

REVIEW

Open Access



# Systematic review: fluid biomarkers and machine learning methods to improve the diagnosis from mild cognitive impairment to Alzheimer's disease

Kevin Blanco<sup>1</sup>, Stefanny Salcidua<sup>2,3</sup>, Paulina Orellana<sup>1,2</sup>, Tania Sauma-Pérez<sup>2</sup>, Tomás León<sup>4,5</sup>, Lorena Cecilia López Steinmetz<sup>2,6,7</sup>, Agustín Ibañez<sup>2,4,8,9</sup>, Claudia Duran-Aniotz<sup>1,2\*</sup> and Rolando de la Cruz<sup>2,3,10\*</sup>

## Abstract

Mild cognitive impairment (MCI) is often considered an early stage of dementia, with estimated rates of progression to dementia up to 80–90% after approximately 6 years from the initial diagnosis. Diagnosis of cognitive impairment in dementia is typically based on clinical evaluation, neuropsychological assessments, cerebrospinal fluid (CSF) biomarkers, and neuroimaging. The main goal of diagnosing MCI is to determine its cause, particularly whether it is due to Alzheimer's disease (AD). However, only a limited percentage of the population has access to etiological confirmation, which has led to the emergence of peripheral fluid biomarkers as a diagnostic tool for dementias, including MCI due to AD. Recent advances in biofluid assays have enabled the use of sophisticated statistical models and multimodal machine learning (ML) algorithms for the diagnosis of MCI based on fluid biomarkers from CSF, peripheral blood, and saliva, among others. This approach has shown promise for identifying specific causes of MCI, including AD. After a PRISMA analysis, 29 articles revealed a trend towards using multimodal algorithms that incorporate additional biomarkers such as neuroimaging, neuropsychological tests, and genetic information. Particularly, neuroimaging is commonly used in conjunction with fluid biomarkers for both cross-sectional and longitudinal studies. Our systematic review suggests that cost-effective longitudinal multimodal monitoring data, representative of diverse cultural populations and utilizing white-box ML algorithms, could be a valuable contribution to the development of diagnostic models for AD due to MCI. Clinical assessment and biomarkers, together with ML techniques, could prove pivotal in improving diagnostic tools for MCI due to AD.

**Keywords** Mild cognitive impairment, Alzheimer's disease, Fluid biomarker, Machine learning, Artificial intelligence

\*Correspondence:

Claudia Duran-Aniotz

Claudia.Duran@uai.cl

Rolando de la Cruz

rolando.delacruz@uai.cl

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Introduction

Mild cognitive impairment (MCI) is defined as a heterogeneous clinical syndrome including cognitive impairments of any cognitive function while maintaining independence [1]. The prevalence rate of MCI ranges from 6% in the population over 60 years of age [2] up to 25% for ages 80–84 [3]. Importantly, MCI is often considered a prodromal stage of dementia, especially considering that neuropathological changes of dementia may develop many years before the diagnosis, presenting both cognitive and behavioral symptoms previously the patients lose their independence [4]. The rates of progression of MCI due to Alzheimer's disease (AD) to dementia have been estimated at between 8 and 15% [5, 6], which increases up to 80–90% after approximately 6 years [7–12]. Timely diagnosis of MCI is essential for identifying patients who are likely to progress to dementia and implementing early interventions to delay the pathological progression. Lifestyle interventions for MCI in its early stages may help to delay the onset of dementia [13]. Therefore, improving the assessment of MCI, including the incorporation of biomarkers in the usual clinical diagnostic procedures, could be critical for developing better diagnostic tools for MCI, particularly in determining its cause. Currently, much of the research on MCI has focused on MCI due to AD. Thus, incorporating biomarkers into clinical diagnosis procedures may help identify patients who are likely to develop AD and facilitate earlier intervention.

### Biomarkers in MCI

Biomarkers are objective measures that evaluate normal biological processes, pathological processes, or pharmacological responses to therapeutic interventions [14]. In the field of AD, neuroimaging and cerebrospinal fluid (CSF) protein analysis are the most widely used biomarkers [15]. Neuroimaging biomarkers include magnetic resonance imaging (MRI) to evaluate brain atrophy [16–18], positron emission tomography (PET) with F18-fluorodeoxyglucose (FDG) to measure glucose metabolism in different regions of the brain showing neuronal loss and neurodegeneration [19–22], and PET with tracers to detect amyloid-beta ( $A\beta$ ) and Tau proteins in vivo [23–25]. CSF protein analysis of  $A\beta$  and Tau in their total and phosphorylated forms is also a validated biomarker for AD [26]. These biomarkers play a crucial role in the early and accurate diagnosis of AD, enabling earlier intervention and improving patient outcomes.

### Fluids biomarkers

Peripheral biomarkers are of great interest because they are less invasive, less expensive, and more accessible. These types of biomarkers include molecules such

as proteins, peptides, nucleic acids, microRNAs (miRNAs), lipids, and metabolites which can be detected in several biological fluids such as plasma, serum, urine, saliva, exosomes, or cellular components [27, 28].  $A\beta$  proteins in their 42 amino acid form, which form amyloid plaques, total tau (T-Tau), which reflects the intensity of neurodegeneration; and phosphorylated tau (p-Tau), which correlates with the production of neurofibrillary tangles, are measured in CSF, and they are currently validated as diagnosis support in AD [15]. Despite the advances in biomarker research, no validated biomarkers are available for the diagnosis of MCI due to AD. Some studies showed lower concentrations of  $A\beta$  ( $A\beta$ 1–40,  $A\beta$ 1–42, and  $A\beta$ 1–42/ $A\beta$ 1–40 ratio) in CSF of MCI patients compared to healthy controls, reflecting higher brain  $A\beta$  concentrations and progressive cognitive impairment [29–34]. Furthermore, the detection of total and phosphorylated Tau (p-Tau) in CSF samples has also been used to diagnose MCI and AD with at least 85% sensitivity and 80% specificity, indicating neuronal damage and predicting progression from MCI to AD [35]. The combination of  $A\beta$ 1–42 and Tau has demonstrated a high sensitivity of 95% and specificity of 83% in predicting the progression of MCI to AD [36].

Although blood levels of  $A\beta$  and tau have been evaluated as potential biomarkers for cognitive impairment, their concentrations are lower in blood compared to CSF, making their detection more challenging. Additionally, studies investigating  $A\beta$ 42 and Tau levels in subjects with cognitive impairment have produced inconsistent results [37].

Recent research has focused on neurofilament light chain (NfL) as a potential biomarker for neurodegenerative diseases, including AD [38–43]. Studies have shown that plasma levels of NfL are significantly higher in patients with AD and MCI compared to controls [44] and are associated with cognitive and neuroimaging features [45–47]. Another potential blood biomarker is microRNAs (miRNAs), which play a role in regulating gene expression in the brain [48, 49]. However, validation studies are needed before miRNAs can be used clinically as a biomarker for MCI [50].

### Genetics in MCI due to AD

Autosomal dominantly inherited forms of AD are present early in life; however, most early cases do not show a clear pattern of inheritance (2–10%); however, genetic predisposition to non-Mendelian inheritance of AD is high, with an estimated heritability of 80% [51]. Three genes including amyloid precursor protein (APP), presenilin 1 (PSEN1), and PSEN2 with fully penetrant mutations have been discovered as a cause of autosomal dominant AD, accounting for 5–10% of the occurrence of early AD. In

addition to the above, the  $\epsilon 4$  allele of the apolipoprotein E (ApoE) gene was identified as a strong risk factor for both early- and late-onset AD, where heterozygous carriers of the  $\epsilon 4$  allele have an estimated threefold risk of developing AD and 15-fold risk in homozygous carriers of this allele [52–57]. In addition, at least 21 genetic risk loci have been identified in genome-wide association studies (GWAS) and mass sequencing that demonstrate how complex and multifactorial AD is in genetic terms [58, 59].

### Machine learning for medical diagnosis

Machine learning (ML) algorithms have been proposed as a promising tool to integrate multiple biomarkers for early detection, diagnosis, and prediction of dementia. ML models can analyze large amounts of data and identify complex patterns that may not be visible to human experts. Furthermore, ML algorithms can integrate data from different sources such as neuroimaging, genetics, and clinical data to develop models that can accurately predict the onset and progression of dementia [60–62]. Studies have shown that ML algorithms can improve the accuracy of dementia diagnosis and prediction compared to traditional methods based on single biomarkers [63].

ML can also help clinicians develop personalized treatment plans based on the individual patient's biomarker profile and disease stage. By analyzing patterns in data from imaging, genetic, and biomarker assays, ML algorithms can identify the best treatment options and predict the effectiveness of specific interventions for individual patients. This can help improve the accuracy and effectiveness of treatment, potentially leading to better outcomes and improved quality of life for patients with Alzheimer's disease and other forms of dementia [64].

In general terms, ML algorithms allow robust enquiries on many datasets to find patterns and relationships among the data [65]. There are some variations of how to define the types of ML algorithms but commonly they can be divided into categories according to their purpose and the main categories are the following: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning [66]. Utilizing ML in large-scale data analysis, and taking into account the Four V's of big data: volume, velocity, variety, and veracity [67, 68], is revolutionizing the production of scientific knowledge, by enabling novel and highly efficient ways of designing and evaluating research [69]. It is important to point out that this efficiency is given by the optimization of the use of resources to collect massive quality data in medical and healthcare contexts [70]. To the extent that the processes involved in the massive generation of scientific data are strategically optimized, as a consequence, ML analyzes and models acquire greater versatility and

are potentially more scalable in their development and continuous improvement [71]. In that sense, one of the interests of developing ML research strategies using fluid biomarkers data is related to exploring ways of adjusting optimization gaps of cost-efficiency in the use of resources to improve the diagnosis from MCI to AD on a large scale.

Without a doubt, ML offers interesting new data research opportunities. However, in the medical community, there is great concern about the development of medical applications based on ML models [72, 73]. This is because as ML algorithms become more advanced, it is more challenging to comprehend and retrace how the algorithm came to a result, which translates into a trust issue due to the lack of explainability that these models have [74]. The whole calculation process used by an ML algorithm is turned into what is commonly referred to as a “black box” that is impossible to interpret. These black box models are created directly from the data, and not even the researchers who create the algorithm can understand or explain what exactly is happening inside them or how the ML algorithm arrived at a specific result [75]. Many of the ML algorithms cannot explain how and why they have issued a given answer or decision [76]. This occurs mainly in modeling approaches based on neural networks (one of the most popular in use) [77]. In the given context, explainable artificial intelligence (XAI) is a rapidly growing research area within the realm of machine learning. It focuses on uncovering the ways in which these algorithms make decisions that are considered “black box,” by examining the measurements and rules at play and assisting in making the modeling process more self-explanatory. In that sense, the XAI becomes increasingly crucial for machine learning-driven applications, especially in medical diagnosis [78]. Essentially, for the successful development of machine learning-based applications to improve the diagnosis from MCI to AD either from fluid biomarkers or multimodal data, it is necessary that the explainability of the model be clear and consistent from all possible perspectives, in coherence with its theoretical and experimental framework [79].

To fully realize the potential of ML in this context, the purpose of this research is to conduct a systematic review of existing studies on the use of machine learning and fluid biomarkers in dementia research. This review aims to identify gaps in our understanding of the relationships between different biomarkers, highlight areas where additional research is needed, and provide guidance for the design of future studies in this field. The combination of machine learning and fluid biomarkers research holds enormous promise for advancing our understanding of dementia pathobiology, and a systematic review of the

existing literature is a crucial step towards realizing this potential.

