

Review

Machine Learning Driven by Magnetic Resonance Imaging for the Classification of Alzheimer Disease Progression: Systematic Review and Meta-Analysis

Gopi Battineni^{1,2}, PhD; Nalini Chintalapudi¹, PhD; Francesco Amenta¹, Prof Dr, MD

¹Clinical Research, Telemedicine and Telepharmacy Centre, School of Medicinal and Health Products Sciences, University Camerino, Camerino, Italy

²Centre for Global Health Research, Saveetha University, Saveetha Institute of Medical and Technical Sciences, Chennai, India

Corresponding Author:

Gopi Battineni, PhD

Clinical Research, Telemedicine and Telepharmacy Centre

School of Medicinal and Health Products Sciences

University Camerino

Via Madonna Delle Carceri 9

Camerino, 62032

Italy

Phone: 39 3331728206

Email: gopi.battineni@unicam.it

Abstract

Background: To diagnose Alzheimer disease (AD), individuals are classified according to the severity of their cognitive impairment. There are currently no specific causes or conditions for this disease.

Objective: The purpose of this systematic review and meta-analysis was to assess AD prevalence across different stages using machine learning (ML) approaches comprehensively.

Methods: The selection of papers was conducted in 3 phases, as per PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) 2020 guidelines: identification, screening, and final inclusion. The final analysis included 24 papers that met the criteria. The selection of ML approaches for AD diagnosis was rigorously based on their relevance to the investigation. The prevalence of patients with AD at 2, 3, 4, and 6 stages was illustrated through the use of forest plots.

Results: The prevalence rate for both cognitively normal (CN) and AD across 6 studies was 49.28% (95% CI 46.12%-52.45%; $P=.32$). The prevalence estimate for the 3 stages of cognitive impairment (CN, mild cognitive impairment, and AD) is 29.75% (95% CI 25.11%-34.84%, $P<.001$). Among 5 studies with 14,839 participants, the analysis of 4 stages (nondemented, moderately demented, mildly demented, and AD) found an overall prevalence of 13.13% (95% CI 3.75%-36.66%; $P<.001$). In addition, 4 studies involving 3819 participants estimated the prevalence of 6 stages (CN, significant memory concern, early mild cognitive impairment, mild cognitive impairment, late mild cognitive impairment, and AD), yielding a prevalence of 23.75% (95% CI 12.22%-41.12%; $P<.001$).

Conclusions: The significant heterogeneity observed across studies reveals that demographic and setting characteristics are responsible for the impact on AD prevalence estimates. This study shows how ML approaches can be used to describe AD prevalence across different stages, which provides valuable insights for future research.

(JMIR Aging 2024;7:e59370) doi: [10.2196/59370](https://doi.org/10.2196/59370)

KEYWORDS

Alzheimer disease; ML-based diagnosis; machine learning; prevalence; cognitive impairment; classification; biomarkers; imaging modalities; MRI; magnetic resonance imaging; systematic review; meta-analysis

Introduction

The progression of Alzheimer disease (AD) affects memory, thinking, and behavioral functions over time [1]. Not only the

individuals affected by the condition but also their families and caregivers, who have to cope with it daily. AD has become a major health concern worldwide because of the aging population in the last 3 decades [2,3]. The majority of cases of AD occur

among older individuals, and increasing evidence suggests that a combination of genetic, lifestyle, and environmental factors is behind it [3,4]. The progression of the disease causes a slow deterioration of memory and cognitive abilities.

AD is represented by different stages of progression such as cognitively normal (CN) [5], significant memory concern (SMC) [6], early mild cognitive impairment (EMCI) [7], mild cognitive impairment (MCI) [8], and late mild cognitive impairment (LMCI) [7,8]. Biomarkers could help detect individuals at risk of AD before symptoms occur. Cerebrospinal fluid (CSF) testing is considered the most reliable marker of progression of AD. Brain neuroimaging like computerized tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET), blood tests, and genetic testing are attracting increasing attention as important markers of this pathology [1,9,10]. CSF biomarkers such as β -amyloid 42 and tau and phosphor tau are key indicators of AD [11]. An MRI or CT scan can reveal structural changes associated with AD, while a PET scan can reveal amyloid plaques and tau tangles in the brain [12]. The early diagnosis of AD can be aided by the identification of novel biomarkers, the identification of hidden data patterns, and the generation of hypotheses [13-16]. Machine learning (ML)-based predictive models can help us detect early signs of AD, improve diagnostic accuracy, and enable timely interventions [16,17].

ML applications in medicine have received significant attention for their potential in disease detection and diagnosis [18]. ML models have been proposed in existing literature to improve diagnostic accuracy for early detection of AD [19-21]. It is said that ML algorithms aid in forecasting outcomes for patients with AD, diagnosing illnesses, and tailoring treatments [15]. ML models have been reported to be able to predict patient readmissions, which allows health care providers to allocate resources more efficiently and improve patient outcomes [15,22]. In addition, deep learning (DL) algorithms can examine medical images, like CT scans or MRIs, to aid in identifying abnormalities [23-25]. The application of DL techniques to conventional MRI could reduce patient burden, risk, and cost when extracting biomarker information [26,27].

DL-based neural networks contribute significantly to AD detection [28,29]. Hierarchical representations can be learned by neural networks and achieve promising results in AD, especially when applied to neuroimaging data [30,31]. Their role includes assisting in the discovery of new AD biomarkers and analyzing large datasets to identify patterns and correlations that are indicative of AD progression [32]. Convolutional neural networks (CNNs) are used in the analysis of AD image data in the form of MRI [33], PET [34], and CT scans [35]. CNNs can automatically extract relevant features from complex imaging data and learn hierarchical representations of subtle AD patterns.

Advanced techniques like Gradient-Weighted Class Activation Mapping after CNN model training highlight important regions of the input MRI brain image [36,37]. The brain areas in these regions are responsible for influencing the model's AD prediction. These techniques bridge the gap between accuracy and interpretability in AD detection. Moreover, recurrent neural networks are capable of analyzing temporal data, such as

longitudinal studies examining cognitive decline over time [38]. Predicting cognitive decline trajectories and future outcomes is possible through the capture of sequential dependencies in data [38,39]. Multimodal data integration can enhance the accuracy of AD detection models, resulting in a more comprehensive view of the patient's condition [40].

The role of ML models in the early diagnosis of AD has not been determined through extensive review of ML algorithms and meta-analysis. The accuracy and efficiency of AD diagnosis can be enhanced by using advanced algorithms and models as well as careful feature selection and extraction. However, the level of reliability of these techniques is a significant factor. The objective of this study is to address the knowledge gap by conducting a systematic review and meta-analysis of ML applications for AD detection, which aim to establish their role in improving diagnostic accuracy and patient outcomes. The main contribution of this study is (1) assessing the role of image feature selection methods in achieving competitive accuracy in AD classification modeling, (2) examining the ML methods that can be used to detect AD with the help of magnetic resonance image modeling, and (3) identifying the best ML classifier based on accuracy metrics.

Methods

This study was conducted by identifying, selecting, and analyzing relevant studies, which included a literature search, screening document inclusion criteria, and tools for risk bias assessment.

Search Strategy

A systematic search was carried out using libraries such as PubMed (MEDLINE), Scopus, and Web of Science. The search followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) 2020 guidelines to maintain transparency, authenticity, and completeness of details of reporting [41]. The PRISMA checklist of this paper can be found in [Multimedia Appendix 1](#). This search was carried out over the last 15 years and was centered on published studies specific to early-stage AD detection and classification (between January 2010 and March 2024). Limiting our review to the last 15 years of publication allowed us to focus on papers reflective of current trends in research.

The search strategy used the following keywords: "Alzheimer's disease," "machine learning," "early detection," "diagnostic accuracy," "diagnosis," "predictive models," "biomarkers," "deep learning," "diagnostic accuracy," "feature selection," "AD biomarkers," and "ML models." The search strategy was ("machine learning" OR "artificial intelligence" OR "classification") AND "Alzheimer's disease" AND "MRI" AND "diagnosis" AND "classification."

Inclusion and Exclusion Criteria

Full-text papers in the English language were considered. We have included in this study only published papers in peer-reviewed journals. The majority of the papers analyzed were centered on MRI data combined with ML models in AD diagnosis. Selected studies included patients diagnosed with early-stage AD and healthy controls. The papers published with

a title or abstract containing at least 1 abovementioned keyword were considered for inclusion.

Papers written in a language other than English were excluded. We excluded studies that were not specifically conducted in the context of AD diagnosis using MRI and were not primarily focused on ML models. Papers published before 2010 were not considered. Studies in which ML in MRI was not explicitly linked to clinical diagnosis, medical training, or initiatives to improve AD diagnosis were excluded. This review excluded studies using PET and CT scans because the primary focus was on ML in MRI, which is specifically linked to clinical diagnosis, medical training, and initiatives to enhance AD diagnosis. The selection process excluded review papers, conference proceedings, and gray literature reports.

Paper Screening

Multiple stages were involved in the paper selection process. The results of the systematic search were documented in a spreadsheet using the above strategy. The selected papers were equally distributed among the authors, and each paper was screened by examining titles and abstracts to identify potentially relevant publications. The selected papers were then reviewed comprehensively according to predefined inclusion and exclusion criteria in the subsequent phase. To facilitate synthesis, relevant information was extracted and organized in a tabular format, covering study design, datasets, performance metrics, model validation, and feature selection. As a result, a summary of each study's main findings to discern trends, patterns, and common themes was done.

Quality and Publication Bias

The Newcastle-Ottawa Scale [42] was used to assess the study quality based on different factors such as selection, comparability, and outcome, providing a structured approach to gauge the risk of bias. In terms of quality, scores ranged from very poor (0-3) to moderate (4-6) to excellent (7-9). The papers meeting the score (Newcastle-Ottawa Scale ≥ 7) were only considered for final review. Two authors (GB and NC) independently assessed the quality, and any discrepancies were resolved through discussion or consultation with a third author (FA).

Statistical Analysis

The statistical tests Egger regression [43] and Begg rank correlation [44] were used to address the potential bias of publications. To assess the strength of our findings against potential biases or variations in study characteristics, sensitivity analyses were performed. Lower methodological quality or different study designs were excluded. To identify the effect size measures and quantify the strength or magnitude of the relationship between variables or the magnitude of differences between AD groups, we applied the "PLOGIT" function to the logit transformation of the proportion [45]. The logit transformation is commonly used when dealing with proportions or probabilities, especially when they are bounded between 0 and 1. An inverse variance method has been applied that

specifies the method for pooling effect sizes. There were 2 types of models considered in the meta-analysis: fixed effects and random effects. Using the fixed effects model when we observed a low level of heterogeneity, the test is not statistically significant.

