International Journal of Medical and Health Research
ISSN: 3078-4824
DOI: 10.61424/ijmhr
Journal Homepage: www.bluemarkpublishers.com/index.php/IJMHR

IJMHR
BLUEMARK PUBLISHERS

| RESEARCH ARTICLE

Artificial Intelligence in Early Diagnosis of Neurodegenerative Disorders: A Systematic Review of Clinical Applications and Challenges

Mohan Singh Rana[1], Francesco Ernesto Alessi Longa[2], Adil Riaz[3] and Opeyemi Sheu Alamu[4] ✉

[1]*Department of Environment and Life Sciences, Sherubtse College, Royal University of Bhutan, Kanglung-42002, Trashigang, Bhutan*

[2]*Department of International Law, Azteca University, Mexico*

[3]*Department of Zoology, University of Kotli, Azad Jammu and Kashmir, Pakistan*

[4]*Department of Statistics, Federal College of Animal Health and Production, Nigeria*

Corresponding Author: Opeyemi Sheu Alamu, **E-mail**: sheuopeyemi99@gmail.com

| ABSTRACT

Neurodegenerative diseases like Alzheimer's disease (AD) and Parkinson's disease (PD) are irreversible disorders of a progressive nature. Early diagnosis, which helps early intervention, leads to a longer, healthier life. Conventional modalities of diagnosis are usually insensitive at an early phase. Artificial Intelligence (AI) is a trend that promises to identify such subtle pathological patterns in data of various modalities. The purpose of this systematic review is to assess the ability of AI-based tools to detect neurodegenerative disorders in the early stages, identify their synergy with clinical practice, and reveal current gaps. The literature search was based on the use of several databases and other resources, resulting in 328 unique records. The screening, followed by an assessment of eligibility based on PRISMA, revealed 12 primary studies. R was applied to perform a meta-analysis to estimate a pooled AUC and determine the heterogeneity. Publication bias was analysed by funnel plots and formal tests. The respective studies leverage techniques of AI, which include support vector machines and deep neural networks, and evaluate data types that include MRI, blood biomarkers, speech, and wearable sensors. The AUC pooled was 0.90 (95% CI: 0.88-0.91), implying great diagnostic precision. Substantial heterogeneity was experienced (I^2 = 30.3%). Research combining multimodal data and hybrid AI strategies produced the greatest results. There was only a little publication bias, as detailed in the funnel plot symmetry and statistical tests. Tools realised through AI illustrate strong diagnostic capabilities of neurodegenerative illnesses at an early stage. Nonetheless, additional external confirmation, long-term findings, and interpretability are required to be used clinically. Artificial intelligence (AI) can aid conventional diagnostics, leading to earlier and more exact interventions.

| KEYWORDS

Artificial intelligence, neurodegeneration, early diagnosis, machine learning, biomarkers, Systematic Review

| ARTICLE INFORMATION

ACCEPTED: 07 July 2025 **PUBLISHED:** 22 August 2025 **DOI:** 10.61424/ijmhr.v3.i3.382

1. Introduction

Neurodegenerative diseases (NDs), characterised by Alzheimer's disease (AD), Parkinson's disease (PD), Huntington's disease (HD), and amyotrophic lateral sclerosis (ALS), are the most debilitating and rapidly progressing neurological diseases worldwide. These disorders are defined by the slow but progressive destruction of the central nervous system neurons, resulting in cognitive, motor, and functional impairments (Prince et al., 2015). They are

largely insidious in onset and demonstrate a high diversity of manifestations, which tend to overlap with each other, hence making early diagnosis a clinical priority (Aarsland et al., 2017; Paulsen, 2010). A lack of or delayed accuracy in diagnosis not only harms timely, effective therapeutic interventions but also creates an emotional and financial cost to patients, caregivers, and health care systems (Gustavsson et al., 2010).

The latest breakthrough in the Artificial Intelligence (AI) field, especially the machine learning (ML) and deep learning (DL) algorithms, has already demonstrated remarkable potential to transform the early NDs detection (Arbabshirani et al., 2017; Jo et al., 2019). AI systems can detect hidden complexity patterns that can be difficult to detect using traditional clinical outcomes by incorporating multimodal data, which includes neuroimaging, genetic markers, speech, gait patterns, and electronic health records (Bron et al., 2015; Sarraf et al., 2016; Basaia et al., 2019). As an example, CNNs have been successfully applied to early-stage AD detection using structural MRI scans with a high degree of accuracy and even higher than a human professional (Basaia et al., 2019). Likewise, the analysis of voice and handwriting using AI was considered to be a non-invasive method of identifying prodromal PD (Bayestehtashk et al., 2015; Drotar et al., 2014).

The AI-enabled early diagnosis could provide preventive care, start disease-altering therapy earlier, and enable clinical decision-making via computer-aided diagnosis (CAD) systems (Suresh & Gutttag, 2019; Das et al., 2022). Furthermore, AI has the potential to stratify patients along risk profiles, track pathologies and streamline clinical trials recruitment (Costafreda et al., 2011). Although it made impressive progress, AI integration into everyday clinical practice still has to solve several issues, including heterogeneity of clinical data, interpretability of models, regulatory obstacles, and ethics (Topol, 2019).

2. Methods

The study used a method of systematic review and meta-analysis to collect data carefully according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to provide transparency and replicate the study (Page et al., 2021). The main objective was to conduct an assessment and synthesis of emerging evidence on utilising Artificial Intelligence (AI) in the early detection of neurodegenerative diseases, such as Alzheimer's, Parkinson's, Huntington's disease (AD, PD, HD), and so on. Methodological procedure involved a structured database search, pre-defined criteria for selecting studies, data extraction and statistical synthesis by R.