**Materials and methods**

For this research, we followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist and statement.

**Identification of studies**

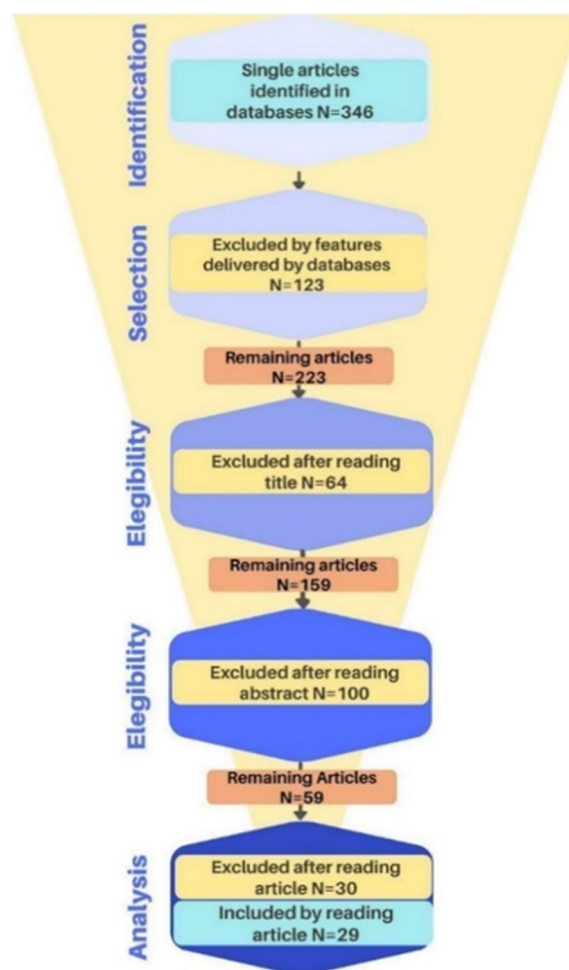
We searched PubMed Central and Scopus databases. PubMed Central is the digital archive of the United States National Institutes of Health which was selected for its wide scope and its relevance in the biomedical and life sciences. Scopus is a wide database containing peer-reviewed abstracts and citations. We performed a sensitive literature search by using specific keywords that were defined in four categories (including synonyms and related words) that allow to maintain an accurate search: (1) MCI—mild cognitive impairment and MCI; (2) Diagnosis—diagnosis and diagnostic; (3) Machine learning—machine learning, artificial intelligence, algorithm, and deep learning; and (4) Fluid biomarkers—Tau, cerebrospinal fluid, amyloid-beta, fluid biomarker, miRNA, microRNA, blood, serum, plasma, urine, saliva, progranulin, and neurofilament. The search queries were carried out defining that at least one of the keywords of each category considered appeared in the title, abstract, or keywords of the article. We focused on articles published between January 1, 2012, and January 1, 2023, to base our analysis on recent studies.

**Selection of studies**

All abstracts were screened by co-authors. The full text of selected abstracts was assessed for eligibility by co-authors, and conflicts were resolved by consensus. Firstly, a search was carried out for the selected keywords using Boolean operators in the two aforementioned databases. Secondly, the articles were checked for the inclusion and exclusion criteria by features found in the databases. After that, we read the titles of all the remaining articles to check if the articles were within the scope of our study and considered at least one biomarker other than neuroimages (fluid, genetic, or clinical data). In the same way, we proceeded to read the abstracts of the remaining articles. Finally, after reading these, we selected the articles that were included in this review. The flowchart in Fig. 1 shows the sequence of actions and the outcomes.

**Inclusion and exclusion criteria**

Studies were eligible if (i) the article described empirical, quantitative, longitudinal studies; follow-up studies; neuroimaging studies; randomized controlled trials;



**Fig. 1** Systematic literature search flow diagram (PRISMA). This diagram starts with the total number of records identified through database searching. From there, the diagram outlines the number of records screened. Then, it indicates the number of records excluded after the initial screening, typically because the titles or abstracts clearly indicate that the studies do not meet the inclusion criteria. Next, the diagram shows the number of full-text articles assessed for eligibility, followed by the number of full-text articles excluded and the reasons for their exclusion. Finally, the diagram presents the number of studies included. This process makes the selection process transparent, which is crucial for the credibility of the systematic review

quasi-randomized controlled trials; and cross-sectional studies, based on human populations all over the world, and (ii) considering samples of mild cognitive impairment in conjunction with machine learning and fluid biomarkers, in which the algorithm has validation.

Studies were excluded if they were (i) review articles, conference abstracts, and studies without a complete set of data or with no algorithm validation.

### Categorization of studies

Specific data from each study was recorded on a table including all the relevant citation information: digital object identifier (DOI), authors, title, and year of publication. Then, the important information in each paper was included: abstract, size, and origin of the cohort, which machine learning algorithms were used, performance of the best algorithm, features, number of features, and validation technique. Thirdly, after reviewing each paper, it was classified by the type of feature. In that sense, all the papers found in this systematic review fall into the category of supervised ML algorithms.

Supervised learning is the most common ML approach in which an algorithm learns to make predictions or decisions based on labeled input–output pairs [80]. In this approach, the learning model is provided with training data, which consists of input features and corresponding output labels. The goal of the algorithm is to learn the relationship between the input features and the output labels, which can then be used to make predictions on new, unseen data. Also, it is important to explain that supervised machine learning methods can be for classification and/or regression outputs. Classification deals with the task of predicting discrete output labels or categories, whereas regression involves predicting continuous output values. Classification models are evaluated using metrics such as accuracy, precision, recall, and F1-score, while regression models are evaluated using metrics like mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and *R*-squared. The main algorithms for classification are as follows: (1) logistic regression—a linear model for classification that uses the logistic function to estimate probabilities [81], (2) decision trees—a tree-like structure that recursively splits the input space based on feature values to make predictions [82], (3) support vector machines—a method that finds the optimal hyperplane to separate the different classes [83], (4) random forest—an ensemble learning method that constructs multiple decision trees and combines their output [84], and (5) neural networks—a method to make predictions by being trained on a labeled dataset, where input–output pairs are provided. One of the most popular supervised learning algorithms for neural networks is backpropagation [85]. On the other hand, the main algorithms for regression are as follows: (1) linear regression—a linear model that predicts the target variable by minimizing the sum of squared errors [86], (2) Lasso regression—a linear model that includes L1 regularization, which helps in feature selection and reducing overfitting [87], (3) ridge regression—a linear model that includes L2 regularization, which helps in reducing overfitting [88], (4) decision trees (for regression)—similar to classification trees but

predicting continuous values instead of classes [82], (5) neural networks (for regression)—similar to classification neural networks but optimized for continuous output predictions [89].

In summary, according to the above definitions, all reviewed articles can be categorized for classification purposes based on supervised ML algorithms (Table 1). The diagram of Fig. 2 represents a supervised learning process for medical diagnosis using biomarkers.

### Results

Our search identified 346 articles published between 1/2012 and 1/2023, of which 123 studies were excluded based on features delivered by databases. Sixty-four studies were excluded after reading the title. One-hundred studies were excluded after reading the abstract. Thirty studies were excluded during full-text screening. Finally, 29 studies meet our criteria for this review of MCI using machine learning as a diagnostic tool.

The systematic review highlights the variety of machine learning algorithms used in diagnosing MCI and AD, with traditional methods being more common in transversal studies and a diverse set of other algorithms used in longitudinal studies. Traditional machine learning methods, such as support vector machine and random forest, are the most common, but other algorithms like logistic regression methods and non-traditional techniques like extreme learning machine are also present. The Alzheimer's Disease Neuroimaging Initiative (ADNI) is the most common cohort source. Sample sizes for the groups HC, MCI, and AD vary significantly across studies. AUC and ACC values also vary across studies, but not all studies report both values. Some of the highest AUC values are found in Redolfi et al. [90] and Sh et al. [91], while high ACC values are reported in Khatri et al. [92] and Barbará-Morales et al. [93]. Table 1 represents a descriptive summary of the review.

### Study types

We identified 2 main groups among the selected papers, those that use cross-sectional data and those that use longitudinal data from the established cohort (Table 1). Almost half of the selected articles correspond to each of the categories mentioned (15 transversal and 14 longitudinal). Also, we found that the Alzheimer's Disease Neuroimaging Initiative (ADNI) is the widely used cohort for these studies. Another cohort is the Oxford Project to Investigate Memory and Aging (OPTIMA), while the rest are their own non-public cohorts (Table 2).

Transversal studies demonstrate a higher number of algorithms focused on logistic regression methods and traditional machine learning methods. The highest reported ACC is 0.86 [94]. Longitudinal studies utilize

**Table 1** Review's descriptive summary

STUDY TYPES:	
Transversal studies:	<b>15</b>
Longitudinal studies:	<b>14</b>

COHORTS	
ADNI	<b>19</b>
Italy	<b>3</b>
Finland	<b>1</b>
Spain	<b>2</b>
UK (OPTIMA)	<b>1</b>
GEO DB	<b>1</b>
USA	<b>1</b>
China	<b>1</b>

AUC (AREA UNDER THE CURVE):	
Ranges from	<b>0.75 to 0.97</b>
Not reported in	<b>14 studies</b>

TYPES OF ALGORITHMS:	
Logistic Regression (Explainable):	<b>4</b>
Support Vector Machine (No Explainable):	<b>8</b>
Ensemble methods (No Explainable):	<b>11</b>
Neural networks (No Explainable):	<b>2</b>
Other algorithms:	<b>4</b>

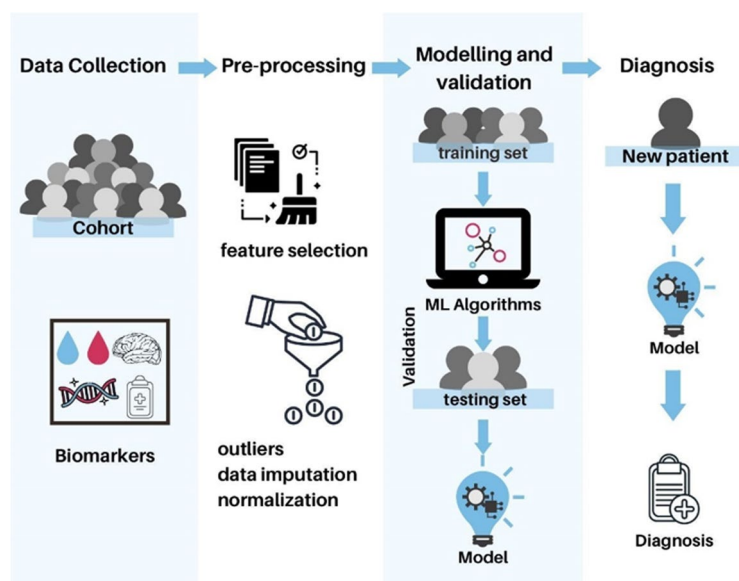
AUC (AREA UNDER THE CURVE):	
Ranges from	<b>0.61 to 0.97</b>
Not reported in	<b>13 studies</b>

DATA INPUTS	
Fluid	<b>29</b>
Neuroimaging:	<b>15</b>
Genetic:	<b>6</b>
Cognitive test / Demographic:	<b>8</b>

SAMPLE SIZES	
HC: Ranges from	<b>14 to 432</b>
MCI: Ranges from	<b>10 to 456</b>
AD: Ranges from	<b>16 to 451</b>



