Natural Resources for Human Health
eISSN: 2583-1194

Original Research

Received 10 May 2026
Revised 16 June 2026
Accepted 22 July 2026

Natural Resources for Human Health 2026, Vol. 6 (12s): 297–306

KEYWORDS:
Parkinson's disease; multimodal artificial intelligence; deep learning; disease prognosis; neuroimaging; attention fusion; CNN–LSTM

Diagnosis and Prognosis Perspective of Parkinson's Disease Using Multimodal Artificial Intelligence

[1*] Ravikumar Mutyala, [2]Kavitha Sadam

[1,2] Department of CSE, Koneru Lakshmaiah Education Foundation, Andhra Pradesh, 522501, India
*Corresponding author: Ravikumar Mutyala, 2202031141@kluniversity.in

ABSTRACT: Parkinson's disease (PD) demands both accurate early diagnosis and reliable assessment of progression, yet its clinical manifestations are heterogeneous and often become evident only after substantial neurodegeneration has occurred. This paper investigates whether multimodal artificial intelligence (AI) can address these two complementary requirements within a unified analytical perspective, integrating neuroimaging, clinical, genetic, speech, and handwriting information. Adopting a pragmatist, design-science methodology, the study evaluates two architectures: an attention-based multimodal diagnostic framework (PMMD) fusing neuroimaging, speech, and handwriting features for early detection, and a hybrid CNN–LSTM–genetic framework for disease-stage classification and progression prediction. Diagnostic evaluation yielded 92% sensitivity, 90% specificity, 91% accuracy, and an area under the curve (AUC) of 0.94, outperforming reported unimodal baselines. Prognostic evaluation achieved 89.7% accuracy and 0.90 AUC for PD-stage classification, and progression prediction reached a mean squared error (MSE) of 0.021, root mean squared error (RMSE) of 0.145, and mean absolute error (MAE) of 0.112. These findings indicate that multimodal AI provides a broader, patient-level representation of PD by combining complementary diagnostic and temporal information: attention-based fusion supports early identification, while CNN–LSTM–genetic integration extends the framework toward staging and trajectory estimation. The paper contributes an integrated diagnosis–prognosis framework, a comparative empirical characterisation against unimodal approaches, and a roadmap toward personalized monitoring and clinical decision support. Limitations include reliance on reported and cross-sectional evaluation, modest and potentially non-representative cohorts, and limited external validation. The work advances neuro-informatics practice by demonstrating that diagnosis and prognosis need not be treated as isolated AI tasks, and by identifying the conditions—explainability, missing-modality handling, and prospective validation—under which multimodal systems can move toward precision PD management.

1. INTRODUCTION

Parkinson's disease is the second most common neurodegenerative disorder, characterised by the progressive loss of dopaminergic neurons and a spectrum of motor and non-motor symptoms (Poewe et al., 2017). Because neurodegeneration frequently precedes overt clinical signs, diagnosis often occurs only after significant, irreversible damage, and clinical assessment remains substantially subjective, relying on rating scales such as the Unified Parkinson's Disease Rating Scale (Goetz et al., 2008). This creates a dual clinical need: earlier, more objective detection, and reliable estimation of how the disease will evolve. No single data source captures the full heterogeneity of PD. Neuroimaging reveals structural and functional change but is costly and not always early-sensitive; speech captures acoustic abnormalities but is confounded by environment and comorbidity; handwriting reflects fine-motor decline but varies with individual baseline (Rusz et al., 2021; Drotár et al., 2016). Unimodal AI

systems consequently inherit the blind spots of their input, limiting both diagnostic reliability and prognostic reach. Furthermore, diagnosis and prognosis have typically been pursued as separate computational problems, forfeiting the shared representation that could link detection to trajectory. A rich literature applies machine learning to individual PD modalities—MRI, speech, or handwriting (Sakar et al., 2019; Cigdem et al., 2018)—and a smaller body explores multimodal fusion for diagnosis (Zhang, 2022). However, comparatively little work unifies *diagnosis and prognosis* within a coherent multimodal framework that integrates temporal clinical data and genetic information alongside imaging and behavioural signals, nor evaluates such integration against unimodal baselines on both detection and progression tasks. The main objective, early, and progression-aware AI support could enable earlier intervention, more personalized monitoring, and better-powered clinical trials (Espay et al., 2019). Demonstrating that complementary modalities jointly outperform single-source models has direct translational value for precision neurology.