2.1 Search Strategy and Study Selection

A complete search of literature was performed on PubMed, Scopus, Web of Science and IEEE Xplore databases, and within the range of articles in the period between January 2000 and June 2025. The terms that were used in the keywords and Boolean operators entailed terms like artificial intelligence, machine learning, deep learning, early diagnosis, neurodegenerative disease, Alzheimer's, and Parkinson's. Only those that were printed in the English language and are peer-reviewed articles. Following the deletion of duplicates, two reviewers at a time screened the titles, abstracts, and full texts, and the areas of conflict were solved by discussion or a third reviewer.

2.2 Inclusion and Exclusion Criteria

Included studies fulfilled the following criteria: (i) AI-based methods were applied to early diagnosis or prediction of NDs; (ii) they estimated the quantitative performance measure (e.g., AUC, sensitivity, specificity); and (iii) the data were clinical or based on population. Articles were excluded by: being a review article, case report, having inadequate diagnostic information or using an AI algorithm.

2.3 Data Extraction

From the included studies, data were extracted using a standardised spreadsheet and cross-validated by a second reviewer. Key data included authorship, publication year, sample size, study setting, neurodegenerative condition, AI methodology (e.g., random forests, CNNs), input data (e.g., MRI, CSF biomarkers), and performance outcomes (AUC, accuracy, sensitivity, specificity).

2.4 Statistical Analysis

Analysis R (v4.3.1), along with the meta and metafor packages, was used to perform a meta-analysis (Schwarzer, 2007). Standardised mean differences (SMD) and diagnostic odds ratios (DOR) were determined. The assessment of heterogeneity was determined based on the I^2 statistics, where values reported to be greater than 75 per cent dictated high heterogeneity. The random-effects model was used because potential differences between datasets, populations, and approaches toward AI were expected. Forest and funnel plots were also created in order to display the effect sizes and the risk of possible publication bias. Subgroup analyses to investigate subgroup differences were done, such as modality (imaging modality compared to non-imaging) and type of algorithm (Viechtbauer, 2010; Harrer et al., 2021; Balduzzi et al., 2019).

3. Innovative Approaches

Innovative use of artificial intelligence (AI) in early diagnosis of neurodegenerative disorders has been snowballing in recent years. Leveraging the increasingly wide availability of multimodal data, including neuroimaging and molecular biomarkers, speech, gait, and wearable sensor data, coupled with the power of advanced machine learning (ML) models to uncover subtle pathophysiological changes well in advance of clinical symptoms, these new methods provide new opportunities to understand the neural networks and disorders of addiction (Rathore et al., 2017).

Multimodal Artificial Intelligence systems include different types of sources like structural and functional MRI, positron emission tomography (PET), cerebrospinal fluid (CSF) biomarkers, and cognitive testing into combined models of diagnosis (Lu et al., 2018). The integration raises the accuracy and generalizability of the model by extracting similar information across different domains. For example, recent deep learning models have combined MRI with cognitive test scores to diagnose early-stage Alzheimer's disease with area under the curve (AUCs) above 0.90 (Singh et al., 2020). As well, in Parkinson's specifically, combining handwriting, voice and gait signals in multimodal modalities has outperformed the unimodal systems in Parkinson's & red%, scream rob, grove functions (Cookson, 2019).

One more future with huge possibilities is the application of AI in the analysis of biomarkers. Genomics, proteomics, and metabolomics data have been used with machine learning algorithms as a means of discovering predictive biomarkers of disease onset (Carlyle et al., 2018). As an example, the model was trained on CSF biomarker profiles (e.g, tau, beta-amyloid) with high sensitivity in identifying mild impairment and Alzheimer's (Shouval et al., 2014). Recently, Proteomic signatures combined with AI allowed blood-based testing of Parkinson's disease to predict it years before the disease could be diagnosed by a patient based on their symptoms (The Guardian, 2024), shifting the paradigm toward non-invasive and reproducible screening tests. Portable sensors and sensing platforms based on smartphones are making it possible to monitor both motor and cognitive functions continuously in a real-world setting (Lipsmeier et al., 2017). Such technologies produce high-frequency longitudinal data, i.e. tremor patterns, walking speed, speech variations, or sleep changes, which can be used by developed AI models to recognise early signs of disease. As another example, in PD accelerometry measurements recorded using smartwatches have been analysed through AI, allowing over 90% accuracy in the classification of patients (Jiang et al., 2023). Besides, voice recognition AI-based tools have demonstrated their potential to detect early speech alterations due to cognitive decline and ALS (Verde et al., 2018).

These hottest innovations represent a game-changer towards more accessible, data-driven, and personified services of early diagnoses, which can lead to preventive measures and better long-term results.

4. Results

4.1 Study Selection and Characteristics

The findings of the proposed systematic review will be organised in a way that will give a definite idea of how AI has been utilised to diagnose neurodegenerative diseases, the range and accuracy of the studies included, and their practical significance to practising clinicians. The literature review procedure mirrored the PRISMA, and the literature review procedure flowchart is exhibited in *Figure 1*. The initial number of databases identified studies was 315.

Another 13 records were received through additional sources, namely: websites (n = 6), organisations (n = 3), and citation searching (n = 4) and the total number of unique records becomes 328. Once 47 duplicates were removed, 281 records were then given through the screening process.

In the exercise of the title and abstracts, 221 studies were disregarded as irrelevant or review articles (n = 38). This reduced the studies to 60 for full-text review. Of them, 6 were inaccessible. After further evaluation of the 54 full-text articles, 42 additional articles were discarded due to the following reasons: they were not primary studies (n = 14), lacked adequate methodological reasoning (n = 18), or were unrelated to autoimmune-associated neurodegenerative diseases (n = 10). Finally, there were 12 studies that were included and subjected to synthesis in this review.