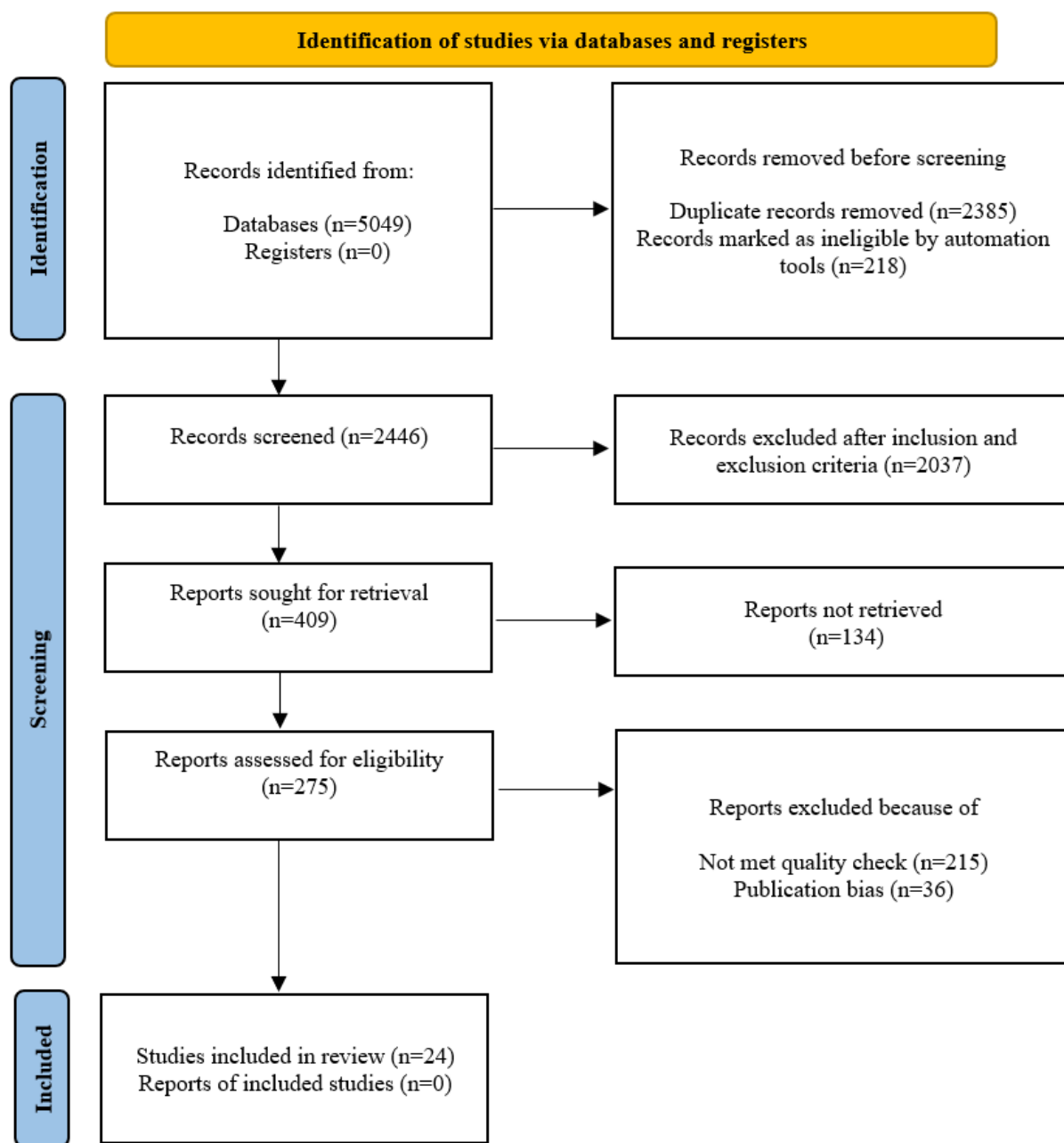
The random effects model (REM) was considered for the heterogeneity test with statistical significance [46]. By calculating T^2 , the amount of heterogeneity between the true effect sizes of different studies was quantified. An estimation method using a restricted maximum likelihood estimator that maximizes the likelihood function while accounting for other parameters of the model was used [47]. I^2 and Cochran Q statistic tests were conducted to assess the heterogeneity among the effect sizes of individual studies [48,49]. The measures of heterogeneity (T^2 and I^2) indicate the variability in AD prevalence estimates across the studies [50].

The prevalence of patients with AD across different subgroups within the overall population was also investigated. Subgroup analysis enables the identification of factors that can influence prevalence estimates and provide insight into the sources of heterogeneity [51]. A subgroup-specific meta-analysis model was used to calculate the pooled prevalence estimates for each subgroup, followed by a comparison of the prevalence estimates across subgroups to assess whether there were any significant differences. Data were subgrouped into 4 category-based AD classifications namely, 2-group classification, 3-group classification, 4-group classification, and 6-group classification. The 2-group classification involved individuals either without dementia (nondemented, ND) or with dementia (demented, AD). The 3-group classification includes CN, MCI, and AD. The 4-group classification comprises ND, mildly demented (MD), moderately demented (MoD), and AD. Meanwhile, the 6-group classification involves CN, SMC, EMCI, MCI, LMCI, and AD. Each subgroup data was recorded separately into a Microsoft Excel spreadsheet, which was further supplied as input to R software (version 4.3.3; R Foundation for Statistical Computing). For prevalence and summary meta-analysis, we used the "meta prop" functions available in the *meta* package.

Results

Search Outcomes

During the identification phase, 5049 records were obtained from 3 major scientific databases using the given search strategy. Following the removal of duplicates ($n=2355$) and the assessment of ineligibility using tools ($n=218$), 2446 records were included in the screening stage. The inclusion and exclusion criteria determined that 2037 records were ineligible. We further screened 409 records, with 134 being excluded due to lack of full-text availability. In total, 251 records from the remaining 275 were excluded due to low-quality scores and publication bias. A total of 24 papers were included in the final analysis. Details on the procedures for selecting papers are summarized in Figure 1.

Figure 1. Paper screening procedure flowchart.

Data Sources

The data collected for this study were collected from various geographical locations and may have included memory clinics

and neurology departments, suggesting a focus on cognitive impairment and related conditions. [Table 1](#) displays the distribution of AD imaging sample data along with data sources.

Table 1. Participants’ data collected from different sources.

| Data source | AD ^a , n/N (%) | Reference |
|---|---------------------------|-----------|
| ADNI ^b | 33/204 (16.17) | [52] |
| Tianjin First Central Hospital, China | 27/56 (48.21) | [53] |
| ADNI and AIBL ^c | 1673/3335 (50.16) | [54] |
| OASIS ^d 3 | 1077/3979 (27.06) | [55] |
| ADNI | 70/210 (30) | [56] |
| ADNI | 193/818 (23.59) | [57] |
| ADNI | 3200/6400 (50) | [58] |
| ADNI | 186/805 (23.10) | [59] |
| Kaggle | 3200/6400 (50) | [60] |
| ADNI | 231/432 (50) | [61] |
| Kaggle | 3200/6400 (50) | [62] |
| Shanghai Pudong New Area People’s Hospital | 55/119 (46.21) | [63] |
| ADNI | 268/1048 (25.57) | [64] |
| ADNI and NACC ^e | 1170/4644 (25.19) | [65] |
| Kaggle and ADNI | 390/1310 (29.77) | [66] |
| ADNI | 584/1421 (41.1) | [67] |
| ADNI | 25/138 (18.11) | [68] |
| OASIS 1 | 78/150 (52) | [69] |
| Kaggle | 3200/6400 (50) | [70] |
| Memory clinic of the neurology department in Nanfang Hospital | 44/180 (24.44) | [71] |
| ADNI | 118/245 (48.16) | [72] |
| ADNI | 24/142 (16.90) | [73] |
| ADNI | 1077/3979 (27.06) | [74] |
| ADNI | 260/560 (46.42) | [75] |

^aAD: Alzheimer disease.
^bADNI: Alzheimer’s Disease Neuroimaging Initiative.
^cAIBL: Australian Imaging Biomarkers and Lifestyle Flagship Study of Ageing.
^dOASIS: Open Access Series of Imaging Studies.
^eNACC: National Alzheimer’s Coordinating Center.

The analyzed studies collected image data from various sources such as Alzheimer’s Disease Neuroimaging Initiative (ADNI) [76], Open Access Series of Imaging Studies (OASIS) [77], Australian Imaging Biomarkers and Lifestyle Flagship Study of Ageing [78], and public domains like Kaggle [79]. ADNI datasets were used more often for image collection [52,54,56-59,61,64-68,72-75,80]. The purpose of ADNI is to develop biomarkers for early detection and AD tracks through a multicenter study involving clinical imaging, genetics, and biochemistry. The studies that use ADNI datasets aim to detect AD at its prime stage. One study jointly applied 2 image datasets from ADNI and Australian Imaging Biomarkers and Lifestyle Flagship Study of Ageing [54].

OASIS brains aim to make it possible for anyone to access neuroimaging datasets of the brain through an initiative known as Open Access to Neuroimaging Datasets. Through this project,

researchers can access and use a variety of brain imaging data for free. This resource assists neuroscience researchers in advancing their research by providing a comprehensive collection of brain imaging datasets. Cross-sectional OASIS 1 data were used by researchers for hypothesis-driven analysis, neuroanatomical atlases, and segmentation algorithms [69]. In another study, OASIS-3 was integrated with longitudinal neuroimaging, clinical, cognitive, and biomarker data [55]. The use of public datasets or participation in Kaggle competitions related to AD research helps as a platform for data science competitions and datasets [70]. Three studies collected data from 3 hospitals in China [53,63,71]. The findings indicate that a diverse dataset from multiple sources, such as clinical settings and publicly available datasets, could provide a comprehensive basis for AD research and analysis.

Study Characteristics

AD Stages

Table 2 presents a summary of various studies, which includes authors, publication year, AD stages, preprocessing techniques, classifiers, validation methods, and the best-performing model. Four studies have examined the progression of AD over 6 stages to gain a better understanding of how diseases develop and change [52,54,68,73]. Seven studies examined 4 groups of AD stages analyzing neurobiological mechanisms behind cognitive decline or exploring nonpharmacological treatments [55,58-60,64,66,70]. Similarly, 7 works associated with 3-stage

classification studies involved patients with CN, MCI, and AD [56,57,62,65,67,71,74]. These studies were mainly focused on the early detection of dementia with subtle differences in biomarkers and cognitive performance. Moreover, the ML models used in the study predicted AD progress in estimating the transition from MCI to dementia. Finally, 6 studies associated a binary or 2-stage classification of AD with ML models to identify biomarkers that predict treatment response or disease progression [53,61,63,69,72,75]. This enables more effective targeted therapies and biomarker-driven clinical trials to be developed.

Table 2. Machine learning models and their characteristics.

| | Author | Year | AD ^a stages | Image preprocessing methods | ML ^b models incorporated | Validation | Diagnosis accuracy (%) | Best model |
|----|------------------------|------|------------------------|--|---|---------------------------------|------------------------|-----------------------------------|
| 1 | Alorf and Khan [52] | 2022 | 6 | Normalization and smoothing | GLMICA ^c | K-fold (10) | 84.03 | BC-GCN ^d |
| 2 | Chen et al [53] | 2017 | 2 | Diffusivity and kurtosis mapping and ROI ^e | SVM ^f | K-fold (10) | 96.23 | SVM with DKI ^g |
| 3 | Mofrad et al [54] | 2021 | 6 | LME ^h for ROI extraction | SVC ⁱ | K-fold (15) | 69-75 | SVC |
| 4 | EL-Geneedy et al [55] | 2023 | 4 | Image normalization | DenseNet121 ^j , ResNet50 ^k , VGG16 ^l , EfficientNetB7, and InceptionV3 | K-fold (10) | 99.68 | Customized CNN ^m model |
| 5 | Hazarika et al [56] | 2022 | 3 | Histogram-based approach | 20 Different DL ⁿ models | K-fold (10) | 90.22 | DenseNet121 |
| 6 | Khan et al [57] | 2022 | 3 | SMOTE ^o | 16 Different ML models | K-fold (10) | 90.24 | SVM with DKI |
| 7 | Sorour et al [58] | 2024 | 4 | Image normalization and labeling | CNN, LSTM ^p , SVM, and VGG16 | K-fold (10) | 99.92 | CNN-LSTM |
| 8 | Abdelaziz et al [59] | 2021 | 4 | Interpolation | CNN | K-fold (10) | 98.22 | CNN |
| 9 | Sharma et al [60] | 2022 | 4 | VGG16 | Neural network with VGG16 feature extractor | K-fold (10) | 90.4 | VGG16 |
| 10 | Nguyen et al [61] | 2022 | 2 | Augmentation | 3D-ResNet, XGB ^q | K-fold (5) | 96.20 | XGB |
| 11 | Saleh et al [62] | 2023 | 3 | CNN feature extraction | DenseNet121, 169, and 201 | K-fold (10) | 96.05 | DenseNet201 |
| 12 | Yang et al [63] | 2022 | 2 | Recursive feature elimination | Recursive random forest (RF) | K-fold (10) | 97 | RF |
| 13 | El-Sappagh et al [64] | 2021 | 4 | SMOTE | SVM, KNN ^r , DT ^s , NB ^t , RF | K-fold (10) | 87.76 | RF |
| 14 | Liu et al [65] | 2022 | 3 | Unified segmentation | 3D CNN | Holdout and external validation | 85.12 | 3D CNN |
| 15 | Elgammal et al [66] | 2022 | 4 | Generalization | KNN | Multifractal geometry | 99.4 | KNN |
| 16 | Das et al [67] | 2021 | 3 | Skull stripping, intensity normalization, corpus callosum segmentation | SVM | K-fold (100) | 90 | SVM |
| 17 | Chelladurai et al [68] | 2023 | 6 | Gray-level co-occurrence matrix | RF, XGB, DT, SVM, MLP ^u | Evaluation metrics | 99.44 | MLP |
| 18 | Battineni et al [69] | 2021 | 2 | Outliers' detection | RF, GNB ^v , LR ^w , SVM, gradient boosting, and Ada boosting | K-fold (10) | 97.58 | Gradient boosting |
| 19 | Sharma et al [70] | 2022 | 4 | Normalization and augmentation | SVM, XGB, GNB | Not mentioned | 89.89 | SVM |