**Fig. 2** Supervised machine learning (ML) process. The first step is when biomarkers are taken from a cohort, then the data is pre-processed, outliers are removed, data is imputed and normalized, then the data is divided between testing and training data sets; the first is used to train the algorithms, and the second test it and validate it, if the model is selected as the best performing algorithm. Finally, the model can be used to diagnose a new patient

**Table 2** Comprehensive synthesis of the systematic review findings

AUTHORS	TYPE OF STUDY	TYPE OF STUDY	ALGORITHM DESCRIPTION	COHORT	SAMPLES SIZE	AUC	ACC	FLUID	NEUROIMAGING	GENE TIC	COGNITIVE TEST / DEMOGRAPHIC
Martinez-Torteya et al, 2015	Transversal	Logistic regression methods	One thousand five-feature logistic regression models	ADNI	HC=48, MCI=98, AD= 48	-	0.86	X	X		X
Ficiarà et al. 2021	Transversal	Logistic regression methods	Logistic regression	Italy	HC=14, MCI=17, AD=16 FTC=22	0.75	-	X			
Jääskeläinen et al., 2020	Transversal	Other algorithms	Disease State Index (DSI)	Finland	N=498	0.81	-	X			X
Olazarán et al., 2015	Transversal	Logistic regression methods	Logistic regression	Spain	HC=93, MCI=58, AD=100	0.83	-	X			
Gray et al., 2013	Transversal	Ensemble methods	Multi modality random forests, combined embedding	ADNI	HC=35, pMCI=34, sMCI=41, AD=37	-	-	X	X		
Zhao et al., 2020	Transversal	Ensemble methods	Random forest	UK (Optima)	HC=51, MCI= 13, AD=32	-	-	X			
Redolfi et al., 2020	Transversal	Ensemble methods	Gradient Boosting	Italy	HC=432, MCI=456, AD=451	0.95	-	X			X
Wang et al., 2018	Transversal	SVM	Support vector machine	ADNI	HC=52, MCI=99	-	0.76	X	X		
Miller et al., 2020	Transversal	SVM	Support vector machine	ADNI	HC=250, MCI=411, AD= 80	0.9	-	X			
Hu et al., 2016	Transversal	SVM	Support vector machine	ADNI	MCI=135, AD=95	-	0.61	X			
Sh et al., 2021	Transversal	SVM	Support vector machine	GEO DB	HC=238, MCI=189, AD=284	0.97	0.92	X			
Yilmaz et al., 2020	Transversal	SVM	Support vector machine	USA	HC=29, MCI=10, AD=20	0.78	-	X			
Peña-Bautista et al., 2019	Transversal	Neural networks	Artificial neural networks	Spain	HC=28, MCI=70	0.88	-	X			X
Dong et al., 2021	Transversal	Other algorithms	hLRR method	ADNI	HC=226, MCI=393, AD=186	0.84	0.85	X	X		
Chang et al., 2021	Transversal	Other algorithms	Naïve bayes	China	HC=31, MCI=21, AD=133	0.82	-	X			
Santangelo et al., 2020	Longitudinal	Logistic regression methods	Stepwise forward logistic regression	Italy	MCI=77	0.88	-	X	X		
Abate et al., 2021	Longitudinal	Ensemble methods	Non-parametric method regression Tree	ADNI	HC=64, sMCI=25, pMCI=111	0.92	-	X		X	X
Lin et al., 2021	Longitudinal	Ensemble methods	Extreme learning machine-based decision tree	ADNI	HC=200, sMCI=208, pMCI=110, AD=105	-	0.67	X	X	X	
Devanarayan et al., 2019	Longitudinal	Ensemble methods	Sequential Bootstrapping and Aggregation of Trees (BATting)	ADNI	HC=86, sMCI=64, pMCI=71, AD= 66	-	0.72	X	X		
Barbará-Morales et al., 2020	Longitudinal	Ensemble methods	Random forest	ADNI	HC=40, sMCI=40, AD=40	-	0.9	X	X		X
Iddi et al., 2019	Longitudinal	Ensemble methods	Random forest	ADNI	HC=154, MCI=414, AD=273	-	0.8	X	X		X
Lin et al., 2020	Longitudinal	Ensemble methods	Extreme learning machine (ELM)-based grading method	ADNI	HC=200, sMCI=205, pMCI=110, AD=102	-	0.85	X	X	X	
Mathotaarachchi et al., 2017	Longitudinal	Ensemble methods	RUSRF	ADNI	sMCI=230, pMCI=43	0.91	0.84	X			
Khatri et al., 2020	Longitudinal	Ensemble methods	Extreme learning machine	ADNI	HC=57, sMCI=35, pMCI=43, AD=53	-	0.97	X	X	X	X
Gupta et al., 2019	Longitudinal	SVM	Novel machine learning-based framework	ADNI	HC=38, sMCI=38, pMCI=46, AD= 38	-	-	X	X	X	
Zhang et al., 2012	Longitudinal	SVM	Support vector machine	ADNI	HC=50, MCI=91, AD=45	-	-	X	X		
Eke et al., 2021	Longitudinal	SVM	Support vector machine	ADNI	HC=54, MCI=136, AD=108	0.8	-	X			
Cheng et al., 2015	Longitudinal	Neural networks	Novel multimodal manifold-regularized transfer learning (M2TL)	ADNI	HC=52, sMCI=58, pMCI=43, AD=51	-	0.8	X	X		
Escudero et al., 2013	Longitudinal	Other algorithms	Local weighted classifier	ADNI	HC=45, sMCI=75, pMCI=12, AD=41	0.78	0.79	X	X	X	

References: Martínez-Torteya et al, 2015 [94], Ficiarà et al. 2021 [95], Jääskeläinen et al., 2020 [96], Olazarán et al., 2015 [97], Gray et al., 2013 [98], Zhao et al., 2020 [99], Redolfi et al., 2020 [90], Wang et al., 2018 [100], Miller et al., 2020 [101], Hu et al., 2016 [102], Sh et al., 2021 [91], Yilmaz et al., 2020 [103], Peña-Bautista et al., 2019 [104], Dong et al., 2021 [105], Chang et al., 2021 [106], Santangelo et al., 2020 [107], Abate et al., 2021 [108], Lin et al., 2021 [109], Devanarayan et al., 2019 [110], Barbará-Morales et al., 2020 [93], Iddi et al., 2019 [111], Lin et al., 2020 [112], Mathotaarachchi et al., 2017 [113], Khatri et al., 2020 [92], Gupta et al., 2019 [114], Zhang et al., 2012 [115], Eke et al., 2021 [116], Cheng et al., 2015 [117], Escudero et al., 2013 [118]

a more diverse set of algorithms, including other algorithms such as extreme learning machine and novel machine learning-based frameworks. The highest reported ACC is 0.97 [92].

**Algorithms inputs**

Regarding the characteristics used as input for the algorithms (Table 1), 10 articles were found to use only fluid biomarkers as input for the algorithm, and the others use it accompanied by other types of biomarkers, specifically 15 use neuroimaging, 8 neuropsychological tests, and 6 genetic information in addition to fluid biomarkers (showing that the tendency is to employ more than one source of information, in a multimodal way).

In fact, we observed that longitudinal studies tend to be multimodal, applying different types of inputs. Importantly, we found that genetic data is the least used feature in ML algorithms while the most used is neuroimaging data, considering that this study only included articles that used neuroimaging in support of fluid biomarkers.

However, only 14 articles did not use neuroimaging as input to the algorithm, corresponding to less than 50% of the included studies. Also, we found that algorithms using only fluid biomarkers as features have reported very good performances.

**Cohorts**

Regarding patients' home country, 28 cohorts include patients from the USA, 4 from Italy, 3 from Spain, and 2 from Korea, and with one study each, considering patients from China, Holland, Finland, and the UK were found (Table 1). In addition, we identified that most of the included studies consider three main classification categories as targets for their patients: healthy control (HC), MCI, and AD. However, we also found three studies that considered only HC and MCI, one that considered MCI and AD and 2 that considered only MCI divided into two subcategories (progressive and stable) (Table 1).

Studies using the ADNI cohort have a broad range of sample sizes and algorithm types. Performance metrics

(AUC and ACC) also vary considerably within this group. Studies from Italy, Spain, and other countries generally report high AUC and ACC values, though sample sizes are often smaller than those in ADNI cohort studies.

**Performance metrics (AUC and ACC)**

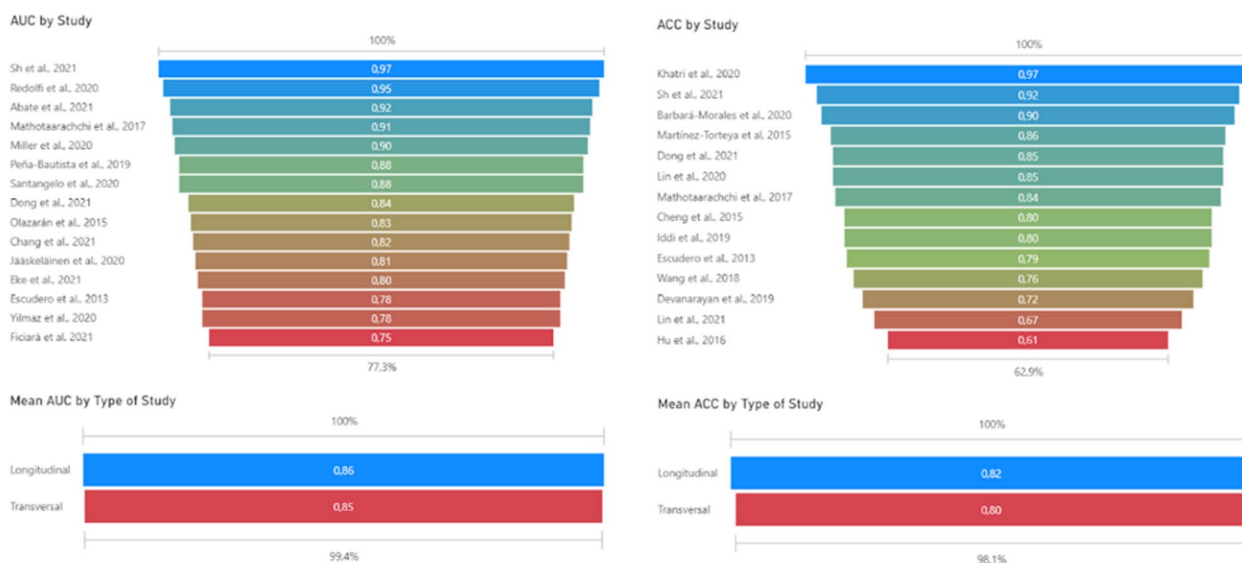
In machine learning, area under the curve (AUC ROC) and accuracy (ACC) are two distinct performance metrics used to evaluate the effectiveness of classification models. AUC refers to the area under the ROC curve, which plots the true-positive rate (sensitivity) against the false-positive rate (1-specificity) at various threshold settings. AUC ranges from 0 to 1, with a higher value indicating better model performance. It is particularly useful when dealing with imbalanced datasets, as it considers both sensitivity and specificity. ACC, on the other hand, is the ratio of correct predictions to the total number of predictions made. It measures the overall performance of a model and is more suitable for balanced datasets. However, ACC can be misleading in the case of imbalanced datasets, as it may not account for the true effectiveness of a classifier. In summary, AUC is a more robust performance metric that considers both sensitivity and specificity, while ACC measures the overall performance but may be less informative for imbalanced datasets [119]. In this review, AUC values are generally high, with most reported values above 0.8. The highest AUC value is 0.97 [91], achieved using a

support vector machine in a transversal study. ACC values also tend to be high, with most reported values above 0.8. The highest ACC value is 0.97 [92], achieved using an extreme learning machine in a longitudinal study (Fig. 3). In some cases, only one of the two performance metrics is reported, making it difficult to comprehensively compare the algorithms' performances. No significant difference is observed between the accuracy achieved between the two groups, nor in the reported AUC. However, longitudinal studies have the potential of predicting the diagnosis of diseases, which could allow early action to use treatments aiming to delay the progression of the disease.