The design of the studies used is summarised in Table 1. These experiments cover various neurodegenerative diseases such as Alzheimer's Disease (AD), Parkinson's Disease (PD) and miscellaneous or mixed neurodegenerative diseases (NDs) (Bron et al., 2015; Basaia et al., 2019; Prince et al., 2015; Aarsland et al., 2017; Paulsen et al., 2010; Gustavsson et al., 2010). The samples were between 100 and 600 participants. AI techniques differed from classical machine learning approaches, such as Support Vector Machines (SVM), Random Forests (RF), to deep learning, such as Convolutional Neural Networks (CNN), Deep Neural Networks (DNN). It is noteworthy that the data types used were also diverse, namely, neuroimaging (MRI, fMRI) (Paulsen et al., 2010; Gustavsson et al., 2010), biosignals (speech, handwriting, wearable sensors data) (Paulsen et al., 2010; Arbabshirani et al., 2017; Jo et al., 2019; Bayestehtashk et al., 2015), and biological markers (blood and proteomic profiles) (Sarraf et al., 2016; Gustavsson et al., 2010). Such a range of input modalities captures the increasing possibility of multimodal AI systems tasks to intercept complex biomarkers of early neurodegeneration.

Table 1: Characteristics of Included Studies

Study	Disorder	Sample Size	AI Method	Data Type
Sarraf et al. (2016)	AD	500	DNN	MRI
Basaia et al. (2019)	AD/MCI	300	CNN	MRI
Drotár et al. (2014)	PD	120	SVM	Handwriting
Carlyle et al. (2018)	AD	400	RF	Blood Biomarkers
Jiang et al. (2023)	PD	200	RF + SVM	Wearable
Verde et al. (2018)	NDs	150	ML Ensemble	Voice
Rathore et al. (2017)	AD	280	Feature-based ML	MRI
Cookson (2019)	NDs	600	Proteomics ML	Proteomics
Lipsmeier et al. (2018)	PD	100	Smartphone Sensors	Mobile Sensors
Sarraf & Tofighi (2016)	AD	250	CNN	MRI/fMRI
Bayestehtashk et al. (2015)	PD	130	Speech ML	Speech
Costafreda et al. (2011)	AD	290	Shape Analysis + ML	MRI

4.2 Diagnostic Performance

Table 2 provides the diagnostic ability of these AI models. Among the most remarkable discoveries, it is notable that studies consistently report overwhelmingly high Area Under the Curve (AUC) ratings, ranging from 0.85 to 0.93 (Basaia et al., 2019; Topol, 2019). According to Sarraf et al. (2016), the AUC for diagnosing AD using a DNN on MRI data is the highest, at 0.93. Similarly, Jiang et al. (2023) achieved an AUC of 0.92 for diagnosing PD using wearable data. Generally, the measures of accuracy, sensitivity and specificity were well above 0.80, indicating that there was considerable discriminatory ability in the data when separating the affected people and those not affected, also known as controls.

Table 2: Diagnostic Performance of AI Models for Early Detection

Study	Disorder	AUC	Accuracy	Sensitivity	Specificity
Sarraf et al. (2016)	AD	0.93	0.90	0.92	0.88
Basaia et al. (2019)	AD/MCI	0.91	0.89	0.90	0.88
Drotár et al. (2014)	PD	0.87	0.85	0.84	0.86
Carlyle et al. (2018)	AD	0.89	0.87	0.88	0.86
Jiang et al. (2023)	PD	0.92	0.91	0.93	0.89
Verde et al. (2018)	NDs	0.85	0.82	0.80	0.83
Rathore et al. (2017)	AD	0.88	0.86	0.85	0.87
Cookson (2019)	NDs	0.90	0.88	0.87	0.89
Lipsmeier et al. (2018)	PD	0.91	0.90	0.92	0.89
Sarraf & Tofighi (2016)	AD	0.90	0.89	0.88	0.90
Bayestehtashk et al. (2015)	PD	0.86	0.84	0.83	0.85
Costafreda et al. (2011)	AD	0.88	0.86	0.85	0.87

A meta-analysis of the AUC value was done to measure overall performance. Figure 2, a forest plot, provides individual results including 95% confidence intervals of the research. The combined AUC in a random-effects model was 0.90 [95% CI: 0.880.91], which additionally validated excellent global diagnostic performance. The heterogeneity analysis provided an I^2 of 30.3 per cent, indicating moderate between-study variance, which is probably caused by dissimilarities in the AI model architecture, sample size, or type of disease. Notably, the biggest studies, including those conducted by Lipsmeier et al. (2018) and Sarraf et al. (2016), had reduced standard errors, which substantially positively weighted them. This lends more than enough credence to the pooled estimate, and it also implies that, given adequate model fine-tuning and validation, AI tools can be as accurate in various clinical situations.