| | Author | Year | AD ^a stages | Image preprocessing methods | ML ^b models incorporated | Validation | Diagnosis accuracy (%) | Best model |
|----|-------------------------|------|------------------------|---|---|---------------|------------------------|-----------------------------------|
| 20 | Long et al [71] | 2023 | 3 | MRMR ^x algorithm in combination with the SFC ^y method | SVM, ANN ^z | K-fold (10) | 80.36 | SVM |
| 21 | Wang et al [72] | 2023 | 2 | Deep features extraction | CNN | K-fold (5) | 98.86 | CNN |
| 22 | Tajammal et al [73] | 2023 | 6 | Augmentation | VGG16, ResNet18, Alex Net, Inception V1, Custom CNN | Not mentioned | 96.2 | Custom CNN |
| 23 | Golovanevsky et al [74] | 2022 | 3 | Unified hyperparameter tuning | Multimodal | K-fold (3) | 96.88 | Multimodal AD diagnosis framework |
| 24 | Li and Yang [75] | 2021 | 2 | Transfer learning | SVM, VGG Net ^{aa} , ResNet | K-fold (5) | 95 | VGG Net, ResNet |

^aAD: Alzheimer disease.
^bML: machine learning.
^cGLMICA: generalized linear model incorporating covariates analysis.
^dBC-GCN: brain connectivity-based graph convolutional network.
^eROI: region of interest.
^fSVM: support vector machine.
^gDKI: diffusion kurtosis imaging.
^hLME: linear mixed-effects model.
ⁱSVC: support vector classifier.
^jDenseNet: dense convolutional network.
^kResNet: residual network.
^lVGG: Visual Geometry Group.
^mCNN: convolutional neural network.
ⁿDL: deep learning.
^oSMOTE: Synthetic Minority Oversampling Technique.
^pLSTM: long short-term memory.
^qXGB: extreme gradient boosting.
^rKNN: k-nearest neighbor.
^sDT: decision tree.
^tNB: Naïve Bayes.
^uMLP: multilayer perceptron.
^vGNB: Gaussian Naïve Bayes.
^wLR: logistic regression.
^xMRMR: minimum redundancy maximum relevance.
^ySFC: sparse functional connectivity.
^zANN: artificial neural network.
^{aa}VGG Net: Visual Geometry Group network.

Feature Engineering Techniques

Feature engineering plays an important contribution in brain image analysis [81]. Various feature techniques were discussed to tackle challenges in AD classification, such as class imbalance, feature extraction, robustness, and generalization. ConvNet or CNN was designed for processing grid-like data, such as images, using convolutional layers to learn spatial hierarchies of features automatically [62]. Visual Geometry Group (VGG16) uses 3×3 convolution filters to construct a

16-layer CNN architecture and is known for its simplicity and high performance in image classification tasks [60]. Models like multilayer perceptron, Dense Net, Efficient Net, and residual network in AD classification lie in their ability to effectively handle deep neural networks for feature extraction and classification, which is crucial in analyzing complex brain magnetic resonance images for AD detection. Support vector machine (SVM) is a supervised learning algorithm used for AD classification, and it constructs hyperplanes in a high-dimensional space to separate different classes. In contrast,

diffusion kurtosis imaging (DKI) is an MRI procedure that captures non-Gaussian diffusion, giving insight into tissue microstructure and facilitating better brain mapping. These techniques range from basic normalization [55,58,70], outlier detection [69], interpolation [59], and transfer learning [75] to more advanced methods such as data augmentation [61,70,73], feature extraction using DL models like VGG16 [60], deep feature extraction [72], ConvNet [62], and statistical modeling for region of interest extraction [54]. Another paper extracted features related to corpus callosum atrophy for AD diagnosis [67]. A single study investigated texture analysis in brain images using the Gabor and gray-level co-occurrence matrix [52]. For feature selection and analysis of functional connectivity patterns, another investigation used the minimum redundancy maximum relevance algorithm alongside the sparse functional connectivity method [55]. Unified hyperparameter tuning was applied to optimize model parameters across algorithms and settings [58].

Classifiers

Supervised models like SVM were used by several studies for classification tasks due to their effectiveness in handling high-dimensional magnetic resonance image data and nonlinear relationships [53,54,58,64,67-69,71,75]. The generalized linear model incorporating covariates analysis was used by Alorf and Khan [52] to assess a model's performance and generalization ability by ensuring that all data points are used during both training and validation, reducing overfitting risk and allowing more reliable model performance estimates. The authors demonstrated that MRI data can be fine-tuned to capture subtle differences in brain morphology associated with AD by using pretrained models [55].

Similarly, to learn discriminative patterns, other models like logistic regression (LR), decision tree, Gaussian Naive Bayes, and k-nearest neighbor (KNN) largely contribute to the MRI-based AD classification. The combination of these multimodal classifiers was adopted among 6 works to leverage AD early diagnosis [63,64,66,68-70]. Alternatively, CNN-based DL models have the capability of autonomous learning and represent complex patterns in magnetic resonance images. In this review were identified 2 studies that used dense convolutional network (DenseNet) [55,62] and Inception [55,73]. In total, 4 studies applied residual network [55,61,73,75], 5 studies used VGG [55,58,60,73,75], and 1 study the EfficientNet [55]. The multimodeling approaches (comparison of 16 and 20 classifiers) of CNN models were incorporated in 2 works [56,57]. Long short-term memory, another DL framework largely used in the context of MRI classification, can be used to analyze sequential data, such as time-series MRI scans, to detect temporal changes in brain structures characteristic of AD progression [58]. One study used a different approach, the multimodal neural networks for analyzing data from multiple sources or modalities [74]. Ensemble learning techniques like extreme gradient boosting (XGB), gradient boosting, and Ada boosting combine weak learners to create a more powerful classification. MRI data in 4 studies were successfully handled by the XGB classifier, which captured nonlinear relationships between features and predicted AD status accurately [61,68-70].

Validation Techniques

K-fold cross-validation is a common method used by most studies, where the dataset is divided into K subsets, and the model is trained and tested for K times. Testing was conducted on each subset, while the remaining ones served as training. This method can be used to assess model performance and generalization across different subsets of data. The K-fold has been used in most studies with varying values of K including 3 [74], 5 [61,72,75], 10 [52,53,55-60,62-64,69,71,72], 15 [54], and 100 [67], indicating that the total partitioning of data varies depending on the level of validation. It is important to take into account the differences between different methods of validation. A recent study used a holdout technique and external validation, dividing the dataset into training and testing sets and performing an additional test on completely new, from-scratch datasets [65]. A unique approach to data analysis that uses multifractal geometries has been introduced by Elgammal et al [66] and is likely to involve characterizing complex patterns in data using fractal-based techniques. The findings above show that many validation methods need to be considered. Therefore, adaptable methodologies are necessary when it comes to datasets and objectives. On the other hand, there are a few mentions of specific evaluation metrics [68]. The use of K-fold cross-validation remains common, but the inclusion of alternative methods such as holdout and multifractal geometry suggests a willingness to explore new approaches to evaluating model performance and ensuring the robustness of ML and data analysis tasks.

Prevalence-Based Participant Pooling

There was no evidence of publication bias with Eggers ($P=.49$) or Begg ($P=.38$) tests. Figures 2-5 present the forest plot with the prevalences of participants with AD for 2, 3, 4, and 6 AD stage subgroups, respectively. Six studies with 1562 participants were identified among disease diagnoses with 2 stages including CN and AD [53,61,63,69,72,75]. The overall pooled prevalence of the REM reported 49.28% (95% CI 46.12%-52.45%; $I^2=15\%$; $P=.32$). Studies do not differ significantly in their estimates of prevalence, and the test of heterogeneity does not reveal substantial differences between them. Seven studies were identified with a total sample of 17,588 patients with AD with 3-stage AD classification including CN, MCI, and AD [56,57,62,65,67,71,73]. The overall prevalence of AD diagnosis is estimated at 29.75% (95% CI 25.11%-34.84%; $I^2=97\%$; $P<.001$). Each study provides an estimate of the AD prevalence among their respective populations with 95% CI. For example, Hazarika et al [56] found AD prevalence at 33.33% (95% CI 27%-40.15%). This indicates that if we were to combine the results of all the studies, this would be the estimated AD prevalence. $I^2=97\%$ indicates that a large proportion of the total variation in prevalence estimates is due to true differences between study populations rather than random error. The significant P value ($<.01$) for the test of heterogeneity indicates that there is substantial variability in AD diagnostic prevalence estimates among the studies.

Figure 2. A forest plot AD diagnosis prevalence (%) among 2-stage classification using random effects model [53,61,63,69,72,75]. AD: Alzheimer disease.

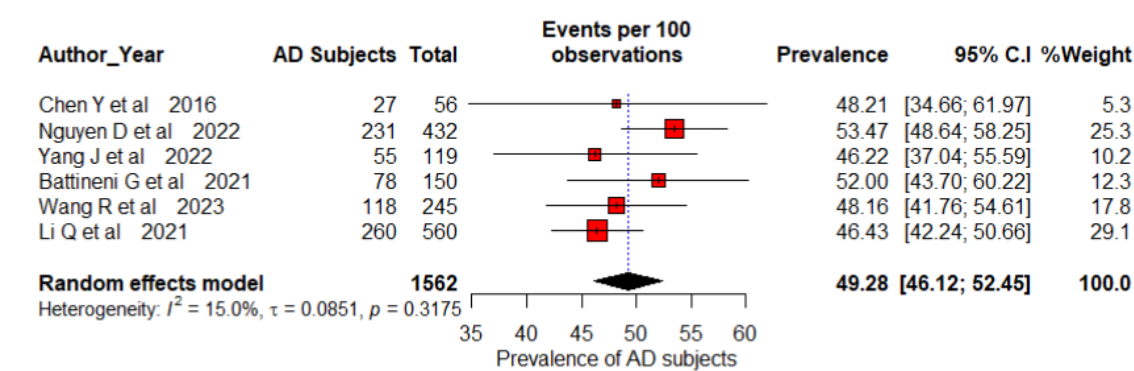


Figure 3. A forest plot AD diagnosis prevalence (%) among 3-stage classification using random effects model [56,57,62,65,67,71,74]. AD: Alzheimer disease.