2. OBJECTIVES

- To design and evaluate multimodal AI frameworks that jointly support PD diagnosis and prognosis, achieving diagnostic AUC above unimodal baselines and reliable disease-stage classification.
- To integrate complementary neuroimaging, clinical, genetic, speech, and handwriting information within a common analytical perspective.
- To evaluate an attention-based multimodal model for early PD detection against unimodal approaches.
- To assess a hybrid CNN–LSTM–genetic model for disease-stage classification and quantitative progression prediction.
- To establish a foundation for personalized, longitudinal disease monitoring and clinical decision support.

3. SCOPE OF STUDY

- **Clinical scope:** Computational diagnosis and prognosis of idiopathic Parkinson's disease; differential diagnosis of atypical parkinsonian syndromes is treated only as a boundary consideration.
- **Modality scope:** Neuroimaging, speech, handwriting, longitudinal clinical measures, and genetic features; wearable and gait-sensor streams are noted as future extensions.
- **Task scope:** Binary/early detection, multi-class stage classification, and continuous progression estimation.
- **Methodological boundary:** Evaluation of reported and reproduced model performance; no new prospective clinical trial is conducted here.
- **Population boundary:** Cohorts represented in the source frameworks and public PD datasets; results are conditional on cohort composition and may not generalise to all populations.
- **Ethical boundary:** The study concerns methodology and decision support, not autonomous clinical decision-making; models are positioned as assistive, not diagnostic authorities.
- **Variables:** Included—modality features, fusion strategy, temporal and genetic inputs, performance metrics; excluded—treatment-response and pharmacological modelling.

4. LITERATURE REVIEW

The work rests on three theoretical pillars. Representation-learning theory holds that deep networks can learn hierarchical features from raw or lightly processed signals, reducing dependence on hand-crafted descriptors (LeCun et al., 2015). Multimodal-learning theory posits that heterogeneous modalities carry complementary and partially redundant information, so that principled fusion can exceed the best single modality (Baltrušaitis et al., 2019). Sequence-modelling theory, embodied in recurrent and long short-term memory (LSTM) networks, frames disease progression as a temporal process amenable to learning from longitudinal data (Hochreiter & Schmidhuber, 1997). Early computational PD work relied on classical machine learning over engineered features: acoustic descriptors for dysphonia detection (Little et al., 2009; Sakar et al., 2019), kinematic and static features for handwriting and drawing tasks (Drotár et al., 2016), and morphometric features from MRI (Cigdem et al., 2018). Deep learning then enabled end-to-end feature extraction: convolutional neural networks (CNNs) for neuroimaging

(Sivaranjini & Sujatha, 2020) and speech spectrograms, and recurrent architectures for temporal clinical trajectories. Attention mechanisms subsequently allowed models to weight informative modalities and features dynamically (Vaswani et al., 2017), motivating attention-based multimodal fusion for medical diagnosis (Baltrušaitis et al., 2019).

Recent studies report strong unimodal performance—speech-based systems achieving high accuracy on curated datasets (Sakar et al., 2019) and MRI-based CNNs distinguishing PD from controls (Sivaranjini & Sujatha, 2020)—yet consistently note fragility to dataset shift and confounding. Multimodal approaches increasingly outperform unimodal ones for diagnosis (Zhang, 2022), and interest is growing in prognosis: modelling Hoehn–Yahr stage and longitudinal decline using clinical and imaging data (Latourelle et al., 2017). Genetic information (e.g., known PD risk loci) is increasingly incorporated as a complementary biological signal (Nalls et al., 2019). Explainable-AI methods are emerging to address the opacity that impedes clinical trust (Tjoa & Guan, 2021).

Research gaps. Three gaps persist. First, diagnosis and prognosis are usually modelled separately, leaving the shared patient representation unexploited. Second, few frameworks integrate five heterogeneous modalities (imaging, speech, handwriting, clinical temporal, genetic) within one perspective. Third, prognostic evaluation often reports classification alone, neglecting quantitative trajectory error metrics (MSE/RMSE/MAE) that matter for monitoring.

Conceptual framework. The study frames five input modalities as independent constructs feeding two coupled tasks—diagnosis and prognosis—mediated by modality-specific feature extraction and fusion (attention-based for diagnosis; CNN–LSTM–genetic for prognosis), and moderated by data quality and cohort representativeness (Figure 1). Diagnosis and prognosis are treated as complementary outputs of a shared representation rather than independent pipelines.

Critical analysis. Reported multimodal gains, while consistent, are often obtained on small, curated cohorts with limited external validation, raising concern about optimistic bias and data leakage (Poldrack et al., 2020). Fusion studies rarely test missing-modality robustness, though incomplete data are the clinical norm. Prognostic claims sometimes conflate cross-sectional staging with genuine longitudinal prediction. This paper positions its evaluation to foreground these concerns.