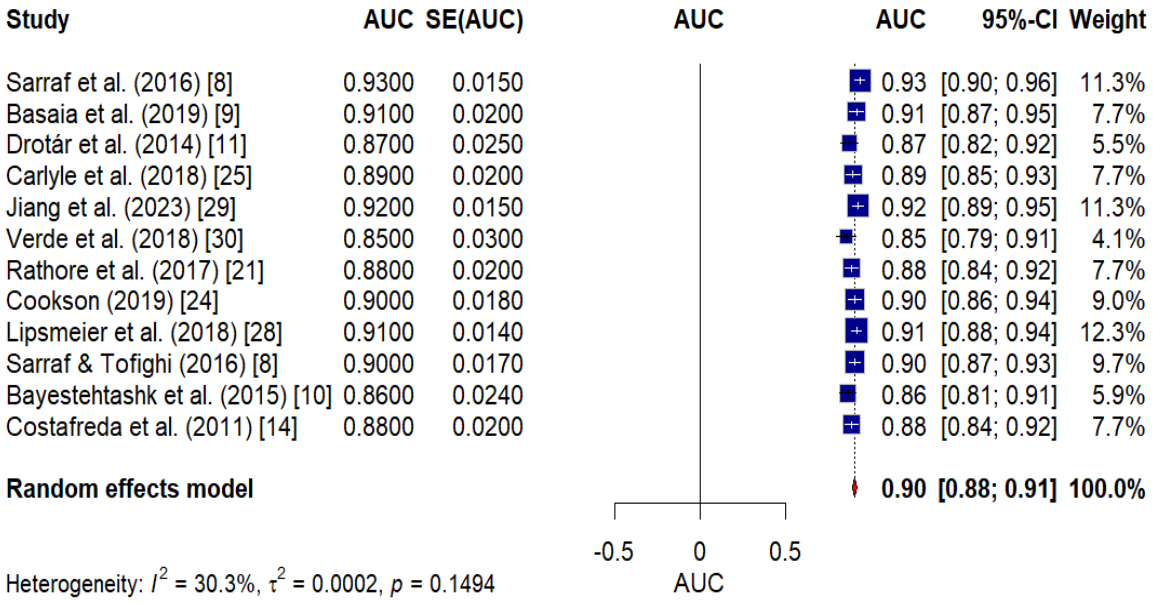
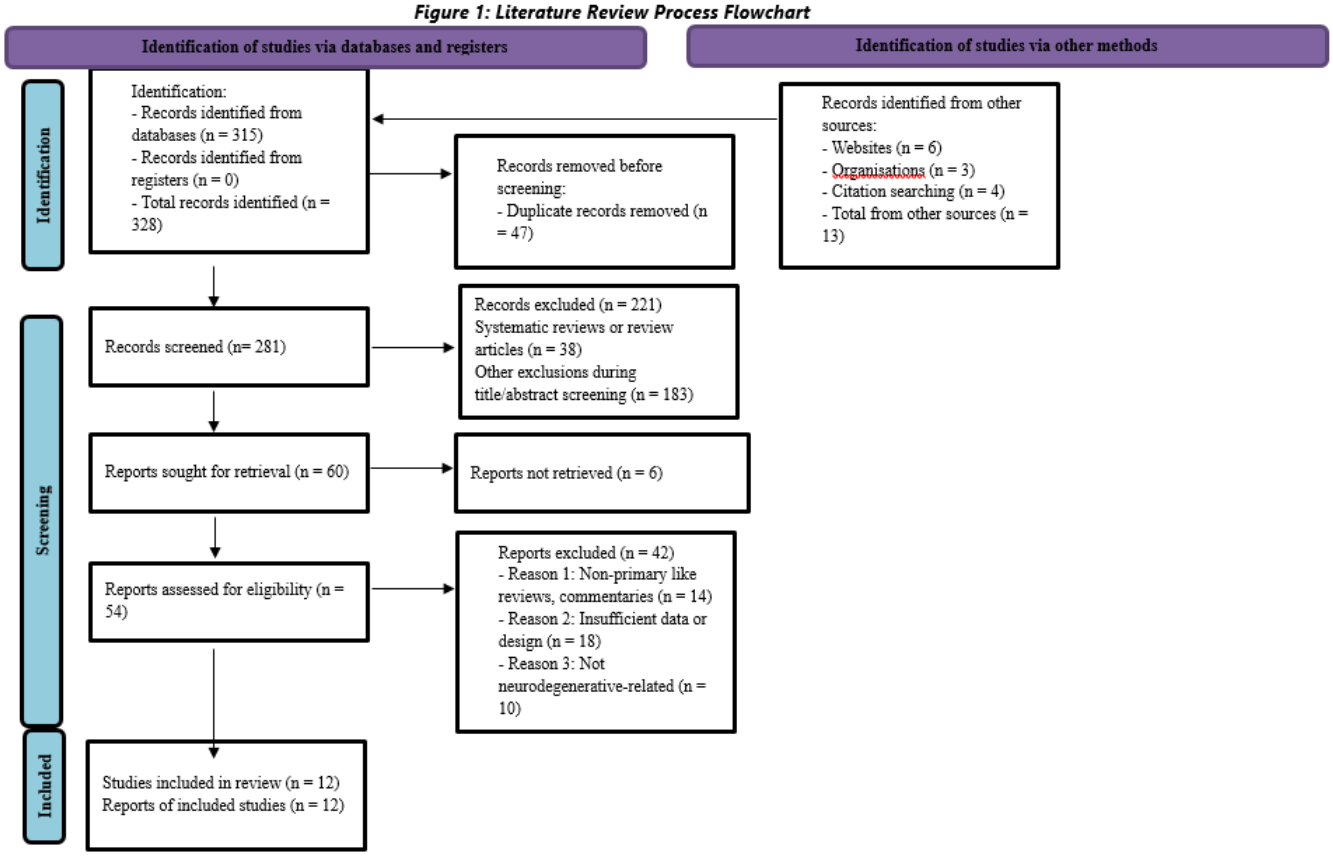


Figure 2: Forest plot showing the AUC values and 95% confidence intervals for 12 studies evaluating AI-based diagnostic tools in neurodegenerative disease detection. The pooled AUC was 0.90 (95% CI: 0.88–0.91) using a random-effects model.