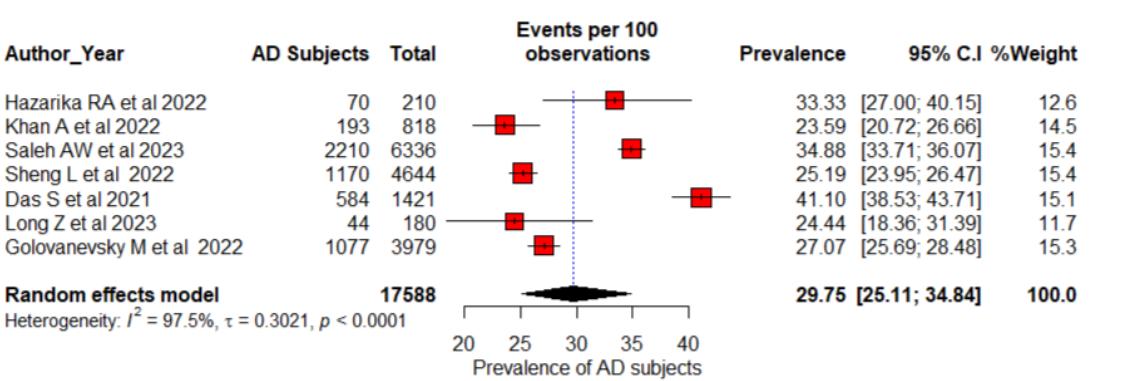


Figure 4. A forest plot AD diagnosis prevalence (%) among 4-stage classification using random effects model [55,59,64,66,70]. AD: Alzheimer disease.

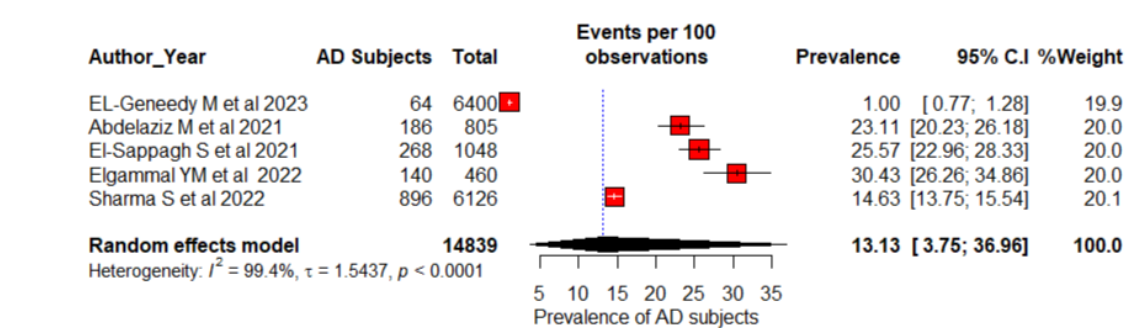
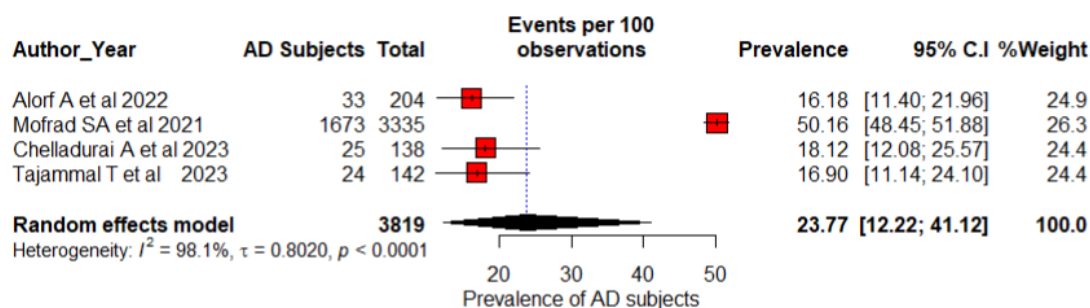


Figure 5. A forest plot AD diagnosis prevalence (%) among 6-stage classification using random effects model [52,54,68,73]. AD: Alzheimer disease.



Five studies with 14,839 participants were included for the meta-analysis of 4-stage AD classifications as ND, MoD, MD, and overt AD [55,59,64,66,70]. This systematic review included 7 studies, but we excluded 2 studies [58,60] because they used the same dataset with 6400 ADNI participants. Overall prevalence estimation with REMs is 13.13% (95% CI 3.75%-36.96%; $I^2=99\%$; $P<.001$). There is significant heterogeneity in the studies based on the high I^2 and significant P value and a considerable variation in the prevalence of AD across these studies, according to these estimates. Different research studies have found prevalence estimates ranging from 1% [55] to 30.43% [66]. The CIs indicate the degree of uncertainty in these estimates. As a result of the high degree of heterogeneity observed in the study, the true prevalence of AD may vary significantly between populations and settings. Four studies with 3819 were considered for the calculation of the overall prevalence of AD diagnosis of 6 stages such as CN, SMC, EMCI, MCI, LMCI, and AD [52,54,68,73]. The estimated prevalence for each study is ranging from 16.18% [52] to 50.16% [54]. The overall estimate of prevalence from the REM stands at 23.77% (95% CI 12.22%-41.12%; $I^2=0.8020$; $P<.001$). One study has a substantially greater estimated proportion of AD prevalence diagnosis than the other studies [54]. Compared to others, it reported the highest prevalence of 50.16% (95% CI 48.45%-51.88%) but does not differ weights (26.3%) significantly from other studies.

Meta-analysis through forest plots provides a comprehensive way of understanding meta-analysis results. It can be argued, however, that forest plots can only display CIs by assuming a fixed significant threshold ($P<.05$). It causes a replication crisis when hypothesis tests are conducted using P values. Based on P value functions, drapery plots were proposed to resolve this problem [82]. Using a drapery plot, an average effect and a confidence curve can be identified. The x-axis shows the effect size metric, and the y-axis shows the assumed P value. Multimedia Appendix 2 presents the drapery plots. There is a red curve showing the overall REM, which shows the P values for various effect sizes. Compared to the CI of pooled effects, the shaded area represents the prediction range. The prediction range is noticeably wider than the CI for the pooled effect. It indicates that the overall pooled effect does not fully capture the variability or uncertainty across different effect sizes.

Discussion

Principal Findings

In this work, we conducted a systematic review and meta-analysis based on the prevalence of patients with AD among different disease progression stages. For the systematic review, 24 studies were selected, among 22 selected for the meta-analysis. Due to their association with the same dataset of ADNI and similar sample size of patients with AD, these 2 studies avoid bias in the analysis [58,60]. The studies included in this review have explored the ML applications for AD diagnosis and intended to provide an understanding of AD progression, potentially with a focus on biomarker identification.

Different preprocessing techniques used to extract relevant features including cortical thickness [83], hippocampal volume [84], and brain activity patterns [85] from magnetic resonance images associated with AD were examined. According to the research objectives and AD stages being investigated, each study applied specific image preprocessing techniques. The progression of AD has been evaluated across multiple stages in our work. An accuracy range of 69%-75% is achieved with linear mixed-effects models that account for region of interest features with interparticipant variability of hierarchical structures [54]. Using image normalization, 1 study classified AD stages with different labeling with 84.03% accuracy by ensuring consistency in intensity and spatial properties [52,86]. Combining DL models with imaging techniques like MRI and PET has shown that structural and functional changes in the brain associated with AD can be detected [87,88]. Water molecules' diffusion properties in brain tissue can be measured using diffusivity and kurtosis mapping. The results provided insight into microstructural changes for a maximum accuracy of 96.23% [53]. By conducting magnetic resonance image normalization, the authors proposed an MRI-based DL technique for 99.68% accurate AD detection [55]. Magnetic resonance images were investigated for pixel intensity distributions to detect AD abnormalities [56].

These techniques encompass diverse methodologies ranging from normalization and smoothing to advanced mapping and feature extraction methods [89-91]. Several approaches have demonstrated high accuracy in identifying AD features, including image normalization, histogram-based approaches,

and diffusion mapping [92-95]. Techniques like recursive feature elimination and outlier detection showcase promising results, emphasizing the importance of feature selection and data quality assessment in enhancing classification performance [57,63]. A similar study analyzed and segmented different tissue types within MRI scans using unified segmentation. A magnetic resonance image of the brain was segmented simultaneously into different tissue types with 85.12% accuracy [65,96]. KNN-trained data can be used to classify AD with 99.4% accuracy using the generalization method [66]. Moreover, the use of advanced DL architectures such as VGG16 [60] and ConvNet [62] for feature extraction underscores the significance of leveraging sophisticated computational tools in AD research. Augmentation methods, interpolation, and transfer learning also emerge as valuable strategies for improving classification accuracy and robustness [73-75].

By integrating statistical and ML algorithms with preprocessing techniques, AD diagnosis research further enhances its interdisciplinary nature. The CNN-long short-term memory model had an accuracy of 99.92%, followed by the multimodal AD diagnosis framework model with a precision of 96.88%. The accuracy of a customized CNN model was 99.68%, SVM with DKI was 96.23%, XGB was 96.20%, and multilayer perceptron was 99.44%. In addition, DenseNet121, CNN, DenseNet201, random forest, and gradient boosting achieved accuracy levels between 90% and 97%. While some models demonstrated higher accuracy, such as 3D CNN and SVM, others demonstrated lower accuracy, 85.12% and 80.36%, respectively.

Many ML modeling techniques have been explored, including SVM, LR, and DenseNet. Ensemble methods like gradient boosting and Ada boosting have highlighted the importance of aggregating multiple models to improve predictive accuracy and robustness, especially when dealing with complex neurological disorders like AD [69,97]. The identification of specific best-performing models further underscores the importance of optimization of methods and model selection to improve diagnostic accuracy. The use of SVM along with DKI or DenseNet201 in different studies illustrates the researchers' tailored approach to leveraging each algorithm's and feature representation's strengths [98-100]. AD diagnosis is a nuanced process, where the choice of ML model can have a significant impact on model reliability and efficacy.

Data from magnetic resonance images have been analyzed using various ML models and validation techniques. To ensure robustness and generalization, the common technique used is K-fold cross-validation. Additionally, some authors have applied specific DL models along with traditional ML techniques, reflecting the diversity of approaches for modeling and validation [72,73]. Different mechanisms and approaches are used in each of these models to detect AD using magnetic resonance images. We have observed that SVM classifiers are

largely used for 2-stage classification such as CN and AD [53,54,67,71]. Similarly, LR classifiers were used in other studies to assess MRI-based AD status interpretation and predictive factors for disease risk assessment. Based on learned discriminative patterns from magnetic resonance images, these models, as well as others mentioned, produce accurate AD detection predictions. Additionally, KNN can be used to identify magnetic resonance images with feature vectors similar to those associated with AD helping to detect patterns.

The meta-analysis shows that there is a great deal of variation between studies when it comes to estimating AD prevalence. The reason for this is probably because the study involved a wide range of diagnostic criteria and populations, not just prevalence rates. The prevalence estimates are diverse due to some studies focusing on specific AD stages while others cover a wider spectrum. The significant P values and I^2 statistics show that the diagnosis of AD is highly heterogeneous and requires a nuanced understanding of its epidemiology. The challenges associated with synthesizing prevalence data from disparate sources are revealed by this analysis. The prevalence of AD is subject to complex and variable research, which leads to wider CIs in some studies. Even after trying to use REMs to account for this heterogeneity, significant variation persists, suggesting that variables like demographics, study design, and diagnostic methodology may play a significant role. The provision of more reliable estimates requires the adoption of standardized protocols and collaboration in future research efforts, which stresses the importance of rigorous methodology and careful interpretation of results.

