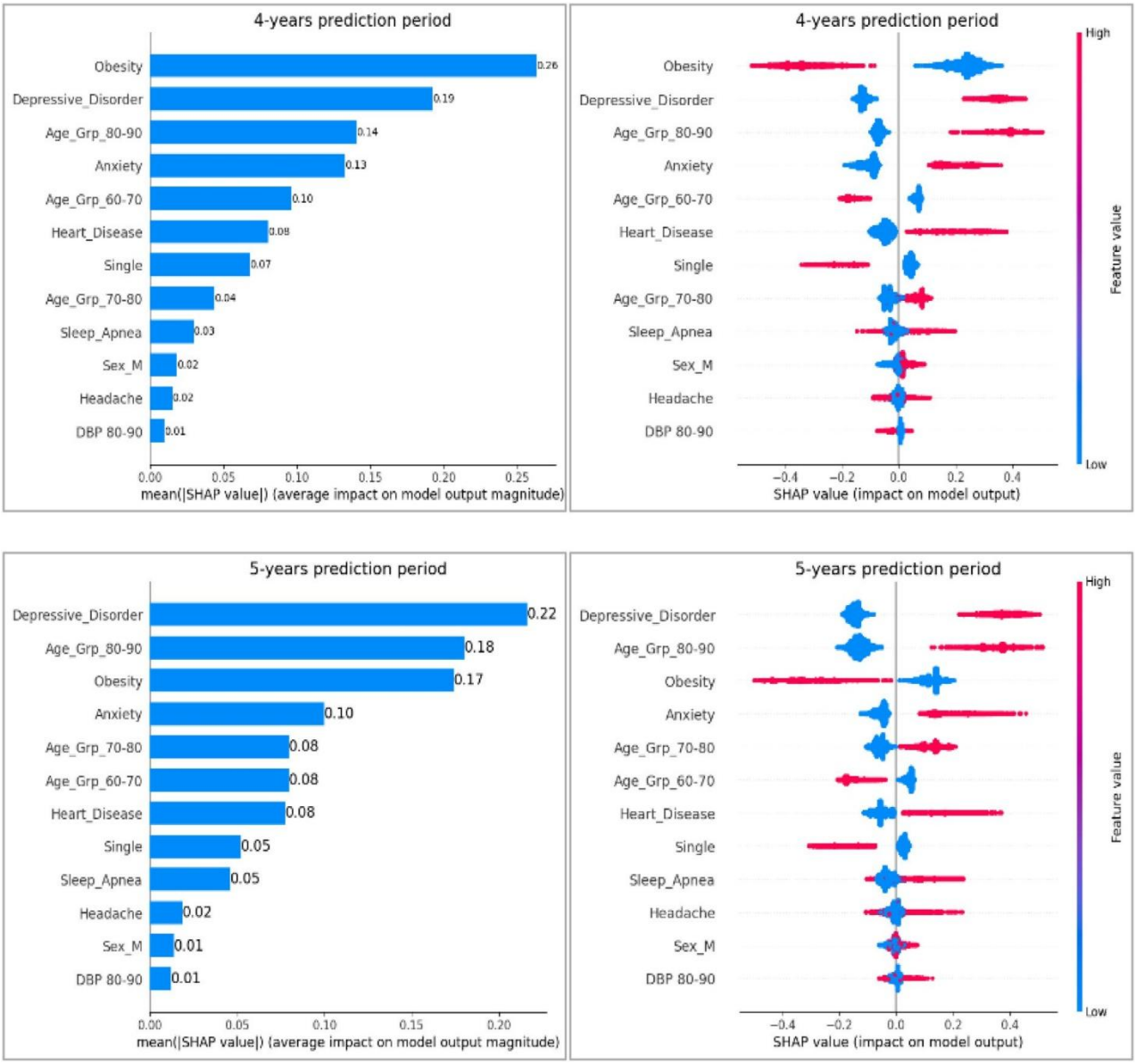

Fig. 3. SHAP plots of the top-12 features for the GBT models (1-year - 5-year prediction windows), Continued

By filling up methodological gaps in earlier research, this study demonstrated the promise of informatics-driven approaches to analyse longitudinal EHR data. We were able to prevent overfitting to short-term fluctuations and capture wider patterns by keeping a constant dataset throughout prediction windows. Scalability and generalisability are supported by the design, which offers a foundation for creating ADRD prediction models in different healthcare systems. By refining how longitudinal data may be used for predictive modelling in chronic illnesses, such improvements advance informatics. By using SHAP analysis to uncover important risk variables for ADRD, our work makes a substantial contribution to clinical science. In keeping with their established connections with cognitive decline, established variables such heart disease, depressive illness, and age groups 80–90 and 70–80 years were verified. According to the SHAP analysis, age is a significant risk factor for ADRD, and its effects differ depending on the group. In line with studies that demonstrate ADRD risk increases beyond age 80 as a result of neurodegeneration, amyloid buildup, and vascular dysfunction, the 80–90-year age group had the greatest favourable impact, followed by the 70–80-year age group. Individuals in the 60–70 age range may benefit from cognitive reserve, which might postpone the diagnosis of ADRD by compensating for early pathogenic alterations. The influence was more varied in the 60–70-year age group, indicating that age-related effects are not consistent. Some people may be in the preclinical stage of ADRD, which is characterised by pathological alterations but no symptoms. Additionally, as older persons (70+) receive more frequent cognitive examinations, which result in early diagnosis, disparities in healthcare utilisation and screening methods may also play a role. These results emphasise the significance of age-stratified risk assessments and the requirement for prediction models that take age-related changes in ADRD risk into account. The ways

in which lifestyle variables and comorbidities alter ADRD risk at various periods of life should be investigated further.