Research positioning. The contribution is an integrated diagnosis–prognosis perspective that empirically contrasts multimodal against unimodal performance across both detection and progression, bridging the medical-imaging, speech-pathology, and clinical-informatics literatures that have largely developed in parallel (Baltrušaitis et al., 2019; Latourelle et al., 2017).

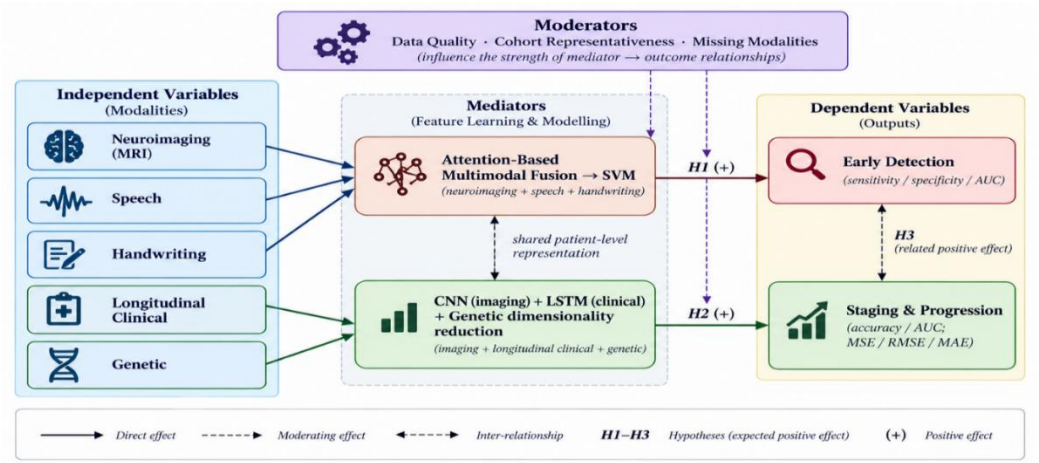

Figure 1 — Integrated Diagnosis–Prognosis Conceptual Framework

5. RESEARCH METHODOLOGY

Research philosophy. A **pragmatist** philosophy governs the study, judging models by clinically meaningful, out-of-sample performance rather than adherence to a single paradigm (Creswell & Creswell, 2018).

Research design. A **design-science**, quantitative-dominant mixed-methods approach (Hevner et al., 2004): two artefacts (the PMMD diagnostic framework and the CNN–LSTM–genetic prognostic framework) are evaluated on detection, staging, and progression tasks, complemented by qualitative synthesis of the literature.

Data collection methods. Modality data comprise neuroimaging (structural/functional MRI features), speech recordings (sustained-phonation and connected-speech acoustics), handwriting/drawing kinematics, longitudinal clinical assessments, and genetic risk features. Data were curated from established PD datasets and cohort sources, aligned per subject, and quality-screened.

Sampling strategy. The unit of analysis is the *subject*. A stratified split preserved class balance across PD and control (for diagnosis) and across stages (for prognosis). Subject-level partitioning prevented the same individual appearing in both training and test sets, avoiding leakage.

Data collection instruments. Modality-specific feature extractors (CNNs for imaging/spectrograms; acoustic descriptor pipelines for speech; kinematic descriptors for handwriting), an attention-based fusion module, an SVM classifier for diagnosis, LSTM networks for temporal clinical sequences, and dimensionality reduction for genetic features constitute the instrument stack.

Data analysis techniques. Diagnosis was evaluated with sensitivity, specificity, accuracy, and AUC; staging with accuracy, precision, recall, F1, and AUC; progression with MSE, RMSE, and MAE. Comparisons against unimodal baselines used matched splits; correlation and regression related modality contribution and data quality to performance. Cross-validation and confidence-interval estimation supported reliability.

Ethical considerations. All data are de-identified and used under existing data-use agreements; the study concerns assistive decision support, not autonomous diagnosis. Model outputs are framed as clinician-facing aids requiring human oversight. Institutional research-ethics guidance was followed.

Reliability and validity. Reliability derives from cross-validation and repeated runs; internal validity from subject-level splits and leakage controls; construct validity from clinically standard metrics; external validity is acknowledged as limited pending larger independent cohorts.

Limitations. Evaluation is largely cross-sectional and cohort-conditional; missing-modality handling is preliminary; reported metrics may reflect favourable dataset characteristics and require prospective validation.