Overall, logistic regression methods generally report high ACC values (0.86 and 0.88), though AUC values are not always available. Traditional ML methods, specifically Support Vector Machines, are widely used in transversal studies with varying ACC values (0.61 to 0.97). Ensemble methods (i.e.: Random Forest algorithms) are more common in longitudinal studies and report ACC values between 0.8 and 0.9. Other algorithms show a wider range of performance, with AUC values from 0.75 to 0.97 and ACC values from 0.67 to 0.97 (Fig. 4).

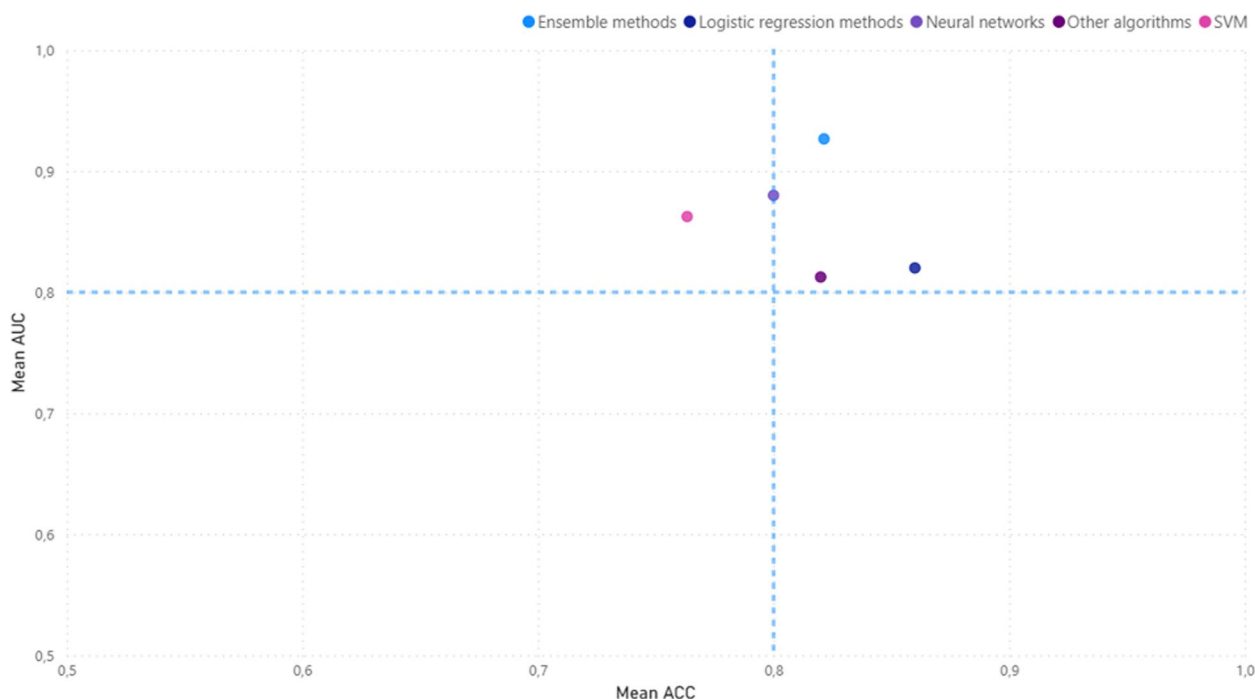
**Black-box algorithms and sample sizes**

In relation to the black-box problem mentioned, of the total of 29 studies reviewed, 11 (38%) of them use algorithms of neural networks, support vector machines, and



**Fig. 3** Funnel plot of algorithm performance by study. This funnel plot is a specialized form of the scatterplot, uniquely tailored for the analysis and visualization of data behavior between minimal and maximum metrics. Its primary function is to assist in identifying anomalies or outliers within the data set. In a funnel plot, data points are depicted as dots and plotted within a funnel-shaped graphical field. The shape of the funnel serves as a visual guideline, delineating the expected range of variation based on statistical norms. Consequently, any data point, or dot, that is plotted outside this funnel shape is classified as an outlier, indicating a substantial deviation from the anticipated pattern or range. In the context of this review, it is noteworthy that all the metrics derived from the studies are plotted within the confines of the funnel. This suggests that there is a consistent pattern in the data, with no significant anomalies or outliers detected. It implies that the metrics of the studies fall within the expected range and adhere to the statistical norms, reinforcing the reliability and validity of the reviewed studies





**Fig. 4** Scatterplot of algorithm performance. Scatterplot representing the performance of the ML algorithms. The x-axis, labeled “ACC,” measures the accuracy of the algorithms. Accuracy is a simple metric for binary classification problems, representing the proportion of true results (both true positives and true negatives) among the total number of cases examined. The y-axis, labeled “AUC,” represents the area under the receiver operating characteristic (ROC) curve. AUC is a popular metric in machine learning for binary classification problems. It measures the tradeoff between a true-positive rate and a false-positive rate. An AUC of 1.0 means the model has a perfect classification, while an AUC of 0.5 implies the model is no better than random guessing. Each point in the scatterplot represents a different machine learning algorithm. The position of the point on the graph shows the performance of the algorithm on both metrics: its accuracy and its AUC score. The scatterplot also features a performance target of 0.8. This could be represented as a line or a highlighted area in the plot, indicating the desired minimum performance level for both the accuracy and AUC. Algorithms that fall within or above this target region are considered to meet or exceed the performance goal. This visual comparison makes it easier to quickly identify which algorithms meet the performance target according to these two key metrics

random forests, which are considered within the black-box methods [120]. Importantly, relative to the sample sizes of the studies included here (Table 1), it is observed that of a total of 18 studies (62%), the sample is less than 100 cases. Of the 15 cross-sectional studies, 8 (53%) of them have a sample size of less than 100. In longitudinal studies, a total of 10 (71%) use a sample of less than 100 cases. A very relevant aspect of modeling based on ML is related to the sampling size and the power of its estimates. Using ML on small-size datasets could present a problem. The smaller the dataset, the less powerful and less accurate the models [121]. The process of ML involves training, validation, and test datasets [122]. This perspective is essential in the development of models and must be considered.

For the cross-sectional studies, the most used types of techniques are the explicable ones with 4 articles and support vector machine (SVM) with 5 articles out of a total of 15, while for the longitudinal studies, the assemblies are the ones that take the lead with 7 out of 14 articles. The category of explainable models includes logistic regression and decision tree.

#### Methodological issues

No article was found in this review that considered the longitudinal dimension within the algorithm itself. However, in all cases, this dimension is delivered to the data label with which the algorithm will be trained later. For the use of longitudinal data, deep learning methods are the most recommended in the literature but considering that these fields the availability of data is limited, and it is also sought that the methodology used be interpretable, which excludes efforts with deep learning that have an advantage in the representation of longitudinal data [123]. The classification targets of the algorithm are relevant for the comparison of their results. It is possible to differentiate between classification targets which include disease progression, which require a longitudinal study to be able to evaluate the individuals at least in two different time points, and classification targets without progression, where the targets are defined on a single data point. In the case of classification targets that include disease progression, they are usually built to represent the longitudinal dimension of the data. For example, in a study,

4 classes were used (healthy control (HC), cognitive impairment without progression to Alzheimer's (sMCI), cognitive impairment with progression to Alzheimer's (cMCI), and Alzheimer's (AD)).

## Discussion

The growing use of artificial intelligence techniques today to work new diagnostic algorithms with current data has an impact on the diagnostic tool, improving its accuracy and helping to predict the status or possible evolution of the patient. Biomarkers can play a very important role [124, 125] which could distinguish between AD and MCI or between MCI and age-related changes [126]. There are different types of biomarkers, some of which are better studied, such as those obtained from cerebrospinal fluid and neuroimaging, and novel types, especially due to their cost-efficiency, such as fluid biomarkers that include those obtained from the blood, urine, and saliva, which correspond to proteins and miRNA, among others [127–129]. On the one hand, they would allow greater access to the population and, on the other, to the investigation of their use for diagnosis, increasing sample sizes, especially when we talk about ML tools, where the sample size becomes relevant in order to achieve greater statistical power.

In this review, it is seen that the sample sizes tend to be low; however, the article by Redolfi et al. stands out with 1339 participants between HC, MCI, and AD for the articles of the cross-sectional category, using the base built from of a Medical Informatics Platform installed across 3 Italian memory clinics and the article by Iddi et al., with 841 participants between controls, MCI, and AD for the longitudinal articles using the ADNI database. Then, articles with a sample size of up to 59 patients were included [103], where an AUC of 78% was obtained, in this case, the MCI category has 10 participants, while HC and AD have 29 and 20 patients, respectively, indicating a very low statistical power.

Today the diagnosis of MCI is based mainly on neuropsychological tests that include cognitive and functional tests. These tests can be influenced by various factors such as age, education, and lifestyle, among others [130, 131], leading to a focus on new forms of detection that are more reliable and ideally easily accessible. Neuropsychological tests can provide complementary information to suggest the etiological diagnosis, but not enough to do it on their own. Here, we carry out a systematic review of studies from the last 10 years that consider ML techniques and use fluid biomarkers for the diagnosis of MCI due to AD. As for the selected articles, most use some type of biomarker in addition to fluid biomarkers, as complementary information. It was found that fluid biomarkers are mainly added to neuroimaging

characteristics, neuropsychological tests, and genetic information. Being neuroimaging is the most used, especially in longitudinal articles, this may be due to the amount of neuroimaging data available to test new ML architectures in databases such as ADNI, compared to the datasets of other biomarkers, which facilitates the implementation and investigation of these techniques.

Another point to highlight is the explainability of the models to be used, in terms of biomedicine and diagnostics in general; it is expected to be closer to white-box than black-box type models, since an important component is the process of the architecture of the model. In this case, there are 4 explainable models for cross-sectional articles, corresponding to different logistic regressions, while for longitudinal articles, there are only two, a logistic regression and a decision tree. While the other models lose explainability to a certain degree, reaching black boxes such as the neural network that only appears in 2 articles in this review [132, 133] where the explainability of the process is completely lost.

In addition, a relevant aspect in the development of diagnostic models for MCI is related to the cultural representativeness gap [134]. There is no single cause of AD, but multiple factors are involved [135]. Among these factors, the socioeconomic and cultural condition is very important for both diagnosis and treatment [136, 137]. Unfortunately, most research in MCI and AD is conducted in the US and European population, where findings and inferences cannot be extrapolated cross-culturally to everyone in appropriate cultural contexts and niches [138–143]. This premise is applied in this specific review of the diagnosis of MCI with the use of ML, where there is no article on the Latin American or African population and the majority includes the population of the USA and/or Europe, leaving a gap in the study of these related populations.