Comparison With Existing Reviews

There have been a few systematic reviews and meta-analyses about the importance of ML models in AD diagnosis. Table 3 summarizes the comparison between our work and the reviews that have already been published. In our analysis, we concentrated on using ML for AD diagnosis, while other studies were focused on using it for dementia forecasting [101]. In a similar study [102], the authors explored the effectiveness of both ML and DL models in AD diagnosis. In this study, the authors did not examine multistage AD cases but only the binary classification of AD. A single study [103] conducted a meta-analysis based on Wilcoxon signed rank tests and discussed multiple imaging modalities, including MRI, PET, and CSF. Despite this, there is a lack of discussion about feature selection techniques and their potential impact on ML accuracy. A prevalence-based meta-analysis on MRI-centered AD discussions is presented in our study along with an in-depth description of subcategories of AD. Our study stands out because it covers all aspects of ML in AD diagnosis, including imaging modalities and stages of AD. We reviewed and analyzed various imaging modalities, talked about feature selection methods, and delved deeper into AD subcategories in our research.

Table 3. Comparison of this review with existing systematic reviews.

| Study | Systematic review | Meta-analysis | Imaging modalities | Feature selection | Alzheimer disease stages |
|-----------|-------------------|---------------|--------------------|-------------------|--------------------------|
| [101] | ✓ | | ✓ | | 4 |
| [102] | ✓ | | ✓ | | 2 |
| [103] | ✓ | ✓ | ✓ | | 6 |
| Our study | ✓ | ✓ | ✓ | ✓ | 6 |

Future Directions and Study Limitations

Data from open-access libraries such as ADNI, Kaggle, and others were used in studies, as evidenced by the analysis of datasets. Prospective validation studies should be carried out in the future to assess the accuracy of ML models for AD diagnosis across diverse populations and clinical settings. The incorporation of multimodal data, including imaging, genetics, and clinical information, into ML models can improve their accuracy and robustness in diagnosing AD and distinguishing it from other brain disorders [89]. To enhance their clinical utility and acceptability, ML models must be interpretable and explainable. It may be possible to use these models to predict the onset and AD progression based on longitudinal studies that track individuals over time [14,101]. Future research must incorporate ML models into diagnostic workflows and assess their influence on patient outcomes and health care delivery.

Despite its comprehensiveness, this study is characterized by some shortcomings. The availability and quality of data are essential for the effectiveness of ML approaches. The outcome of the meta-analysis may have been influenced by the limitations in access to complete datasets with different levels of quality. The potential for publication bias, in which studies with positive findings are more likely to be published, may lead to an overestimation of the effectiveness of ML approaches for diagnosing AD. The included studies may have experienced heterogeneity due to variations in study designs, patient populations, imaging modalities, and ML algorithms, making it difficult to draw definitive conclusions. Despite our best efforts to conduct a thorough review, some relevant studies may have been mistakenly excluded, potentially creating gaps in the

analysis. The generalizability of ML models for AD diagnosis may be limited by their development and validation on specific datasets.

Conclusions

A summary comparison of current literature on ML approaches in AD diagnosis, along with a systematic review and meta-analysis, helps to understand the prevalence of disease at different stages. Our analysis of 24 relevant papers shows a significant difference in AD prevalence estimates, as individuals progress from CN to MCI and ultimately to overt AD. We observed a pooled prevalence of 49.28% during the CN to AD transition. This was followed by 29.75% for CN, MCI, and AD, 13.13% for CN, MoD, MD, and AD, and 23.75% for CN, SMC, EMCI, MCI, LMCI, and AD. Our analysis reveals the importance of adjusting diagnostic and management strategies to minimize the impact of demographic and setting characteristics on AD prevalence estimates. Due to the heterogeneity observed across studies, it is necessary to consider various factors to accurately estimate the prevalence of AD. Our study is different from other studies by comparing it to existing systematic reviews and meta-analyses, which provide an original contribution to the topic under evaluation. Unlike previous studies that have focused on imaging modalities and AD stages, our study has comprehensively analyzed ML in AD diagnosis. Multiple imaging modalities were reviewed and analyzed, feature selection techniques were discussed, and AD subcategories were explored, focusing particularly on MRIs. Although none of the biomarkers currently available can provide a precise diagnosis of AD, using ML approaches to identify prevalence patterns across disease stages will lead to progress in AD diagnosis.

Acknowledgments

This study was supported by institutional grants of the University of Camerino.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) checklist.
[\[PDF File \(Adobe PDF File\), 57 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Drapery plots.
[\[DOCX File , 62 KB-Multimedia Appendix 2\]](#)

References

1. Hodson R. Alzheimer's disease. *Nature*. 2018;559(7715):S1. [doi: [10.1038/d41586-018-05717-6](https://doi.org/10.1038/d41586-018-05717-6)] [Medline: [30046078](https://pubmed.ncbi.nlm.nih.gov/30046078/)]
2. Jucker M, Walker LC. Alzheimer's disease: from immunotherapy to immunoprevention. *Cell*. 2023;186(20):4260-4270. [FREE Full text] [doi: [10.1016/j.cell.2023.08.021](https://doi.org/10.1016/j.cell.2023.08.021)] [Medline: [37729908](https://pubmed.ncbi.nlm.nih.gov/37729908/)]
3. Weller J, Budson A. Current understanding of Alzheimer's disease diagnosis and treatment. *F1000Res*. 2018;7:1161. [FREE Full text] [doi: [10.12688/f1000research.14506.1](https://doi.org/10.12688/f1000research.14506.1)] [Medline: [30135715](https://pubmed.ncbi.nlm.nih.gov/30135715/)]
4. Beata BK, Wojciech J, Johannes K, Piotr L, Barbara M. Alzheimer's disease—biochemical and psychological background for diagnosis and treatment. *Int J Mol Sci*. 2023;24(2):1059. [FREE Full text] [doi: [10.3390/ijms24021059](https://doi.org/10.3390/ijms24021059)] [Medline: [36674580](https://pubmed.ncbi.nlm.nih.gov/36674580/)]
5. Subramanyam Rallabandi VP, Seetharaman K. Classification of cognitively normal controls, mild cognitive impairment and Alzheimer's disease using transfer learning approach. *Biomed Signal Process Control*. 2023;79:104092. [doi: [10.1016/j.bspc.2022.104092](https://doi.org/10.1016/j.bspc.2022.104092)]
6. Swinford CG, Risacher SL, Charil A, Schwarz AJ, Saykin AJ. Memory concerns in the early Alzheimer's disease prodrome: regional association with tau deposition. *Alzheimers Dement (Amst)*. 2018;10(1):322-331. [FREE Full text] [doi: [10.1016/j.dadm.2018.03.001](https://doi.org/10.1016/j.dadm.2018.03.001)] [Medline: [29780876](https://pubmed.ncbi.nlm.nih.gov/29780876/)]
7. Lin SY, Lin PC, Lin Y, Lee Y, Wang C, Peng S, et al. The clinical course of early and late mild cognitive impairment. *Front Neurol*. 2022;13:685636. [FREE Full text] [doi: [10.3389/fneur.2022.685636](https://doi.org/10.3389/fneur.2022.685636)] [Medline: [35651352](https://pubmed.ncbi.nlm.nih.gov/35651352/)]
8. Ilardi CR, Chieffi S, Iachini T, Iavarone A. Neuropsychology of posteromedial parietal cortex and conversion factors from mild cognitive impairment to Alzheimer's disease: systematic search and state-of-the-art review. *Aging Clin Exp Res*. 2022;34(2):289-307. [FREE Full text] [doi: [10.1007/s40520-021-01930-y](https://doi.org/10.1007/s40520-021-01930-y)] [Medline: [34232485](https://pubmed.ncbi.nlm.nih.gov/34232485/)]
9. Ogbodo JO, Agbo CP, Njoku UO, Ogugofor MO, Egba SI, Ihim SA, et al. Alzheimer's disease: pathogenesis and therapeutic interventions. *Curr Aging Sci*. 2022;15(1):2-25. [doi: [10.2174/1874609814666210302085232](https://doi.org/10.2174/1874609814666210302085232)] [Medline: [33653258](https://pubmed.ncbi.nlm.nih.gov/33653258/)]
10. Dubois B, von Arnim CAF, Burnie N, Bozeat S, Cummings J. Biomarkers in Alzheimer's disease: role in early and differential diagnosis and recognition of atypical variants. *Alzheimers Res Ther*. 2023;15(1):175. [FREE Full text] [doi: [10.1186/s13195-023-01314-6](https://doi.org/10.1186/s13195-023-01314-6)] [Medline: [37833762](https://pubmed.ncbi.nlm.nih.gov/37833762/)]
11. Gunes S, Aizawa Y, Sugashi T, Sugimoto M, Rodrigues PP. Biomarkers for Alzheimer's disease in the current state: a narrative review. *Int J Mol Sci*. 2022;23(9):4962. [FREE Full text] [doi: [10.3390/ijms23094962](https://doi.org/10.3390/ijms23094962)] [Medline: [35563350](https://pubmed.ncbi.nlm.nih.gov/35563350/)]
12. Shaw LM, Arias J, Blennow K, Galasko D, Molinuevo JL, Salloway S, et al. Appropriate use criteria for lumbar puncture and cerebrospinal fluid testing in the diagnosis of Alzheimer's disease. *Alzheimers Dement*. 2018;14(11):1505-1521. [FREE Full text] [doi: [10.1016/j.jalz.2018.07.220](https://doi.org/10.1016/j.jalz.2018.07.220)] [Medline: [30316776](https://pubmed.ncbi.nlm.nih.gov/30316776/)]
13. Ahsan MM, Luna SA, Siddique Z. Machine-learning-based disease diagnosis: a comprehensive review. *Healthcare (Basel)*. 2022;10(3):541. [FREE Full text] [doi: [10.3390/healthcare10030541](https://doi.org/10.3390/healthcare10030541)] [Medline: [35327018](https://pubmed.ncbi.nlm.nih.gov/35327018/)]
14. Kumar S, Oh I, Schindler S, Lai A, Payne P, Gupta AM. Machine learning for modeling the progression of Alzheimer disease dementia using clinical data: a systematic literature review. *JAMIA Open*. 2021;4(3):ooab052. [FREE Full text] [doi: [10.1093/jamiaopen/ooab052](https://doi.org/10.1093/jamiaopen/ooab052)] [Medline: [34350389](https://pubmed.ncbi.nlm.nih.gov/34350389/)]
15. Chang CH, Lin CH, Lane HY. Machine learning and novel biomarkers for the diagnosis of Alzheimer's disease. *Int J Mol Sci*. 2021;22(5):2761. [FREE Full text] [doi: [10.3390/ijms22052761](https://doi.org/10.3390/ijms22052761)] [Medline: [33803217](https://pubmed.ncbi.nlm.nih.gov/33803217/)]
16. Wang N, Chen J, Xiao H, Wu L, Jiang H, Zhou Y. Application of artificial neural network model in diagnosis of Alzheimer's disease. *BMC Neurol*. 2019;19(1):154. [FREE Full text] [doi: [10.1186/s12883-019-1377-4](https://doi.org/10.1186/s12883-019-1377-4)] [Medline: [31286894](https://pubmed.ncbi.nlm.nih.gov/31286894/)]
17. Świetlik D, Białowas J. Application of artificial neural networks to identify Alzheimer's disease using cerebral perfusion SPECT data. *Int J Environ Res Public Health*. 2019;16(7):1303. [FREE Full text] [doi: [10.3390/ijerph16071303](https://doi.org/10.3390/ijerph16071303)] [Medline: [30979022](https://pubmed.ncbi.nlm.nih.gov/30979022/)]
18. Rahmani AM, Yousefpoor E, Yousefpoor MS, Mehmood Z, Haider A, Hosseinzadeh M, et al. Machine learning (ML) in medicine: review, applications, and challenges. *Mathematics*. 2021;9(22):2970. [doi: [10.3390/math9222970](https://doi.org/10.3390/math9222970)]
19. Borchert RJ, Azevedo T, Badhwar AP, Bernal J, Betts M, Bruffaerts R, et al. Artificial intelligence for diagnostic and prognostic neuroimaging in dementia: a systematic review. *Alzheimers Dement*. 2023;19(12):5885-5904. [doi: [10.1002/alz.13412](https://doi.org/10.1002/alz.13412)] [Medline: [37563912](https://pubmed.ncbi.nlm.nih.gov/37563912/)]
20. Hansson O, Blennow K, Zetterberg H, Dage J. Blood biomarkers for Alzheimer's disease in clinical practice and trials. *Nat Aging*. 2023;3(5):506-519. [FREE Full text] [doi: [10.1038/s43587-023-00403-3](https://doi.org/10.1038/s43587-023-00403-3)] [Medline: [37202517](https://pubmed.ncbi.nlm.nih.gov/37202517/)]
21. Skolariki K, Exarchos TP, Vlamos P. Computational models for biomarker discovery. *Adv Exp Med Biol*. 2023;1424:289-295. [doi: [10.1007/978-3-031-31982-2_33](https://doi.org/10.1007/978-3-031-31982-2_33)] [Medline: [37486506](https://pubmed.ncbi.nlm.nih.gov/37486506/)]
22. Qiu S, Miller MI, Joshi PS, Lee JC, Xue C, Ni Y, et al. Multimodal deep learning for Alzheimer's disease dementia assessment. *Nat Commun*. 2022;13(1):3404. [FREE Full text] [doi: [10.1038/s41467-022-31037-5](https://doi.org/10.1038/s41467-022-31037-5)] [Medline: [35725739](https://pubmed.ncbi.nlm.nih.gov/35725739/)]
23. Amiri Z, Heidari A, Navimipour NJ, Unal M, Mousavi A. Adventures in data analysis: a systematic review of deep learning techniques for pattern recognition in cyber-physical-social systems. *Multimed Tools Appl*. 2023;83(8):22909-22973. [doi: [10.1007/s11042-023-16382-x](https://doi.org/10.1007/s11042-023-16382-x)]
24. Klang E. Deep learning and medical imaging. *J Thorac Dis*. 2018;10(3):1325-1328. [FREE Full text] [doi: [10.21037/jtd.2018.02.76](https://doi.org/10.21037/jtd.2018.02.76)] [Medline: [29708147](https://pubmed.ncbi.nlm.nih.gov/29708147/)]
25. Li M, Jiang Y, Zhang Y, Zhu H. Medical image analysis using deep learning algorithms. *Front Public Health*. 2023;11:1273253. [FREE Full text] [doi: [10.3389/fpubh.2023.1273253](https://doi.org/10.3389/fpubh.2023.1273253)] [Medline: [38026291](https://pubmed.ncbi.nlm.nih.gov/38026291/)]