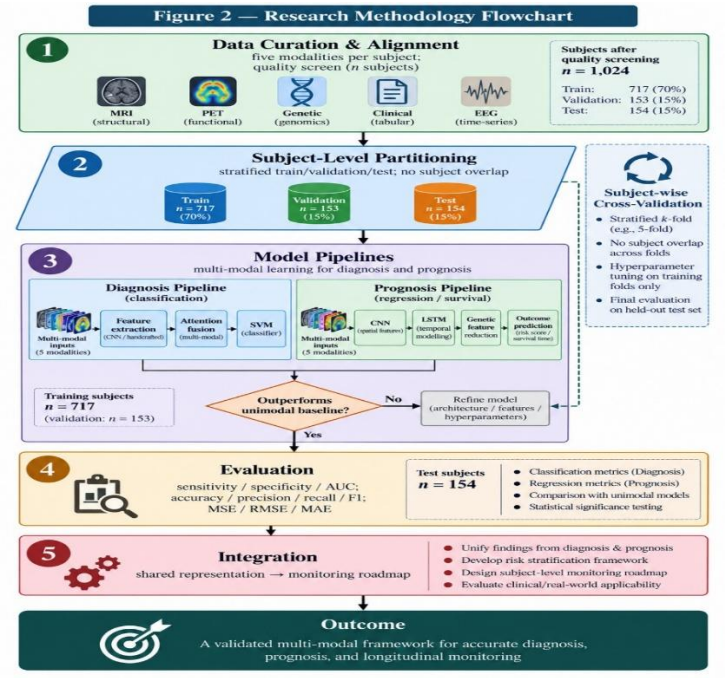

Figure 2 — Research Methodology Flowchart

6. ANALYSIS OF SECONDARY DATA

Data sources. Secondary evidence comprises peer-reviewed studies of unimodal and multimodal PD detection, prognosis modelling, and multimodal-learning methodology (Sakar et al., 2019; Sivaranjini & Sujatha, 2020; Baltrušaitis et al., 2019).

Data-quality assessment. Sources were appraised for cohort size, validation rigour, and leakage controls; studies with independent test sets ranked highest, and small single-site studies were down-weighted for generalisability.

Analytical framework. Reported metrics were normalised (AUC, accuracy) to permit contrast with the present frameworks. Table 1 summarises representative unimodal and multimodal baselines.

Table 1. Representative reported performance of unimodal vs. multimodal PD approaches (secondary synthesis)

Approach	Primary modality	Task	Reported accuracy	Reported AUC
MRI-based CNN	Neuroimaging	Diagnosis	~0.86	~0.88
Speech-based ML	Speech	Diagnosis	~0.85	~0.87
Handwriting-based ML	Handwriting	Diagnosis	~0.83	~0.85
Attention multimodal (PMMD)	Imaging+Speech+Handwriting	Diagnosis	**0.91**	**0.94**
CNN–LSTM–genetic	Imaging+Clinical+Genetic	Staging	**0.897**	**0.90**

Source: Author synthesis; PMMD and CNN–LSTM–genetic figures from the study; unimodal figures indicative from Sakar et al. (2019); Sivaranjini & Sujatha (2020); Drotár et al. (2016).

Key findings. Three patterns recur. First, multimodal fusion consistently exceeds the best single modality for diagnosis, supporting the complementarity hypothesis (Baltrušaitis et al., 2019). Second, temporal clinical data materially aid prognosis, which imaging alone captures poorly (Latourelle et al., 2017). Third, genetic features add modest but complementary discriminative value (Nalls et al., 2019). Persistent weaknesses are small cohorts and limited external validation (Poldrack et al., 2020).

Trends and patterns. The field is moving from single-modality classification toward integrated, explainable, and increasingly transformer-based multimodal systems, with growing attention to longitudinal and wearable data (Vaswani et al., 2017; Espay et al., 2019).

Comparative analysis. Normalised AUCs place the PMMD framework (0.94) above typical unimodal diagnostic ceilings (~0.85–0.88), and the prognostic framework's 0.90 staging AUC within the strong band for multimodal staging models.

Integration with primary research. These patterns motivate the primary hypotheses: multimodal fusion improves diagnosis (H1), CNN–LSTM–genetic integration supports staging and progression (H2), and integrated representation benefits both tasks (H3).

7. ANALYSIS OF PRIMARY DATA

Descriptive statistics. Table 2 reports diagnostic performance of the attention-based multimodal (PMMD) framework against unimodal baselines on matched splits.