The evolution of ADRD is influenced by heart disease's effects on vascular health and depression's reduction of cognitive reserve. Furthermore, our results emphasise headache and sleep apnoea as new predictors, highlighting the possible contribution of chronic pain and sleep disturbances to elevated ADRD risk. These revelations cast doubt on long-held beliefs and provide up fresh possibilities for enhancing clinical screening and intervention techniques.
The study's inclusion of gender-specific risk variables improves our clinical comprehension of ADRD. For instance, among males, insomnia was the most common predictor, which is consistent with studies that relate sleep disruptions to cognitive impairment. Furthermore, the correlation between ADRD risk and gender disparities in healthcare prevalence was highlighted, indicating that sex-specific biological pathways could affect the beginning of illness. These results support customised tactics for ADRD prevention rather than generic ones. It's interesting to note that, in contrast to other research, diabetes was not a significant predictor in our population. This disparity highlights the necessity of context-specific analysis and the heterogeneity of risk variables among communities. Our results suggest that cardiac disease and depression may predominate in some groups, which calls for more research into the ways that clinical and sociodemographic traits influence the risk of ADRD. The use of SHAP analysis, which improves model interpretability by determining feature significance, is another example of how informatics has advanced. SHAP bridges the gap between sophisticated analytical and clinical utility by offering actionable insights into the reasons behind certain forecasts, in contrast to standard black-box machine learning models. For instance, the discovery of headache, sleep apnoea, and depression as important predictors highlights how ML-driven technologies may reveal correlations in EHR data that were previously unknown. This openness promotes clinician trust and opens the door for ML models to be incorporated into standard medical treatment.

Clinically, preventive measures to slow the course of the illness are made possible by early ADRD prediction. ADRD risk can be considerably decreased by changing one's lifestyle to include better sleep hygiene, controlling cardiovascular risk factors, and taking care of mental health concerns. Our results support the multifaceted character of ADRD, which involves intricate interactions between demographic, metabolic, vascular, and psychological variables. For example, whereas obesity has long been thought of as a risk factor for cognitive decline, our study found a less clear correlation between obesity and ADRD, indicating that the metabolic health role in dementia may include unidentified processes that require more investigation.
Our results have significant ramifications for public health and therapeutic treatments. The necessity for integrated mental and physical health management in ageing populations is underscored by the constant association between elevated ADRD risk and depression and cardiovascular disease, which have been identified as major predictors. In a similar vein, amyloid buildup and compromised cognitive performance have been linked to sleep disorders such sleep apnoea, which has become a new risk factor. These revelations aid in the creation of focused screening initiatives and behavioural treatments. Designing age and gender-sensitive preventative efforts can also be aided by knowledge of demographic diversity in ADRD risk, such as variances by sex and higher relationships in older persons. Lastly, there are several ways in which clinical decision support systems that incorporate machine learning-based prediction tools might benefit practitioners. For instance, better patient outcomes may result from modifying prescription use, such as avoiding anticholinergic medications, which are linked to an increased risk of dementia. Additionally, early high-risk patient identification can provide tailored therapies such prompt supportive care or supportive care for patients with ADRD.

This study has ramifications for healthcare policy and practice. Early specialist referrals might be made possible by the use of automated prediction models in clinical settings, allowing for prompt diagnosis and individualised treatment strategies. These methods also give researchers a reliable way to find high-risk participants for clinical trials, which improves recruiting effectiveness and trial results dependability. Furthermore, by addressing various stages of ADRD development, merging short-term and long-term prediction models might improve patient care and provide a thorough framework for disease monitoring.
By presenting evidence for new risk variables that were not given priority in earlier research, such headache and sleep apnoea, this study also lays the groundwork for the advancement of clinical science. By incorporating these findings into future study, the risk of ADRD may be reduced by developing targeted interventions, such as treatments for sleep problems or chronic pain management. Furthermore, our model's capacity to generalise across prediction windows without enlarging the dataset provides other institutions looking to use EHR data for chronic illness prediction with a reproducible method. This study offers important insights into the complex nature of illness development by considering a variety of ADRD risk variables. For future studies aiming at improving the accuracy and use of ML models in clinical practice, the results highlight the significance of utilising predictive models to direct early identification and intervention.