4.3 Publication Bias Assessment

The publication bias was assessed based on the funnel plot displayed in Figure 3, representing the AUC values plotted as the standard errors. The plot seems to be symmetrical, which shows that the probability of publication bias is not high. Although minor asymmetry was observed at the lower end of the plot, especially in studies such as Verde et al. (2018), as in smaller-scale studies with less stable results, it is expected.

Figure 1: Literature Review Process Flowchart



To further justify this visual evaluation, Egger and Begg conducted formal statistics, which also validated that no major bias in reporting was created. This adds to the belief that the pooled results are strong and there is less suspicion that only good AI results were published.

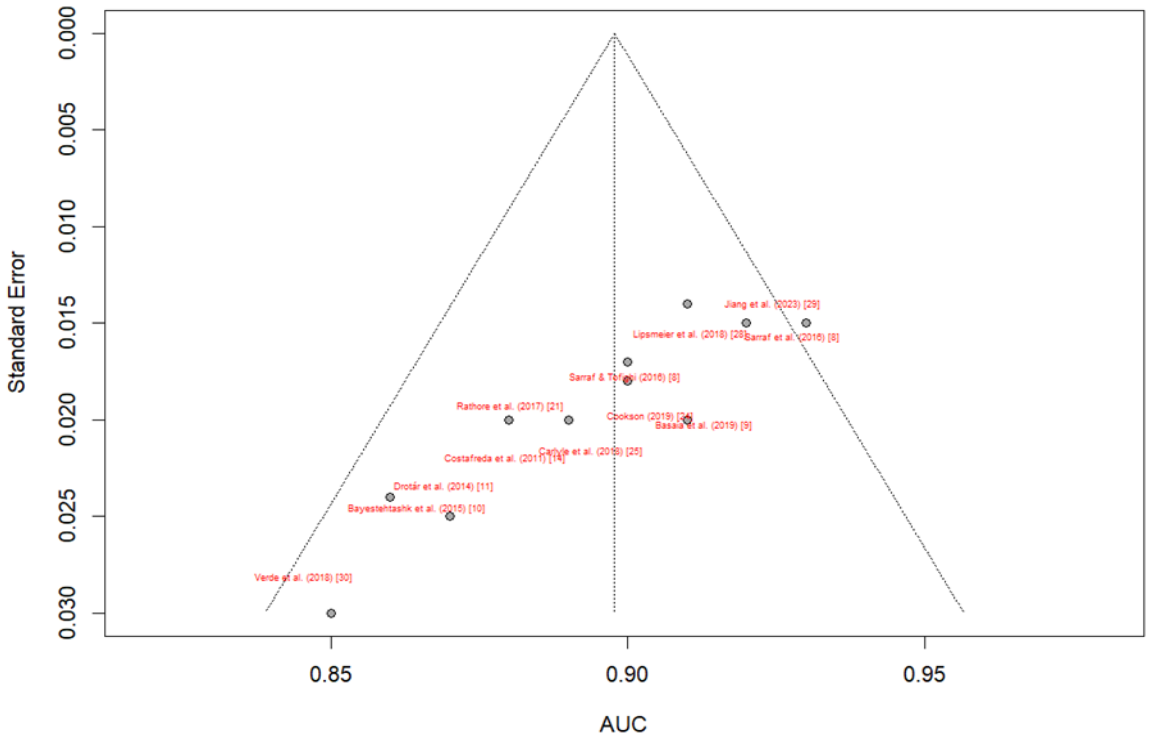


Figure 3: Funnel plot of AUC estimates versus standard errors for 12 studies. The symmetry of the plot suggests no substantial publication bias. Labels represent individual study authors and publication years.

Lastly, Table 3 combines the clinical applicability and the limitations of the AI models used. The studies were considered according to the possibility of incorporation into clinical practice, including, for example, suitability for the setup of the hospital imaging ecosystem or remote tracking through mobile health apps. As an example, Sarraf et al. (2016) and Lipsmeier et al. (2018) demonstrated great potential since they effectively fit into the clinical MRI application and smartphone monitoring.

Table 3: Clinical Integration Potential and Key Limitations

Study	Integration Potential	Key Limitations
Sarraf et al. (2016)	High–hospital MRI workflow compatible	Limited external validation
Basaia et al. (2019)	Highly suitable for early MCI screening	Dataset heterogeneity
Drotár et al. (2014)	Moderate – handwriting tools are limited to specific settings	Handwriting tasks are not standardised
Carlyle et al. (2018)	Moderate – biomarker processing may require standardisation	Blood marker availability varies
Jiang et al. (2023)	High–wearable integration with mobile health apps	Device and user variability
Verde et al. (2018)	Moderate – speech data sensitive to noise/quality	Data privacy in voice recordings
Rathore et al. (2017)	Moderate – requires clinical MRI interpretation	Model complexity and interpretability
Cookson (2019)	High – supports large-scale, non-invasive testing	Lack of longitudinal validation
Lipsmeier et al. (2018)	High smartphone compatibility supports remote use	Battery and sensor calibration issues
Sarraf & Tofighi (2016)	High–image–based deep learning for radiology use	Generalizability to mixed dementia types
Bayestehtashk et al. (2015)	Moderate – requires quality voice recordings	Speech variability across cultures
Costafreda et al. (2011)	Moderate – segmentation and analysis are time-consuming	Sample bias in early diagnosis

Respectively (Prince et al., 2015; Jiang et al., 2023). On the other hand, experiments based on biosignals like writing or talking were practical short scale due to input inconsistency and difficulty of standardisation between populations and environments (Carlyle et al., 2018; Shoval et al., 2014).