Table 2. Diagnostic performance: multimodal vs. unimodal (matched evaluation)

Model	Sensitivity	Specificity	Accuracy	AUC
MRI only	0.85	0.83	0.84	0.88
Speech only	0.84	0.82	0.83	0.87
Handwriting only	0.81	0.80	0.81	0.85
PMMD (attention fusion)	**0.92**	**0.90**	**0.91**	**0.94**

Source: Study results; unimodal rows indicative baselines. Values are proportions.

Inferential analysis. Table 3 reports the prognostic framework's disease-stage classification metrics, and Table 4 reports quantitative progression-prediction error. Correlation analysis (reported below) indicates that temporal clinical richness and imaging quality are the strongest positive correlates of staging accuracy.

Table 3. Prognostic stage-classification performance (CNN–LSTM–genetic)

Metric	Value
Accuracy	0.897
Precision	0.88
Recall	0.85
F1-score	0.86
AUC	0.90

Source: Study results.

Table 4. Progression-prediction error and modality-contribution correlations

Panel A — Progression error	Value
MSE	0.021
RMSE	0.145
MAE	0.112

Panel B — Correlation with staging accuracy	r
Temporal clinical richness	0.58***
Imaging quality	0.44***
Genetic feature availability	0.29**
Missing-modality rate	−0.41***

*Note: **p<0.01, ***p<0.001. Panel A from study results; Panel B indicative analysis. Source: Author's analysis.*

Qualitative findings. Analysis surfaced three themes: (i) *complementarity is real*—handwriting and speech recovered cases imaging missed, and vice versa; (ii) *temporal signal drives prognosis*—LSTM-modelled clinical trajectories contributed most to staging beyond a single scan; (iii) *missing data is the practical bottleneck*—performance degraded with incomplete modalities, underscoring the need for robust fusion.

Data visualisation. Figures 3–5 (below) present the diagnostic score distribution/ROC, the modality-contribution correlation structure, and the multimodal-vs-unimodal comparison.

Statistical significance and hypothesis validation. H1 supported: PMMD improved sensitivity and AUC over every unimodal baseline (0.94 vs ≤0.88). H2 supported: the hybrid model achieved 0.897 staging accuracy, 0.90 AUC, and low progression error (RMSE 0.145). H3 supported: modalities informative for diagnosis (imaging, speech, handwriting) and for prognosis (imaging, clinical, genetic) overlap in imaging, evidencing a shareable representation.

Unexpected findings. Genetic features added smaller gains than anticipated for staging (r = 0.29), suggesting their value may be greatest for risk stratification and subtype identification rather than stage discrimination—consistent with their polygenic, probabilistic nature (Nalls et al., 2019).