On the other hand, it is important to highlight that most of the ML approaches mentioned in this review are based on associative inference. Primarily, these methods propose a diagnostic framework that establishes correlations among various factors including symptoms, neuroimaging data, neuropsychological tests, genetic information, and other relevant variables. However, associative inference is the simplest in a hierarchy of possible inference schemes because it does not allow causal explanations to be attributed to the data [144]. Instead, an approach based on causal and/or counterfactual inference modeling would allow it [145]. There is evidence that a highly accurate medical diagnosis can be performed based as a counterfactual inference approach using ML methods such as probabilistic graphical model (PGM), Bayesian networks, and noisy-OR algorithms [146]. Under this approach, it is argued that diagnosis is

the process of finding causal explanations for a patient's symptoms. This implies promoting causal reasoning to show that the probability of occurrence of an effect B has really been caused by cause A. This leads to developing a diagnostic measure to classify the probability that a disease X is causing the symptoms of a patient given the evidence. In general, the criticism leveled at associative inference is the fact that not separation of correlation from causation places strong constraints on the accuracy of associative diagnostic algorithms, sometimes leading to suboptimal diagnostic results [147]. In this case, these methodological perspectives of causal inference would be very helpful for the development of MCI and AD diagnostic models based on ML algorithms.

### Conclusions and future directions

The latest advances in neuroimaging, laboratory analysis, genetic, and ML techniques have led to a progressive change in the diagnosis of neurodegenerative diseases. Overall, it is not possible to determine with our systematic review a group of techniques or features which achieves better results than others, since the metrics reported vary widely. However, it is possible to see that the following points address the major themes and challenges that appear in the article and would be essential considerations for any researcher planning to embark on a similar research journey in the realm of neurodegenerative diseases using ML techniques:

- **Multi-modal approach:** The growing trend towards a multi-modal approach in utilizing neuroimaging, laboratory, genetic, and ML techniques from cohorts like the Alzheimer's Disease Neuroimaging Initiative, Neuroimaging in Frontotemporal Dementia, and UNITED Consortium [63] requires careful planning and integration. Researchers need to understand how to synergize various data types and technologies, leveraging them for more accurate diagnoses.
- **Longitudinal data challenges:** The lack of sufficient data, short monitoring time, and a need for interpretable results in longitudinal studies present significant challenges. When planning research, it is essential to ensure that the methodology allows for long-term monitoring and that the tools used can handle sparse or incomplete data.
- **Cultural representativeness and diverse population sampling:** The absence of studies in underrepresented populations like Latin American, Caribbean, and African regions necessitates planning for more inclusive research [137, 142, 148, 149]. This could involve considering different cultural, socioeconomic, and demographic factors that may influence the disease progression and diagnosis.

- **Use of white-box models:** The emphasis on using white-box type machine learning algorithms reflects a need for transparency and interpretability in the models. Researchers need to carefully choose or design algorithms that not only perform well but also provide insights into how and why they are making specific predictions.
- **Cost-effective large-scale studies:** Achieving long-term multimodal longitudinal monitoring in a cost-effective manner is crucial. Planning must include budget considerations, the incorporation of large-scale sample sizes, and a strategic approach to collect and analyze the large quantities of data required.

In summary, the intricate and varied array of techniques within neuroimaging, laboratory analysis, genetics, and machine learning alludes to a captivating yet demanding trajectory in neurodegenerative disease research. Successful amalgamation of these methodologies demands thorough planning, inclusiveness, and transparency, thereby establishing the foundation for a groundbreaking era in diagnosing and comprehending these ailments. Furthermore, the establishment of a multidisciplinary task force is imperative to rectify diagnostic accuracy.

### Abbreviations

MCI	Mild cognitive impairment
AD	Alzheimer's disease
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analysis
CSF	Cerebrospinal fluid
PET	Positron emission tomography
FDG	F18-fluorodeoxyglucose
NFL	Neurofilament light chain
ML	Machine learning

### Authors' contributions

CDA and RdIC designed the proposal. KB, SS and TS performed PRISMA analysis. KB and SS wrote the drafts and discussed contributions from all co-authors. All authors searched the literature, participated in discussing the contents of the paper, contributed to editing, and approved the final version of the article.

### Funding

We thank María F. Aguirre-Pinto for the technical support. CDA is supported by ANID/FONDECYT Regular 1210622. AI is supported by grants from ANID/FONDECYT Regular (1210195 and 1210176 and 1220995), ANID/FONDAP/15150012, FONDEF ID20110152 and ID22110029, ANID/FONDAP 15150012, Takeda CW2680521, and the MULTI-PARTNER CONSORTIUM TO EXPAND DEMENTIA RESEARCH IN LATIN AMERICA [ReDLat, supported by National Institutes of Health, National Institutes of Aging (R01 AG057234), Alzheimer's Association (SG-20-725707), Rainwater Charitable foundation – Tau Consortium, and Global Brain Health Institute]). AI, CDA, and RdIC are supported by grant ANID/PIA/ANILLO ACT210096. The contents of this publication are solely the responsibility of the authors and do not represent the official views of these institutions.

### Availability of data and materials

All data generated or analyzed during this study are included in this published article.

### Declarations

#### Ethics approval and consent to participate

Not applicable.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare no competing interests.

**Author details**

<sup>1</sup>Center for Social and Cognitive Neuroscience (CSCN), School of Psychology, Universidad Adolfo Ibáñez, Diagonal Las Torres 2640, Peñalolén, Santiago, Chile. <sup>2</sup>Latin American Institute for Brain Health (BrainLat), Universidad Adolfo Ibáñez, Santiago, Chile. <sup>3</sup>Faculty of Engineering and Sciences, Universidad Adolfo Ibáñez, Diagonal Las Torres 2700, Building D, Peñalolén, Santiago, Chile. <sup>4</sup>Global Brain Health Institute, Trinity College, Dublin, Ireland. <sup>5</sup>Memory and Neuropsychiatric Center (CMYN) Neurology Department, Hospital del Salvador and Faculty of Medicine, University of Chile, Santiago, Chile. <sup>6</sup>Technische Universität Berlin, Berlin, Deutschland. <sup>7</sup>Instituto de Investigaciones Psicológicas (IIPsi), Universidad Nacional de Córdoba (UNC) y Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Córdoba, Argentina. <sup>8</sup>Global Brain Health Institute, University of California San Francisco (UCSF), San Francisco, CA, USA. <sup>9</sup>Cognitive Neuroscience Center (CNC), Universidad de San Andrés, & National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina. <sup>10</sup>Data Observatory Foundation, ANID Technology Center No. DO210001, Santiago, Chile.

Received: 2 June 2023 Accepted: 15 September 2023

Published online: 14 October 2023

**References**

- Winblad B, Palmer K, Kivipelto M, Jelic V, Fratiglioni L, Wahlund LO, Nordberg A, Bäckman L, Albert M, Almkvist O, Arai H, Basun H, Blennow K, de Leon M, DeCarli C, Erkinjuntti T, Giacobini E, Graff C, Hardy J, Jack C, et al. Mild cognitive impairment—beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. *J Internal Med.* 2004;256(3):240–6.
- Sachdev PS, Lipnicki DM, Kochan NA, Crawford JD, Thalamuthu A, Andrews G, Brayne C, Matthews FE, Stephan BC, Lipton RB, Katz MJ, Ritchie K, Carrière I, Ancelin ML, Lam LC, Wong CH, Fung AW, Guaita A, Vaccaro R, Davin A, et al. Cohort Studies of Memory in an International Consortium (COSMIC). The prevalence of mild cognitive impairment in diverse geographical and ethnocultural regions: the COSMIC Collaboration. *PLoS ONE.* 2015;10(11).
- Livingston G, Sommerlad A, Orgeta V, Costafreda SG, Huntley J, Ames D, Ballard C, Banerjee S, Burns A, Cohen-Mansfield J, Cooper C, Fox N, Gitlin LN, Howard R, Kales HC, Larson EB, Ritchie K, Rockwood K, Sampson EL, Samus Q, et al. Dementia prevention, intervention, and care. *Lancet (London, England).* 2017;390(10113):2673–734.
- Hulette CM, Welsh-Bohmer KA, Murray MG, Saunders AM, Mash DC, McIntyre LM. Neuropathological and neuropsychological changes in “normal” aging: evidence for preclinical Alzheimer disease in cognitively normal individuals. *J Neuropathol Exp Neurol.* 1998;57(12):1168–74.
- Grill JD, Raman R, Ernstrom K, Aisen P, Karlawish J. Effect of study partner on the conduct of Alzheimer disease clinical trials. *Neurology.* 2013;80(3):282–8.
- Mitchell AJ, Shiri-Feshki M. Rate of progression of mild cognitive impairment to dementia—meta-analysis of 41 robust inception cohort studies. *Acta Psychiatr Scand.* 2009;119(4):252–65.
- Petersen RC, Doody R, Kurz A, Mohs RC, Morris JC, Rabins PV, Ritchie K, Rossor M, Thal L, Winblad B. Current concepts in mild cognitive impairment. *Arch Neurol.* 2001;58(12):1985–92.
- Bruscoli M, Lovestone S. Is MCI really just early dementia? A systematic review of conversion studies. *Int Psychogeriatr.* 2004;16(2):129–40.
- Petersen RC. Mild cognitive impairment as a diagnostic entity. *J Intern Med.* 2004;256(3):183–94.
- Panza F, D'Introno A, Colacicco AM, Capurso C, Del Parigi A, Caselli RJ, Pilotto A, Argentieri G, Scapicchio PL, Scafato E, Capurso A, Solfrizzi V. Current epidemiology of mild cognitive impairment and other pre-dementia syndromes. *Am J Geriatr Psychiatry.* 2005;13(8):633–44.
- Pinto C, Subramanyam AA. Mild cognitive impairment: the dilemma. *Indian J Psychiatry.* 2009;51(Suppl 1):S44–51.
- DeCarli C. Mild cognitive impairment: prevalence, prognosis, aetiology, and treatment. *Lancet Neurol.* 2003;2(1):15–21.
- Chun CT, Seward K, Patterson A, Melton A, MacDonald-Wicks L. Evaluation of available cognitive tools used to measure mild cognitive decline: a scoping review. *Nutrients.* 2021;13(11):3974.
- Biomarkers Definitions Working Group, Atkinson AJ, Colburn WA, DeGruttola VG, DeMets DL, Downing GJ, Hoth DF, Oates JA, Peck CC, Schooley RT, Spilker BA, Woodcock J, Zeger SL. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther.* 2001;69(3):89–95.
- Teunissen CE, Verberk IMW, Thijssen EH, Vermunt L, Hansson O, Zetterberg H, van der Flier WM, Mielke MM, Del Campo M. Blood-based biomarkers for Alzheimer's disease: towards clinical implementation. *Lancet Neurol.* 2022;21(1):66–77.
- Bobinski M, de Leon MJ, Convit A, De Santi S, Wegiel J, Tarshish CY, Saint Louis LA, Wisniewski HM. MRI of entorhinal cortex in mild Alzheimer's disease. *Lancet (London, England).* 1999;353(9146):38–40.
- Zarow C, Vinters HV, Ellis WG, Weiner MW, Mungas D, White L, Chui HC. Correlates of hippocampal neuron number in Alzheimer's disease and ischemic vascular dementia. *Ann Neurol.* 2005;57(6):896–903.
- Bobinski M, de Leon MJ, Wegiel J, Desanti S, Convit A, Saint Louis LA, Rusinek H, Wisniewski HM. The histological validation of postmortem magnetic resonance imaging-determined hippocampal volume in Alzheimer's disease. *Neuroscience.* 2000;95(3):721–5.
- Phelps ME. Positron emission tomography provides molecular imaging of biological processes. *Proc Natl Acad Sci USA.* 2000;97(16):9226–33.
- Silverman DH, Phelps ME. Application of positron emission tomography for evaluation of metabolism and blood flow in human brain: normal development, aging, dementia, and stroke. *Mol Genet Metab.* 2001;74(1–2):128–38.
- Drzezga A, Lautenschlager N, Siebner H, Riemenschneider M, Willoch F, Minoshima S, Schwaiger M, Kurz A. Cerebral metabolic changes accompanying conversion of mild cognitive impairment into Alzheimer's disease: a PET follow-up study. *Eur J Nucl Med Mol Imaging.* 2003;30(8):1104–13.
- de Leon MJ, Convit A, Wolf OT, Tarshish CY, DeSanti S, Rusinek H, Tsui W, Kandil E, Scherer AJ, Roche A, Imossi A, Thorn E, Bobinski M, Caraos C, Lesbre P, Schlyer D, Poirier J, Reisberg B, Fowler J. Prediction of cognitive decline in normal elderly subjects with 2-[<sup>18</sup>F]fluoro-2-deoxy-D-glucose/positron-emission tomography (FDG/PET). *Proc Natl Acad Sci USA.* 2001;98(19):10966–71.
- Koychev I, Gunn RN, Firouzi A, Lawson J, Zamboni G, Ridha B, Sahakian BJ, Rowe JB, Thomas A, Rochester L, Ffytche D, Howard R, Zetterberg H, MacKay C, Lovestone S, Deep and Frequent Phenotyping study team. PET tau and amyloid- $\beta$  burden in mild Alzheimer's disease: divergent relationship with age, cognition, and cerebrospinal fluid biomarkers. *J Alzheimer's Dis.* 2017;60(1):283–93.
- Ossenkopppele R, Iaccarino L, Schonhaut DR, Brown JA, La Joie R, O'Neil JP, Janabi M, Baker SL, Kramer JH, Gorno-Tempini ML, Miller BL, Rosen HJ, Seeley WW, Jagust WJ, Rabinovici GD. Tau covariance patterns in Alzheimer's disease patients match intrinsic connectivity networks in the healthy brain. *NeuroImage Clin.* 2019;23: 101848.
- Aschenbrenner AJ, Gordon BA, Benzinger TLS, Morris JC, Hassenstab JJ. Influence of tau PET, amyloid PET, and hippocampal volume on cognition in Alzheimer disease. *Neurology.* 2018;91(9):e859–66.
- Blennow K, Zetterberg H. Biomarkers for Alzheimer's disease: current status and prospects for the future. *J Intern Med.* 2018;284(6):643–63.
- Lista S, Faltraco F, Prvulovic D, Hampel H. Blood and plasma-based proteomic biomarker research in Alzheimer's disease. *Progr Neurobiol.* 2013;101–102:1–17.
- Zetterberg H, Wilson D, Andreasson U, Minthon L, Blennow K, Randall J, et al. Plasma tau levels in Alzheimer's disease. *Alzheimer's Res Ther.* 2013;5:9.
- Herukka SK, Hallikainen M, Soininen H, Pirttilä T. CSF A $\beta$ 42 and tau or phosphorylated tau and prediction of progressive mild cognitive impairment. *Neurology.* 2005;64:1294–7.
- Blennow K, Hampel H. Review CSF markers for incipient Alzheimer's disease CSF markers for incipient AD. *Lancet.* 2003;2:605–13.
- Parnetti L, Chiasserini D, Eusebi P, Giannandrea D, Bellomo G, de Carlo C, et al. Performance of A $\beta$ 1–40, A $\beta$ 1–42, total tau, and phosphorylated tau