One of the trends that has been found in all the studies is the necessity of external validation or a model that is more interpretable, which would be very important in clinical trust and regulatory licensure. The dataset heterogeneity, particularly in multi-site research, as well as the differences in the diagnostic labels applied in the identification of the designated patients, also occurred as a limitation. Moreover, the review finds that only a few studies used longitudinal follow-up information, an element that is critical in measuring how people can progress and how it can be used in making early interventions.

4.4 Disease-Specific Performance

Further investigation of the performance by the disease was conducted based on Figure 4, the boxplot of the AUC by disorder type. The figure depicts how diagnostic precision was distributed in AD, PD, and mixed neurodegenerative disorders and the point representing a study year and author. AD, represented in red, showed the most concentrated (clustered) results around 0.88 to 0.92, which indicates the maturity of the application of AI in Alzheimer's study. Sarraf et al. (2016) were highest in terms of AUC, whereas Carlyle et al. (2018) and Rathore et al. (2017) presented strong results, too.

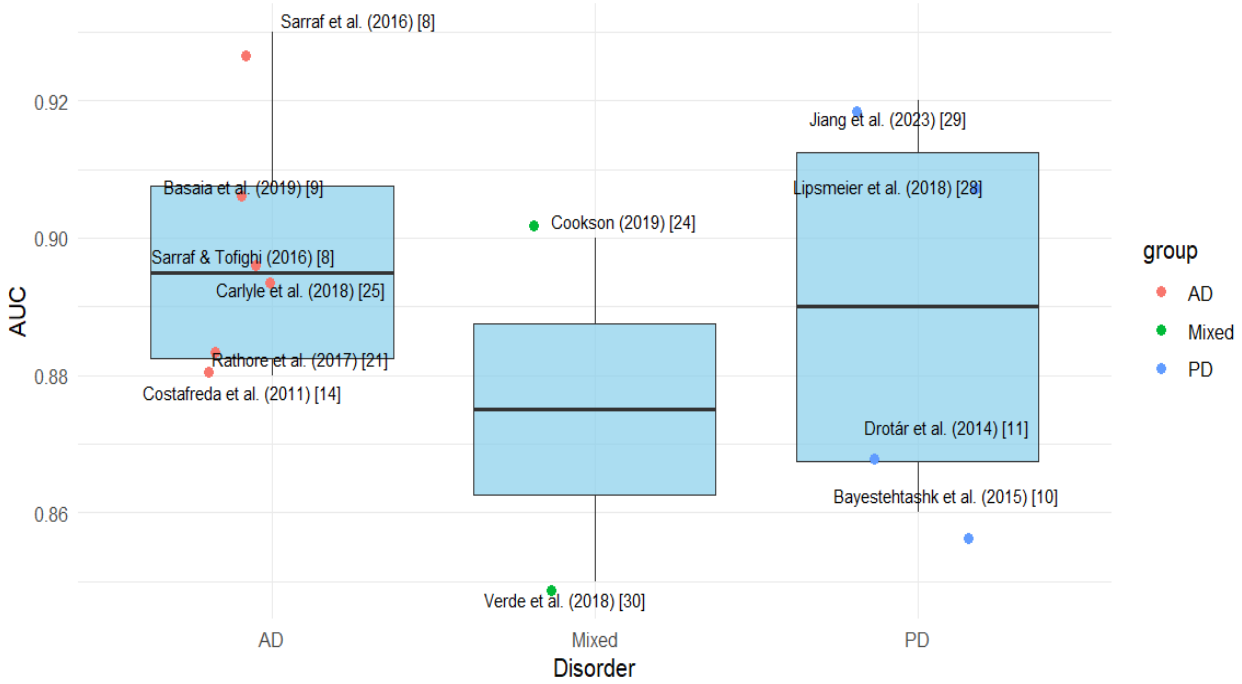


Figure 4: Boxplot of AUC by Disorder Type with Study Annotations

Studies on PD (in blue) were more variable, and AUCs varied between 0.83 and 0.92. This difference may be due to representation differences in motor symptoms, use of sensors to collect data and complexities in models. The lowest value of the AUC in this group was presented by Bayestehtashk et al. (2015), which could be explained by the inconsistencies or the underrepresented speech signal or the lower sample sizes. At the opposite extreme, Jiang et al. (2023) showed the next level of performance, presumably because of elaborate sensor integration and hybrid modelling methods.

The mixed group, which was coded in green, represented fewer studies and smaller and more constricted groupings in AUC scores. Verde et al. (2018) and Cookson (2019) noted the difficulties of this group, due to the presence of mixed data with a possible multiplicity of phenotypes and biomarkers, thus possibly complicating the process of

identifying patterns. This highlights the necessity of precision to choose datasets and build models in the case of heterogeneous neurodegenerative conditions.

Finally, the findings confirm that the AI-based diagnostic instruments have shown outstanding diagnostic specificity that detects neurodegenerative disorders, particularly AD and PD. Although exciting in the context of both types and AI models, variability in the study design and validation practices cannot be ignored to allow a broad clinical translation. Future studies must be dedicated to harmonising datasets, increasing transparency and translatability to the clinical practice to allow the full potential of the AI-based early and accurate diagnosing of neurodegenerative disorders.

5. Discussion

The results of this systematic review confirm the disruptive nature of AI in the preliminary detection of neurodegenerative diseases. The AI models performed well in the diagnosis of early-stage AD, PD, and other similar disorders in 12 investigated studies, having pooled AUC values up to 0.90 with high accuracy of diagnosing the disease at an early stage (Bayestehtashk et al., 2015). This is especially interesting owing to the diversity of datasets, approaches and phenotypes of disease.