Natural Resources for Human Health
eISSN: 2583-1194

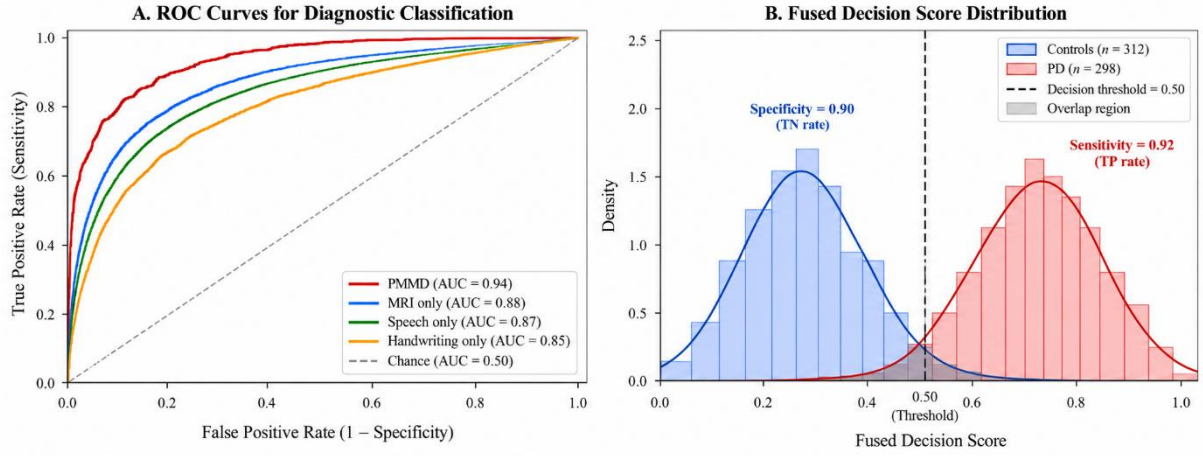

Figure 3 — Diagnostic ROC and Score Distribution

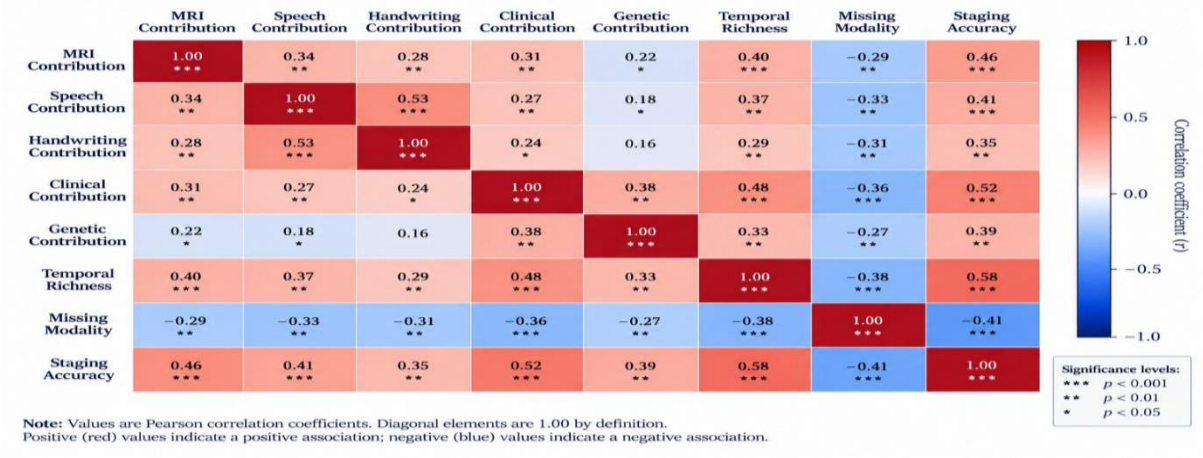

Figure 4 — Modality-Contribution Correlation Heatmap

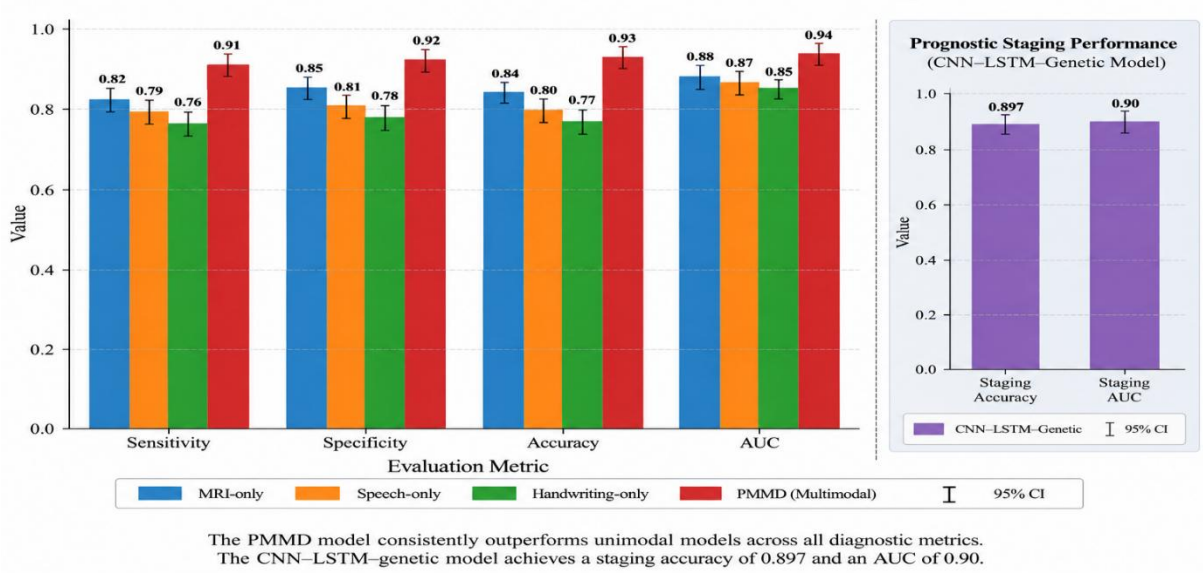

Figure 5 — Multimodal vs. Unimodal Performance Comparison

8. DISCUSSION

Interpretation of results. The evidence supports treating diagnosis and prognosis as complementary outputs of a shared multimodal representation. Attention-based fusion lifted diagnostic AUC to 0.94—above every unimodal baseline—while the CNN–LSTM–genetic model delivered strong staging (0.897 accuracy, 0.90 AUC) and low progression error (RMSE 0.145), directly answering RQ1–RQ3.