- as predictors of dementia in a cohort of patients with mild cognitive impairment. *J Alzheimer's Dis.* 2012;29:229–38.
32. Okonkwo OC, Alosco ML, Griffith HR, Mielke MM, Shaw LM, Trojanowski JQ, et al. Cerebrospinal fluid abnormalities and rate of decline in everyday function across the dementia spectrum: normal aging, mild cognitive impairment, and Alzheimer disease. *Archiv Neurol.* 2010;67:688–96.
  33. Hansson O, Seibyl J, Stomrud E, Zetterberg H, Trojanowski JQ, Bittner T, et al. CSF biomarkers of Alzheimer's disease concord with amyloid- $\beta$  PET and predict clinical progression: a study of fully automated immunoassays in BioFINDER and ADNI cohorts. *Alzheimer's Dement.* 2018;14:1470–81.
  34. Forlenza OV, Radanovic M, Talib LL, Aprahamian I, Diniz BS, Zetterberg H, et al. Cerebrospinal fluid biomarkers in Alzheimer's disease: diagnostic accuracy and prediction of dementia. *Alzheimer's Dement.* 2015;1:455–63.
  35. Trojanowski JQ, Vanderstichele H, Korecka M, Clark CM, Aisen PS, Petersen RC, et al. Update on the biomarker core of the Alzheimer's disease neuroimaging initiative subjects. *Alzheimer's Dement.* 2010;6:230–8.
  36. Hansson O, Zetterberg H, Buchhave P, Londos E, Blennow K, Minthon L. Association between CSF biomarkers and incipient Alzheimer's disease in patients with mild cognitive impairment: a follow-up study. *Lancet Neurol.* 2006;5:228–34.
  37. Chen YX, Liang N, Li XL, Yang SH, Wang YP, Shi NN. Diagnosis and treatment for mild cognitive impairment: a systematic review of clinical practice guidelines and consensus statements. *Front Neurol.* 2021;12: 719849.
  38. Gaetani L, Blennow K, Calabresi P, Di Filippo M, Parnetti L, Zetterberg H. Neurofilament light chain as a biomarker in neurological disorders. *J Neurol Neurosurg Psychiatry.* 2019;90(8):870–81.
  39. Ashton NJ, Hye A, Rajkumar AP, Leuzy A, Snowden S, Suárez-Calvet M, Karikari TK, Schöll M, La Joie R, Rabinovici GD, Höglund K, Ballard C, Hortobágyi T, Svenningsson P, Blennow K, Zetterberg H, Aarsland D. An update on blood-based biomarkers for non-Alzheimer neurodegenerative disorders. *Nat Rev Neurol.* 2020;16(5):265–84.
  40. Palmqvist S, Janelidze S, Quiroz YT, Zetterberg H, Lopera F, Stomrud E, Su Y, Chen Y, Serrano GE, Leuzy A, Mattsson-Carlsson N, Strandberg O, Smith R, Villegas A, Sepulveda-Falla D, Chai X, Proctor NK, Beach TG, Blennow K, Dage JL, et al. Discriminative accuracy of plasma phospho-tau217 for Alzheimer disease vs other neurodegenerative disorders. *JAMA.* 2020;324(8):772–81.
  41. Moscoso A, Grothe MJ, Ashton NJ, Karikari TK, Lantero Rodríguez J, Snellman A, Suárez-Calvet M, Blennow K, Zetterberg H, Schöll M, Alzheimer's Disease Neuroimaging Initiative. Longitudinal associations of blood phosphorylated Tau181 and neurofilament light chain with neurodegeneration in Alzheimer disease. *JAMA Neurol.* 2021;78(4):396–406.
  42. Quiroz YT, Zetterberg H, Reiman EM, Chen Y, Su Y, Fox-Fuller JT, Garcia G, Villegas A, Sepulveda-Falla D, Villada M, Arboleda-Velasquez JF, Guzmán-Vélez E, Vila-Castelar C, Gordon BA, Schultz SA, Protas HD, Ghisays V, Giraldo M, Tirado V, Baena A, et al. Plasma neurofilament light chain in the presenilin 1 E280A autosomal dominant Alzheimer's disease kindred: a cross-sectional and longitudinal cohort study. *Lancet Neurol.* 2020;19(6):513–21.
  43. Illán-Gala I, Lleo A, Karydas A, Staffaroni AM, Zetterberg H, Sivasankaran R, Grinberg LT, Spina S, Kramer JH, Ramos EM, Coppola G, La Joie R, Rabinovici GD, Perry DC, Gorno-Tempini ML, Seeley WW, Miller BL, Rosen 6 HJ, Blennow K, Boxer AL, ... Rojas JC. Plasma tau and neurofilament light in frontotemporal lobar degeneration and Alzheimer disease. *Neurology.* 2021;96(5), e671–e683.
  44. Mattsson N, Andreasson U, Zetterberg H, Blennow K, Alzheimer's Disease Neuroimaging Initiative. Association of plasma neurofilament light with neurodegeneration in patients with Alzheimer disease. *JAMA Neurol.* 2017;74(5):557–66.
  45. Baldacci F, Lista S, Manca ML, Chiesa PA, Cavado E, Lemercier P, Zetterberg H, Blennow K, Habert MO, Potier MC, Dubois B, Vergallo A, Hampel H, INSIGHT-preAD study group, & Alzheimer Precision Medicine Initiative (APMI). Age and sex impact plasma NFL and t-Tau trajectories in individuals with subjective memory complaints: a 3-year follow-up study. *Alzheimer's Res Ther.* 2020;12(1):147.
  46. Chatterjee S, Mudher A. Alzheimer's disease and type 2 diabetes: a critical assessment of the shared pathological traits. *Front Neurosci.* 2018;12:383.
  47. Giacomucci G, Mazzeo S, Bagnoli S, Ingannato A, Leccese D, Berti V, Padiglioni S, Galdo G, Ferrari C, Sorbi S, Bessi V, Nacmias B. Plasma neurofilament light chain as a biomarker of Alzheimer's disease in subjective cognitive decline and mild cognitive impairment. *J Neurol.* 2022;269(8):4270–80.
  48. Cogswell JP, Ward J, Taylor IA, Waters M, Shi Y, Cannon B, Kelnar K, Kempainen J, Brown D, Chen C, Prinjha RK, Richardson JC, Saunders AM, Roses AD, Richards CA. Identification of miRNA changes in Alzheimer's disease brain and CSF yields putative biomarkers and insights into disease pathways. *J Alzheimer's Dis.* 2008;14(1):27–41.
  49. Takousis P, Sadlon A, Schulz J, Wohlers I, Dobricic V, Middleton L, Lill CM, Pernecky R, Bertram L. Differential expression of microRNAs in Alzheimer's disease brain, blood, and cerebrospinal fluid. *Alzheimer's Dement.* 2019;15(11):1468–77.
  50. Ogonowski N, Salcidua S, Leon T, Chamorro-Veloso N, Valls C, Avalos C, Bisquertt A, Rentería ME, Orellana P, Duran-Aniotz C. Systematic review: microRNAs as potential biomarkers in mild cognitive impairment diagnosis. *Front Aging Neurosci.* 2022;13: 807764.
  51. Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, Fiske A, Pedersen NL. Role of genes and environments for explaining Alzheimer disease. *Arch Gen Psychiatry.* 2006;63(2):168–74.
  52. Saunders AM, Strittmatter WJ, Schmechel D, George-Hyslop PH, Pericak-Vance MA, Joo SH, Rosi BL, Gusella JF, Crapper-MacLachlan DR, Alberts MJ. Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. *Neurology.* 1993;43(8):1467–72.
  53. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, Roses AD, Haines JL, Pericak-Vance MA. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science (New York, NY).* 1993;261(5123):921–3.
  54. Neu SC, Pa J, Kukull W, Beekly D, Kuzma A, Gangadharan P, Wang LS, Romero K, Arneric SP, Redolfi A, Orlandi D, Frisoni GB, Au R, Devine S, Auerbach S, Espinosa A, Boada M, Ruiz A, Johnson SC, Kosciak R, et al. Apolipoprotein E genotype and sex risk factors for Alzheimer disease: a meta-analysis. *JAMA Neurol.* 2017;74(10):1178–89.
  55. Xu X, Zhang B, Wang X, Zhang Q, Wu X, Zhang J, Bai Y, Gu X. A meta-analysis of Alzheimer's disease's relationship with human ApoE gene variants. *Am J Translat Res.* 2021;13(9):9974–82.
  56. Qin W, Li W, Wang Q, Gong M, Li T, Shi Y, Song Y, Li Y, Li F, Jia J. Race-related association between APOE genotype and Alzheimer's disease: a systematic review and meta-analysis. *J Alzheimer's Dis.* 2021;83(2):897–906.
  57. Lumsden AL, Mulugeta A, Zhou A, Hyppönen E. Apolipoprotein E (APOE) genotype-associated disease risks: a phenome-wide, registry-based, case-control study utilising the UK Biobank. *EBioMedicine.* 2020;59: 102954.
  58. Seshadri S, Fitzpatrick AL, Ikram MA, DeStefano AL, Gudnason V, Boada M, Bis JC, Smith AV, Carassquillo MM, Lambert JC, Harold D, Schrijvers EM, Ramirez-Lorca R, Debette S, Longstreth WT Jr, Janssens AC, Pankratz VS, Dartigues JF, Hollingworth P, Aspelund T, et al. EAD11 Consortium. Genome-wide analysis of genetic loci associated with Alzheimer disease. *JAMA.* 2010;303(18):1832–40.
  59. Kamboh MI, Demirci FY, Wang X, Minster RL, Carrasquillo MM, Pankratz VS, Younkin SG, Saykin AJ, Alzheimer's Disease Neuroimaging Initiative, Jun G, Baldwin C, Logue MW, Buros J, Farrer L, Pericak-Vance MA, Haines JL, Sweet RA, Ganguli M, Feingold E, Dekosky ST, et al. Genome-wide association study of Alzheimer's disease. *Transl Psychiatry.* 2012;2(5):e117.
  60. Bachli MB, Sedeño L, Ochab JK, Piguet O, Kumfor F, Reyes P, Torralva T, Roca M, Cardona JF, Campo CG, Herrera E, Slachevsky A, Matallana D, Manes F, García AM, Ibáñez A, Chialvo DR. Evaluating the reliability of neurocognitive biomarkers of neurodegenerative diseases across countries: a machine learning approach. *Neuroimage.* 2020;208: 116456.
  61. Prado P, Birba A, Cruzat J, Santamaría-García H, Parra M, Moguilner S, Tagliazucchi E, Ibáñez A. Dementia ConnEEGtome: towards multicentric harmonization of EEG connectivity in neurodegeneration. *Int J Psychophysiol.* 2022;172:24–38.