26. Vrahatis AG, Skolariki K, Krokidis MG, Lazaros K, Exarchos TP, Vlamos P. Revolutionizing the early detection of Alzheimer's disease through non-invasive biomarkers: the role of artificial intelligence and deep learning. *Sensors (Basel)*. 2023;23(9):4184. [FREE Full text] [doi: [10.3390/s23094184](https://doi.org/10.3390/s23094184)] [Medline: [37177386](https://pubmed.ncbi.nlm.nih.gov/37177386/)]
27. Feng X, Provenzano FA, Small SA. A deep learning MRI approach outperforms other biomarkers of prodromal Alzheimer's disease. *Alzheimers Res Ther*. 2022;14(1):45. [FREE Full text] [doi: [10.1186/s13195-022-00985-x](https://doi.org/10.1186/s13195-022-00985-x)] [Medline: [35351193](https://pubmed.ncbi.nlm.nih.gov/35351193/)]
28. Saleem TJ, Zahra SR, Wu F, Alwakeel A, Alwakeel M, Jeribi F, et al. Deep learning-based diagnosis of Alzheimer's disease. *J Pers Med*. 2022;12(5):815. [FREE Full text] [doi: [10.3390/jpm12050815](https://doi.org/10.3390/jpm12050815)] [Medline: [35629237](https://pubmed.ncbi.nlm.nih.gov/35629237/)]
29. Alsubaie MG, Luo S, Shaikat K. Alzheimer's disease detection using deep learning on neuroimaging: a systematic review. *Mach Learn Knowl Extr*. 2024;6(1):464-505. [doi: [10.3390/make6010024](https://doi.org/10.3390/make6010024)]
30. Basaia S, Agosta F, Wagner L, Canu E, Magnani G, Santangelo R, et al. Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *Neuroimage Clin*. 2019;21:101645. [FREE Full text] [doi: [10.1016/j.nicl.2018.101645](https://doi.org/10.1016/j.nicl.2018.101645)] [Medline: [30584016](https://pubmed.ncbi.nlm.nih.gov/30584016/)]
31. Liu J, Li M, Luo Y, Yang S, Li W, Bi Y. Alzheimer's disease detection using depthwise separable convolutional neural networks. *Comput Methods Programs Biomed*. 2021;203:106032. [doi: [10.1016/j.cmpb.2021.106032](https://doi.org/10.1016/j.cmpb.2021.106032)] [Medline: [33713959](https://pubmed.ncbi.nlm.nih.gov/33713959/)]
32. Abdelwahab MM, Al-Karawi KA, Semary HE. Deep learning-based prediction of Alzheimer's disease using microarray gene expression data. *Biomedicines*. 2023;11(12):3304. [FREE Full text] [doi: [10.3390/biomedicines11123304](https://doi.org/10.3390/biomedicines11123304)] [Medline: [38137524](https://pubmed.ncbi.nlm.nih.gov/38137524/)]
33. Ebrahimi A, Luo S, Alzheimer's Disease Neuroimaging Initiative. Convolutional neural networks for Alzheimer's disease detection on MRI images. *J Med Imaging*. 2021;8(02):024503. [doi: [10.1117/1.jmi.8.2.024503](https://doi.org/10.1117/1.jmi.8.2.024503)]
34. Liu M, Cheng D, Yan W. Classification of Alzheimer's disease by combination of convolutional and recurrent neural networks using FDG-PET images. *Front Neuroinform*. 2018;12:35. [FREE Full text] [doi: [10.3389/fninf.2018.00035](https://doi.org/10.3389/fninf.2018.00035)] [Medline: [29970996](https://pubmed.ncbi.nlm.nih.gov/29970996/)]
35. Lien WC, Yeh CH, Chang CY, Chang C, Wang W, Chen C, et al. Convolutional neural networks to classify Alzheimer's disease severity based on SPECT images: a comparative study. *J Clin Med*. 2023;12(6):2218. [FREE Full text] [doi: [10.3390/jcm12062218](https://doi.org/10.3390/jcm12062218)] [Medline: [36983226](https://pubmed.ncbi.nlm.nih.gov/36983226/)]
36. Guluwadi S, Mohamed MM, Mahesh TR, Vinoth KV. Enhancing brain tumor detection in MRI images through explainable AI using Grad-CAM with Resnet 50. *BMC Med Imaging*. 2024;24(1):107. [FREE Full text] [doi: [10.1186/s12880-024-01292-7](https://doi.org/10.1186/s12880-024-01292-7)] [Medline: [38734629](https://pubmed.ncbi.nlm.nih.gov/38734629/)]
37. Yang C, Rangarajan A, Ranka S. Visual explanations from deep 3D convolutional neural networks for Alzheimer's disease classification. *AMIA Annu Symp Proc*. 2018;2018:1571-1580. [Medline: [30815203](https://pubmed.ncbi.nlm.nih.gov/30815203/)]
38. Cui R, Liu M, Alzheimer's Disease Neuroimaging Initiative. RNN-based longitudinal analysis for diagnosis of Alzheimer's disease. *Comput Med Imaging Graph*. 2019;73:1-10. [doi: [10.1016/j.compmedimag.2019.01.005](https://doi.org/10.1016/j.compmedimag.2019.01.005)] [Medline: [30763637](https://pubmed.ncbi.nlm.nih.gov/30763637/)]
39. Giorgio J, Landau SM, Jagust WJ, Tino P, Kourtzi Z. Modelling prognostic trajectories of cognitive decline due to Alzheimer's disease. *Neuroimage Clin*. 2020;26:102199. [FREE Full text] [doi: [10.1016/j.nicl.2020.102199](https://doi.org/10.1016/j.nicl.2020.102199)] [Medline: [32106025](https://pubmed.ncbi.nlm.nih.gov/32106025/)]
40. Chen Y, Wang L, Ding B, Shi J, Wen T, Huang J, et al. Automated Alzheimer's disease classification using deep learning models with soft-NMS and improved ResNet50 integration. *J Radiat Res Appl Sci*. 2024;17(1):100782. [doi: [10.1016/j.jrras.2023.100782](https://doi.org/10.1016/j.jrras.2023.100782)]
41. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. [FREE Full text] [doi: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)] [Medline: [33782057](https://pubmed.ncbi.nlm.nih.gov/33782057/)]
42. Lo CKL, Mertz D, Loeb M. Newcastle-Ottawa Scale: comparing reviewers' to authors' assessments. *BMC Med Res Methodol*. 2014;14:45. [doi: [10.1186/1471-2288-14-45](https://doi.org/10.1186/1471-2288-14-45)] [Medline: [24690082](https://pubmed.ncbi.nlm.nih.gov/24690082/)]
43. Lin L, Chu H. Quantifying publication bias in meta-analysis. *Biometrics*. 2018;74(3):785-794. [FREE Full text] [doi: [10.1111/biom.12817](https://doi.org/10.1111/biom.12817)] [Medline: [29141096](https://pubmed.ncbi.nlm.nih.gov/29141096/)]
44. Gjerdevik M, Heuch I. Improving the error rates of the Begg and Mazumdar test for publication bias in fixed effects meta-analysis. *BMC Med Res Methodol*. 2014;14(1):109. [doi: [10.1186/1471-2288-14-109](https://doi.org/10.1186/1471-2288-14-109)] [Medline: [25245217](https://pubmed.ncbi.nlm.nih.gov/25245217/)]
45. Data Analysis in the Geosciences. URL: <http://stratigrafia.org/8370/rtips/proportions.html> [accessed 2024-03-27]
46. Bollen KA, Brand JE. A general panel model with random and fixed effects: a structural equations approach. *Soc Forces*. 2010;89(1):1-34. [FREE Full text] [doi: [10.1353/sof.2010.0072](https://doi.org/10.1353/sof.2010.0072)] [Medline: [21769157](https://pubmed.ncbi.nlm.nih.gov/21769157/)]
47. Nusrang M, Annas S, Heinonen J, Clarke J. Performa restricted maximum likelihood and maximum likelihood estimators on small area estimation. *J Phys Conf Ser*. 2018;1028:012234. [FREE Full text] [doi: [10.1088/1742-6596/1028/1/012234](https://doi.org/10.1088/1742-6596/1028/1/012234)]
48. von Hippel PT. The heterogeneity statistic I(2) can be biased in small meta-analyses. *BMC Med Res Methodol*. 2015;15(1):35. [FREE Full text] [doi: [10.1186/s12874-015-0024-z](https://doi.org/10.1186/s12874-015-0024-z)] [Medline: [25880989](https://pubmed.ncbi.nlm.nih.gov/25880989/)]
49. Thorlund K, Imberger G, Johnston BC, Walsh M, Awad T, Thabane L, et al. Evolution of heterogeneity (I2) estimates and their 95% confidence intervals in large meta-analyses. *PLoS One*. 2012;7(7):e39471. [FREE Full text] [doi: [10.1371/journal.pone.0039471](https://doi.org/10.1371/journal.pone.0039471)] [Medline: [22848355](https://pubmed.ncbi.nlm.nih.gov/22848355/)]
50. Aslam M. Cochran's Q test for analyzing categorical data under uncertainty. *J Big Data*. 2023;10(1):1-10. [doi: [10.1186/s40537-023-00823-3](https://doi.org/10.1186/s40537-023-00823-3)]