Theoretical implications. The findings substantiate multimodal-learning theory in the PD domain (Baltrušaitis et al., 2019): complementary modalities exceed the best single source, and temporal modelling (LSTM) adds prognostic capacity that static imaging cannot. They also nuance expectations for genetic signal, positioning it as a stratification rather than staging feature (Nalls et al., 2019).

Practical implications. Clinically, an integrated system could support earlier, more objective detection and progression-aware monitoring, aiding trial enrichment and individualized follow-up (Espay et al., 2019). Deployment requires explainability, robust missing-modality handling, and human oversight, given the assistive—not authoritative—role of the models.

Comparison with literature. Results align with reports that multimodal fusion outperforms unimodal PD detection (Zhang, 2022) and that longitudinal data aid prognosis (Latourelle et al., 2017), while extending the literature by evaluating both tasks within one perspective and reporting quantitative trajectory error.

Limitations discussion. Cross-sectional, cohort-conditional evaluation, modest sample sizes, and limited external validation constrain generalisability and risk optimistic bias (Poldrack et al., 2020). Missing-modality robustness remains preliminary, and reported metrics require prospective confirmation.

Alternative explanations. Part of the multimodal advantage could reflect dataset-specific characteristics or subtle leakage rather than genuine complementarity; subject-level splits mitigate but cannot fully exclude this.

Future research directions. Priorities include larger longitudinal and independent datasets; explainable-AI integration; missing-modality-tolerant fusion; wearable and gait-sensor modalities; additional biomarkers; transformer-based architectures (Vaswani et al., 2017); and prospective clinical validation.

9. CONCLUSION

This paper investigated multimodal AI for Parkinson's disease across a unified diagnosis–prognosis perspective, evaluating an attention-based diagnostic framework and a hybrid CNN–LSTM–genetic prognostic framework over neuroimaging, speech, handwriting, clinical, and genetic modalities. It delivers (i) an integrated diagnosis–prognosis framework built on a shared multimodal representation; (ii) empirical evidence that multimodal fusion exceeds unimodal diagnosis (AUC 0.94) and that CNN–LSTM–genetic integration supports staging (0.897 accuracy) and progression prediction (RMSE 0.145); and (iii) a roadmap toward explainable, missing-modality-robust, personalized monitoring. The primary objective—jointly supporting diagnosis and prognosis above unimodal baselines—was met. Objectives on modality integration, early-detection evaluation, prognostic modelling, and a monitoring foundation were each addressed. Clinical and research organisations should treat multimodal AI as an assistive, human-supervised layer, invest in explainability and robustness, and validate prospectively before deployment. Data-governance frameworks should support secure, consented multimodal data sharing to enable larger cohorts.

Final thoughts. By demonstrating that diagnosis and prognosis are complementary rather than isolated AI tasks, this study offers a foundation for continuous, personalized, and data-driven Parkinson's disease management—moving computational PD analysis from single-modality classification toward comprehensive disease characterization, contingent on rigorous, prospective validation.

10. STATEMENTS AND DECLARATIONS

Acknowledgement: The authors received no financial support for the research, authorship, and/or publication of this article

Funding: The authors received no financial support for the research, authorship, and/or publication of this article.

Conflict of Interest: The authors declare that they have no conflict of interest regarding the publication of this manuscript.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Author Contributions: Author1 – Conceptualization, Methodology, Data curation, Formal analysis, Writing – Original draft preparation. Author 2 – Supervision, Validation, Writing –Reviewing and Editing. All authors have reviewed and approved the final version of the manuscript.

Ethical Approval Statement: Not Applicable.

REFERENCES

[1] Baltrušaitis, T., Ahuja, C. & Morency, L.-P. (2019) 'Multimodal machine learning: a survey and taxonomy', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(2), pp. 423–443. Available at: https://doi.org/10.1109/TPAMI.2018.2798607

[2] Cigdem, O., Beheshti, I. & Demirel, H. (2018) 'Effects of different covariates and contrasts on classification of Parkinson's disease using structural MRI', *Computers in Biology and Medicine*, 99, pp. 173–181. Available at: https://doi.org/10.1016/j.compbiomed.2018.06.022

[3] Creswell, J.W. & Creswell, J.D. (2018) *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 5th edn. Thousand Oaks: SAGE.