62. Moguilner S, Birba A, Fittipaldi S, Gonzalez-Campo C, Tagliazucchi E, Reyes P, Matallana D, Parra MA, Slachevsky A, Fariás G, Cruzat J, García A, Eyre HA, La Joie R, Rabinovici G, Whelan R, & Ibáñez A. Multi-feature computational framework for combined signatures of dementia in underrepresented settings. *J Neural Eng.* 2022;19(4). <https://doi.org/10.1088/1741-2552/ac87d0>.
63. Moguilner S, Whelan R, Adams H, Valcour V, Tagliazucchi E, Ibáñez A. Visual deep learning of unprocessed neuroimaging characterises dementia subtypes and generalises across non-stereotypic samples. *EBioMedicine.* 2023;90: 104540.
64. Davatzikos C. Machine learning in neuroimaging: Progress and challenges. *NeuroImage.* 2019;197:652–6.
65. Rebala G, Ravi A, Churiwala S. An Introduction to Machine Learning. 1st ed. Springer Publishing Company, Incorporated. 2019.
66. Sarker IH. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Comput Sci.* 2021;2(3):160.
67. Ristevski B, Chen M. Big data analytics in medicine and healthcare. *J Integr Bioinform.* 2018;15(3):20170030.
68. Hariri RH, Fredericks EM, Bowers KM. Uncertainty in big data analytics: survey, opportunities, and challenges. *J Big Data.* 2019;6:44.
69. Leonelli S. Data-Centric Biology: A Philosophical Study. London: University of Chicago Press; 2016.
70. Partington SN, Papakroni V, Menzies T. Optimizing data collection for public health decisions: a data mining approach. *BMC Public Health.* 2014;14:593.
71. Zhang A, Xing L, Zou J, Wu JC. Shifting machine learning for healthcare from development to deployment and from models to data. *Nat Biomed Eng.* 2022;6(12):1330–45.
72. Habehh H, Gohel S. Machine learning in healthcare. *Curr Genomics.* 2021;22(4):291–300.
73. Javaid M, Haleem A, Singh RP, Suman R, Rab S. Significance of machine learning in healthcare: features, pillars and applications. *Int J Intell Netw.* 2022;3:58–73.
74. Reddy S. Explainability and artificial intelligence in medicine. *Lancet Digital health.* 2022;4(4):e214–5.
75. Bogdanovic B, Eftimov T, Simjanoska M. In-depth insights into Alzheimer's disease by using explainable machine learning approach. *Sci Rep.* 2022;12(1):6508.
76. Hayashi Y. The right direction needed to develop white-box deep learning in radiology, pathology, and ophthalmology: a short review. *Front Robotics AI.* 2019;6:24.
77. Carvalho DV, Pereira EM, Cardoso JS. Machine learning interpretability: a survey on methods and metrics. *Electronics.* 2019;8(8):832.
78. Yang G, Ye Q, Xia J. Unbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: a mini-review, two showcases and beyond. *Int J Inf Fusion.* 2022;77:29–52.
79. Krishnan M. Against interpretability: a critical examination of the interpretability problem in machine learning. *Philos Technol.* 2020;33:487–502.
80. Shanthamallu US, Spanias A. Machine and Deep Learning Applications. In: Machine and Deep Learning Algorithms and Applications. Synthesis Lectures on Signal Processing. Cham: Springer; 2022.
81. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, inference, and prediction. 2nd ed. Stanford: Stanford University; 2009.
82. Breiman L, Friedman J, Olshen R, Stone C. Classification and regression trees. Wadsworth: Chapman and Hall; 1984.
83. Cortes C, Vapnik V. Support-vector networks. *Mach Learn.* 1995;20:273–97.
84. Breiman L. Random forests. *Mach Learn.* 2001;45:5–32.
85. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015;521:436–44.
86. Fisher RA. On the mathematical foundations of theoretical statistics. *Philos Trans Royal Soc London, Ser A.* 1922;222:309–68.
87. Tibshirani R. Regression shrinkage and selection via the Lasso. *J R Stat Soc Se B (Methodol).* 1996;58(1):267–88.
88. Hoerl A, Kennard R. Ridge regression: biased estimation for nonorthogonal problems. *Technometrics.* 1970;12:55–67.
89. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature.* 1986;323:533–6.
90. Redolfi A, De Francesco S, Palesi F, Galluzzi S, Muscio C, Castellazzi G, Tiraboschi P, Savini G, Nigri A, Bottini G, Bruzzone MG, Ramusino MC, Ferraro S, Gandini Wheeler-Kingshott CAM, Tagliavini F, Frisoni GB, Rylvlin P, Demonet JF, Kherif F, Cappa SF, et al. Medical Informatics Platform (MIP): a pilot study across clinical Italian cohorts. *Front Neurol.* 2020;11:1021.
91. Sh Y, Liu B, Zhang J, Zhou Y, Hu Z, Zhang X. Application of artificial intelligence modeling technology based on fluid biopsy to diagnose Alzheimer's disease. *Front Aging Neurosci.* 2021;13: 768229.
92. Khatri U, Kwon GR. An efficient combination among sMRI, CSF, cognitive score, and APOE ε4 biomarkers for classification of AD and MCI using extreme learning machine. *Comput Intell Neurosci.* 2020;2020:8015156.
93. Barbará-Morales E, Pérez-González J, Rojas-Saavedra KC, Medina-Bañuelos V. Evaluation of brain tortuosity measurement for the automatic multimodal classification of subjects with Alzheimer's disease. *Comput Intell Neurosci.* 2020;2020:4041832.
94. Martínez-Torteya A, Treviño V, Tamez-Peña JG. Improved diagnostic multimodal biomarkers for Alzheimer's disease and mild cognitive impairment. *Biomed Res Int.* 2015;2015: 961314.
95. Ficià E, Boschi S, Ansari S, D'Agata F, Abollino O, Caroppo P, Di Fede G, Indaco A, Rainero I, Guiot C. Machine learning profiling of Alzheimer's disease patients based on current cerebrospinal fluid markers and iron content in biofluids. *Front Aging Neurosci.* 2021;13: 607858.
96. Jääskeläinen O, Hall A, Tiainen M, van Gils M, Lötjönen J, Kangas AJ, Helisalmi S, Pikkarainen M, Hallikainen M, Koivisto A, Hartikainen P, Hiltunen M, Ala-Korpela M, Soininen P, Soininen H, Herukka SK. Metabolic profiles help discriminate mild cognitive impairment from dementia stage in Alzheimer's disease. *J Alzheimer's Dis.* 2020;74(1):277–86.
97. Olazarán J, Gil-de-Gómez L, Rodríguez-Martín A, Valentí-Soler M, Frades-Payo B, Marín-Muñoz J, Antúnez C, Frank-García A, Acedo-Jiménez C, Morlán-Gracia L, Petidier-Torregrossa R, Guisasaola MC, Bermejo-Pareja F, Sánchez-Ferro Á, Pérez-Martínez DA, Manzano-Palomo S, Farquhar R, Rábano A, Calero M. A blood-based, 7-metabolite signature for the early diagnosis of Alzheimer's disease. *J Alzheimer's Dis.* 2015;45(4):1157–73.
98. Gray KR, Aljabar P, Heckemann RA, Hammers A, Rueckert D, Alzheimer's Disease Neuroimaging Initiative. Random forest-based similarity measures for multi-modal classification of Alzheimer's disease. *Neuroimage.* 2013;65:167–75.
99. Zhao X, Kang J, Svetnik V, Warden D, Wilcock G, David Smith A, Savage MJ, Laterza OF. A machine learning approach to identify a circulating MicroRNA signature for Alzheimer disease. *J Appl Lab Med.* 2020;5(1):15–28.
100. Wang, Bing & Lu, Kun & 龙红明, Hong-ming & Zhou, Yuming & Zheng, Chun-Hou & Zhang, Jun & Chen (陈鹏), Peng. (2018). Early stage identification of Alzheimer's disease using a two-stage Ensemble classifier. *Curr Bioinf.* 13. <https://doi.org/10.2174/1574893613666180328093114>.
101. Miller JB, Kauwe JSK. Predicting clinical dementia rating using blood RNA levels. *Genes.* 2020;11(6):706.
102. Hu WT, Watts KD, Taylor P, Nguyen TP, Howell JC, Lee RC, Seyfried NT, Gearing M, Hales CM, Levey AI, Lah JJ, Lee EK, Alzheimer's Disease Neuro-Imaging Initiative. CSF complement 3 and factor H are staging biomarkers in Alzheimer's disease. *Acta Neuropathol Commun.* 2016;4:14.
103. Yilmaz A, Ugur Z, Bisgin H, Akyol S, Bahado-Singh R, Wilson G, Imam K, Maddens ME, Graham SF. Targeted metabolic profiling of urine highlights a potential biomarker panel for the diagnosis of Alzheimer's disease and mild cognitive impairment: a pilot study. *Metabolites.* 2020;10(9):357.
104. Peña-Bautista C, Durand T, Oger C, Baquero M, Vento M, Cháfer-Pericás C. Assessment of lipid peroxidation and artificial neural network models in early Alzheimer disease diagnosis. *Clin Biochem.* 2019;72:64–70.
105. Dong A, Li Z, Wang M, Shen D, Liu M. High-order Laplacian regularized low-rank representation for multimodal dementia diagnosis. *Front Neurosci.* 2021;15: 634124.
106. Chang CH, Lin CH, Liu CY, Huang CS, Chen SJ, Lin WC, Yang HT, Lane HY. Plasma d-glutamate levels for detecting mild cognitive impairment and Alzheimer's disease: machine learning approaches. *J Psychopharmacol (Oxford, England).* 2021;35(3):265–72.
107. Santangelo R, Masserini F, Agosta F, Sala A, Caminiti SP, Cecchetti G, Caso F, Martinelli V, Pinto P, Passerini G, Perani D, Magnani G, Filippi M. CSF p-tau/Aβ42 ratio and brain FDG-PET may reliably detect