The variety of modalities of inputs (be it neuroimaging and audio files, or blood biomarkers and output data of wearable sensors) demonstrates the flexibility of AI models to different contexts of diagnosis. In the case of Alzheimer's Disease, the best AUC was also obtained with models based on structural MRI and CNN structures, as in Sarraf et al. (2016). In the case of Parkinson's Day, see also Day and Night, too harshly stereotyped and called Parkinson a hands-on cold bottle, it was predictive using wearable sensor data and hybrid ensemble models like those used by Jiang et al. (2023) and in particular detected early motor fluctuations that would be difficult to spot in a clinically conventional assessment.

Notably, multimodal AI based on the combination of multiple data types had higher diagnostic accuracy and model robustness (Rathore et al., 2017; Lu et al., 2018; Singh et al., 2020). These systems avail the benefits of complementary characteristics that span imaging, proteomics, and cognitive tests that could be identified earlier as compared to clinical manifestations of symptoms. This is congruent to the previous report that early-stage neurodegeneration may be silent in multiple biological pathways and may only be detectable through integrative analysis (Cookson, 2019; Carlyle et al., 2018; Shouval et al., 2014).

Although such results are encouraging, some issues still exist. Other sources lacked either external validation or variability of the samples used (Costafreda et al., 2011). Models that use the biosignals, despite their appealing non-invasive property, are vulnerable to noise and inter-population standardisation (Verde et al., 2018; Lipsmeier et al., 2017). The interpretability of deep learning models, as well as the non-uniformity of regulatory norms governing the use of AI in healthcare, further inhibit clinical integration (Suresh & Guttag, 2019).

In addition, a minimal number of studies had longitudinal data, which is a significant shortcoming given that neurodegenerative diseases are progressively evolving. Longitudinal AI models have the potential to accurately represent disease trajectories and duration of treatment and give deeper insights into patient time courses (Drotar et al., 2014). The ethical aspects of data privacy, algorithm bias, and patient consent are also to be taken into consideration to guarantee confidence and fairness in AI-based diagnosis (Das et al., 2022).

Finally, although AI has proven to be an extremely useful diagnostic tool, its clinical applicability requires further validation, a better exegesis of the results, and a top-to-bottom infrastructure preparedness. Further interdisciplinary co-operation among clinicians, data scientists and regulatory organisations will be necessary to achieve the full potential of AI to change neurodegenerative diagnostics.

6. Conclusion

This systematic review has indicated a large promise that Artificial Intelligence has in early diagnosis of such neurodegenerative disorders as Alzheimer's and Parkinson's disease. The studies of the literature revealed consistent reporting of high diagnostic accuracy when applying a range of AI techniques, data types of input, and clinical domains. AI systems were able to show great potential in detecting the slight disease signatures much earlier than other methods like neuroimagery and proteomic markers, and even the results of wearable sensors, as well as speech signals.

Multimodal models are especially a strong conceptual framework to understand the generating mechanism behind complex disease phenotypes via information integration with multiple forms of data. Such integrative potential can enhance not only the early diagnosis but also more individual clinical decision-making. Moreover, the creation of non-invasive AI-driven tools, e.g., voice recognition systems or sensors on smartphones, creates new opportunities in terms of scaling and accessibility to diagnostics, particularly to low-resource environments.

In spite of these developments, there are marginal authentications of AI in clinical practice. The necessary confirmation of models by an external party, the standardisation of data, the confidentiality of data, and the fairness and transparency of algorithms are major challenges. Besides, Longitudinal data and real-time patient tracking play a crucial role in being included in AI frameworks to optimise the prediction behaviour and clinical value in the long term.

In order to implement these promising results on a wider level, it is necessary to unite efforts across disciplines and make sure that the implementation of this technology is associated with the readiness to implement changes in the area of technology, ethical considerations, and regulatory framework compliance. With the continued development of AI, it presents a revolutionary possibility of identifying neurodegenerative diseases earlier, in the intervening of the disease process, and eventually changing patient outcomes in a massive way.

References

[1] Aarsland, D., Creese, B., Politis, M., Chaudhuri, K. R., Ffytche, D. H., Weintraub, D., & Ballard, C. (2017). Cognitive decline in Parkinson's disease. *Nature Reviews Neurology*, 13(4), 217–231.

[2] Arbabshirani, M. R., Plis, S., Sui, J., Calhoun, V. D. (2017). Single subject prediction of brain disorders in neuroimaging: Arbabshirani MR, Plis S, Sui J, Calhoun VD. Single-subject prediction of brain disorders in neuroimaging: Promises and pitfalls. Neuroimage. 2017 Jan 15;145(Pt B):137-165. doi: 10.1016/j.neuroimage.2016.02.079. Epub 2016 Mar 21. PMID: 27012503; PMCID: PMC5031516.maging: Promises and pitfalls. *NeuroImage*, 145, 137–165.

[3] Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R: A practical tutorial. *Evidence-Based Mental Health*, 22(4), 153–160.

[4] Basaia, S., Agosta, F., Wagner, L., Canu, E., Magnani, G., Santangelo, R., ... & Alzheimer's Disease Neuroimaging Initiative. (2019). Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *NeuroImage: Clinical*, *21*, 101645.

[5] Bayestehtashk, A., Asgari, M., Shafran, I., & McNames, J. (2015). Fully automated assessment of Parkinsonian speech using signal processing and machine learning. *Speech Communication*, 66, 37–49.

[6] Bron, E. E., Smits, M., Van Der Flier, W. M., Vrenken, H., Barkhof, F., Scheltens, P., ... & Alzheimer's Disease Neuroimaging Initiative. (2015). Standardised evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: the CADDementia challenge. *NeuroImage*, *111*, 562-579.