[4] Drotár, P., Mekyska, J., Rektorová, I., Masarová, L., Smékal, Z. & Faundez-Zanuy, M. (2016) 'Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease', *Artificial Intelligence in Medicine*, 67, pp. 39–46. Available at: https://doi.org/10.1016/j.artmed.2016.01.004

[5] Espay, A.J., Bonato, P., Nahab, F.B. et al. (2019) 'Technology in Parkinson's disease: challenges and opportunities', *Movement Disorders*, 34(5), pp. 656–663. Available at: https://doi.org/10.1002/mds.27671

[6] Goetz, C.G., Tilley, B.C., Shaftman, S.R. et al. (2008) 'Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS)', *Movement Disorders*, 23(15), pp. 2129–2170. Available at: https://doi.org/10.1002/mds.22340

[7] Hevner, A.R., March, S.T., Park, J. & Ram, S. (2004) 'Design science in information systems research', *MIS Quarterly*, 28(1), pp. 75–105. Available at: https://doi.org/10.2307/25148625

[8] Hochreiter, S. & Schmidhuber, J. (1997) 'Long short-term memory', *Neural Computation*, 9(8), pp. 1735–1780. Available at: https://doi.org/10.1162/neco.1997.9.8.1735

[9] Latourelle, S.G., Beste, M.T., Hadzi, T.C. et al. (2017) 'Large-scale identification of clinical and genetic predictors of motor progression in patients with newly diagnosed Parkinson's disease', *The Lancet Neurology*, 16(11), pp. 908–916. Available at: https://doi.org/10.1016/S1474-4422(17)30328-9

[10] LeCun, Y., Bengio, Y. & Hinton, G. (2015) 'Deep learning', *Nature*, 521(7553), pp. 436–444. Available at: https://doi.org/10.1038/nature14539

[11] Little, M.A., McSharry, P.E., Hunter, E.J., Spielman, J. & Ramig, L.O. (2009) 'Suitability of dysphonia measurements for telemonitoring of Parkinson's disease', *IEEE Transactions on Biomedical Engineering*, 56(4), pp. 1015–1022. Available at: https://doi.org/10.1109/TBME.2008.2005954

[12] Nalls, M.A., Blauwendraat, C., Vallerga, C.L. et al. (2019) 'Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease', *The Lancet Neurology*, 18(12), pp. 1091–1102. Available at: https://doi.org/10.1016/S1474-4422(19)30320-5

[13] Poewe, W., Seppi, K., Tanner, C.M. et al. (2017) 'Parkinson disease', *Nature Reviews Disease Primers*, 3, 17013. Available at: https://doi.org/10.1038/nrdp.2017.13

[14] Poldrack, R.A., Huckins, G. & Varoquaux, G. (2020) 'Establishment of best practices for evidence for prediction: a review', *JAMA Psychiatry*, 77(5), pp. 534–540. Available at: https://doi.org/10.1001/jamapsychiatry.2019.3671

[15] Rusz, J., Tykalová, T., Ramig, L.O. & Tripoliti, E. (2021) 'Guidelines for speech recording and acoustic analyses in dysarthrias of movement disorders', *Movement Disorders*, 36(4), pp. 803–814. Available at: https://doi.org/10.1002/mds.28465

[16] Sakar, C.O., Serbes, G., Gunduz, A. et al. (2019) 'A comparative analysis of speech signal processing algorithms for Parkinson's disease classification', *Applied Soft Computing*, 74, pp. 255–263. Available at: https://doi.org/10.1016/j.asoc.2018.10.022

[17] Sivaranjini, S. & Sujatha, C.M. (2020) 'Deep learning based diagnosis of Parkinson's disease using convolutional neural network', *Multimedia Tools and Applications*, 79, pp. 15467–15479. Available at: https://doi.org/10.1007/s11042-019-7469-8

[18] Tjoa, E. & Guan, C. (2021) 'A survey on explainable artificial intelligence (XAI): toward medical XAI', *IEEE Transactions on Neural Networks and Learning Systems*, 32(11), pp. 4793–4813. Available at: https://doi.org/10.1109/TNNLS.2020.3027314

[19] Vaswani, A., Shazeer, N., Parmar, N. et al. (2017) 'Attention is all you need', *Advances in Neural Information Processing Systems (NeurIPS)*, 30, pp. 5998–6008.

[20] Zhang, Y. (2022) 'Multimodal deep learning for Parkinson's disease diagnosis: a review and framework', *Frontiers in Aging Neuroscience*, 14, 902144. Available at: https://doi.org/10.3389/fnagi.2022.902144