5. CONCLUSION

This study underscores the effectiveness of machine learning models—particularly Gradient Boosted Trees—in predicting the onset of Alzheimer’s Disease and Related Dementias across different temporal windows using real-world EHR data. By integrating diverse risk factors such as demographic details, clinical comorbidities, and behavioural metrics, our approach

achieved high predictive accuracy and provided interpretability through SHAP analysis. The identification of depression, cardiovascular disorders, sleep apnoea, and specific age brackets as key predictors reinforces the multi-dimensional nature of ADRD risk. Importantly, the model's robustness across varying prediction windows demonstrates its utility for both short- and long-term clinical forecasting. These findings contribute significantly to informatics-driven public health strategies and support the development of personalized, preventative care frameworks. Furthermore, the study advocates for the integration of interpretable ML models into clinical workflows, paving the way for early screening, targeted interventions, and improved patient outcomes in ADRD care

REFERENCES

[1] Mattson MP. Pathways towards and away from Alzheimer's disease. Nature 2004;430 (7000):631–9.

[2] Kavitha C, Mani V, Srividhya SR, Khalaf OI, Tavera Romero CA. Early-stage Alzheimer's disease prediction using machine learning models. Front Public Heal 2022;10(March):1–13.

[3] Šerý O, Povová J, Míšek I, Pešák L, Janout V. Molecular mechanisms of neuropatho- logical changes in Alzheimer's disease: a review. Folia Neuropathol 2013;51(1):1–9.

[4] Jack CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. NIA-AA Research Framework: toward a biological definition of Alzheimer's disease. Alzheimer's Dement [Internet] 2018;14(4):535–62 Disponible a. doi: 10.1016/j.jalz.2018.02.018.

[5] Hammond TC, Xing X, Wang C, Ma D, Nho K, Crane PK, et al. β-amyloid

[6] and tau drive early Alzheimer's disease decline while glucose hypometabolism drives late decline. Commun Biol [Internet] 2020;3(1):1–13 Disponible a. doi:10.1038/s42003-020-1079-x.

[7] Chang CH, Lin CH, Lane HY. Machine learning and novel biomarkers for the diag- nosis of alzheimer's disease. Int J Mol Sci 2021;22(5):1–12.

[8] Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the Na- tional Institute on Aging-Alzheimer's Association workgroups on diagnostic guide- lines for Alzheimer's disease. Alzheimer's Dement 2011;7(3):280–92.

[9] Villemagne VL, Burnham S, Bourgeat P, Brown B, Ellis KA, Salvado O, et al. Amy- loid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. Lancet Neurol [Internet] 2013;12(4):357–67 Disponible a. doi:10.1016/S1474-4422(13)70044-9.

[10] McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the Na- tional Institute on Aging-Alzheimer's Association workgroups on diagnostic guide- lines for Alzheimer's disease. Alzheimer's Dement 2011;7(3):263–9.

[11] Petersen RC, Caracciolo B, Brayne C, Gauthier S, Jelic V, Fratiglioni L. Mild cognitive impairment: a concept in evolution. J Intern Med 2014;275(3):214–28.

[12] Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The di- agnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diag- nostic guidelines for Alzheimer's disease. Alzheimer's Dement 2011;7(3):270–9.

[13] Alzheimer's Association Report2022 Alzheimer's disease facts and figures. Alzheimer's Dement 2022;18(4):700–89.

[14] Alzheimer's Association Report2023 Alzheimer's disease facts and figures. Alzheimer's Dement 2023;19(4):1598–695.

[15] Hurd MD, Martorell P, Delavande A, Mullen KJ, Langa KM. Monetary costs of de- mentia in the United States. N Engl J Med 2013;368(14):1326–34.

[16] Zhang R, Simon G, Yu F. Advancing Alzheimer's research: a review of big data promises. Int J Med Inform [Internet] 2017; 106:48–56 June 2016Disponible a. doi: 10.1016/j.ijmedinf.2017.07.002.

[17] Wimo A, Guerchet M, Ali GC, Wu YT, Prina AM, Winblad B, et al. The world- wide costs of dementia 2015 and comparisons with 2010. Alzheimer's Dement 2017;13(1):1–7.

[18] Cummings JL, Morstorf T, Zhong K. Alzheimer's disease drug-development pipeline: few candidates, frequent failures. Alzheimers Res Ther 2014;6(4):37.

[19] Hampel H, Shaw LM, Aisen P, Chen C, Lleó A, Iwatsubo T, et al. State-of-the-art of lumbar puncture and

its place in the journey of patients with Alzheimer's disease. Alzheimer's Dement 2022;18(1):159–77.

[20] Alzheimer's Association Report 2020Alzheimer's disease facts and figures 2020. Alzheimer's Dement 2020;16(3):391–460.

[21] Barthold D., Joyce G., Ferido P., Drabo E.F., Marcum Z.A., Gray S.L., et al. Pharma- ceutical treatment for Alzheimer' s disease and related dementias: utilization and disparities. 2020; 76:579–89.

[22] Yiannopoulou KG, Papageorgiou SG. Current and future treatments for Alzheimer's disease. Ther Adv Neurol Disord 2013;6(1):19–33.

.