51. Barendregt JJ, Doi SA, Lee YY, Norman RE, Vos T. Meta-analysis of prevalence. *J Epidemiol Community Health*. 2013;67(11):974-978. [doi: [10.1136/jech-2013-203104](https://doi.org/10.1136/jech-2013-203104)] [Medline: [23963506](https://pubmed.ncbi.nlm.nih.gov/23963506/)]
52. Alorf A, Khan MUG. Multi-label classification of Alzheimer's disease stages from resting-state fMRI-based correlation connectivity data and deep learning. *Comput Biol Med*. 2022;151(Pt A):106240. [FREE Full text] [doi: [10.1016/j.compbiomed.2022.106240](https://doi.org/10.1016/j.compbiomed.2022.106240)] [Medline: [36423532](https://pubmed.ncbi.nlm.nih.gov/36423532/)]
53. Chen Y, Sha M, Zhao X, Ma J, Ni H, Gao W, et al. Automated detection of pathologic white matter alterations in Alzheimer's disease using combined diffusivity and kurtosis method. *Psychiatry Res Neuroimaging*. 2017;264:35-45. [FREE Full text] [doi: [10.1016/j.psychres.2017.04.004](https://doi.org/10.1016/j.psychres.2017.04.004)] [Medline: [28448817](https://pubmed.ncbi.nlm.nih.gov/28448817/)]
54. Mofrad SA, Lundervold A, Lundervold AS. A predictive framework based on brain volume trajectories enabling early detection of Alzheimer's disease. *Comput Med Imaging Graph*. 2021;90:101910. [FREE Full text] [doi: [10.1016/j.compmedimag.2021.101910](https://doi.org/10.1016/j.compmedimag.2021.101910)] [Medline: [33862355](https://pubmed.ncbi.nlm.nih.gov/33862355/)]
55. EL-Geneedy M, Moustafa HED, Khalifa F, Khater H, Abdelhalim E. An MRI-based deep learning approach for accurate detection of Alzheimer's disease. *Alexandria Eng J*. 2023;63:211-221. [doi: [10.1016/j.aej.2022.07.062](https://doi.org/10.1016/j.aej.2022.07.062)]
56. Hazarika RA, Kandar D, Maji AK. An experimental analysis of different deep learning based models for Alzheimer's disease classification using brain magnetic resonance images. *J King Saud Univ - Comput Inf Sci*. 2022;34(10):8576-8598. [doi: [10.1016/j.jksuci.2021.09.003](https://doi.org/10.1016/j.jksuci.2021.09.003)]
57. Khan A, Zubair S. Development of a three tiered cognitive hybrid machine learning algorithm for effective diagnosis of Alzheimer's disease. *J King Saud Univ - Comput Inf Sci*. 2022;34(10):8000-8018. [doi: [10.1016/j.jksuci.2022.07.016](https://doi.org/10.1016/j.jksuci.2022.07.016)]
58. Sorour SE, El-Mageed AAA, Albarrak KM, Alnaim AK, Wafa AA, El-Shafeiy E. Classification of Alzheimer's disease using MRI data based on deep learning techniques. *J King Saud Univ - Comput Inf Sci*. 2024;36(2):101940. [doi: [10.1016/j.jksuci.2024.101940](https://doi.org/10.1016/j.jksuci.2024.101940)]
59. Abdelaziz M, Wang T, Elazab A. Alzheimer's disease diagnosis framework from incomplete multimodal data using convolutional neural networks. *J Biomed Inform*. 2021;121:103863. [FREE Full text] [doi: [10.1016/j.jbi.2021.103863](https://doi.org/10.1016/j.jbi.2021.103863)] [Medline: [34229061](https://pubmed.ncbi.nlm.nih.gov/34229061/)]
60. Sharma S, Guleria K, Tiwari S, Kumar S. A deep learning based convolutional neural network model with VGG16 feature extractor for the detection of Alzheimer disease using MRI scans. *Meas Sens*. 2022;24:100506. [doi: [10.1016/j.measen.2022.100506](https://doi.org/10.1016/j.measen.2022.100506)]
61. Nguyen D, Nguyen H, Ong H, Le H, Ha H, Duc NT, et al. Ensemble learning using traditional machine learning and deep neural network for diagnosis of Alzheimer's disease. *IBRO Neurosci Rep*. 2022;13:255-263. [FREE Full text] [doi: [10.1016/j.ibneur.2022.08.010](https://doi.org/10.1016/j.ibneur.2022.08.010)] [Medline: [36590098](https://pubmed.ncbi.nlm.nih.gov/36590098/)]
62. Saleh AW, Gupta G, Khan SB, Alkhalidi NA, Verma A. An Alzheimer's disease classification model using transfer learning Densenet with embedded healthcare decision support system. *Decis Anal J*. 2023;9:100348. [doi: [10.1016/j.dajour.2023.100348](https://doi.org/10.1016/j.dajour.2023.100348)]
63. Yang J, Sui H, Jiao R, Zhang M, Zhao X, Wang L, et al. Random-forest-algorithm-based applications of the basic characteristics and serum and imaging biomarkers to diagnose mild cognitive impairment. *Curr Alzheimer Res*. 2022;19(1):76-83. [FREE Full text] [doi: [10.2174/1567205019666220128120927](https://doi.org/10.2174/1567205019666220128120927)] [Medline: [35088670](https://pubmed.ncbi.nlm.nih.gov/35088670/)]
64. El-Sappagh S, Alonso JM, Islam SMR, Sultan AM, Kwak KS. A multilayer multimodal detection and prediction model based on explainable artificial intelligence for Alzheimer's disease. *Sci Rep*. 2021;11(1):2660. [FREE Full text] [doi: [10.1038/s41598-021-82098-3](https://doi.org/10.1038/s41598-021-82098-3)] [Medline: [33514817](https://pubmed.ncbi.nlm.nih.gov/33514817/)]
65. Liu S, Masurkar AV, Rusinek H, Chen J, Zhang B, Zhu W, et al. Generalizable deep learning model for early Alzheimer's disease detection from structural MRIs. *Sci Rep*. 2022;12(1):17106. [FREE Full text] [doi: [10.1038/s41598-022-20674-x](https://doi.org/10.1038/s41598-022-20674-x)] [Medline: [36253382](https://pubmed.ncbi.nlm.nih.gov/36253382/)]
66. Elgammal YM, Zahran MA, Abdelsalam MM. A new strategy for the early detection of Alzheimer disease stages using multifractal geometry analysis based on K-Nearest Neighbor algorithm. *Sci Rep*. 2022;12(1):22381. [FREE Full text] [doi: [10.1038/s41598-022-26958-6](https://doi.org/10.1038/s41598-022-26958-6)] [Medline: [36572791](https://pubmed.ncbi.nlm.nih.gov/36572791/)]
67. Das S, Panigrahi P, Chakrabarti S. Corpus callosum atrophy in detection of mild and moderate Alzheimer's disease using brain magnetic resonance image processing and machine learning techniques. *J Alzheimers Dis Rep*. 2021;5(1):771-788. [doi: [10.3233/adr-210314](https://doi.org/10.3233/adr-210314)]
68. Chelladurai A, Narayan DL, Divakarachari PB, Loganathan U. fMRI-based Alzheimer's disease detection using the SAS method with multi-layer perceptron network. *Brain Sci*. 2023;13(6):893. [FREE Full text] [doi: [10.3390/brainsci13060893](https://doi.org/10.3390/brainsci13060893)] [Medline: [37371371](https://pubmed.ncbi.nlm.nih.gov/37371371/)]
69. Battineni G, Hossain MA, Chintalapudi N, Traini E, Dhulipalla VR, Ramasamy M, et al. Improved Alzheimer's disease detection by MRI using multimodal machine learning algorithms. *Diagnostics (Basel)*. 2021;11(11):2103. [FREE Full text] [doi: [10.3390/diagnostics11112103](https://doi.org/10.3390/diagnostics11112103)] [Medline: [34829450](https://pubmed.ncbi.nlm.nih.gov/34829450/)]
70. Sharma S, Gupta S, Gupta D, Altameem A, Saudagar AKJ, Poonia RC, et al. HTLML: hybrid AI based model for detection of Alzheimer's disease. *Diagnostics (Basel)*. 2022;12(8):1833. [doi: [10.3390/diagnostics12081833](https://doi.org/10.3390/diagnostics12081833)] [Medline: [36010183](https://pubmed.ncbi.nlm.nih.gov/36010183/)]
71. Long Z, Li J, Fan J, Li B, Du Y, Qiu S, et al. Identifying Alzheimer's disease and mild cognitive impairment with atlas-based multi-modal metrics. *Front Aging Neurosci*. 2023;15:1212275. [FREE Full text] [doi: [10.3389/fnagi.2023.1212275](https://doi.org/10.3389/fnagi.2023.1212275)] [Medline: [37719872](https://pubmed.ncbi.nlm.nih.gov/37719872/)]