- MCI "imminent" converters to AD. *Eur J Nucl Med Mol Imaging*. 2020;47(13):3152–64.
108. Abate G, Vezzoli M, Polito L, Guaita A, Albani D, Marizzoni M, Garrafa E, Marengoni A, Forloni G, Frisoni GB, Cummings JL, Memo M, Uberti D. A conformation variant of p53 combined with machine learning identifies Alzheimer disease in preclinical and prodromal stages. *J Pers Med*. 2020;11(1):14.
109. Lin W, Gao Q, Du M, Chen W, Tong T. Multiclass diagnosis of stages of Alzheimer's disease using linear discriminant analysis scoring for multimodal data. *Comput Biol Med*. 2021;134: 104478.
110. Devanarayan P, Devanarayan V, Llano DA, Alzheimer's Disease Neuroimaging Initiative. Identification of a simple and novel cut-point based cerebrospinal fluid and MRI signature for predicting Alzheimer's disease progression that reinforces the 2018 NIA-AA Research Framework. *J Alzheimer's Dis*. 2019;68(2):537–50.
111. Iddi S, Li D, Aisen PS, Rafii MS, Thompson WK, Donohue MC, Alzheimer's Disease Neuroimaging Initiative. Predicting the course of Alzheimer's progression. *Brain informatics*. 2019;6(1):6.
112. Lin W, Gao Q, Yuan J, Chen Z, Feng C, Chen W, Du M, Tong T. Predicting Alzheimer's disease conversion from mild cognitive impairment using an extreme learning machine-based grading method with multimodal data. *Front Aging Neurosci*. 2020;12:77.
113. Mathotaarachchi S, Pascoal TA, Shin M, Benedet AL, Kang MS, Beaudry T, Fonov VS, Gauthier S, Rosa-Neto P, Alzheimer's Disease Neuroimaging Initiative. Identifying incipient dementia individuals using machine learning and amyloid imaging. *Neurobiol Aging*. 2017;59:80–90.
114. Gupta Y, Lama RK, Kwon GR, Alzheimer's Disease Neuroimaging Initiative. Prediction and classification of Alzheimer's disease based on combined features from apolipoprotein-E genotype, cerebrospinal fluid, MR, and FDG-PET imaging biomarkers. *Front Comput Neurosci*. 2019;13:72.
115. Zhang D, Shen D, Alzheimer's Disease Neuroimaging Initiative. Multimodal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease. *Neuroimage*. 2012;59(2):895–907.
116. Eke CS, Jammeh E, Li X, Carroll C, Pearson S, Ifeachor E. Early detection of Alzheimer's disease with blood plasma proteins using support vector machines. *IEEE J Biomed Health Inform*. 2021;25(1):218–26.
117. Cheng B, Liu M, Suk HI, Shen D, Zhang D, Alzheimer's Disease Neuroimaging Initiative. Multimodal manifold-regularized transfer learning for MCI conversion prediction. *Brain Imaging Behav*. 2015;9(4):913–26.
118. Escudero J, Ifeachor E, Zajicek JP, Green C, Shearer J, Pearson S, Alzheimer's Disease Neuroimaging Initiative. Machine learning-based method for personalized and cost-effective detection of Alzheimer's disease. *IEEE Trans Biomed Eng*. 2013;60(1):164–8.
119. Fawcett T. An Introduction to ROC Analysis. *Pattern Recognit Lett*. 2006;27:861–74.
120. Dinov ID. Black Box Machine-Learning Methods: Neural Networks and Support Vector Machines. In: *Data Science and Predictive Analytics*. Cham: Springer; 2018.
121. Kokol P, Kokol M, Zagoranski S. Machine learning on small size samples: a synthetic knowledge synthesis. *Sci Prog*. 2022;105(1):368504211029777.
122. Wang H, Zheng H. Model testing, machine learning. In: Dubitzky W, Wolkenhauer O, Cho KH, Yokota H, editors. *Encyclopedia of Systems Biology*. New York: Springer; 2013.
123. Bi Y, Abrol A, Fu Z, Chen J, Liu J, Calhoun V. Prediction of gender from longitudinal MRI data via deep learning on adolescent data reveals unique patterns associated with brain structure and change over a two-year period. *J Neurosci Methods*. 2023;384: 109744.
124. Campbell NL, Unverzagt F, LaMantia MA, Khan BA, Boustani MA. Risk factors for the progression of mild cognitive impairment to dementia. *Clin Geriatr Med*. 2013;29:873–93.
125. Dunne RA, Aarsland D, O'Brien JT, Ballard C, Banerjee S, Fox NC, et al. Mild cognitive impairment: the manchester consensus. *Age Ageing*. 2021;50:72–80.
126. Bonilla-Santos J, Zea-Romero E, González-Hernández A, Cala-Martínez D. Cognitive, biological, anatomical and behavioral markers of mild cognitive impairment and Alzheimer's disease. A systematic review. *Ecuat Neurol*. 2021;30(2):57–67.
127. Doecke JD, Laws SM, Faux NG, Wilson W, Burnham SC, Lam CP, Mondal A, Bedo J, Bush AI, Brown B, De Ruycck K, Ellis KA, Fowler C, Gupta VB, Head R, Macaulay SL, Pertile K, Rowe CC, Rembach A, Rodrigues M, et al. Australian Imaging Biomarker and Lifestyle Research Group. Blood-based protein biomarkers for diagnosis of Alzheimer disease. *Arch Neurol*. 2012;69(10):1318–25.
128. Weinstein D, Leininger J, Hamby C, Safai B. Diagnostic and prognostic biomarkers in melanoma. *J Clin Aesthetic Dermatol*. 2014;7(6):13–24.
129. Wang J, Knol MJ, Tiulpin A, Dubost F, de Bruijne M, Vernooij MW, Adams HHH, Ikram MA, Niessen WJ, Roshchupkin GV. Gray matter age prediction as a biomarker for risk of dementia. *Proc Natl Acad Sci USA*. 2019;116(42):21213–8.
130. O'Driscoll C, Shaikh M. Cross-cultural applicability of the Montreal Cognitive Assessment (MoCA): a systematic review. *J Alzheimer's Dis*. 2017;58(3):789–801.
131. Gagnon LG, Belleville S. Working memory in mild cognitive impairment and Alzheimer's disease: contribution of forgetting and predictive value of complex span tasks. *Neuropsychology*. 2011;25(2):226–36.
132. Peña-Bautista C, Baquero M, Ferrer I, Hervás D, Vento M, García-Blanco A, Cháfer-Pericás C. Neuropsychological assessment and cortisol levels in biofluids from early Alzheimer's disease patients. *Exp Gerontol*. 2019;123:10–6.
133. Cheng L, Doecke JD, Sharples RA, Villemagne VL, Fowler CJ, Rembach A, Martins RN, Rowe CC, Macaulay SL, Masters CL, Hill AF, Australian Imaging, Biomarkers and Lifestyle (AIBL) Research Group. Prognostic serum miRNA biomarkers associated with Alzheimer's disease shows concordance with neuropsychological and neuroimaging assessment. *Mol Psychiatry*. 2015;20(10):1188–96.
134. Ardila A. Cross-cultural neuropsychology: history and prospects. *RUDN J Psychol Pedagogics*. 2020;17(1):64–78.
135. 2021 Alzheimer's disease facts and figures. *Alzheimers Dement*. 2021;17(3):327–406.
136. Rosselli M, Uribe IV, Ahne E, Shihadeh L. Culture, ethnicity, and level of education in Alzheimer's disease. *Neurotherapeutics*. 2022;19(1):26–54.
137. Parra MA, Orellana P, Leon T, Victoria CG, Henriquez F, Gomez R, Avalos C, Damian A, Slachevsky A, Ibañez A, Zetterberg H, Tijms BM, Yokoyama JS, Piña-Escudero SD, Cochran JN, Matallana DL, Acosta D, Allegri R, Arias-Suárez BP, Barra B, et al. Biomarkers for dementia in Latin American countries: gaps and opportunities. *Alzheimer's Dement*. 2023;19(2):721–35.
138. Luria AR, Vygotsky LS. Ape, primitive man and child, 1930/1992. *Great Britain: Harvester Wheatsheaf*; 1930/1992.
139. Vygotsky LS. "Psikhologija i uchenije o localizacii psikhicheskikh funktsii" in L.S. Vygotsky. *Sobranije sochinenii*. Vol. 1 *Voprosy teorii i istorii psikhologii*. eds. A. R. Luria and Jaroshevskii (Moscow: Pedagogika), 168–174. (Original work published in 1934); 1934/1982.
140. Poortinga YH. Equivalence of cross-cultural data: an overview of basic issues. *Int J Psychol*. 1989;24(6):737–56.
141. Pollet TV, Tybur JM, Frankenhuys WE, Rickard JJ. What can cross-cultural correlations teach us about human nature? *Hum Nat (Hawthorne, NY)*. 2014;25(3):410–29.
142. Parra MA, Garcia AM, Ibanez A Sr, LAC-CD. Addressing dementia challenges through international networks: evidence from the Latin American and Caribbean Consortium on Dementia (LAC-CD). *Alzheimer's Dement*. 2021;17(Suppl 8): e055106.
143. Luria AR. *The human brain and psychological processes*. New York: Harper & Row; 1966.
144. Gigerenzer G, Marewksi JN. Surrogate science: the idol of a universal method for scientific inference. *J Manag*. 2015;41(2):421–40.
145. Pearl J. Theoretical impediments to machine learning with seven sparks from the causal revolution. In: *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining (WSDM '18)*. New York: Association for Computing Machinery; 2018. p. 3.
146. Richens JG, Lee CM, Johri S. Improving the accuracy of medical diagnosis with causal machine learning. *Nat Commun*. 2020;11(1):3923.
147. McLachlan S, Dube K, Hitman GA, Fenton NE, Kyrimi E. Bayesian networks in healthcare: distribution by medical condition. *Artif Intell Med*. 2020;107: 101912.
148. Parra MA, Baez S, Sedeño L, Gonzalez Campo C, Santamaría-García H, Arahamian I, Bertolucci PH, Bustin J, Camargos Bicalho MA, Cano-Gutierrez C, Caramelli P, Chaves MLF, Cogram P, Beber BC, Court FA, de Souza LC, Custodio N, Damian A, de la Cruz M, Diehl Rodriguez R, et al.

Dementia in Latin America: paving the way toward a regional action plan. *Alzheimer's Dement.* 2021;17(2):295–313.

149. Parra MA, Baez S, Allegri R, Nitrini R, Lopera F, Slachevsky A, Custodio N, Lira D, Piguet O, Kumfor F, Huepe D, Cogram P, Bak T, Manes F, Ibanez A. Dementia in Latin America: assessing the present and envisioning the future. *Neurology.* 2018;90(5):222–31.

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