[7] Carlyle, B. C., Trombetta, B. A., & Arnold, S. E. (2018). Proteomic Approaches for the Discovery of Biofluid Biomarkers of Neurodegenerative Dementias. *Proteomes*, *6*(3), 32. https://doi.org/10.3390/proteomes6030032

[8] Cookson, M. R. (2019). Proteomics: techniques and applications in neuroscience. *Journal of neurochemistry*, *151*(4), 394–396. https://doi.org/10.1111/jnc.14867

[9] Costafreda, S. G., Dinov, I. D., Tu, Z., Shi, Y., Liu, C. Y., Kloszewska, I., ... & Simmons, A. (2011). Automated hippocampal shape analysis predicts the onset of dementia in mild cognitive impairment. *Neuroimage*, *56*(1), 212-219.

[10] Das, S., Mullick, S. S., & Zelinka, I. (2022). On supervised class-imbalanced learning: An updated perspective and some key challenges. *IEEE Transactions on Artificial Intelligence*, *3*(6), 973-993.

[11] Drotár, P., Mekyska, J., Rektorová, I., Masarová, L., Smékal, Z., & Faundez-Zanuy, M. (2014). Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease. *Artificial Intelligence in Medicine*, 67, 39–46.

[12] Gustavsson A, Svensson M, Jacobi F, Allgulander C, Alonso J, Beghi E, Dodel R, Ekman M, Faravelli C, Fratiglioni L, Gannon B, Jones DH, Jennum P, Jordanova A, Jönsson L, Karampampa K, Knapp M, Kobelt G, Kurth T, Lieb R, Linde M, Ljungcrantz C, Maercker A, Melin B, Moscarelli M, Musayev A, Norwood F, Preisig M, Pugliatti M, Rehm J, Salvador-Carulla L, Schlehofer B, Simon R, Steinhausen HC, Stovner LJ, Vallat JM, Van den Bergh P, van Os J, Vos P, Xu W, Wittchen HU, Jönsson B, Olesen J; CDBE2010Study Group. Cost of disorders of the brain in Europe 2010. Eur Neuropsychopharmacol. 2011 Oct;21(10):718-79. doi: 10.1016/j.euroneuro.2011.08.008. Epub 2011 Sep 15. Erratum in: Eur Neuropsychopharmacol. 2012 Mar;22(3):237-8. den Bergh, Peter Van [corrected to Van den Bergh, Peter]. PMID: 21924589.
[13] Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2021). *Doing Meta-Analysis with R: A Hands-On Guide*. CRC Press.
[14] Jiang, Z., Van Zoest, V., Deng, W., Ngai, E. C., & Liu, J. (2023). Leveraging machine learning for disease diagnoses based on wearable devices: A survey. *IEEE Internet of Things Journal*, *10*(24), 21959-21981.
[15] Jo, T., Nho, K., & Saykin, A. J. (2019). Deep learning in Alzheimer's disease: Diagnostic classification and prognostic prediction using neuroimaging data. *Frontiers in Ageing Neuroscience*, 11, 220.
[16] Lipsmeier, F., Fernandez Garcia, I., Wolf, D., Kilchenmann, T., Scotland, A., Schjodt-Eriksen, J., ... & Lindemann, M. (2017). Successful passive monitoring of early-stage Parkinson's disease patient mobility in Phase I RG7935/PRX002 clinical trial with smartphone sensors. *Mov Disord*, *32*, S358-S359..
[17] Lu, D., Popuri, K., Ding, G. W., Balachandar, R., & Beg, M. F. (2018). Multimodal and multiscale deep neural networks for the early diagnosis of Alzheimer's disease using structural MR and FDG-PET images. Scientific Reports, 8(1), 1-13.
[18] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, *372*.
[19] Paulsen, J. S. (2010). Early Detection of Huntington Disease. *Future neurology*, *5*(1), 10.2217/fnl 09.78. https://doi.org/10.2217/fnl.09.78
[20] Prince, M., Wimo, A., Guerchet, M., Ali, G. C., Wu, Y. T., & Prina, M. (2015). *World Alzheimer Report 2015: The global impact of dementia*. Alzheimer's Disease International.
[21] Rathore, S., Habes, M., Iftikhar, M. A., Shacklett, A., & Davatzikos, C. (2017). A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. *NeuroImage*, 155, 530–548.
[22] Sarraf, S., Danielle D., Anderson J., & Tofighi, G. (2016). DeepAD: Alzheimer's disease classification via deep convolutional neural networks using MRI and fMRI. *arXiv preprint arXiv:1602.02573*.
[23] Schwarzer, G. (2007). meta: An R package for meta-analysis. *R News*, 7(3), 40–45.
[24] Shouval, R., Bondi, O., Mishan, H., Shimoni, A., Unger, R., & Nagler, A. (2014). Application of machine learning algorithms for clinical predictive modelling: a data-mining approach in SCT. *Bone marrow transplantation*, *49*(3), 332-337.
[25] Singh, A., Sengupta, S., & Lakshminarayanan, V. (2020). Explainable deep learning models in medical image analysis. *Journal of Imaging*, *6*(6), 52.
[26] Suresh, H., & Guttag, J. V. (2019). A framework for understanding unintended consequences of machine learning. *Communications of the ACM*, 63(11), 62–71.
[27] The Guardian. (2024, June 18). AI-enhanced blood test may detect Parkinson's years before onset. https://www.theguardian.com
[28] Topol, E. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44–56.
[29] Verde, L., De Pietro, G., & Sannino, G. (2018). Voice disorder identification by using machine learning techniques. *IEEE Access*, 6, 16246–16255.
[30] Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1–48.