72. Wang R, He Q, Han C, Wang H, Shi L, Che Y. A deep learning framework for identifying Alzheimer's disease using fMRI-based brain network. *Front Neurosci.* 2023;17:1177424. [FREE Full text] [doi: [10.3389/fnins.2023.1177424](https://doi.org/10.3389/fnins.2023.1177424)] [Medline: [37614342](https://pubmed.ncbi.nlm.nih.gov/37614342/)]
73. Tajammal T, Khurshid SK, Jaleel A, Qayyum Wahla S, Ziar RA. Deep learning-based ensembling technique to classify Alzheimer's disease stages using functional MRI. *J Healthc Eng.* 2023;2023:6961346. [FREE Full text] [doi: [10.1155/2023/6961346](https://doi.org/10.1155/2023/6961346)] [Medline: [37953911](https://pubmed.ncbi.nlm.nih.gov/37953911/)]
74. Golovanevsky M, Eickhoff C, Singh R. Multimodal attention-based deep learning for Alzheimer's disease diagnosis. *J Am Med Inform Assoc.* 2022;29(12):2014-2022. [FREE Full text] [doi: [10.1093/jamia/ocac168](https://doi.org/10.1093/jamia/ocac168)] [Medline: [36149257](https://pubmed.ncbi.nlm.nih.gov/36149257/)]
75. Li Q, Yang MQ. Comparison of machine learning approaches for enhancing Alzheimer's disease classification. *PeerJ.* 2021;9:e10549. [doi: [10.7717/peerj.10549](https://doi.org/10.7717/peerj.10549)] [Medline: [33665002](https://pubmed.ncbi.nlm.nih.gov/33665002/)]
76. About ADNI. Alzheimer's Disease Neuroimaging Initiative. URL: <https://adni.loni.usc.edu/about/> [accessed 2024-03-15]
77. Marcus DS, Wang TH, Parker J, Csernansky JG, Morris JC, Buckner RL. Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. *J Cogn Neurosci.* 2007;19(9):1498-1507. [doi: [10.1162/jocn.2007.19.9.1498](https://doi.org/10.1162/jocn.2007.19.9.1498)] [Medline: [17714011](https://pubmed.ncbi.nlm.nih.gov/17714011/)]
78. The AIBL Study. URL: <https://aibl.org.au/> [accessed 2024-03-15]
79. Alzheimer's MRI brain scan images (augmented). Kaggle. URL: <https://www.kaggle.com/datasets/vencerlanz09/alzheimers-mri-brain-scan-images-augmented> [accessed 2024-03-15]
80. Crystal O, Maralani PJ, Black S, Fischer C, Moody AR, Khademi A. Detecting conversion from mild cognitive impairment to Alzheimer's disease using FLAIR MRI biomarkers. *Neuroimage Clin.* 2023;40:103533. [FREE Full text] [doi: [10.1016/j.nicl.2023.103533](https://doi.org/10.1016/j.nicl.2023.103533)] [Medline: [37952286](https://pubmed.ncbi.nlm.nih.gov/37952286/)]
81. Opfer R, Suppa P, Kepp T, Spies L, Schippling S, Huppertz H. Atlas based brain volumetry: how to distinguish regional volume changes due to biological or physiological effects from inherent noise of the methodology. *Magn Reson Imaging.* 2016;34(4):455-461. [doi: [10.1016/j.mri.2015.12.031](https://doi.org/10.1016/j.mri.2015.12.031)] [Medline: [26723849](https://pubmed.ncbi.nlm.nih.gov/26723849/)]
82. Infanger D, Schmidt-Trucksäss A. P value functions: an underused method to present research results and to promote quantitative reasoning. *Stat Med.* 2019;38(21):4189-4197. [doi: [10.1002/sim.8293](https://doi.org/10.1002/sim.8293)] [Medline: [31270842](https://pubmed.ncbi.nlm.nih.gov/31270842/)]
83. Pettigrew C, Soldan A, Zhu Y, Wang M, Brown T, Miller M, et al. Cognitive reserve and cortical thickness in preclinical Alzheimer's disease. *Brain Imaging Behav.* 2017;11(2):357-367. [FREE Full text] [doi: [10.1007/s11682-016-9581-y](https://doi.org/10.1007/s11682-016-9581-y)] [Medline: [27544202](https://pubmed.ncbi.nlm.nih.gov/27544202/)]
84. Li A, Li F, Elahifasae F, Liu M, Zhang L. Hippocampal shape and asymmetry analysis by cascaded convolutional neural networks for Alzheimer's disease diagnosis. *Brain Imaging Behav.* 2021;15(5):2330-2339. [doi: [10.1007/s11682-020-00427-y](https://doi.org/10.1007/s11682-020-00427-y)] [Medline: [33398778](https://pubmed.ncbi.nlm.nih.gov/33398778/)]
85. Horikawa T, Tamaki M, Miyawaki Y, Kamitani Y. Neural decoding of visual imagery during sleep. *Science.* 2013;340(6132):639-642. [doi: [10.1126/science.1234330](https://doi.org/10.1126/science.1234330)] [Medline: [23558170](https://pubmed.ncbi.nlm.nih.gov/23558170/)]
86. Zou M, Zhou Q, Li R, Hu M, Qian L, Yang Z, et al. Image quality using synthetic brain MRI: an age-stratified study. *Acta Radiol.* 2023;64(5):2010-2023. [doi: [10.1177/02841851231152098](https://doi.org/10.1177/02841851231152098)] [Medline: [36775871](https://pubmed.ncbi.nlm.nih.gov/36775871/)]
87. Arafa DA, Moustafa HED, Ali HA, Ali-Eldin AMT, Saraya SF. A deep learning framework for early diagnosis of Alzheimer's disease on MRI images. *Multimed Tools Appl.* 2023;83(2):3767-3799. [doi: [10.1007/s11042-023-15738-7](https://doi.org/10.1007/s11042-023-15738-7)]
88. Amini M, Pedram MM, Moradi A, Jamshidi M, Ouchani M. GC-CNNnet: diagnosis of Alzheimer's disease with PET images using genetic and convolutional neural network. *Comput Intell Neurosci.* 2022;2022:7413081. [FREE Full text] [doi: [10.1155/2022/7413081](https://doi.org/10.1155/2022/7413081)] [Medline: [35983158](https://pubmed.ncbi.nlm.nih.gov/35983158/)]
89. Salvi M, Loh HW, Seoni S, Barua PD, García S, Molinari F, et al. Multi-modality approaches for medical support systems: a systematic review of the last decade. *Inf Fusion.* 2024;103:102134. [doi: [10.1016/j.inffus.2023.102134](https://doi.org/10.1016/j.inffus.2023.102134)]
90. Catalino MP, Yao S, Green DL, Laws ER, Golby AJ, Tie Y. Mapping cognitive and emotional networks in neurosurgical patients using resting-state functional magnetic resonance imaging. *Neurosurg Focus.* 2020;48(2):E9. [FREE Full text] [doi: [10.3171/2019.11.FOCUS19773](https://doi.org/10.3171/2019.11.FOCUS19773)] [Medline: [32006946](https://pubmed.ncbi.nlm.nih.gov/32006946/)]
91. Frost MA, Esposito F, Goebel R. Improved correspondence of resting-state networks after macroanatomical alignment. *Hum Brain Mapp.* 2014;35(2):673-682. [FREE Full text] [doi: [10.1002/hbm.22191](https://doi.org/10.1002/hbm.22191)] [Medline: [23161519](https://pubmed.ncbi.nlm.nih.gov/23161519/)]
92. Tang X, Qin Y, Zhu W, Michael MMH. Surface-based vertexwise analysis of morphometry and microstructural integrity for white matter tracts in diffusion tensor imaging: with application to the corpus callosum in Alzheimer's disease. *Hum Brain Mapp.* 2017;38(4):1875-1893. [FREE Full text] [doi: [10.1002/hbm.23491](https://doi.org/10.1002/hbm.23491)] [Medline: [28083895](https://pubmed.ncbi.nlm.nih.gov/28083895/)]
93. Katabathula S, Wang Q, Xu R. Predict Alzheimer's disease using hippocampus MRI data: a lightweight 3D deep convolutional network model with visual and global shape representations. *Alzheimers Res Ther.* 2021;13(1):104. [FREE Full text] [doi: [10.1186/s13195-021-00837-0](https://doi.org/10.1186/s13195-021-00837-0)] [Medline: [34030743](https://pubmed.ncbi.nlm.nih.gov/34030743/)]
94. Shaikh T, Ali R. Automated atrophy assessment for Alzheimer's disease diagnosis from brain MRI images. *Magn Reson Imaging.* 2019;62:167-173. [doi: [10.1016/j.mri.2019.06.019](https://doi.org/10.1016/j.mri.2019.06.019)] [Medline: [31279772](https://pubmed.ncbi.nlm.nih.gov/31279772/)]
95. Zhang T, Wu S, Zhang X, Dai Y, Wang A, Zhang H, et al. Spatial normalization and quantification approaches of PET imaging for neurological disorders. *Eur J Nucl Med Mol Imaging.* 2022;49(11):3809-3829. [doi: [10.1007/s00259-022-05809-6](https://doi.org/10.1007/s00259-022-05809-6)] [Medline: [35624219](https://pubmed.ncbi.nlm.nih.gov/35624219/)]

96. Ashburner J, Friston KJ. Unified segmentation. *Neuroimage*. 2005;26(3):839-851. [doi: [10.1016/j.neuroimage.2005.02.018](https://doi.org/10.1016/j.neuroimage.2005.02.018)] [Medline: [15955494](#)]
97. Shojaie M, Cabrerizo M, DeKosky ST, Vaillancourt DE, Loewenstein D, Duara R, et al. A transfer learning approach based on gradient boosting machine for diagnosis of Alzheimer's disease. *Front Aging Neurosci*. 2022;14:966883. [doi: [10.3389/fnagi.2022.966883](https://doi.org/10.3389/fnagi.2022.966883)] [Medline: [36275004](#)]
98. Atik C, Kut RA, Yilmaz R, Birant D. Support vector machine chains with a novel tournament voting. *Electronics*. 2023;12(11):2485. [doi: [10.3390/electronics12112485](https://doi.org/10.3390/electronics12112485)]
99. Oliveira M, Wilming R, Clark B, Budding C, Eitel F, Ritter K, et al. Benchmarking the influence of pre-training on explanation performance in MR image classification. *Front Artif Intell*. 2024;7:1330919. [FREE Full text] [doi: [10.3389/frai.2024.1330919](https://doi.org/10.3389/frai.2024.1330919)] [Medline: [38469161](#)]
100. Amaya-Tejera N, Gamarra M, Vélez JI, Zurek E. A distance-based kernel for classification via Support Vector Machines. *Front Artif Intell*. 2024;7:1287875. [FREE Full text] [doi: [10.3389/frai.2024.1287875](https://doi.org/10.3389/frai.2024.1287875)] [Medline: [38469159](#)]
101. Javeed A, Dallora AL, Berglund JS, Ali A, Ali L, Anderberg P. Machine learning for dementia prediction: a systematic review and future research directions. *J Med Syst*. 2023;47(1):17. [FREE Full text] [doi: [10.1007/s10916-023-01906-7](https://doi.org/10.1007/s10916-023-01906-7)] [Medline: [36720727](#)]
102. Arya AD, Verma SS, Chakarabarti P, Chakrabarti T, Elngar AA, Kamali AM, et al. A systematic review on machine learning and deep learning techniques in the effective diagnosis of Alzheimer's disease. *Brain Inform*. Jul 14, 2023;10(1):17. [FREE Full text] [doi: [10.1186/s40708-023-00195-7](https://doi.org/10.1186/s40708-023-00195-7)] [Medline: [37450224](#)]
103. Odusami M, Maskeliūnas R, Damaševičius R, Misra S. Machine learning with multimodal neuroimaging data to classify stages of Alzheimer's disease: a systematic review and meta-analysis. *Cogn Neurodyn*. 2024;18(3):775-794. [FREE Full text] [doi: [10.1007/s11571-023-09993-5](https://doi.org/10.1007/s11571-023-09993-5)] [Medline: [38826669](#)]

Abbreviations

AD: Alzheimer disease
ADNI: Alzheimer's Disease Neuroimaging Initiative
CN: cognitively normal
CNN: convolutional neural network
CSF: cerebrospinal fluid
CT: computerized tomography
DenseNet: dense convolutional network
DKI: diffusion kurtosis imaging
DL: deep learning
EMCI: early mild cognitive impairment
KNN: k-nearest neighbor
LMCI: late mild cognitive impairment
LR: logistic regression
MCI: mild cognitive impairment
MD: mildly demented
ML: machine learning
MoD: moderately demented
MRI: magnetic resonance imaging
ND: nondemented
OASIS: Open Access Series of Imaging Studies
PET: positron emission tomography
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis
REM: random effects model
SMC: significant memory concern
SVM: support vector machine
VGG: Visual Geometry Group
XGB: extreme gradient boosting

Edited by J Wang; submitted 10.04.24; peer-reviewed by U Snehalatha, Z Wu; comments to author 24.05.24; revised version received 12.06.24; accepted 25.09.24; published 23.12.24

Please cite as:

Battineni G, Chintalapudi N, Amenta F

Machine Learning Driven by Magnetic Resonance Imaging for the Classification of Alzheimer Disease Progression: Systematic Review and Meta-Analysis

JMIR Aging 2024;7:e59370

URL: <https://aging.jmir.org/2024/1/e59370>

doi: [10.2196/59370](https://doi.org/10.2196/59370)

PMID:

©Gopi Battineni, Nalini Chintalapudi, Francesco Amenta. Originally published in JMIR Aging (<https://aging.jmir.org>), 23.12.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on <https://aging.jmir.org>, as well as this copyright and license information must be included.